ISSN (edición electrónica) 2007-4719

ACTA DE INVESTIGACIÓN PSICOLÓGICA

PSYCHOLOGICAL RESEARCH RECORDS

Volumen 4, Número 3, Diciembre 2014.

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE PSICOLOGÍA

Acta de Investigación Psicológica

Acta de Investigación Psicológica, Año 4, No. 12, septiembre-diciembre 2014, es una publicación cuatrimestral editada por la Universidad Nacional Autónoma de México, Cd. Universitaria, Coyoacán, C.P. 04510, México, D.F., a través de la Facultad de Psicología, Av. Universidad 3004, Col. Copilco–Universidad, Del. Coyoacán, CP. 04510, México, D.F., Tel/Fax. (55)56222305 y (55)56222326, http://www.psicologia.unam.mx/pagina/es/155/acta-de-investigacion-psicologica, actapsicologicaunam@gmail.com, Editor responsable: Dr. Rolando Díaz Loving, Reserva de derechos al uso exclusivo N° 04-2011-040411025500-203, ISSN 2007-4719, Responsable de la última actualización de este número: Unidad de Planeación, Facultad de Psicología, Lic. Augusto A. García Rubio Granados, Av. Universidad 3004, Col. Copilco–Universidad, Del. Coyoacán, C.P. 04510, México, D.F., fecha de última modificación, 3 de noviembre de 2014.

Sistema de índices y resúmenes: AIP se encuentra en Latindex, CLASE y ScIELO
Abstracting and Indexing: PRR is abstracted or indexed in Latindex, CLASE and ScIELO

Página dejada intencionalmente en blanco

Índice
Index

Diciembre 2014
December 2014

Volumen 4
Volume 4

Número 3
Issue 3

Índice
Index

Diciembre 2014
December 2014

Volumen 4
Volume 4

Número 3
Issue 3

Prólogo

Acta de Investigación Psicológica (AIP) es una revista ecléctica cuya misión es difundir trabajos de investigación psicológica original. El número especial inaugural de AIP, publicado en abril de 2011, consistió de una colección de trabajos sobre Análisis Experimental de la Conducta. Carlos Bruner y yo fuimos honrados con la invitación para ensamblar dicha colección con la ayuda de contribuyentes internacionales. AIP me volvió a honrar al invitarme a compilar el presente número especial, que está dedicado al Análisis Conductual Aplicado (ACA). Dado que nuestro número especial previo se enfocó en investigación básica, esta colección de trabajos sobre ACA puede considerarse como un complemento a ese primer número especial. El principal propósito del presente número es introducir el ACA a los lectores de la AIP, que consiste de un grupo heterogéneo de psicólogos y estudiantes. Espero que la diseminación del conocimiento y de la filosofía del ACA atraiga a los lectores hacia este enfoque. Para lograr mi meta, le pedí al Dr. Kennon A. Lattal que me recomendara a analistas conductuales aplicados prominentes cuyo trabajo cubriera una gran variedad de tópicos de aplicación con niños. Desde luego, le di preferencia a la originalidad, a la investigación empírica, a la facilidad de lectura y a trabajos que tuvieran un atractivo intuitivo. Decidí dedicar este número a aplicaciones enfocadas a la conducta de los niños debido a que este grupo promete la mejor inversión de tiempo y esfuerzo, dado que muchas de las conductas problema predicen ya sea un empeoramiento de los mismos problemas o el desarrollo de conductas problemáticas diferentes y más severas.

La Facultad de Psicología de la Universidad Nacional Autónoma de México publica AIP no sólo para el beneficio de su comunidad académica, sino para el beneficio de otros psicólogos y estudiantes en todo el país. Dada esta población heterogénea, parece conveniente mencionar algunas características del ACA descritas por Baer, Wolf, y Risley (1968). El ACA consiste en la aplicación de diferentes procedimientos para resolver una variedad de problemas conductuales de interés humano inmediato. Algunos de esos procedimientos se derivan de investigación básica en el laboratorio ya sea con animales o con humanos. No obstante, las variables independientes del ACA que prometen efectividad no se limitan a las generadas por la investigación básica, sino que pueden derivarse de múltiples fuentes, incluyendo la observación casual de la conducta en ambientes sociales y investigación anterior en la se mostró la efectividad de intervenciones similares o relacionadas. Una característica prominente del ACA es que es un enfoque pragmático a la solución de problemas; es decir, cualquier variable independiente que pruebe su efectividad para cambiar la conducta será objeto de mayor estudio. El utilizar la expresión de “intervenciones basadas en los resultados” es otra forma de referirse al valor que el ACA le confiere a los métodos probadamente efectivos que se usan en la intervención conductual. Otra característica del ACA es la evaluación de la generalidad de una variable independiente efectiva. Es decir, el ACA busca responder preguntas como si el mismo procedimiento es efectivo con diferentes individuos y en diferentes situaciones sociales. Preguntas relacionadas con las anteriores son la

permanencia del cambio conductual en el tiempo y la inducción del cambio a otras instancias de conducta deseable. La evaluación de la generalidad de la variable independiente incluye la descripción de las condiciones suficientes y necesarias para replicar el cambio conductual. Otra característica más del ACA es el intento para integrar y sistematizar las aplicaciones exitosas de sus variables independientes a un cuerpo de conocimientos establecido. Un punto importante mencionado por Baer et al. es que la diferencia entre investigación básica y aplicada no se refiere a cuál de las dos descubre principios básicos y cuál los aplica, dado ambos tipos de investigación buscan averiguar los parámetros que controlan la conducta. Si acaso, la única diferencia entre investigación básica y aplicada es que esta última se lleva a cabo en ambientes sociales.

Para finalizar, me gustaría mencionar que existen múltiples medios para difundir el conocimiento del ACA, principalmente en la forma de revistas. Entre estas últimas destacan el *Journal of Applied Behavior Analysis* y en nuestro país la *Revista Mexicana de Análisis de la Conducta* (fundadas en 1968 y en 1975, respectivamente). El hecho de que existan fuentes abundantes de información del ACA no hace que la presente colección de artículos sea redundante, dado que las fuentes establecidas son especializadas y son principalmente consultadas por lectores sofisticados que ya trabajan bajo este enfoque. El presente número especial, publicado en una revista ecléctica, permitirá la diseminación del conocimiento de ACA entre estudiantes y psicólogos no especialistas, que aún no están familiarizados con el éxito de nuestro campo.

Referencias

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97. doi: 10.1901/jaba.1968.1-91

Bruner, C. A., & Acuña, L. (2011) (Eds.). Número especial en análisis de la conducta. *Acta de Investigación Psicológica, 1*, 1-179.

Laura Acuña
Universidad Nacional Autónoma de México
Editor Invitado

Preface

Acta de Investigación Psicológica (*Psychological Research Records*, PRR; in English) is an eclectic journal with the mission to disseminate original psychological research. Its inaugural special issue, printed in April 2011, consisted of a collection of papers on the Experimental Analysis of Behavior. Carlos Bruner and I were honored with the invitation to assemble such collection with the help of international contributors. I was again honored when PRR invited me to compile the present special issue devoted to Applied Behavior Analysis (ABA). Given that our previous special issue focused on basic research, this collection on ABA can be seen as a complement to our preceding special issue. The main purpose of the present collection is to introduce ABA to the readers of PRR, who consist of an heterogeneous group of psychologists and students. I hope that the dissemination of ABA knowledge and philosophy may attract readers to this enterprise. To accomplish my goal I asked Dr. Kennon A. Lattal to recommend prominent applied behavior analysts whose work cover a wide range of topics of application with children. Of course, I preferred originality, empirical research, easy reading and intuitive appeal. I also decided to concentrate on applications to the behavior of children because this group promises the best return of the time and effort invested, since many of their problem behaviors predict either a worsening of the same problems or the development of different and more severe behavior problems.

The School of Psychology of the National Autonomous University of Mexico publishes PRR not only for the benefit of its academic community but also for the benefit of other psychologists and students across the country. Given this heterogeneous population, it seems convenient to sketch some characteristics of ABA, as described by Baer, Wolf, and Risley (1968). ABA consists in the application of different procedures that promise to solve a variety of behavioral problems of immediate human interest. Some of these procedures are derived from basic research in the laboratory with either animal or human subjects. However, the promising independent variables of ABA are not limited to the later but instead may be suggested from multiple sources, including the casual observation of behavior in social settings and even in previous research on the outcome of similar or related interventions. A prominent characteristic of ABA is that it is a pragmatic approach to problem solving; that is to say that whatever independent variable that proves effective for behavior change merits further study. The qualification of “outcome-based interventions” is another way of referring to the value conferred by ABA to proven methods of behavior intervention. In addition to the former, another characteristic of ABA is the assessment of the generality of the effective independent variable. That is to say it seeks to answer questions concerning whether the same procedure is effective across different individuals and different social situations. Related questions concern the permanence of behavior change and even the induction of other instances of desirable behavior. The assessment of independent-variable generality includes the description of the necessary and sufficient conditions to replicate behavioral change. Still another

characteristic of ABA is an attempt to integrate and systematize the successful application of its independent variables within an established body of knowledge. An important point raised by Baer et al. is that the difference between basic and applied behavior analysis does not refer to which of the two discovers and which one applies the scientific principles, given that both types of research aim at discovering the variables that control behavior. If at all, the only difference between basic and applied research is that the later occurs in the social environment.

Finally, I would like to mention that there are multiple outlets of ABA knowledge mainly in the form of journals, notably the *Journal of Applied Behavior Analysis* and in our country the *Mexican Journal of Behavior Analysis* (founded in 1968 and in 1975, respectively). The fact that there are abundant sources of information on ABA does not make the present collection of articles redundant, the reason being that the established sources are consulted mainly by sophisticated readers already working within this framework. The current special issue, published in an eclectic journal allows the dissemination of behavior analytic knowledge among non-specialized readers, many of whom may be students that are not yet familiar with the fruitfulness of our field.

References

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97. doi: 10.1901/jaba.1968.1-91

Bruner, C. A., & Acuña, L. (2011) (Eds.). Número especial en análisis de la conducta. *Acta de Investigación Psicológica, 1*, 1-179.

Laura Acuña
National Autonomous University of Mexico
Invited Editor

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1621 - 1641

The Effects of the Establishment of Adult Faces and/or Voices as Conditioned Reinforcers for Children with ASD and Related Disorders

Jacqueline Maffei*, Jessica Singer-Dudek*[1] & Dolleen-Day Keohane **
*Columbia University Graduate School of Arts and Sciences, **Nicholls State University

Abstract

We tested the effects of the establishment of conditioned reinforcement for observing human faces and/or voices on the rate of learning, observing responses, and verbal operant emissions for four children, ages 4-5 years, with autism spectrum disorders (ASD) and related disorders. We used a non-concurrent, delayed probe design across participants with pre and post-intervention measures. The intervention included a conjugate stimulus-stimulus pairing procedure. Results demonstrated that as a function of the intervention, faces were conditioned for three out of three participants and voices were conditioned for two out of two participants for whom either was lacking respectively prior to the intervention (both faces and voices were conditioned for one participant). Post-intervention probes demonstrated increases in rate of learning, observing responses, and verbal operants for all four participants.

Keywords: Verbal Developmental Cusps, Conjugate Reinforcement, Stimulus-stimulus Pairings, Conditioned Reinforcement, Faces, Voices.

El Efecto del Establecimiento de Caras y/o Voces de Adultos como Reforzadores Condicionados para Niños con TEA y Desórdenes Relacionados

Resumen

Se probaron los efectos del establecimiento del reforzamiento condicionado al observar caras o voces de humanos sobre la tasa de aprendizaje, de la tasa de respuestas de observación y sobre la emisión de operantes verbales de cuatro niños de 4 a 5 años con trastorno del espectro autista (ASD, por sus siglas en inglés) y de trastornos relacionados. Se utilizó un diseño no concurrente demorado entre participantes con mediciones pre y post intervención. La intervención incluyó un procedimiento de apareamiento conjugado estímulo-estímulo. Los resultados mostraron que, como función de la intervención, las caras se condicionaron para tres de tres participantes y las voces se condicionaron para dos de dos participantes, para quienes dicho condicionamiento no estaba presente antes de la intervención (ambas, las caras y las voces ya estaban condicionadas para uno de los participantes). Los sondeos post-intervención demostraron un incremento en la tasa de aprendizaje, de las respuestas de observación y de las operantes verbales para los cuatro participantes.

Keywords: Cúspides de Desarrollo Verbales, Reforzamiento Conjugado, Apareamientos Estímulo-Estímulo, Reforzamiento Condicionado, Caras, Voces.

Original recibido / Original received: 15/07/2014 Aceptado / Accepted: 25/09/2014

[1] This study was conducted as part of the first author's doctoral dissertation under the mentorship of the second and third authors.

Correspondence: Programs in Teaching as Applied Behavior Analysis, Teachers College, Columbia University, Box 22, 525 West 120th Street, New York, NY 10027, E-mail: js408@tc.columbia.edu

Observing the human eyes, or face, is one of the most crucial early observing responses and what many propose is the first step to becoming verbal (Arnold, Semple, Beale, & Fletcher-Flinn, 2000; Baron-Cohen, Baldwin, & Crowson, 1997; Cleveland, Kobiella, & Striano, 2006; Kleinke, 1986). Others agree that eye contact is important for infant-adult interaction, socialization, and cognitive development (Senju, Kikuchi, Hasegawa, Tojo, & Osanai, 2008; Symons, Hains, & Muir, 1998). While most research has focused on eye gaze or eye contact, other studies have shown that the movement of the lips, jaw, face, and tongue also select out observing responses and aid in communication and learning (Kleinke, 1986; Massaro & Bosseler, 2006; Mirenda, Donnellan, & Yoder, 1983; Striano & Bertin, 2004). Neuro-typical infants 7 to 11 weeks old were found to scan the eye area of faces more intently when voices were introduced (Haith, Bergman, & Moore, 1977). For children with autism, this critical observing response is often missing (Baron-Cohen et al., 1997; Ellsworth, Muir, & Hains, 1993; Hains & Muir, 1996; Senju et al., 2008). Recent research indicates that deficits in eye contact can be detected in infants as young as 2 to 6 months of age and may be indicative of a later diagnosis of autism (Jones & Klin, 2013).

Researchers focusing on verbal behavior development have identified several components that appear foundational to becoming verbal. These include conditioned reinforcement for observing adult faces, listening to adult voices, and observing two- and three-dimensional stimuli in the environment (Greer, Pistoljevic, Cahill, & Du, 2011; Keohane, Luke, & Greer, 2008; Keohane, Pereira Delgado, & Greer, 2009; Pereira Delgado, Greer, Speckman, & Goswami, 2009). Observing people and objects in the environment provides a context for individuals to participate in verbal exchanges with one another. These observing responses are operants, selected out by their reinforcers; thus, the stimuli that are observed must be conditioned reinforcers (Dinsmoor, 1983). Therefore, it is the establishment of the reinforcer for observing that is the critical foundation for verbal development.

When observation of the human face is missing from an individual's community of reinforcers, the individual will likely not respond to or even observe the presence of another individual, let alone verbal antecedents delivered by another (speaker). A child who lacks conditioned reinforcement for human faces and/or voices does not orient toward others, whether they are speaking or not speaking, and is not likely to respond to greetings or instructions from a speaker. A child at this level of verbal capability will most likely present at a pre-listener level of verbal behavior (Greer, 2002; Greer & Keohane, 2005; Greer & Ross, 2008; Skinner, 1957).

When individuals lack observing responses for human faces and/or voices they are not able to contact reinforcement from the presence of other individuals, as a speaker or a listener, and subsequently opportunities for contacting other social contingencies are limited. Conditioned reinforcement for observing faces and voices is foundational to increases in the complexity of verbal development and when it is missing further verbal development is not possible.

The observation of human faces and other observing responses meet the definition of what Rosales-Ruiz and Baer (1997) call behavior developmental cusps. These behavior developmental cusps, which include observing responses

as well as behaviors such as crawling and walking, are important developmental stages that, once attained, allow children to progress in ways they could not prior to their attainment. Once established, the individual is afforded opportunities to contact new environmental contingencies, and new reinforcers, and as a result learn new skills that support the performance of more complex tasks (Greer & Keohane, 2005; Greer & Ross, 2008; Greer & Speckman, 2009). Therefore, children can learn things they could not before because they can contact new contingencies. Children can also learn new things faster due to accelerated establishment of stimulus-response relations, or stimulus control. In summary, verbal cusps allow children to contact new conditioned reinforcers that subsequently lead to accelerated rates of learning; identifying and inducing missing verbal developmental cusps is crucial for the development of complex levels of verbal behavior.

The establishment of new conditioned reinforcers often takes place via stimulus-stimulus pairings. The stimulus-stimulus pairing procedure has been used to expand children's community of reinforcers by conditioning non-preferred stimuli as reinforcers, resulting in new responses. Such responses include looking at books, playing with toys, observing two-dimensional stimuli, responding to human voices, and the emission of new vocal sounds (Greer, Becker, Saxe, & Mirabella, 1985; Greer, Dorow, Wachhaus, & White, 1973; Longano & Greer, 2006; Miguel, Carr, & Michael, 2002; Nuzzolo-Gomez, Leonard, Ortiz, Rivera, & Greer, 2002; Pereira Delgado et al., 2009; Rheingold, Gerwirtz, & Ross, 1959; Smith, Michael, & Sundberg, 1996; Sundberg, Michael, Partington, & Sundberg, 1996; Tsai & Greer 2006; Yoon & Bennett, 2000).

According to several researchers, verbal development begins in the womb (Spence & DeCasper, 1987). Following birth, DeCasper and Fifer (1980) found that newborn infants prefer their mother's voices to those of other females and DeCasper and Spence (1987) found that newborn infants, two-three days old, demonstrated a preference for a passage that had been read to them by their mothers every day for six weeks prior to birth compared to a novel passage. Not only do infants prefer their mother's voices, they prefer familiar sounds. Theories about why this is so are related to conditioned reinforcement that begins prior to birth.

One explanation is that the sounds of the mother's voice are heard in utero and are paired with primary reinforcers present in the womb (e.g., warmth, nourishment, movement) and thus the mother's voice is conditioned as a reinforcer prior to birth. After birth, those pairings continue, with nourishment, touch, and the mother's face that is now paired with her voice. As a result of these pairings, the mother's face, and soon others, become conditioned reinforcers almost immediately after birth. Meltzoff and Moore (1983) found that newborn infants can imitate facial gestures within hours after birth, suggesting that it may also be the novelty of the face and the facial movements that act as primary reinforcers in selecting out the infants' observing and responding.

Conditioned reinforcement for listening to voices is also a necessary cusp for the development of both listener and speaker skills (Greer et al., 2011; Keohane, Luke, & Greer, 2008). Greer et al. (2011) conditioned voices as

reinforcers via a stimulus-stimulus pairing procedure. Results showed that all three participants' rate of learning accelerated, two children's observing responses increased, and two children's stereotypy decreased while their attention to a story read aloud by an adult increased. Keohane et al. (2008) implemented a rotated protocol package that included conditioning faces, voices, two- and three-dimensional stimuli, matching across the senses, and generalized imitation for three elementary students with ASD. Results of this treatment package demonstrated increased rates of learning and increased observing responses for all three children. While conditioning faces was one of the protocols implemented, it is impossible to isolate the effects of this intervention alone.

In the present study we used a conjugate stimulus-stimulus pairing procedure to condition adult faces and/or voices as reinforcers for four children with ASD. According to White (1971), conjugate reinforcement refers to "a schedule of reinforcement in which reinforcement is continuously present (e.g., the opportunity to eat) as long as a specified response is maintained at a criterion rate. Failure to maintain responding results in the discontinuance of reinforcement (e.g., the removal of the food dish) until responding again at criterion level" (p. 137). Conjugate reinforcement has resulted in conditioning novel stimuli as reinforcers (Cotter & Spradlin, 1971; Dunst, Storch, Hutto, & Snyder, 2007; Lindsley, 1956; Lovitt, 1968; Rovee & Rovee, 1969).We tested the effects of the conditioning procedure on the rate of acquisition of curricular objectives, emission of verbal operants, and observing responses to the presence of adults in the environment. In addition, we sought to determine whether adult faces or adult voices functioned as conditioned reinforcers, prior to or as a result of the intervention, so we conducted pre- and post-intervention probes of the reinforcing effects of adult faces and voices as well.

Method

Participants

We selected four males with developmental disabilities, ranging in age from four to eight years, based on classroom observations that indicated that adult faces and/or voices were not conditioned reinforcers. All participants emitted mands (requests) e.g., "I want jelly bean please" and tacts (object, event, or condition names) in complete sentences. All participants had a limited community of social reinforcers. In addition, the participants emitted low numbers of correct responses to learn units (response opportunities) across speaker and listener programs, low levels of observing responses, and low levels of verbal operants across three non-instructional settings, as confirmed by pre-intervention measures.

Participant A was a five-year-old male diagnosed with Autism Spectrum Disorder (ASD). Participant B was a four-year-old male diagnosed with a developmental disability (unspecified). Participant C was a five-year-old male diagnosed with Pervasive Developmental Disorder. Participant D was an 8-year-old male diagnosed with ASD.

Setting

The experiment took place in a private publicly funded preschool 20 miles outside of a major city and a classroom in a public suburban elementary school 40 miles outside of a major city that implemented the CABAS® (Comprehensive Application of Behavior Analysis to Schooling, Greer, 2002) model. All instruction was conducted in the participants' classroom, as part of their normal classroom instruction. All pre- and post- intervention probe sessions and the intervention sessions were conducted outside of the participants' classrooms in an empty part of a hallway, where it was quiet and the walls were bare. The hallway setting contained a child-sized desk, a child-sized chair, and chairs for the experimenter and independent observer, when present.

For Participant D all pre- and post intervention sessions and the intervention were conducted in his home in a quiet room. The room consisted of a large dining table with eight chairs and bare walls. In addition, pre- and post observing response and verbal operant probe sessions were conducted in Participant D's classroom.

Materials

During the conditioned reinforcement for listening to adult voices probe sessions the materials consisted of three child-sized chairs, a timer that counted forward, and two electronic Pal Pad (Adaptivation, Inc.) pressure-activated membrane switches connected to a tape recorder (see Figure 1). When a specific electronic switch was depressed, the tape recorder was activated and played a recorded voice reading a children's story. When the other electronic switch was depressed, no sound was emitted. During the stimulus-stimulus pairing procedure to condition faces as a reinforcer (the intervention) the materials consisted of two child-sized chairs, and a timer that counted forward. Pre and post-intervention measures of rate of acquisition of tacts required five sets of four 2-dimensional tact stimuli. These stimuli included pictures of animals (e.g., lizard, frog), flowers (e.g., lily, tulip), and musical instruments (e.g., trumpet, harp) printed in color, laminated, and affixed to 7.5cm X 12.5cm index cards. Other materials included data sheets and black pens to record the data.

Figure 1. The two electronic Pal Pad (Adaptivation, Inc.) pressure-activated membrane switches connected to a tape recorder used during the pre- and post-intervention probe sessions for voices as conditioned reinforcers. The switches were rotated periodically so that the participant had to find the switch that activated the voice recording.

Dependent Variables

We tested the effects of conditioning adult faces as reinforcers on three dependent variables: 1) the rate of acquiring objectives across two broad curricular areas (listener responses and speaker responses), 2) observing responses to the presence of adults in the environment, and 3) verbal operants emitted across three non-instructional settings (i.e., lunch, art, and recess). To assess rate of learning, we took 1000 learn units consisting of listener responses (following 40 single-step instructions) and 1000 speaker learn units (20 tact stimuli) and divided them by the number of instructional objectives achieved. A learn unit consists of an instructional antecedent, the response from the child, and a consequence that functions to either reinforce future correct responses or a correction that functions to occasion future correct responses (Greer, 2002). The listener learn units included a total of 36 single-step commands (e.g., touch your toes, clap your hands) and nine "nonsense" commands (e.g., "la la la"). The 45 commands were grouped into nine sets of five commands each—four single-step instructions and one nonsense command. Each set was taught separately in blocks of 20 learn units, so that each command was presented four times per session.

In order to assess observing responses to the presence of adults in the environment, we measured whether the child oriented toward an adult (speaking or not speaking) across ten different scenarios (see Table 1). In order to assess the number of verbal operants emitted by the participants we conducted 10-min probes in three non-instructional settings (i.e., lunch, free play, and art) in which we measured the total number of mands, tacts, sequelics, and conversational units emitted (see Table 2 for a complete definition of each). In addition, as tests of the independent variable, we measured whether adult faces and voices functioned as conditioned reinforcers prior to and following the intervention.

Table 1
Verbal operants measured during pre- and post-intervention probes conducted during 10-min sessions across lunch, art, and recess.

Verbal Operants:
Mand: A Mand specifies its reinforcer, and is produced in the presence of the item under deprivation without vocal antecedent and results in the delivery of the item
Tact: Production of a vocal response to a stimulus without vocal antecedent under the control of generalized social listener reinforcement
Sequelic: A verbal operant that occurs when an individual responds as a listener and speaker to intraverbals
Conversational Unit: An exchange that involves a listener and speaker in which each acts as speaker and listener to each other's intraverbals at least twice in an exchange

Table 2
Observing Responses Measured during Pre- and Post-Intervention Probe Sessions

Antecedent and Opportunity for Observing Responses
Participant orients toward a speaker when his name is called in a moderate, but detectable volume from a distance of 0.5-1.5 meters. "Orients" refers to the participant making eye contact or looking at the face of the experimenter or 3rd party for a minimum of 1 s.
Participant orients toward a speaker when name is called in a moderate, but detectable volume from 1.5-2.5 meters.
Participant orients toward a speaker when the child is given a 1-step direction in a moderate, but detectable volume from 0.5-1.5 meters.
Participant orients toward a speaker when the child is given a 1-step direction in a moderate, but detectable volume from 1.5-2.5 meters.
Participant orients toward speaker when the child is spoken to in a moderate, but detectable volume from a distance of 0.5-1.5 meters.
Participant orients toward a speaker when another child is spoken to in a moderate, but detectable volume from 1.5-2.5 meters.
Participant orients toward an adult rearranging the child's materials on desk.
Participant orients toward an adult removing the child's materials from desk.
Participant orients toward an adult entering the room who is speaking in a moderate, but detectable volume.
Participant orients toward an adult entering room who is not speaking.

Intervention: Face Conditioning

We used a conjugate stimulus-stimulus pairing procedure during the intervention to condition adult faces as reinforcers. This procedure involved the experimenter getting the participant to orient to her face using non-vocal sounds (e.g., smacking lips, making loud kissing sounds) and not, for example, calling the child's name or saying, "look at me." Immediately upon the participant orienting toward and observing her face the experimenter delivered vocal, visual, and, in some cases, tactile reinforcement in the form of animated expressions, speaking, singing, and sometimes touching the face, head, or arms of the participant. If at any time the participant looked away from the experimenter's face for longer than one second then the experimenter ceased and attempted to regain the participant's attention to her face using non-vocal sounds. The definition of looking at the experimenter's face included the participant looking at any part of the experimenter's face (e.g., forehead, hair, eyes, cheeks, chin, mouth, etc.).

Procedures and Data Collection

Learn Units-to-Criterion. In the listener instruction, target commands and nonsense commands were selected and divided into sets of five responses each (four commands and one nonsense command). For each of the nine sets (five operants each) of listener learn units, the commands were presented four times each during a 20 learn unit session and each set was taught separately. The experimenter delivered the vocal antecedent, e.g., "clap your hands," without giving any visual cues. In other words, we ensured that the participant only responded to

the auditory properties of the antecedent. If the participant responded correctly to the command within three seconds, the experimenter delivered vocal praise or preferred edibles. For the single-step instructions, if the participant emitted an incorrect response or no response, the experimenter delivered a correction. Corrections involved re-presentations of the antecedent followed by an opportunity for the participant to respond again. In some cases, the experimenter provided a physical prompt in order for the participant to emit the correct responses. Correct responses that followed corrections were not reinforced. For the nonsense commands, the absence of any response was reinforced and incorrect responses were ignored: the experimenter paused and looked away for two seconds and then presented the next learn unit. Criterion consisted of the participants emitting 90% accuracy across two sessions or 100% accuracy for one session.

For tact (speaker responses) instruction each of the five sets of tacts was taught separately. The experimenter held up a stimulus and got the participant's attention. Once the participant looked at the stimulus he was given three seconds to emit the correct response, e.g., "harp." The experimenter delivered vocal praise and attention for correct responses and a correction for incorrect or no responses. For corrections, the experimenter re-presented the antecedent stimulus, provided the correct response, and gave the participant the opportunity to echo the correct response. Correct responses that followed corrections were not reinforced. Criterion consisted of the participants emitting 90% accuracy across two consecutive sessions or 100% accuracy in one session.

Observing Response Probes

During the pre and post-intervention observing response probe sessions, we measured the number of times out of ten opportunities that the participants looked at or in the direction of the approaching or speaking adult across the ten observing response scenarios (see Table 1), for a total of 100 response opportunities. We used different adults, both familiar and unfamiliar to the participants, and provided opportunities that were spaced to provide the most natural non-contrived setting. A plus (+) was recorded when the participant looked at or in the direction of the adult within one second and a minus (-) was recorded if the participant did not look at or in the direction of the adult within one second. When opportunities across all of the observing response scenarios were completed, the cumulative number of observing responses was tallied.

Verbal Operant Probes

During the pre and post-intervention verbal operant probe sessions, verbal operants were measured during three non-instructional settings (i.e., lunch, art, and recess). During these probe sessions two experimenters simultaneously but independently recorded the total number of verbal operants emitted by the participants during three 10-min sessions. At the end of each session, the verbal operants were tallied and categorized to derive a total number of each type (i.e., mands, tacts, sequelics, and conversational units).

Tests of the Independent Variable—Faces and Voices as Conditioned Reinforcers

Probes for Adult Faces as Conditioned Reinforcers. In order to assess conditioned reinforcement for observing adult faces, we conducted a 5-min probe using 5-s partial interval recording during which we measured whether the participant observed the experimenter's face while she moved her mouth and face in animated expressions without making vocal sounds. The experimenter mouthed the words to a poem or a passage from a book, for example, without using her voice. If at any moment in the 5-s interval the participant looked at or in the direction of the experimenter a plus (+) was recorded on the data sheet. If during the 5-s interval the participant did not look at or look in the direction of the experimenter a minus (-) was recorded on the data sheet. The participant was not required to observe the experimenter's face for the entire 5-s interval. At the end of the 5-min session, the number of pluses and minuses were tallied. In order for adult faces to be considered conditioned reinforcers the participant had to emit observing responses for a total of 45 intervals out of 60 (75%).

Probes for Adult Voices as Conditioned Reinforcers. During the conditioned reinforcement for listening to adult voices probe session the participant was required to depress a specified switch that activated an adult voice reading a children's story. If the participant depressed a second switch, no sound was emitted. The two switches were necessary in order to determine if the participant preferred listening to the voice on the recording; they were periodically rotated. Once the participant depressed the correct switch and activated the story, the experimenter started the countdown timer, which was set for five minutes. During the 5-s whole interval recording the experimenter recorded a plus (+) if the participant depressed the switch for the entire 5-s interval and recorded a minus (-) if the participant did not depress the switch for the entire 5-s interval. In order to control for passive depressing of the switch, the experimenter rotated the position of the switches after every ten intervals. The participant then needed to find the switch that again activated the voice recording. At the end of the 5-min session, the pluses and minus were tallied. In order for adult voices to be considered conditioned reinforcers for the participant, he had to depress the switch for a total of 45 intervals out of 60 (75%).

Intervention. During the intervention, we implemented the conjugate stimulus-stimulus pairing procedure. The first step was getting the participant to look at the experimenter's face. The experimenter used vocal sounds or musical instruments to get the participant's attention. For example, the experimenter chewed gum and blew bubbles that popped loudly, clicked her tongue, blew bubbles with her lips, stuck out her tongue and blew, rolled her tongue, or made sounds such as "da da da" or "la la la." At times she also played instruments such as a kazoo or harmonica. It is important to note that the experimenter emitted these sounds until the participant looked at her face. The sounds the experimenter produced were continuously changing, thus varying from moment to moment.

Once the participant oriented to the experimenter's face she immediately started a timer and she delivered two kinds of reinforcement, either separately or simultaneously, contingent upon the participant looking at her face. These included vocal reinforcement and/or tactile reinforcement.

Vocal reinforcement consisted of the experimenter singing animatedly or softly, reciting a nursery rhyme, or delivering vocal praise while the participant was looking at her face. The reinforcement was continually changing; the conjugate procedure involved novel sounds, facial expressions, and movements from the experimenter. As soon as the participant looked away for one second the experimenter stopped, and the trial ended. If the participant looked back at the experimenter within one second she continued with varied sounds and expressions. The experimenter made moment-to-moment decisions as to what sounds, expressions, or touches were reinforcing, and which ones the participant appeared not to like. Some participants preferred soft voices and touches, others preferred loud and exaggerated voices and expressions. Therefore, it was important for the experimenter to respond flexibly and determine immediately what was reinforcing for the participant in order to maintain the participant's observation of her face. The session continued until a total of 20 trials were completed or the participant met the criterion for the intervention. Once 20 trials were completed, the experimenter calculated the sum (e.g., 1+2+1+3+1+10+9+20+5+8+9+5+6+4+1+8+7+6+5+4=115 cumulative s). The intervention continued until the participant emitted 160 cumulative s of observing the experiment's face across 20 trials. One session of the intervention was run each day.

Design

We used a delayed non-concurrent probe design across participants in order to control for maturation and history. We conducted probes or gathered relevant learn unit data immediately prior to and following the intervention to condition adult faces and/or voices as reinforcers. Following the intervention, we repeated the probes and gathered the learn unit data for comparison with pre intervention data.

Interobserver agreement

Probes. Interobserver agreement (IOA) was collected during all pre and post-intervention probe sessions for each participant. IOA for Participant A was conducted for 37% of sessions with a mean agreement of 94% with a range of 89-100%. IOA for B was conducted for 39% of sessions with a mean agreement of 95% and a range of 87-98%. IOA for Participant C was conducted for 34% of sessions with a mean agreement of 99% with a range of 95-100%. IOA for Participant D was conducted for 64% of sessions with a mean agreement of 97% with a range of 93-100%.

I*ntervention*. During intervention, IOA was collected for Participants A, B, C, and D. IOA was calculated on a point-to-point basis for 100% of the intervention sessions for Participant A with a mean of 95% and a range of 87-98%. IOA was calculated on a point-to-point basis for 43% of the sessions for Participant B with a mean agreement of 99% and a range of 99-100%. IOA was calculated on a point-to-point basis for 57% of the intervention sessions for Participant C with a mean of 98% and a range of 94-100%. IOA was calculated on a point-to-point

basis for 50% of the intervention sessions for Participant D with a mean of 89% and a range of 85-93%.

Results

Pre- and Post-Intervention Probes. Figure 2 shows the results from the pre and post-intervention tests of conditioned reinforcement for observing human faces and conditioned reinforcement for listening to adult voices for Participants A, B, C, and D. Prior to intervention, Participant A demonstrated a total of 38 intervals out of 60 (5-min total probe with 5-s partial interval recording) for observing faces without voices. This is equivalent to saying that the participant looked at the experimenter during 63% of the intervals in 5-min probe session when the experimenter moved her face in animated ways or moved her lips while talking but without sound. During the conditioned reinforcement for human voices pre-intervention probe, Participant A emitted a total of 58 intervals out of 60 (5-min total probe with 5-s whole interval recording), or 97%. Therefore, prior to the conditioning intervention faces did not function as conditioned reinforcers for this participant, but voices did. Participant A required five sessions to meet criterion for observing faces during the intervention. Following the intervention, the conditioned reinforcement for observing human faces and conditioned reinforcement for listening to adult voices probes were repeated. The results showed an increase to 52 intervals out of 60 in total (87%) for observing faces without voices and 60 intervals out of 60 in total (100%) for listening to adult voices. Participant A demonstrated criterion-level responding to faces, thus indicating that the intervention functioned to condition faces as reinforcers.

Prior to the intervention, Participant B emitted a total of 56 intervals out of 60 (93%) during the conditioned reinforcement for observing faces probe and 5 intervals out of 60 (8%) during the listening to adult voices probe. Prior to the intervention, faces functioned as conditioned reinforcers for Participant B, but adult voices did not. Participant B required seven sessions to achieve the mastery criterion for observing faces and listening to voices during the intervention. Following the intervention, Participant B emitted 49 intervals out of 60 (82%) during the conditioned reinforcement for faces probe and 46 out of 60 intervals (77%) for probes of conditioned reinforcement for listening to voices. Participant B demonstrated criterion-level responding to voices probe, thus indicating that adult voices became conditioned reinforcers as a result of the intervention.

Prior to the intervention, Participant C emitted a total of 9 intervals out of 60 (15%) during the conditioned reinforcement for observing faces probe and 49 intervals out of 60 (82%) for the conditioned reinforcement for listening to adult voices probe. Therefore, adult faces did not function as conditioned reinforcers for Participant C, but listening to adult voices did. Participant C required seven sessions to achieve the mastery criterion for observing faces during the intervention. Following the intervention, Participant C emitted a total of 50 intervals out of 60 in total (83%) during the conditioned reinforcement for observing faces probe and he emitted a total of 50 intervals out of 60 (83%) during the conditioned reinforcement for listening to adult voices probe. The results indicated that the intervention functioned to condition faces as reinforcers for Participant C.

Prior to the intervention, Participant D emitted a total of 26 intervals out of 60 (43%) for the conditioned reinforcement for observing adult faces probe and 20 intervals out of 60 (33%) during the conditioned reinforcement for listening to voices probe. Participant D required four sessions to meet criterion for observing faces during the intervention.

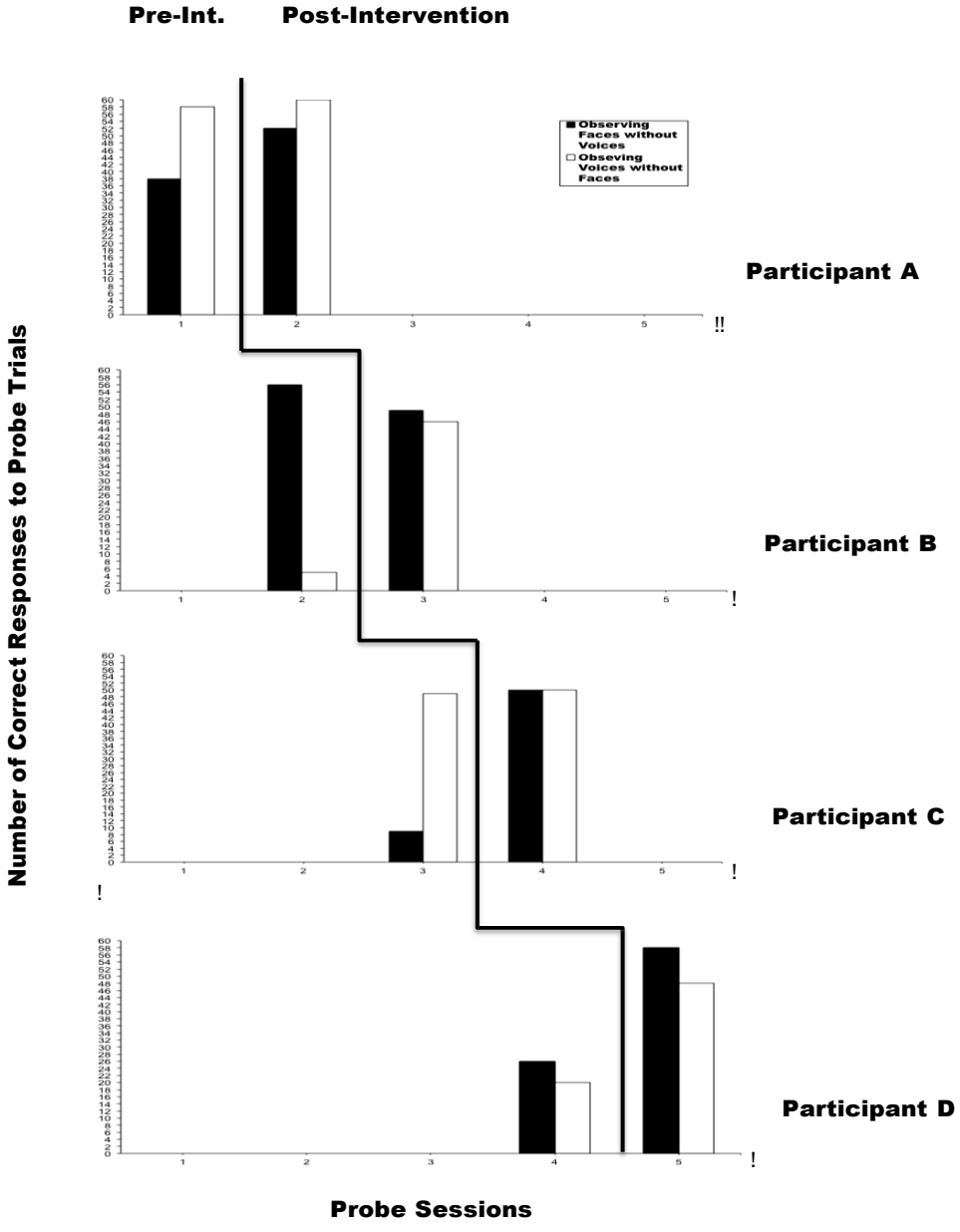


Figure 2. Number of correct responses to pre-and post-intervention probe trials for test of conditioned reinforcement for faces and voices for Participants A-D. The solid black line indicates the intervention. Arrows indicate 0 responses.

Following the intervention, Participant D's observing responses increased to 58 intervals out of 60 in total (97%) for observing faces and 48 intervals out of 60 (80%) for listening to adult voices. Therefore, Participant D acquired both observing adult faces and listening to voices as conditioned reinforcers as a function of the intervention.

Rate of Learning. Figure 3 represents the rate of learning, as indicated by the number of learn units-to-criterion, for Participants A, B, C and D prior to and after the conditioning intervention. As a function of the conditioning intervention, Participant A's learn units-to-criterion for speaker operants decreased from 167 to 111. Participant A demonstrated a listener repertoire at the outset of the study. Participant B's learn units-to-criterion for speaker operants decreased from 143 to 111 as a function of the intervention. Participant B demonstrated a listener repertoire at the outset of the study. Participant C demonstrated a decrease from 333 to 100 learn units-to-criterion for listener responses and 143 to 200 learn units-to-criterion for speaker responses as a function of the conditioning intervention. Participant D's learn units-to-criterion decreased from 200 to 90 for speaker operants and 143 to 77 learn units-to- criterion for listener responses as a function of the intervention.

Observing Responses. Results from the observing responses probes indicated that Participant A's observing responses to the presence of adults increased from 24 to 39 out of a total of 100 opportunities following the intervention. Participants B's observing responses increased from 18 to 35 following the intervention, Participant C's observing responses increased from 12 to 48 following the intervention, and Participant D's observing responses increased from 32 during the pre-intervention probe to 100 out of a possible 100 opportunities following the intervention (Figure 4).

Verbal Operant Probes. Results from the verbal operant probes indicated that prior to the intervention, Participant A emitted a cumulative total of 4 mands, 13 tacts, 12 sequelics, and 0 conversational units across all three settings for a total duration of 30 minutes. During post-intervention probe, Participant A emitted 1 mand, 6 tacts, 9 sequelics, and 5 conversational units. Prior to the intervention, Participant B emitted 2 mands, 1 tact, and 0 sequelics and conversational units. During the post-intervention probe session, Participant B emitted 2 mands, 20 tacts, 4 sequelics, and 0 conversational units. During the pre-intervention probe Participant C emitted 4 mands, 0 tacts, 0 sequelics, and 0 conversational units. Following the intervention, he emitted 50 tacts, 46 mands, 19 sequelics, and 0 conversational units. Prior to the intervention Participant D emitted a cumulative total of 3 mands, 3 tacts, 1 sequelic, and 0 conversational units. During the post-interventionprobe session, Participant D emitted a cumulative total of 12 mands, 1 tact, 0 sequelics, and 0 conversational units (Figure 5).

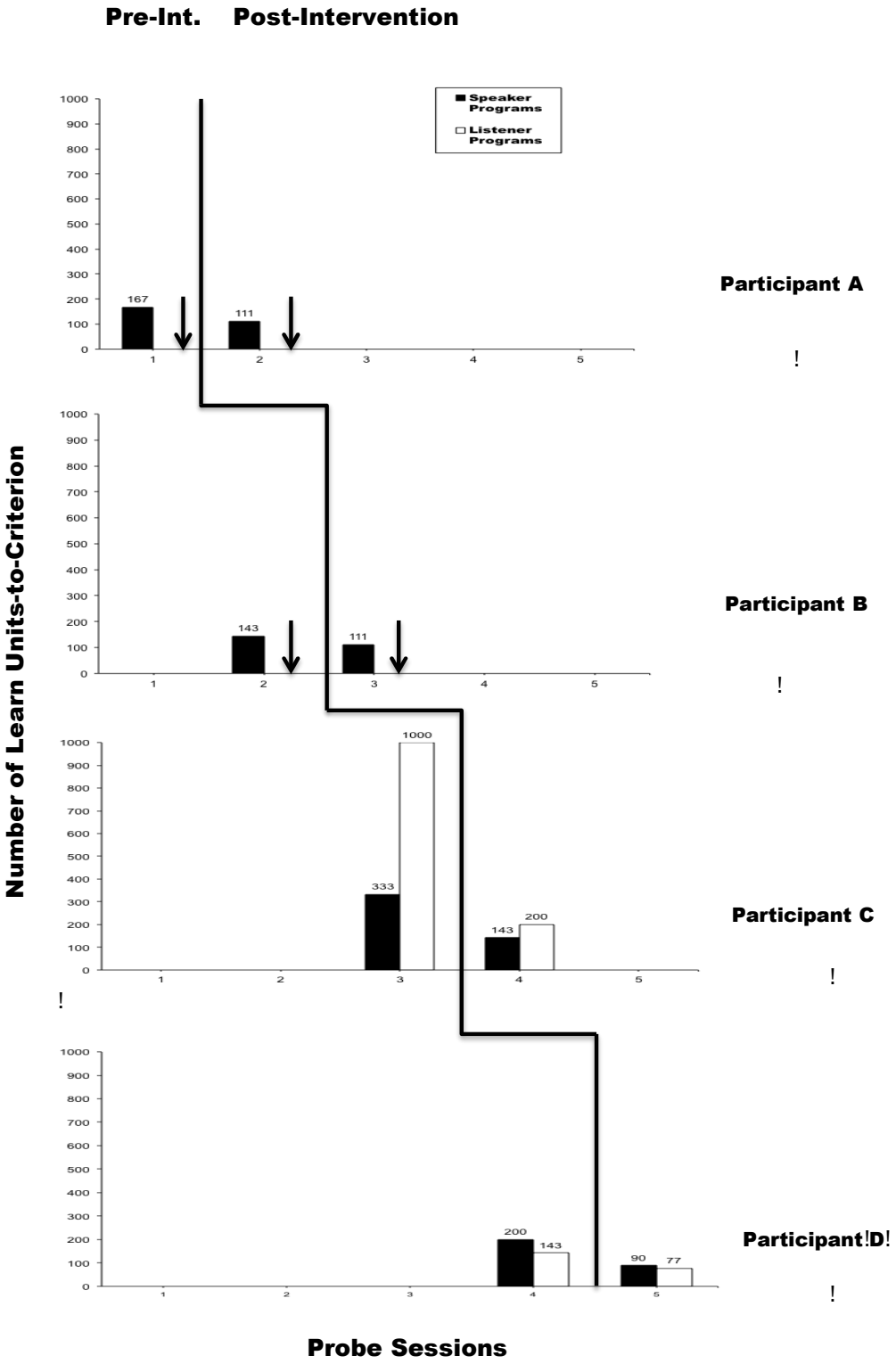


Figure 3. Number of pre-and post-intervention learn units-to-criterion for programs targeting speaker operants (tacts) for Participants A-D and listener operants for Participants C and D. The solid black line indicates the intervention. Arrows Indicate 0 responses

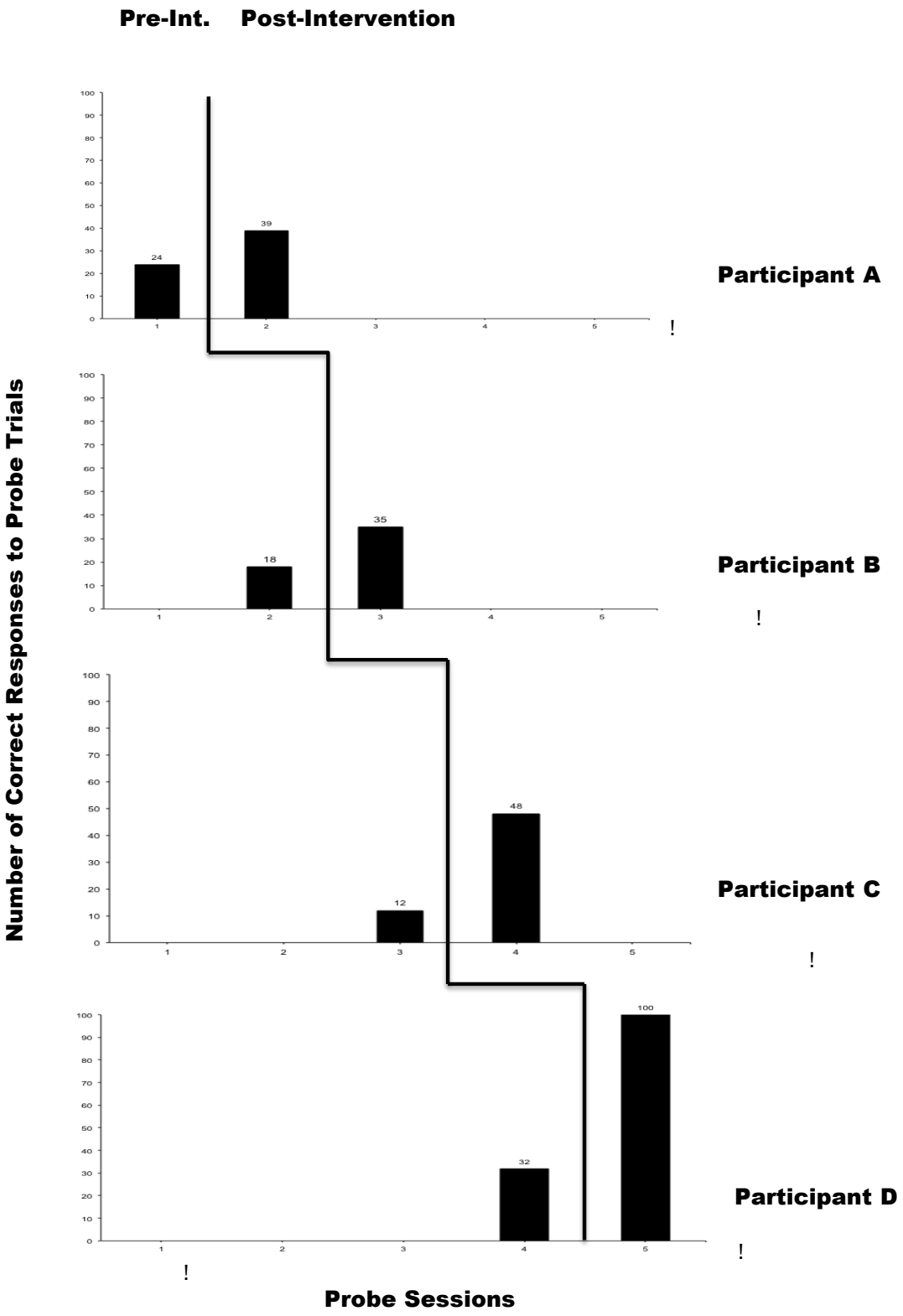


Figure 4. Number of correct responses to pre-and post-intervention probe trials for observing responses for Participants A-D—The Solid black line indicates the intervention. Arrows indicate 0 responses.

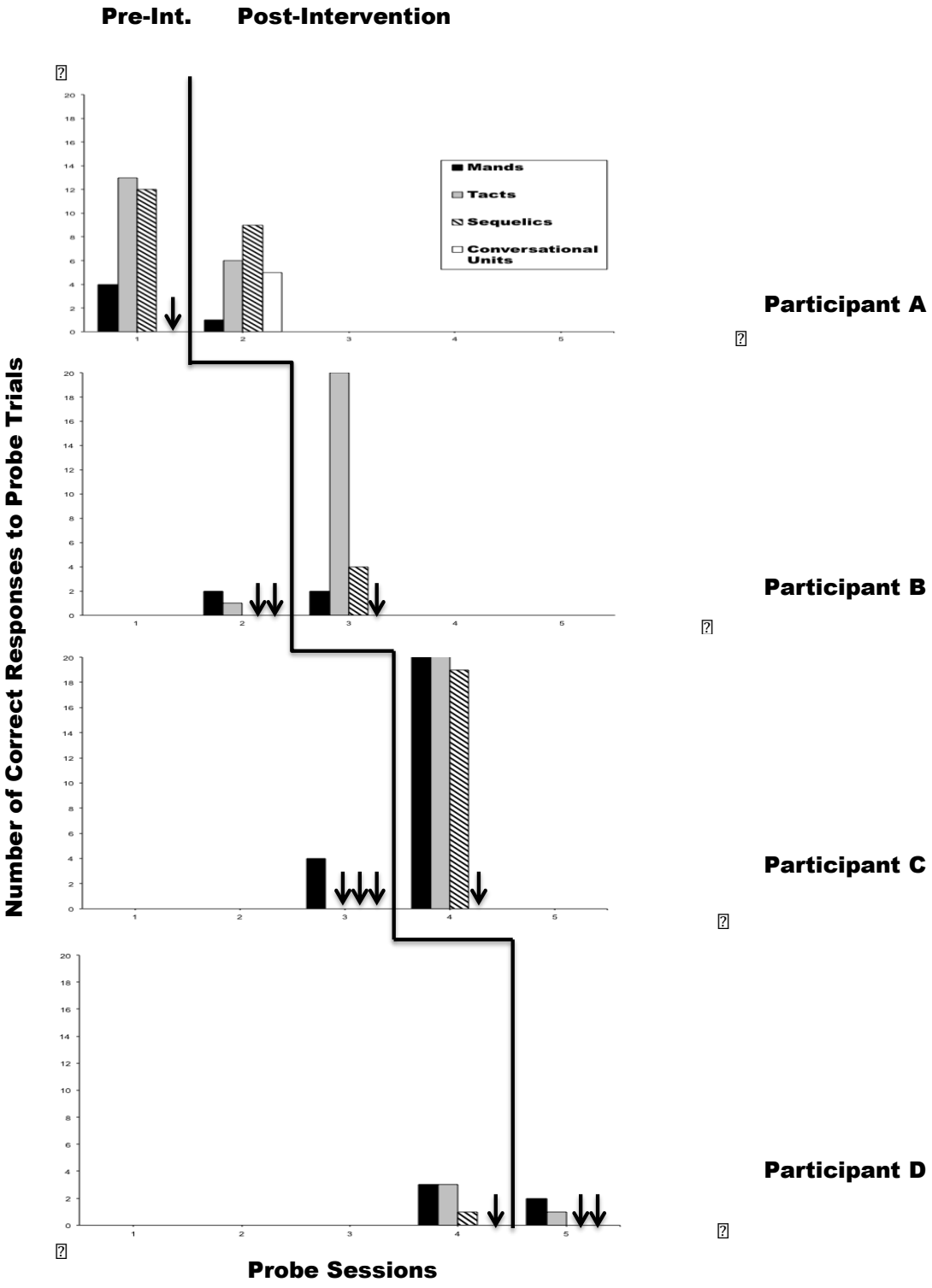


Figure 5. Number of correct responses to pre-and post-intervention probe trials for mands, tacts, sequelics, and conversational units for Participants A-D. The solid black line indicates the intervention. Arrows indicate 0 responses.

Discussion

The results of the present study support the theory that the acquisition of conditioned reinforcement for observing the human face and/or listening to human voices are necessary pre-verbal developmental cusps, as theorized by the VBDT (Greer & Keohane, 2005; Greer & Ross, 2008; Greer & Speckman, 2009). As a result of the intervention, all four of our participants demonstrated 1) accelerated rates of learning, as measured by the number of learn units to criterion, 2) increases in the emission of tacts and mands (for Participants B, C, and D), as measured by the verbal operant probes, and, for three out of four of the participants, the emergence of higher order verbal operants (sequelics for Participants B and C and conversational units for Participant A), and 3) increased attention to the presence of the adults (speaking or not speaking), as measured by the observing response probes.

As a result of the acquisition of reinforcement for observing faces and/or voices, all four of our participants demonstrated increased attention to the presence of a potential speaker or listener, looked at a speaker more often, and listened and responded to instructional antecedents more readily, resulting in increased rates of learning across both listener (for Participants C and D) and speaker operants.

These findings are further supported by the results of the verbal operant probes, which were measures of social interaction. Tacts, sequelics, and conversational units are characteristically reinforced by a response from a listener. They are social repertoires, with social reinforcers. Mands, too, are mediated by a listener, but the reinforcer is the item or condition manded. Participants A, B, and C all demonstrated increases in verbal operants that had social reinforcers. Although Participant A's post-intervention probes indicated that his number of mands, tacts, and sequelics decreased, conversational units emerged, which are higher order verbal operants. The individual alternates responding as both a speaker and a listener in a series of exchanges. This is a higher-order verbal operant. Participant D's tact and sequelic operants decreased following the intervention, but his mands increased. However, Participant D demonstrated 100% of observing responses in post-intervention probes, indicating that he was much more aware of the presence of adults in his environment.

The procedure we used during the intervention resulted in the conditioning of either one or two pre-verbal foundational cusps. Two of our participants (Participants A and C) had voices as conditioned reinforcers prior to the intervention, but faces did not function as reinforcers for observing. Both participants acquired faces as conditioned reinforcers as a function of the intervention. Participant B had faces as conditioned reinforcers prior to the intervention, but voices did not function as reinforcers. The intervention functioned to condition voices as reinforcers for Participant B. For Participant D, neither faces nor voices functioned as reinforcers prior to the intervention. Post-intervention results indicated that both of these cusps were established for Participant D.

We attribute these results, particularly the conditioning of voices in addition to faces as reinforcers, to the ever-changing, novel, conjugate stimulus-stimulus pairings that occurred during the intervention. The moment-to-moment

responsiveness by the experimenter to the participant was critical. In some cases the experimenter's loud tone appeared aversive to the participant, so she lowered her voice and spoke in softer tones while in other cases the participant seemed to prefer louder, more exaggerated tones. The experimenter's continual changes in response to the participant's observed reactions to her voice and facial expressions, animations, and/or tactile touch likely led to the success of this procedure. Consequently, those who will implement this procedure in the future need to constantly observe and respond to the participant and adjust their volume, proximity, intensity, and/or touch in order to ensure that the procedure is in fact conditioning the face and/or voice as a reinforcer.

In summary these results suggest that the acquisition of the human face and/or human voice as conditioned reinforcers are critical in language development and social behavior. Establishing these two pre-verbal cusps is the foundation to the development of language, and their importance is indicated for children with ASD.

References

Adaptivation, Inc. Sioux Falls, SD: http://adaptivation.com/product_detail.php?ID=44

Arnold, A., Semple, R. J., Beale, I., & Fletcher-Flinn, C. M. (2000). Eye contact in children's social interactions: What is normal behaviour?. *Journal of Intellectual and Developmental Disability, 25*, 207-216. doi:10.1080/13269780050144271

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development, 68,* 48-57. doi: 10.2307/1131924

Cleveland, A., Kobiella, A. & Striano, T. (2006). Intention or expression? Four month-olds' reactions to a sudden still-face. *Infant Behavior and Development*, *29*,299-307.doi: 10.1016/j.infbeh.2005.12.004

Cotter, V. W., & Spradlin, J. E. (1971). A nonverbal technique for studying music preference. *Journal of Experimental Child Psychology*, *11*, 357-365. doi: 10.1016/0022-0965(71)90040-3

DeCasper, A. J. and Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208,* (4448), 1174-1176. doi: 10.1126/science.7375928

DeCasper, A. J., & Spence, M. J. (1987). Prenatal maternal speech influencesnewborns' perception of speech sounds. *Infant Behavior and Development*, *9*, 133-150.doi: 10.1016/0163-6383(86)90025-1

Dinsmoor, J. (1983). Observing and conditioned reinforcement. *Behavior and Brain Sciences*, *6*, 693-704. doi: 10.1017/S0140525X00017969

Dunst, C. J., Storch, A. J., Hutto, M. D., & Snyder, D. (2007). Relative effectiveness of episodic and conjugate reinforcement on child operant learning. *Winterberry Research Syntheses*, *1*, 1-15.

Ellsworth, C. P., Muir, D., & Hains, S. M. J. (1993). Social competence and person-object differentiation: An analysis of the still-face effect. *Developmental Psychology, 29,* 63-73. doi: 10.1037/0012-1649.29.1.63

Greer, R. D. (2002). *Designing teaching strategies: An applied behavior analysis*

system approach. New York, U.S.A.: Academic Press.

Greer, R. D., Becker, B. J., Saxe, C. D., & Mirabella, R. F. (1985). Conditioning histories and setting stimuli controlling engagement in stereotypy or toy play. *Analysis and Intervention in Developmental Disabilities, 5,* 269-284.doi: 10.1016/0270-4684(85)90015-1

Greer, R. D., Dorow, L. G., Wachhaus, G., & White, E. R. (1973). Adult approval and students' music selection behavior. *Journal of Research in Music Education*, *21*, 345-354.doi: 10.2307/3344908

Greer, R. D., & Keohane, D. D. (2005). The evolutions of verbal behavior in children. *Behavioral Development Bulletin, 1,* 31-47. Reprinted in 2006 in the *Journal of Speech and Language Pathology: Applied Behavior Analysis, 1*, 111-140. Retrieved from: http://www.baojournal.com/SLP-ABA%20WEBSITE/SLP-ABA-VOL-1/SLP-ABA-1-2.pdf

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing new capabilities and expanding communication repertoires for children.* Boston, MA: Allyn & Bacon.

Greer, R. D., & Speckman, J. (2009). The integration of speaker and listener responses: A theory of verbal development. *The Psychological Record, 59,* 449-488.

Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Conditioning voices as reinforcers for listening responses in preschoolers with autism. *The Analysis of Verbal Behavior, 27*, 103–124.

Hains, S. M., & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child Development,* 67, 1940-1951. doi: 10.2307/1131602

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, *198*, 853-855.doi:10.1126/science.918670

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, *504*, 427-431. doi:10.1038/nature12715

Keohane, D. D., Luke, N. & Greer, R. D. (2008). The things we care to see: The effects of rotated protocol immersion on the emergence of early observing responses. *Journal of Early Intensive Behavior Intervention, 5*(1), 23-39*.* Retrieved from: http://www.baojournal.com/JEIBI/JEIBI-VOL-5/JEIBI-5-1.pdf

Keohane, D. D., Pereira Delgado, J. A., & Greer, R. D. (2009). Observing responses: Foundations of higher order verbal operants. In R. A. Rehfeldt, & Barnes-Holmes, (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 41-62). Oakland, CA: New Harbinger Publications.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*, 78-100. doi: 10.1037/0033-2909.100.1.78

Lindsley, O. R. (1956). Operant conditioning methods applied to research in Chronic schizophrenia. *Psychiatric Research Reports*, *5*, 118-139.

Longano, J., & Greer, R. D. (2006). The effects of a stimulus-stimulus pairing procedure on the acquisition of conditioned reinforcement for observing and manipulating stimuli by young children with autism. *Journal of Early and Intensive Behavior Interventions, 3,* 135-150. Retrieved from:

http://www.baojournal.com/JEIBI/JEIBI-VOL-3/JEIBI-3-1.pdf

Lovitt (1968). Free-operant assessment of musical preference. *Journal of Experimental Child Psychology*, *6*, 361-367. doi: 10.1016/0022-0965(68)90117-3

Massaro, D. W., & Bosseler, A. (2006). Read my lips. *Autism, 10,* 495-510. doi:10.1177/1362361306066599

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, *54*, 702-709. doi: 10.2307/1130058

Miguel, C. F., Carr, J. E., & Michael, J. (2002). The effects of a stimulus-stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *TheAnalysis of Verbal Behavior*, *18*, 3-13.

Mirenda, P. L., Donnellan, A. M., & Yoder, D. E. (1983). Gaze behavior: A new look at an old problem. *Journal of Autism and Developmental Disorders*, *13*, 397-409.doi:10.1007/BF01531588

Nuzzolo-Gomez, R., Leonard, M. A., Ortiz, E., Rivera, C. M., & Greer, R. D. (2002). Teaching children with autism to prefer books or toys over stereotypy or passivity. *Journal of Positive Behavior Interventions, 4,* 80-87. doi: 10.1177/109830070200400203

Pereira-Delgado, J., Greer, R. D., Speckman, J., & Goswami, A. (2009). Effects of conditioning reinforcement for print stimuli on match-to-sample responding in preschoolers. *Journal of Speech Language Pathology and Applied Behavior Analysis, 3.2-3.3,* 60-77. Retrieved from http://www.baojournal.com/SLP-ABA%20WEBSITE/SLP-ABA-VOL-3/SLP-ABA-Consolidated-3-2-3-3.pdf

Rheingold, H. L., Gewirtz, J. L., & Ross, H. W. (1959). Social conditioning of vocalizations in the infant. *Journal of Comparative and General Psychology*,*52*, 68-73.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis, 30*(3)*,* 533-544. doi:10.1901/jaba.1997.30-533

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratorybehavior. *Journal of Experimental Child Psychology*, *8*, 33-39. doi:10.1016/0022-0965(69)90025-3

Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain and Cognition, 67,* 127-139. doi:10.1016/j.bandc.2007.12.001

Skinner, B. F. (1957). *Verbal Behavior.* Acton, MA: Copley Publishing Group and the B. F. Skinner Foundation.

Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcement and automatic punishment in infant vocal behavior. *Analysis of Verbal Behavior*,*13*, 39–48.

Spence, M., & DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development, 10,* 133-142. doi:10.1016/0163-6383(87)90028-2

Striano, T., & Bertin, E. (2004). Contribution of facial and vocal cues in the still-

Face-response of 4-month-old infants. *Infant Behavior & Development, 27,* 499-508.doi:10.1016/j.infbeh.2004.06.002

Sundberg, M. L., Michael, J., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior*, *13*, 21-37.

Symons, L. A., Hains, S. M. J., & Muir, D. W. (1998). Look at me: Five -month-oldinfants' sensitivity to very small deviations in eye-gaze during social interactions.*Infant Behavior & Development*, *21*, 531-537. doi: 10.1016/S0163-6383(98)90026-1

Tsai, H., & Greer, R. D. (2006). Conditioned preference for books and faster acquisition of textual responses by preschool children. *Journal of Early and Intensive Behavioral Interventions 3,* 35-60.Retrieved from http://www.baojournal.com/JEIBI/JEIBI-VOL-3/JEIBI-3-1.pdf

White, O. (1971). *A glossary of behavioral terminology.* Champaign, Illinois: Research Press Company.

Yoon. S. Y., & Bennett, G. (2000). Effects of a stimulus-stimulus pairing procedure on conditioning vocal sounds as reinforces. *Analysis of Verbal Behavior, 17,* 75- 88.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1642 - 1491

Methods for Assessing Social Validity of Behavioral Intervention Plans for Children with Attention Deficit Hyperactivity Disorder[1]

Regina A. Carroll[2] & Claire C. St. Peter
West Virginia University

Abstract

Although behavioral interventions are powerful tools for parents and teachers, they are unlikely to result in lasting change if the intervention agents find them unacceptable. After developing effective behavior intervention plans for classroom use, we compared social validity of those interventions using three measures: concurrent-chains selections from the intervention consumer (students), verbal report of the intervention agent (teachers), and maintenance of the intervention over time. All three measures of social validity identified an intervention that was acceptable to the intervention consumer and intervention delivery agent. These findings are discussed in terms of applied implications for assessing social validity.

Keywords: Attention Deficit Hyperactivity Disorder, Behavior Intervention Plans; Choice; Concurrent-Chains Procedure; Social Validity.

Métodos para Evaluar la Validez Social de Planes de Intervención Conductual con Niños con Desorden por Déficit de Atención e Hiperactividad

Resumen

A pesar de que las intervenciones conductuales son herramientas poderosas para padres y maestros, es posible que no representen un cambio duradero si los agentes de la intervención consideran que no son aceptables. Después de desarrollar planes de intervención efectivos para ser utilizados en el aula, se comparó la validez social de dichas intervenciones usando tres medidas: elección de cadenas concurrentes por el consumidor de la intervención (estudiantes), reportes verbales del agente de la intervención (maestros) y mantenimiento de la intervención a lo largo del tiempo. Las tres medidas de validación social identificaron una intervención que era aceptable tanto para el consumidor como para el agente de la intervención. Los resultados se discuten en términos de las implicaciones aplicadas para evaluar la validez social.

Palabras Clave: Desorden de Atención e Hiperactividad, Planes de Intervención Conductual, Elección, Procedimiento Encadenado Concurrente, Validez Social.

Original recibido / Original received: 11/06/2014 Aceptado / Accepted: 13/09/2014

[1] We would like to thank Jennie Cox and Kara Samaj for their assistance with various aspects of data collection
[2] Address correspondence to Regina A. Carroll, West Virginia University, Department of Psychology, 53 Campus Drive, Morgantown, WV 26506-6040, Email: Regina.Carroll@mail.wvu.edu

Social validity, or the extent to which consumers of our science and practice believe that we are making valuable contributions, has been measured in behavior-analytic work since the 1970's (Kazdin, 1977; Wolf, 1978). Despite this long history, social validity remains an understudied area of behavior analysis, in part because of its relatively subjective measurement. Most systematic measures of social validity consist of rating scales (e.g., the Intervention Rating Profile; Witt & Elliot, 1985) and questionnaires (e.g., Gresham & Lopez, 1996). These scales directly measure consumers' verbal behavior only, which may be problematic if the consumers are not accurate reporters. Additionally, measuring social validity through verbal report alone may not predict the extent to which behavior-analytic procedures are acceptable solutions to addressing social problems.

To address these potential limitations, several authors have argued for the use of direct measurement of social validity (Hanley, 2010; Kennedy, 2002). This direct measurement can take at least two forms. One direct measure of social validity is the extent to which consumers maintain behavior-analytic interventions over time (Kennedy). Unlike measures of verbal report, examining maintenance as a direct measure of social validity may help us to identify common features of procedures that are likely to be adopted and persist in a specific environment.

Another direct measure of social validity is the extent to which consumers choose our interventions. Measurements of choice have been used to allow direct consumers (those personally experiencing the intervention), particularly consumers with limited or no verbal skills, to select which procedure they prefer (e.g., Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997). Consumer preference for interventions has typically been assessed using a modified concurrent-chains procedure. During the initial link of the procedure, consumers select between stimuli that were previously associated with each intervention option. The consumer then experiences the selected intervention during the terminal link of the chain. This kind of modified concurrent-chains procedure effectively evaluated consumer preference for different reinforcement schedules (e.g., Hanley et al., 1997), teaching procedures (e.g., Slocum & Tiger, 2011), and other intervention components.

There are several possible benefits to choice-based measures of social validity with direct consumers. First, it may allow consumers to select an option that best meets their momentary needs, even if those needs change over time. Choice procedures may allow consumers to select the intervention components that are most valuable to them in the moment, thus accounting for shifts in preference or motivating operations. Second, children may prefer situations in which they are permitted to choose over situations that are adult-directed (Fenerty & Tiger, 2010; Schmidt, Hanley, & Layer, 2009; Tiger, Hanley, & Hernandez, 2006; Tiger, Toussaint, & Roath, 2010). Allowing consumers to choose the interventions they experience may dignify the treatment process by allowing input from the client (Bannerman, Sheldon, Sherman, & Harchik, 1990).

There may be benefits to evaluating social validity of interventions with the behavior-change agents (indirect consumers) in addition to the direct consumers who experience the intervention. Allowing indirect consumers to participate in the social validity process provides those individuals with a way to select against

procedures that they do not find acceptable (Hanley, 2010). Establishing social validity with indirect consumers is important because treatment implementation may be unlikely to continue if those responsible for implementing the intervention do not also find the procedures acceptable.

To date, few studies have evaluated the social validity of interventions with both direct and indirect consumers, and studies have not evaluated the use of consumer choice and maintenance data to assess the validity of Behavior Intervention Plans (BIP). Additionally, direct measurement of social validity has not been extended to children with ADHD and their teachers. Yet, improving the acceptability or validity of intervention plans may improve the extent to which teachers implement those plans with fidelity (Mautone, DuPaul, Jitendra, Tresco, Vilejundo, & Volpe, 2009), thereby improving student outcomes (St. Peter Pipkin, Vollmer, & Sloman, 2010). To address this gap in the literature, we evaluated the social validity of two multicomponent BIPs using three measures: student choice for procedures, teachers' verbal reports, and maintenance of intervention over time.

Method

Participants and Setting

Three students who attended an alternative education program and two classroom teachers participated in this study. Zane and Kelvin were diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) and were 6 and 7 years old, respectively. Harmony was an 8-year-old girl diagnosed with mild intellectual disability, ADHD, Post-Traumatic Stress Disorder (PTSD), and phonological disorder. All three participants used complex sentences to communicate, and had an extensive history of engaging in chronic and severe problem behavior that was resistant to intervention. The two classroom teachers, Jamie and Stacy, each had a Master's degree in Elementary Education with certifications in both general and special education and were Board Certified Behavior Analysts. Jamie had been teaching for 13 years and Stacy had been teaching for seven years. Both teachers had been teaching in the alternative education program for approximately two years.

Prior to the start of this study, all three students had participated in an evaluation comparing the efficacy of two different BIPs on problem behavior (i.e., aggression, disruption, inappropriate language, and noncompliance). Both plans were multicomponent interventions that addressed multiple functions of problem behavior, and both BIPs produced similar reductions in problem behavior. Table 1 shows core components of each BIP.

All sessions were conducted in the students' classroom within the alternative education program. During each session, up to two teachers and eight students were present in the classroom. Sessions lasted the entire school day. The teachers were responsible for implementing the BIPs throughout the study.

Table 1
General Components of Each Behavior Intervention Plan

BIP 1	BIP 2
Materials: Timer, student specific academic materials, dry erase marker, point card, bin with 4 high preference toys, bin with 4 moderate-to-low preference toys, prize box with small trinkets (e.g., stickers, sucker, eraser, pencils)	*Materials:* Timer, student specific academic materials, picture of the students face with a magnetic back, a magnetic dry erase board split in half with the word "work" written on one side of the board and the word "break" written on the other side, dry erase marker, I-Pad, break area with preferred toys and activities
1. At the start of each day the teacher met with the target student and read a script outlining the rules for the specific BIP (see Appendix)	1. Same
2. The teacher started a timer counting down from 7 min, timing the work interval. During the work interval the students worked on academic programs either one-on-one with a teacher, or in a small group (e.g., 2-5 students).	2. The teacher placed the student's picture card on the "break" side of the magnetic dry erase board and started a timer counting down from 7 min, timing the work interval. During the work interval the students worked on academic programs either one-on-one with a teacher, or in a small group (e.g., 2-5 students).
3. At the end of the work interval the teacher met with the student and assigned up to three smiles on the point card, based on the absence of problem behavior during the 7-min work interval. Students earned a smile for being: "Safe" if he/she did not engage in aggression or disruptions, and stayed in his/her assigned area; "Respectful" if he/she did not engage in any inappropriate language or negative vocalizations; and "Responsible" if he/she did not engage in noncompliance.	If the student engaged in aggression, disruption, or left his/her assigned area, then the teacher moved the student's picture from the "break" side of the board to the "work" side of the board. When moving the student's picture, the teacher did not say anything to the student. 3. At the end of the work interval if the student's picture was on the "break" side of the board, then he/she was allowed to take a 3-min break with a variety of toys and activities in a designated break area in the classroom.
If the student earned three smiles, then he/she was able to take a 3-min break at his/her desk with a bin with 4 high preference toys, and a bin with 4 moderate-to-low preference toys. If the student earned 2 smiles, then he/she was able to take a 3-min break at his/her desk with a bin with 4 moderate-to-low preference toys only. If the student earned 1 or no smiles, then he/she continued to work on academic tasks at his/her desk during the break interval.	Following each break earned, the teacher drew a tally above the student's picture. If the student had two or more tallies, then he/she had the option to play with an I-Pad during the break. If at any time the student's picture was moved to the "work" side of the board, he/she lost all of the tallies, and had to start back at zero when his/her picture was moved back to the "break" side of the board. 4. At the end of the work interval if the student's picture was on the "work" side of the board, then he/she continued to work on academic tasks at his/her desk during the break interval. At the start of the break interval the student's picture was moved back to the "break" side of the board to signal to the student that he/she was now eligible for earning the next break, if he/she did not engage in any of the target problem behaviors.
5. At the end of the day the teacher met with the student and counted the total number of smiles earned during the day. If the student earned enough smiles to reach his/her daily goal, then they were able to pick a small prize from the prize box.	

Procedures

Student choice. We used a concurrent-chains procedure (Hanley et al., 1997) to evaluate each student's relative preference for the two BIPs. The teachers were trained to implement each of the plans before the start of the study. Both BIPs were associated with specific materials and all three students were familiar with these materials. We selected one item from each of the BIPs to represent that BIP during choice trials. We used a point card for BIP 1 and a picture card for BIP 2. We selected these items because they were relatively salient stimuli associated with the plans, were approximately the same size and shape, and the teachers thought that they were unlikely to be differentially preferred independent of the BIP with which they were associated.

Prior to evaluating students' preference for the different BIPs, the teacher conducted two forced-choice sessions (one session for each BIP) to expose students to the different BIPs associated with selecting each card, and to ensure that students had recent experience with each of the BIPs. During forced-choice sessions, the teacher placed the two cards (the point card and picture card) in front of the student. The teacher pointed to each card and read a script (available from the first author) that briefly described the main components of each BIP. Next, the teacher randomly selected one of the BIPs and instructed the student to hand her the card associated with that intervention. The teacher then implemented that BIP for the rest of the school day (approximately 5 hrs). The next day this procedure was repeated with the other BIP.

After the two forced-exposure days, students were allowed to select the BIP that would be implemented for the day. During student-choice sessions, the teacher presented the two cards to the student, read the script describing the main components of each BIP, and then instructed the student to choose a card. The student selected a BIP by handing the associated card to the teacher. Once the student selected a BIP, the teacher implemented that BIP for the remainder of the school day. If the student had attempted to select both cards, the teacher would have re-presented the cards and asked the student to select only one. However, this never occurred.

During each session, the teachers collected data on students' BIP choices, defined as selecting the card associated with a specific BIP and handing it to the teacher. We calculated the cumulative number of selections for each BIP by adding the total number of selections across sessions. Student-choice sessions continued until the student selected the same BIP across five consecutive school days. After the fifth consecutive selection of the same BIP, the teachers adopted that BIP as part of the student's Individualized Education Plan.

We collected treatment integrity data on the teachers' correct implementation of BIPs as a secondary measure. Treatment integrity data were collected during an average of 23% of the student-choice sessions across participants. Each observation was divided into six 10-min intervals. At the end of each interval, we scored the implementation of each component of a BIP as either correct or incorrect. We calculated treatment integrity by taking the number of BIP components implemented correctly and dividing it by the total number of

components implemented correctly plus the number of components implemented incorrectly, and multiplying by 100.

Teacher Report. We assessed the extent to which teachers found both the choice procedure and the child-selected BIP to be acceptable. Teacher acceptability was measured immediately after the student-choice phase concluded. Each teacher reported on the child or children with whom she worked most often. Jamie reported on the extent to which she found the choice procedure and BIP acceptable for Harmony. Stacy reported on the acceptability of the choice procedure and BIP for Zane and Kelvin. Teachers were provided one week to complete the social validity measures, and were asked to complete the measures independently of each other.

Each teacher was asked to complete two measures. The first measure we used was an open-ended questionnaire based on the one described by Gresham and Lopez (1996) to determine the acceptability of the choice procedures (see Table 2). The questionnaire asked teachers how they felt about allowing children to choose a BIP, the aspects of the procedure they liked the most, the aspects of the procedure they liked the least, how the procedure could be made better or easier, the negative side effects that children might experience, and the efficacy of the choice procedure for reducing problem behavior. The second measure was a modified version of the Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliot, & Darveaux, 1985), which we used to determine how acceptable the teachers found the child-selected BIP (see Table 3). Teachers rated the child-selected BIP in 15 areas, using a 1 to 6 Likert scale, with 1 indicating “strongly disagree” and 6 indicating “strongly agree.”

Intervention maintenance. One month following the completion of the student-choice sessions, we conducted a maintenance observation in the classroom. During this observation, we collected treatment integrity data on the teachers’ implementation of the BIP that the student chose most often during the choice sessions. The purpose of this observation was to evaluate the extent to which teachers continued to (a) implement the BIP selected by students during the choice sessions, and (b) implement the components of the BIP accurately.

Table 2
Teacher Responses to Open-Ended Questions about Validity of Choice Procedure

Jamie	Stacy
Q1. How do you feel about allowing students to choose which behavior support plan they will experience?	
I think it helps them buy into the plan.	If both plans are equally effective, I support allowing students to choose which behavior support plan they will experience.
Q2. Which aspects of the choice procedure did you like the most? Why did you like these aspects?	
That the kid had some control over their reinforcement because they seemed to be happier when they had some control.	It was easy to implement, the script was very easy and quick to read.
Q3. Which aspects of the choice procedure did you like the least? Why did you not like these aspects?	
Having to run a plan that I didn't feel was best or didn't appear to maintain appropriate behavior as successfully, forced choice when the student didn't like one or the other plan.	It was difficult at times because one plan gave feedback regarding noncompliance and one did not give specific feedback for noncompliance. It was not choice itself that appeared difficult at times.
Q4. How could we change the choice procedure to make it better, more acceptable, or easier to implement?	
I think it was easy to implement.	The procedure was easy to implement.
Q5. What negative side effects might giving students choices about behavior support plans have, both for the students given the choices and for other students in your classroom?	
Other students want to do what the student picked, they want to change their choice as soon as they don't like something about the plan. If they don't like it [the plan], forced choice was a challenge.	Other students engaged in problem behavior because they did not have the same materials associated with the specific plans.
Q6. Do you think that giving your student a choice about which behavior support plan should be in place was more effective in solving your student's problem behavior?	
It seemed on day they could make the choice they were "happier"	For some students

Table 3
Teacher Responses to the Intervention Rating Profile-15 Regarding the Selected Intervention

Item	Student			
	Zane	Kelvin	Harmony	Average
This was an acceptable intervention for the child's problem behavior	5	4	6	5.0
Most teachers would find this intervention appropriate for other behavior problems	4	4	5	4.3
This intervention should prove effective in changing the child's problem behavior	5	3	6	4.7
I would suggest the use of this intervention to other teachers	5	5	6	5.3
The child's behavior problem is severe enough to warrant this intervention	5	6	6	5.7
Most teachers would find this intervention suitable for the behavior problem	4	4	5	4.3
I would be willing to use this intervention in the classroom setting	6	5	6	5.7
This intervention would be appropriate for a variety of children	5	5	6	5.3
This intervention would *not* result in negative side effects for the child	3	3	6	4.0
This intervention is consistent with those I have used in classroom settings	4	5	6	5.0
This intervention was a fair way to handle the child's problem behavior	5	6	6	5.7
This intervention was reasonable for the problem behavior	5	6	6	5.7
I liked the procedures used in this intervention	5	5	6	5.3
This intervention was a good way to handle this child's problem behavior	5	4	6	5.0
Overall, this intervention was beneficial for the child	5	4	6	5.0

Interobserver Agreement

Teachers collected data on the student's selection of an intervention by writing the selection on a data sheet provided to them for that purpose. A secondary observer (one of the study authors) independently scored BIP selections during an average of 23% of the choice sessions across students. We compared the primary and secondary observers' data on a session-by-session basis and calculated IOA for students' BIP selections by taking the number of sessions with an agreement on a student selection divided by the total number of sessions and multiplied by 100. We scored an agreement if both observers scored the same BIP selection during a session, and a disagreement if both observers scored a different BIP selection for a given session. The IOA scores on BIP selections were 100% for all students.

Two researchers independently scored the IRP-15 measures. The researchers agreed on each teacher rating provided on the IRP-15 for each student (IOA = 100%), and ensured that IRP-15 and questionnaire results were transcribed accurately.

Results

The results of student choices are shown in Figure 1. All students showed a strong preference for one of the BIPs (BIP 1). Data for Zane are shown in the top graph. Zane selected BIP 2 during only the third choice period. Kelvin's data are shown in the second graph. Like Zane, Kelvin selected BIP 2 during the third session; notably, Zane's third session and Kelvin's third session were not conducted on the same day. Harmony's data are shown in the bottom graph. Harmony always selected BIP 1.

The results of the teachers' verbal reports are summarized in Tables 2 and 3. Table 2 shows the teachers' responses to the six open-ended questions that assessed their acceptability of the student-choice procedure. The teachers reported both positive and negative aspects of the choice procedure. Some of the positive aspects of the choice procedure included ease of implementation and that the students seemed to be happier because the choice procedure gave them some control over reinforcement. Some negative aspects of the choice procedure identified by the teachers included problem behavior that occurred when students could not select the BIP. This problem behavior was reported to occur both for the student participants (e.g., on forced-choice days) as well as other students in the classroom who were not participating in the evaluation. In general, both teachers seemed to find the choice procedure acceptable.

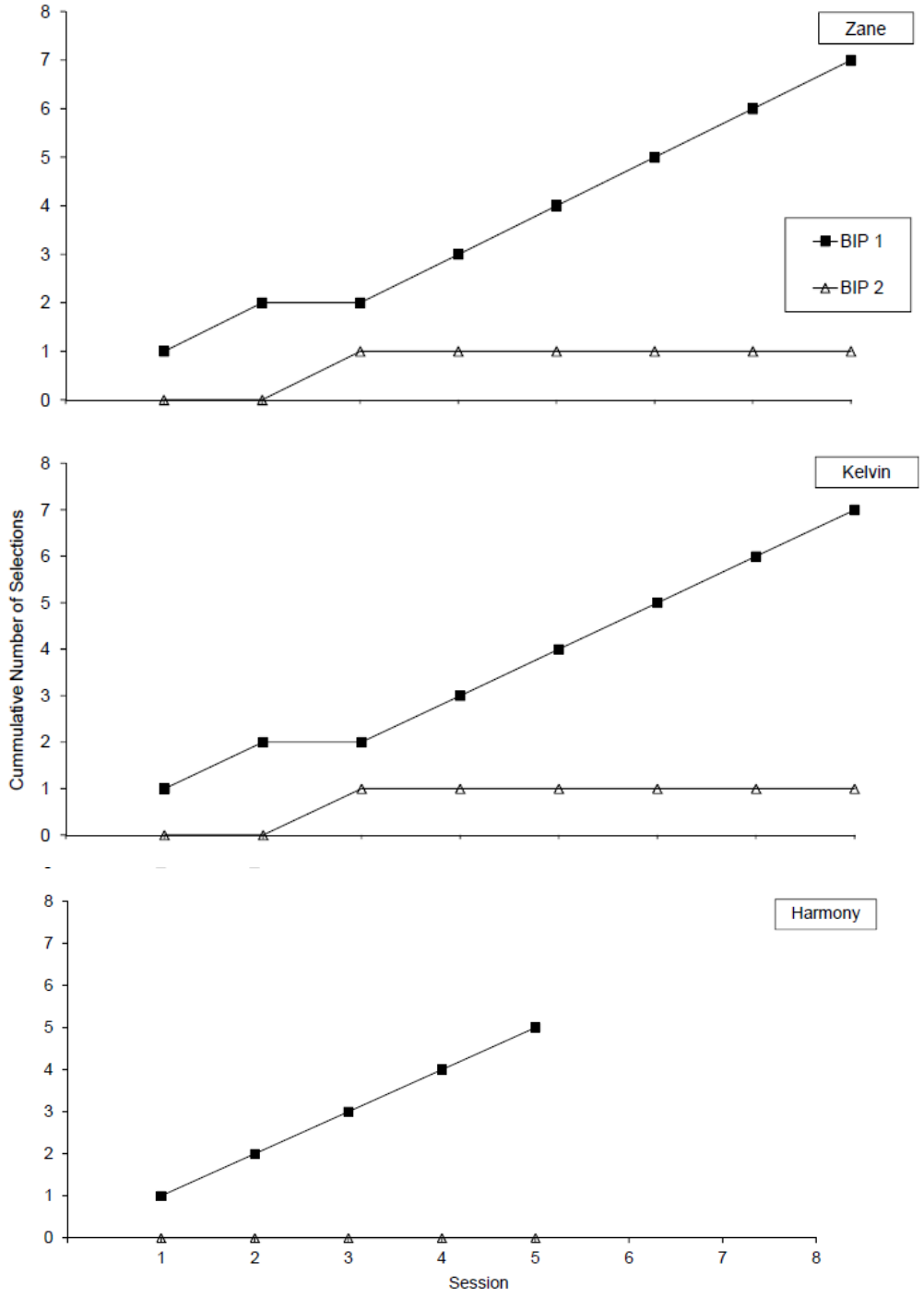


Figure 1. Cumulative number of BIP selections for Zane (first panel), Kelvin (second panel), and Harmony (third panel).

Table 3 summarizes the teachers' ratings of treatment acceptability and perceived effectiveness of BIP 1 across all three children. In general, the teachers' ratings of the BIP were positive. The mean rating across all questions and all students was 5.1 (range, 3.0 to 6.0). Across all three students, the teachers slightly agreed to strongly agreed that the BIP was an acceptable intervention for the child's problem behavior (M = 5.0). The teachers also agreed to strongly agreed that the intervention would be appropriate for a variety of children and that they liked the procedures used in the BIP (M = 5.3). Finally, the teachers slightly agreed to strongly agreed that the overall BIP was beneficial for the child (M = 5.0).

During the one-month follow-up observations, the teachers continued to implement the BIP selected by the students, and mean treatment integrity during these observations was above 90% (range, 91% to 97%) for all three students. Thus, teacher nonverbal behavior corresponded to their verbal behavior regarding the acceptability of the intervention. That is, immediately following the choice phase, both teachers reported that the BIP selected by the child was acceptable, and these verbal reports were confirmed by maintenance of the intervention over time.

Discussion

In the current study, we used three different methods of social validity to assess the acceptability of a BIP with both direct and indirect consumers. All three measures of social validity (student selection, teacher verbal report, and teacher maintenance of implementation) identified an intervention that was acceptable to all parties. To our knowledge, this is the first study to incorporate three different measures of social validity, and one of the only studies to evaluate social validity of procedures that were implemented across an entire school day. When the students were given a choice between two BIPs, all three children reliably selected one of the interventions, and this intervention was reportedly acceptable to the teachers. The teachers reported finding the choice procedure acceptable and manageable for classroom use. The nonverbal behavior of the teachers also suggested that the BIP selected by the students was acceptable because they continued to implement the intervention with integrity over time.

The results of the current study showed that high levels of treatment acceptability for a BIP were obtained across all three measures of social validity. Future studies should evaluate the consistency with which multiple measures of social validity converge. If multiple measures consistently converge, then the easiest or most efficient method of assessing social validity may be sufficient to ensure the acceptability of interventions. For example, assessing social validity through the teachers' verbal report required the least amount of time when compared to the other social validity measures we assessed. Thus, verbal reports may be a preferable measure if they are found to consistently match other, direct measures of acceptability.

There are currently no guidelines regarding how to select interventions when multiple measures of social validity do not converge, or when the most acceptable

intervention is not the most efficacious. We had existing evidence that both interventions were equally efficacious for the students. However, clinicians may not always have a priori information about the efficacy of potential interventions. Alternatively, an intervention known to be less effective may be more preferred by one or more stakeholder. When social validity and efficacy do not align, careful consideration must be given to the context in which the intervention will be implemented. Whenever possible, effective treatments should be developed that incorporate components with high social validity.

Previous research has typically relied on verbal reports as a measure of social validity (Spear, Strickland-Cohen, Romer, & Albin, 2013). In the current study, the verbal reports of teachers were confirmed by direct measure of treatment integrity over time. Direct measurement of integrity may be a useful addition to the literature on measurement of social validity. This addition may be particularly important because individual's verbal reports may not match their observed behavior. In the current study, both the direct and indirect measures of social validity indicated that the teachers found the behavioral interventions acceptable. Future studies should examine the extent to which these measures correspond when acceptability on one measure is low.

Our measurement of social validity for the teachers was limited in at least two ways. First, our only direct measure of teachers' social validity was through the continued use of the intervention over time. We could have also directly measured social validity by replicating our concurrent-chains procedure with the teachers. Future studies could evaluate the extent to which teachers' choices for BIPs matches those of the students. Second, we only had teachers rate the acceptability of the BIP selected by the students. Although the teachers rated the selected BIP highly, it is possible that teachers would have found both interventions to be equally acceptable.

Quantifying agreement between different measures of social validity also warrants further investigation. We obtained global agreement between different measures of social validity, but found it difficult to quantitatively compare across the measures. For example, how much treatment integrity must be maintained over time for the results of this kind of social validity to be said to correspond with high ratings on a social validity questionnaire? How much endorsement is needed on an indirect measure for the intervention to be considered valid? Across what timespan should direct measures of social validity be collected to be an accurate indicator of the acceptability of the treatment?

We obtained only a direct measurement of social validity from the students. Yet, indirect social validity measures have been developed for use with children (e.g., Children's Intervention Rating Profile; Witt & Elliot, 1985). To our knowledge, there are no direct comparisons of children's verbal reports of treatment acceptability and nonverbal selections. However, previous research suggests that there may be a high degree of correspondence between verbal and nonverbal measures of stimulus preference for children who have age-appropriate language (e.g., Northup, Jones, Broussard, & George, 1995), suggesting that correspondence between direct and indirect measures of validity is possible with young informants. Future studies may wish to directly evaluate the extent to which

students' verbal reports of social validity correspond to their choices in a concurrent-chains arrangement.

Overall, the results of the current study suggest that direct measures of social validity can be applied to complex behavior intervention plans for elementary students who engage in chronic and severe challenging behavior. Our results suggest that direct measures of social validity may be possible as part of classroom procedures for special-education students, and that such measures can incorporate both student and teacher responses. Despite these promising initial outcomes, there is still much work to be done to determine best practice for evaluating social validity in complex educational environments.

References

Bannerman, D. J., Sheldon, J. B., Sherman, J. A., & Harchik, A. E. (1990). Balancing the right to habilitation with the right to personal liberties: The rights of people with developmental disabilities to eat too many doughnuts and take a nap. *Journal of Applied Behavior Analysis, 23,* 79-89. doi:10.1901/jaba.1990.23-79

Fenerty, K. A., & Tiger, J. H. (2010). Determining preschoolers' preferences for choice-making opportunities: Choice of task versus choice of consequence. *Journal of Applied Behavior Analysis, 43*, 503-507. doi: 10.1901/jaba.2010.43-503

Gresham, F. M., & Lopez, M. F. (1996). Social validation: A unifying construct for school-based consultation research and practice. *School Psychology Quarterly, 11,* 204-227. doi: 10.1037/h0088930

Hanley, G. P. (2010). Toward effective and preferred programming: A case for the objective measurement of social validity with recipients of behavior-change programs. *Behavior Analysis in Practice*, *3*, 13-21.

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis, 30*, 459-473. doi: 10.1901/jaba.1997.30-459

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, *1*, 427-452. doi:10.1177/014544557714001

Kennedy, C. H. (2002). The maintenance of behavior change as an indicator of social validity. *Behavior Modification*, *26*, 594-604. doi:10.1177/014544502236652

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice, 16,* 191-198. doi: 10.1037/0735-7028.16.2.191

Mautone, J. A., DuPaul, G. J., Jitendra, A. K., Tresco, K. E., Vilejundo, R., & Volpe, R. J. (2009). The relationship between treatment integrity and acceptability of reading interventions for children with attention-deficit/hyperactivity disorder. *Psychology in the Schools, 46*, 919-931. doi: 10.1002/pits

Northup, J., Jones, K., Broussard, C., & George, T. (1995). A preliminary comparison of reinforcer assessment methods for children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, *28*, 99-100. doi:10.1901/jaba.1995.28-99

Schmidt, A. C., Hanley, G. P., & Layer, S. A. (2009). A further analysis of the value of choice: Controlling for illusory discriminative stimuli and evaluating the effects of less preferred items. *Journal of Applied Behavior Analysis*, *42*, 711-716. doi:10.1901/jaba.2009.42-711

Slocum, S. K., & Tiger, J. H. (2011). An assessment of the efficiency of and child preference for forward and backward chaining. *Journal of Applied Behavior Analysis, 44,* 793-805. doi: 10.1901/jaba.2011.44-793

Spear, C. F., Strickland-Cohen, M., Romer, N., & Albin, R. W. (2013). An examination of social validity within single-case research with students with emotional and behavioral

disorders. *Remedial and Special Education*, *34*, 357-370. doi:10.1177/0741932513490809

St. Peter Pipkin, C., Vollmer, T. R., & Sloman, K. N. (2010). Effects of treatment integrity failures during differential reinforcement of alternative behavior: A translational model. *Journal of Applied Behavior Analysis, 43*, 47-70. doi: 10.1901/jaba.2010.43-47.

Tiger, J. H., Hanley, G. P., & Hernandez, E. (2006). An evaluation of the value of choice with preschool children. *Journal of Applied Behavior Analysis*, *39*, 1-16. doi:10.1901/jaba.2006.158-04

Tiger, J. H., Toussaint, K. A., & Roath, C. T. (2010). An evaluation of the value of choice-making opportunities in single-operant arrangements: Simple fixed- and progressive-ratio schedules. *Journal of Applied Behavior Analysis, 43*, 519-524. doi: 10.1901/jaba.2010.43-519

Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed), *Advances in school psychology* (Vol. 4, pp. 251–288). Mahwah, NJ: Erlbaum.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*, 203-214. doi:10.1901/jaba.1978.11-203

Appendix

BIP 1 Daily Script

At the beginning of the day the teacher should meet with the target student, show him/her the point card and provide the following instructions:

"While you are working you will earn smiles for being safe, respectful, and responsible."

"You can earn a smile for being safe if you have safe hands, you stay in your area (i.e., you stay inside the taped area), and you use the materials in the classroom appropriately (for example, you keep your desk on the ground, and you keep your papers and pencils on your desk)."

"You can earn a smile for being respectful if you use nice words when talking to your teachers and friends and you have a quiet voice while you work."

"You can earn a smile for being responsible if you do your work and you follow your teacher's directions."

"If you get all 3 smiles by being safe, respectful, and responsible, then you will get to have a break with items in the bin with either 2 or 3 smile faces!"

"If you get 2 smiles, then you will get to have a break with the items in the bin with 2 smiles."

"If you get less than 2 smiles, then you will not get a break, and will have to continue working at your desk and try to work for smiles for the next break time for being safe, responsible, and respectful."

If you get a total of (goal number) smiles by the end of the day, then you will be able to pick a prize from the prize box.

At the end of the day the teacher should meet with the target student, show him/her the point card and review his/her goal for that day and the number of smiles earned:

"Today you earned ____smiles for being safe, respectful, and responsible while you were working. Your goal was to earn ____smiles."

"Great job reaching your goal! I am so proud of you and you can pick a prize from the prize box!"

"You did not reach your goal today so you do not get to pick a prize, but you can work hard tomorrow to reach your goal."

BIP 2 Daily Script

At the beginning of the day the teacher should meet with the target student, show him/her the break/work board and his/her picture and provide the following instructions:

"At the start of work, your picture will be on the break side of the board."

"If you have safe hands, you stay in your area (i.e., you stay inside the taped area), and use the materials in the classroom appropriately (for example, you keep your desk on the ground, and you keep your papers and pencils on your desk), then your picture will stay on the break side of the board."

"When the timer goes off if your picture is still on the break side of the board, then you will get to take a break in the break area. For every break you get in a row, you will get a tally above your picture, and if you have 2 or more tallies above your picture, you can play with an I-pad on your break."

"If while you are working you do not have safe hands, you do not stay in your area (i.e., you step outside of the tape), or you do not use materials appropriately your picture will be moved to the work side of the board. And any tallies above your picture will be erased."

"When the timer goes off if your picture is on the work side of the board, then you will have to stay at your desk and work, and you can try and earn the next break."

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1657 - 1672

Gathering Evidence for Distance Education[1]

Philip N. Chase[2]
University of Massachusetts Medical School

Abstract

A technology for evaluating computer-based distance education curricula for children and people working with children is described. The technology originated from a model of evaluation described by Markle (1967). The components were elaborated through data-based decisions reported in technical reports for a reading acquisition program, two math programs, a curriculum for people with autism, and a professional development program for clinicians working with children and adolescents. The article integrates single-case and group evaluation strategies, and draws attention to the need for better data in evidence-based decisions, and the use of data in continuous improvement efforts. Details concerning the individual learner at the developmental level of evaluation are emphasized, including an illustration of an e-learning rubric assisting this level of evaluation.

Keywords: E-learning, Distance Education, Evidence-based curricula, Evaluation, Single-case Research, Rubrics, Expert Review

Juntando Evidencia para la Educación a Distancia

Resumen

Se describe una tecnología por computadora para evaluar curricula para la educación a distancia para niños y personas que trabajan con niños. La tecnología se originó de un modelo de evaluación descrito por Markle (1967). Los componentes se elaboraron a través de decisiones basadas en datos, publicadas en reportes técnicos sobre un programa de adquisición de la lectura, dos programas de matemáticas, un curriculum para personas con autismo y en el desarrollo de un programa para clínicos que estaban trabajando con niños y adolescentes. El artículo integra estrategias de evaluación de un solo caso y de grupos y hace hincapié en la necesidad de obtener mejores datos para la toma de decisiones basada en evidencia y para el continuo mejoramiento de los esfuerzos. Se enfatizan los detalles relativos al aprendiz individual a un cierto nivel de desarrollo y evaluación, incluyendo una ilustración de una rúbrica de un e-aprendiz asistiendo este nivel de evaluación.

Palabras Clave: E-aprendizaje, Educación a Distancia, Curricula Basada en Evidencia, Investigación de un Solo Caso, Rúbricas, Revisión por Expertos

Original recibido / Original received: 18/06/2014 Aceptado / Accepted: 01/10/2014

[1] The author wishes to acknowledge the following people for their assistance in developing the methods described in this article: Robert Collins, Chata Dickson, Charles Hamad, T.V. Joe Layng, Andrew Lightner, Harold Lobo, Kristin Mayfield, John Rochford, Janet Twyman, and Vennessa Walker. Support for developing these methods came from grant #10009793-1003772R from iLearn, Inc. to West Virginia University; and Interagency Service Agreement # CT EHS 8UMSCANSISA0000001CB between the University of Massachusetts Medical School and the Children's Behavioral Health Initiative of the Executive Office of Health and Human Services, Commonwealth of Massachusetts.

[2] Correspondence: Philip N. Chase, Ph.D., E.K. Shriver Center, University of Massachusetts Medical School, 465 Medford Street, Charlestown, MA 02129

This article describes issues that have arisen while developing a technology for evaluating computer-based distance education curricula for children and people working with children. The article extends a series of editorials I wrote for the *Current Repertoire,* the newsletter for the Cambridge Center for Behavioral Studies between the winter of 2008 and the spring of 2010 (Cambridge Center for Behavioral Studies, 2014). The article also uses data collected and reported in technical reports for a reading acquisition program, two math programs, a curriculum for people with autism, and a professional development program for clinicians working with children and adolescents. My goal is to report on best practices for evaluating e-learning from a behavior analytic perspective.

A New Dawn for Behavior Analysis

A new dawn has risen for behavior analysts. We have a wonderful opportunity to accomplish many things today because so many people are responding positively to our science. Parents, pediatricians, psychologists, and teachers opt for behavioral treatment plans for people with autism and other developmental disabilities. Zoos and pet owners hire behavior analysts to solve significant problems related to human interaction with other species. Managers, front-line supervisors, workers, and unions recognize the importance of behavioral safety. Record numbers of people attending behavioral conferences attest to these positive reactions from the culture at large. These successes have positioned behavior analysis to have an impact on other areas of human concern involving learning, like the development of e-learning or distance education.

I suggest that we should tread carefully. Behavior analysts have squandered their influence on education before. The history of two significant educational innovations by behavior analysts, Programmed Instruction (PI) and the Personalized System of Instruction (PSI) are informative (Bernstein & Chase, 2012). Both PI and PSI were successful for short periods of time in the main culture. Despite the best efforts of researchers and curriculum designers like Donald Cook, Francis Mechner, Susan Markle, James Holland, Beth Sulzer-Azaroff, and others, quality control lapsed, and so did PI. Similarly, despite the work of many who showed repeatedly that PSI was superior to lectures (e.g., Johnson & Ruskin, 1977; Kulik, Kulik, & Cohen, 1979), adopting the structure of PSI without integrating thorough evaluation did not change the modal method of teaching in universities: we still lecture. Even many forms of distance education try to maintain features of the lecture method, e.g., Harvard's HBX Live (Lavelle & Ziomek, 2013).

I submit that our greatest care should come from assuring that we do not give short shrift to quality control: collecting and communicating the evidence behind our successes. One of the primary technologies of behavior analysis is the technology of gathering evidence about behavior change. In what follows, I will address evidence-based practices in the development of curricula. I will describe some of the general strategies with details--the tactics being developed through our work -- that turn practices into technological solutions to curriculum problems. Like most technologies, behavioral technologies are tied to the critical feedback

provided by scientific methods. Without this feedback, the enterprise collapses. Behavior analytic solutions, like those from any field, are only as good as their last evaluation, and evaluation is only as good as the methods used. What are these methods?

Methods of Evaluation

Behavior analysts typically use experimental methods to evaluate their work. An experiment involves the manipulated comparison of a phenomenon under two or more conditions to minimize plausible alternative explanations (internal validity) and test the generality of results across contexts (external validity), while demonstrating reliability of measurement and replication of procedure. An experimental analysis allows investigators to gather strong evidence that can support an educational practice or show that it does not work.

The current standard for experimental analyses used by educators is the random control experiment or trial (RCT). Educators widely accept the RCT because the logical coherence of a random controlled experiment is exceedingly simple to understand. Random sampling from the population suggests that findings will apply to members of the defined population. Dividing a sample into two or more groups and randomly assigning members to the comparison conditions minimizes alternative explanations for the results. One can answer questions concerning the external and internal validity of a particular e-learning curriculum in a few well-designed studies if one uses an RCT strategy effectively.

Random controlled experiments, however, are only one kind of experiment. Discuss the field of developmental disabilities for a nanosecond and one encounters the concept of functional analysis. As an experimental tactic, functional analysis evaluates which consequences are likely to support a problem behavior. In the simplest case, a clinician manipulates one of the purported reinforcers, for example, attention. Across repeated manipulations, if one finds an increased likelihood in a problem behavior when attention is presented contingent on the problem behavior and no increased likelihood in the problem behavior when attention is presented non-contingently on the problem behavior or contingent on another behavior, then the clinician may conclude that attention functions as a reinforcer. The logic of a single-case experiment illustrates its clarity. The repeated manipulation of a variable across time with one individual helps us understand a functional variable for this individual. This logic suggests why single-case experiments became a powerful part of the technology of evaluation for behavior analysts.

Rather than seeing RCTs and single-case experiments as two parts of an experimental strategy to gather evidence, however, behavior analysts and other educational scientists often have butted heads over experimental tactics. “Us vs. Them” arguments have delayed an integrated approach to evaluation. Yes, many bad decisions have been made using inferential statistics poorly to back up the findings of an RCT (e.g., Branch, 1999) and yes, there are practical problems with RCT’s in schools (e.g., the large number of schools, teachers, and students create administrative road blocks). Single-case experiments also have been criticized for

potential subject biases, lack of generality, and lack of standards for evaluating results (Horner & Spaulding, in press). These problems have been addressed and educators have finally agreed on standards for using single-case methods as well as RCT's to evaluate educational practices. (Horner & Spaulding, in press; Kratochwill, Hitchcock, Horner, Levin, Odom, Rindskopf, & Shadish, 2010). I will not discuss these issues further. Instead I will focus on how a combination of single-case experiments, group experiments, and other forms of evaluation can provide strong evidence of what works for developing e-learning curricula for children and professionals working with children.

The model for curriculum evaluation that my colleagues and I have used as the foundation of our evaluations of curricula, particularly those used in distance education, has been available for many years. Markle (1967) described key components of this evaluation model in a chapter on programmed instruction. Markle described three levels of evaluation that synthesized single-case, small n experiments, and large n group methods. She named these levels developmental, validation, and field-testing. They map well onto current stages of evaluation described by the U.S. Department of Education as development, validation, and scale-up. What follows illustrates these levels of evaluations from our recent work.

Developmental Evaluation. First, intensive individual interactions evaluate the development of an educational practice and document its effectiveness. Markle refers to this level of evaluation as developmental testing. For developmental testing, the curriculum designer/evaluator examines communication problems, learning problems, and motivation problems. Because this level of evaluation has received less attention than single-case and group experimental methods, my colleagues and I designed our own two-pronged tactic. One prong implements a rubric that a trained instructional designer uses to check the accordance of online instruction with the best practices of applied behavior analysis (Bernstein & Chase, 2012), universal design (Universal Design, 2012) and accessibility (WEBAIM, 2012). The other prong of our developmental testing involves frequent interaction with the learner as they progress through the material. We iterate between these tactics while evaluating curricula.

Our use of a rubric for evaluating instructional practices began with selecting tools to assist in writing computer-based instruction (Chase, 1985). Figure 1 illustrates the general characteristics of a rubric that has evolved since 1985. Educators use rubrics to score complex behavior. Rubrics typically involve at least two dimensions, a list of features and a scoring guide. In our case, we developed the rubric to score and track the complex outcome of developing e-learning curricula. The 8 domains of the rubric are listed on the left of Figure 1: Learning and Motivation, Data Collected and Reported, Plain Language and Readability, Use of Updated Technology, Transformability, Multi-modal, Focus and Structure, and Assistive Technology.

Domains	Comments	Absent	Weak	Adequate	Strong	Excellent
1. Learning and motivation						
2. Data collected and reported						
3. Plain language and readability						
4. Updated use of technology						
5. Transformability						
6. Multi-modal						
7. Focus and structure						
8. Assistive technology						

Figure 1. Instructional design rubric with general domains as rows and comments and ratings as columns.

Each domain on the rubric can be commented upon and rated on the five-point scale listed on the top of the tool. We also expand each row or domain on the rubric to a set of features and rate and comment on each of these features. Figure 2 shows a representative subset of the 17 features of the Learning and Motivation domain**.** The features of Learning and Motivation come from a very strong tradition of experimental evidence. As we audit the instruction we ask: Does the instruction provide sufficient examples to test for discrimination between classes and generalization within classes of stimuli? Does the instruction include immediate, frequent, and differential consequences?

1. Learning and motivation	Comments	Absent	Weak	Adequate	Strong	Excellent
8. Mastery requirements						
12. Discrimination among classes assessed						
13. Generalization among classes assessed						
14. Immediate consequences						
15. Frequent consequences						
16. Differential consequences						

Figure 2. Instructional design rubric: Representative features of the Learning and Motivation domain with illustrative features listed in the rows, and comments and ratings in the columns.

We used the rubric as part of an evaluation of Headsprout.com, an English language reading acquisition program. My colleagues and I conducted an expert or peer review of Headsprout as well as an experimental evaluation of it in two kindergarten classes (Walker & Chase, 2006). We also evaluate two math curricula from iLearn.com: iPass and Thinkfast (Chase, Dickson, Alligood, Lobo, Walker, & Cook, 2007; Chase, Dickson, Alligood, Lobo, & Walker, 2008). Again our evaluation included a review using a version of the rubric and an experimental analysis in our lab with children from the community. A team of experts also evaluated the Autism Curriculum Encyclopedia® (ACE) curriculum from the New England Center for Children using a version of the rubric (Chase, Alai-Rosales, Smith, & Twyman, 2012). ACE is a web-based toolkit providing special educators with an evidence-based program to effectively assess, teach, and evaluate individuals with autism. Most recently we have used the rubric to review the Child and Adolescent Needs and Strength (CANS) Training program for the state of Massachusetts (Bondardi, Chase, Hall, Lauer, & Nubrett-Dutra, 2013). I will use our evaluation of CANS to illustrate the tool.

CANS is a communication and care coordination instrument. It supports decision-making, facilitates quality improvement initiatives, and helps monitor the outcomes of behavioral health services for children and youth. Any clinician who provides behavioral health care to a client under the age of 21 and receives funding from Mass Health in Massachusetts is required to use CANS. Mass Health is the public health insurance program for low- to medium-income residents of Massachusetts.

Clinicians must be certified to use CANS with clients. Our evaluation focused on the e-learning certification training designed by a team from the University of Massachusetts Medical School. Our review required frequent iteration between the training and testing materials and the features of the rubric. We surveyed as many components of the training as possible from beginning to end. Then we developed questions to examine various features of the training. We returned to the beginning of the training and read, watch, and listened to each component, attempting to answer questions generated from the survey as well as creating further questions.

During the audit, the features prompted by the tool were noted qualitatively, rated, and the notes and ratings became the substance of the review. The rubric was examined frequently to assure a thorough review of all its features. For example, one of the accessibility features included checking for delivery in multiple modalities. As we examined the training we questioned whether critical components included text, audio, and rich media. We tested these features. And then we completed the comments and rated them before moving on to other features that were being checked.

The review for the domain of multi-modal is shown in Figure 3 for the CANS certification training.

6. Multi-modal	Comments	Absent	Weak	Adequate	Strong	Excellent
1. Content in multiple mediums	Videos not included for some critical learning			x		
2. Video and audio alternatives	Alternatives throughout, though not inspirational				x	
3. Text alternatives	A little confusing in placement				x	
4. Closed captioning						x
5. Illustration, diagrams, icons, and animations used to convey complex information	Inconsistent in placement			x		
6. Pair icons, graphics, etc. with text	Some alerts				x	

Figure 3. Instructional design rubric with the Multi-modal domain completed for CANS certification training.

A sample of the strengths and recommendations indicating how the rubric is translated into a review (Bondardi, Chase, et al., 2013) is provided below:

> A simple, dignified, no-nonsense, well-designed interface allows a straightforward navigation through the materials. Simple language use with jargon and abbreviations kept to a minimum help the leaner understand the material. Intermittent tasks for learners are provided to check their understanding with clear, immediate, and frequent feedback on their responses. Training ends with full case practice examples (vignettes) that helps integrate learning. In addition, multi-modal training is used throughout with closed captions and transcripts for voice and videos that helped focus training on some of the more difficult domains and items within domains.
>
> As currently designed, however, the training and testing do not provide sufficient interactions with a range of examples to teach discriminations between some of the most difficult items and generalization within these items. Further, if the learner responds incorrectly to the questions provided, there is no chance for the learner to recheck learning with a new question on the same item. (Bondardi, Chase, et al., p. 5).

In sum, the rubric helps set standards for peer/expert reviews of online /distance education. It prompts us to examine various features of the curriculum and how it is presented online. It helps to make reviews and critiques efficient. Most importantly, it synthesizes what we know from behavioral education with what we know from accessibility into one set of standards that we can apply to online instruction and training.

Peer or expert review, however, is not sufficient even when standardized as we have done. Like the problems with computer software that arise when checked and tested only by software engineers, end-users (typical learners) should test educational programs. The end-user evaluation we conducted for the iPASS math curriculum illustrates how we interacted with the students as they progressed through the curriculum. iPASS is a web-enabled mathematics curriculum for middle-school students used in several states in the US (iLearn.com, 2014). Teachers and students use iPASS as either a primary or supplementary source of mathematics instruction. The software automates many aspects of the instructional process, including placement, assessment, instruction, remediation, and tracking of student performance.

In addition to an expert team reviewing and critiquing iPASS using the rubric, students of iPASS completed a report-on-problems that asked them to detail anything they found to be problematic or frustrating with the iPASS program. The form prompted the student to describe problems and identify where in the curriculum the problem occurred. It also asked them for their level of agreement with a general question about their understanding of the exercises they had completed. The form was place on the desk next to their computer key-board and they were asked to complete the form as they worked through iPASS. At the end of each session, experimenters questioned the students further and then summarized

these reports into a spreadsheet. Part of the students' payment contingency was to suggest changes. Along with problems the evaluators found during their expert review and from student performance data, reports from students were included in the developmental review of the iPASS curriculum (Chase et al., 2007).

Validation Evaluation. One problem with the developmental level evaluation described here concerns basing evidence simply on a combination of expert opinion and simple data from a few students. Such comparisons hardly minimize alternative explanations for the results and they certainly do not provide confidence in external validity. They do not answer questions such as: Could the results be an artifact of something else? Plausible alternatives can be eliminated using Markle's second level of evaluation: the validation stage. Validation testing investigates the extent to which the curriculum meets its own goals under controlled circumstances. Validation testing evaluates internal validity. Some questions concerning external validity can be answered as well, but these questions will be discussed in more detail in the third or field-testing stage. Our validation stage evaluation starts with an experimental analysis.

For the purposes of internal validity we need to control for alternative explanations. One experimental strategy we have used is the single-case design called a Multiple Baseline Achievement Test design (MBLAT) (Miller & Weaver, 1972). It is an example of a multiple-baseline across behaviors design. As such it involves frequent assessment of the multiple behaviors of at least one student changing from baseline to treatment to minimize alternative explanations. Repeatedly examining baseline to treatment changes with baselines of different lengths of time, treatments of different lengths of time, and replication across behaviors, assesses the contribution of the treatment. The treatment in our use of the MBLAT is exposure to an e-learning curriculum. The behavior of individuals in the experiment is the target of evaluation. The changes examined include those of level, trend, and variability of performance across phases.

Figure 4 provides an example with hypothetical data using iPASS as the example of a treatment. Parallel or identical tests are given repeatedly over time (the successive tests are displayed on the x-axis in Figure 5) and each test assesses the same material from a curriculum or a component of a curriculum. In a hypothetical case, the component of the curriculum consists of three math units: one on addition, one on multiplication, and one on word problems involving addition and multiplication. We also divide the test into sections of items that are aligned to the units of the course. The y-axis records the dependent variable (e.g., % correct on each section of the test). The baselines refer to performances on items aligned with each unit **before** the lessons for a unit are provided. The treatment refers to performances on items aligned with each unit **after** treatment (e.g., iPASS math instruction). If training is effective and the test items for each unit are independent of each other, then one should see changes in test performance related to each unit only when the unit material has been taught (post treatment). As illustrated, the changes evidenced were clear changes in level as percent correct performance always increased after instruction and never before. Changes in trend and variability across phases were not evident.

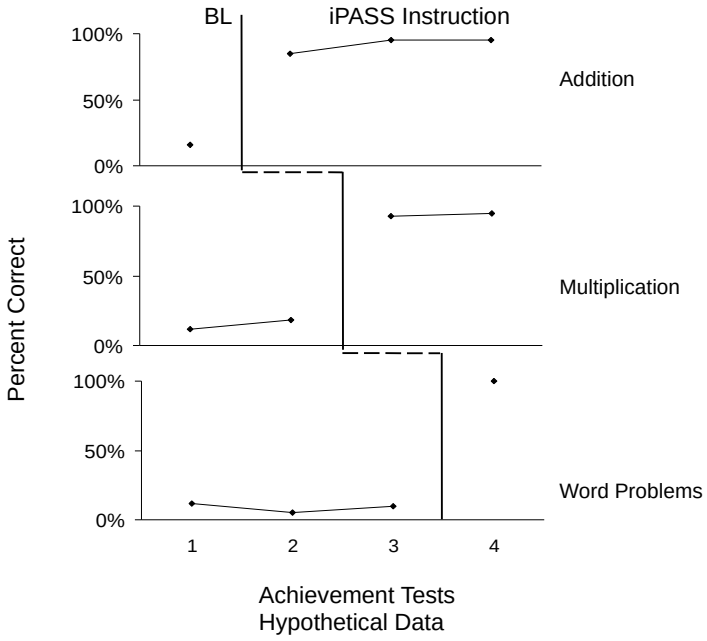


Figure 4. Hypothetical MBLAT across four achievement tests with iPASS Instruction as the treatment, percent correct as the dependent measure, and solving three types of math problems as the behavior.

What are the advantages of the MBLAT design for establishing the validity of a curriculum? First, it is a practical design. Behavioral educators may use the design with many different kinds of curriculum, at many different times in the school year, and with many students or group of students. The design does not depend on random assignment to draw conclusions about the internal validity of the evidence and minimizes the practical problems of using random assignment in ongoing school environments. Students may be randomly selected to participate in the experiment, which strengthens the conclusions one can draw from the MBLAT.

The following list of threats to internal validity adapted from Kazdin (2003) can be examined with the MBLAT: subjects, history, maturation, attrition, selection biases, settings, measurement, instruments, and the adequacy of independent variable (IV) descriptions and definitions (e.g., special treatment in experimental vs. control conditions or diffusion of treatment across conditions). A threat is any known variable that co-varies with the treatment and thus could be a plausible alternative explanation for any changes seen in behavior. To evaluate threats we ask: do subjects, settings, etc., co-vary with changes in conditions.

Many threats to internal validity related to the participants or subjects can be examined with the MBLAT design because the subjects receive both baseline and treatment conditions. If the participant's history concurrent with and outside of the experiment has an impact on their behavior, then it can be evaluated by the staggered introduction of the treatment. Although often described differently,

biological maturation can be described as a plausible change in history. Training research often involves repeated measures over long time periods, and therefore, the participants may experience maturational changes during their time in the experiment that affect learning. These maturational changes can be assessed by the staggered introduction of the treatment. For example, IPASS is a year-long curriculum for children between the ages of 10 and 15, and biological maturation could conceivably affect their behavior. The staggered baseline and treatment phases of the MBLAT allow the researcher to examine whether maturation, the repeated measures themselves, or other variables that occur over the history of the experiment might affect performance. If such historic variables have an impact, the effect would be seen at times other than the introduction of the treatment.

Historic differences among participants prior to the experiment also do not co-vary with the treatment and therefore these threats to internal validity are handled directly by the MBLAT design. If the participants' history with math prior to the experiment allowed them to perform well on the tests, the baseline conditions reveal this. We then assess further changes after treatment. Although differences in baseline performance often occur among participants attributed to historic variables, it is the change in level, trend, and variability of performance from baseline to treatment, replicated across sets of behavior that the MBLAT helps us examine.

Other subject threats, like attrition, frequently pose problems for educational research. Students leave school, transfer classes, and move from school-to-school, so when evaluators use a group design they have to assure that attrition for students who receive a treatment does not differ from those who receive a comparison condition. The MBLAT manages attrition threats again by having each participant as his or her own control. If they leave the experiment it might be costly to the evaluation, but one cannot attribute treatment effects to differential attrition. Attrition affects on external validity are not handled by the MBLAT design as will be described below. Additionally, whether the treatment contributes to attrition cannot be assessed within a participant using the MBLAT, but can be assessed across participants. We can examine attrition as a dependent variable as we add participants: what proportion of the participants leave the experiment during baseline compared to those who leave during treatment conditions?

Other threats such as setting, measurement, and independent variable definitions are handled by assuring that each subject receives the same tests, in the same settings, across all conditions. For example, iPASS uses computers to present the curriculum, therefore, we used computers during baseline instruction. We also used the same tests during baseline and treatment, the same computers were used for testing, and the same people administered testing across conditions. Diffusion between conditions exemplifies problems related to the definition of the independent variable if the baseline and treatment conditions are too similar to each other or influence each other. Special treatment creates another problem with the independent variable if the teachers pay more attention to the kids during treatment than they do during baseline. Controls for threats like test, setting, and IV are managed as in any carefully designed research by the use of highly specified

protocols, and data collected on whether the protocols are carried out as planned and reported as treatment integrity data.

Selection bias is a special case of historic variability. Selection bias refers to the possibility that participants selected for the study have characteristics that make it more likely that they will be affected by the treatment. Selection bias is an identifiable characteristic(s) of the students selected for the study that contributes to the effect of the curriculum. Have we selected students whose special histories allow them to do well in the curriculum? For example, social economic status (SES) factors may influence how the motivation components of a curriculum work. If the students in a study all come from financially privileged families, they may be affected by motivational variables differently than children from less financially advantaged families. Evaluators manage selection bias in RCTs, like all historical variables, by random assignment to conditions. Evaluators manage selection bias in single-case studies because the bias does not co-vary with the treatment. We select the students for the experiment and then we test them under both baseline and treatment conditions. Like other subject threats, selection bias cannot be attributed to the subjects used in the experiment, they are who they are, and if the treatment successfully changes their behavior we should see changes in level, trend, and/or variability. Again, the level of evaluation is individual behavior.

Selection bias, however, does require a little more discussion, a discussion that highlights one aspect of history that MBLAT designs do not eliminate: the interaction of variables with a treatment that could affect efficacy. Selection bias can be a threat if characteristics of the participants interact with the treatment to produce the effects found. But this is a problem of external validity true for almost all the threats I have discussed so far. For example, have we assured that the tests used are not biased toward the curriculum-for example, the problem of teaching to the test? Likewise, have we assured that characteristics of the tested students did not influence the results? As described earlier, randomly selecting students from a population would help to minimize the interaction of selection with the treatment even in a single-case design. In general, however, threats concerning interactions will be considered next under problems of external validity. The MBLAT design allows for a good evaluation of the internal validity for the children who were in the experiment as long as the children, settings, instruments, and tests do not change at the same time as the treatment. The MBLAT does not necessarily evaluate whether these changes will occur across children, settings, instruments, and tests.

Field Evaluation. External validity questions concern whether our results hold up across students, schools, teachers, tests, and other characteristics of the study. Does the study draw conclusions about other students? If so, has the experimenter arranged to test the curriculum with representative participants? Does the study draw conclusions about other environments? If so, has the evaluator arranged to test the curriculum with a representative range of environments? Does the study draw conclusions about the generality to other teachers or staff? If so, has the experimenter arranged to test the curriculum with representative teachers? Does the study draw conclusions about other tests/measures than used? If so, has the evaluator arranged to test whether the curriculum is successful with different measures?

Markle (1967) described how questions of external validity are answered with a field test. She stated that evaluations should be conducted to assess the effectiveness of the curriculum in a variety of settings and with a variety of students. More recently descriptions of such evaluations state that evaluators test the curriculum at scale, the process of “scaling up”. A series of MBLAT experiments can be designed to test various questions of generality. One of the most important questions, because of the single-case nature of the MBLAT, is whether the participants are representative of the population that might use the curriculum. A series of such studies may not be practical, however. For example, one practical problem with single-case designs is the difficulty of examining interactions, so even the combination of highly controlled studies to establish internal validity and data collection from representative samples of students, settings, and teachers to establish external validity, may not be sufficient. Once we have established internal validity with a few well-designed MBLAT evaluations, it might be more efficient to use group designs to test for external validity. Discussion of the appropriate designs to use is beyond the scope of this article, but various sources including the IES What Works Clearing House (http://ies.ed.gov/ncee/wwc/) provide useful guidelines.

One important question of external validity has arisen from our work on making distance education accessible to the widest group of students. There are many websites, curricula, and other forms of e-learning that are not accessible to a large population of people. A highly influential medical information website that we examined has flashing ads, pop-ups, cycling banners, multiple columns, and packed information all of which make it difficult for people to access the information. Using the WEBAIM Wave Program (WEBAIM, 2012), which is design to detect accessibility problems primarily related to visual difficulties, we found 17 accessibility errors on the home page of this website. While this site tries to provide a good service, important audiences cannot use it. Who am I talking about? The following list adapted from the Web Accessibility Initiative (Eichner & Dullabh, 2007) is a good start: A mother with color blindness who seeks information for her child with autism, a reporter with repetitive stress injury, an accountant with blindness, a classroom student with attention deficit hyperactivity disorder and dyslexia, a retiree with low vision, hand tremor, and mild short-term memory loss, or a supermarket assistant with Down syndrome.

According to the U.S. Census figures for 2000, 20% of Americans have a disability that impairs access to websites and Internet content. According to a 2011 report on disability from the World Health Organization, 56 million people in the U.S. were identified as having such disability. Multiply these numbers x-fold for a worldwide population of people with intellectual, cognitive, visual, and age-related disabilities who cannot access information and instruction from the internet. These people need access to online instruction and information. How can we design and evaluate online instruction that works for them? The features of evaluation that I have described throughout this article help. We directly address many issues of accessibility with the rubric used during developmental testing. We further address accessibility through field-testing. We collect data across students, across schools, and across tests for a particular curriculum, assuring that we also have evaluated

e-learning with students across critical demographic groups. Once we do so we should have sufficient evidence to establish the external validity of the curriculum's effectiveness. If we combine these data with the data from our MBLAT and the data from our developmental testing, and all of these levels of evaluation demonstrate the effectiveness of a curriculum, do we need additional evidence?

I always return to what I learned from methodologists. Have we addressed alternate explanations, issues of generality, and the practical concerns of gathering evidence? I think we can check these off if we have data that show a curriculum to be effective across students, settings, and tests as well as having internal validity from prior experimentation. If the data show that the curriculum is not effective with some particular set of students, or in some settings, or on some type of test or outcome, then the data suggest further experimentation with the variables correlated with lack of success.

But even if we do achieve internal and external validity in the manner suggested, some might still ask whether the results could be achieved faster or more reliably with another curriculum. This is a consumer driven question. A series of MBLAT studies comparing curricula across phases and with a counterbalance of the order of receiving the two curricula across students can be used to address such questions. The counterbalancing of order has some practical problems related to aligning curricula with tests, but it can be done in many situations. Of course, another solution would be to conduct a random controlled experiment that focuses on comparing the curricula of interest.

Conclusion

The idea of a major industrial concern turning out 10 percent superior products, 20 percent good products, and 50 percent average products, with the remainder classified as disposable is so ludicrous. Markle, (1967), p.104

This quote struck me when I first saw it and it still rings true today for most of what passes for educational technology: the typical measures of success of educational enterprises are absurd for those interested in replicable procedures. Why do we continue to accept them? As indicated at the beginning, Programmed instruction and PSI both made progress. I am humbled when I read a classic like Glaser (1965) or a biography, like Mechner (in press) about how much was done and how many people were educated through Programmed Instruction. So why have we not made more progress since then?

It comes down to what behavior analysts know best: whether contingencies of reinforcement support behavior required of progress. Educators have not had to demonstrate a high level of effectiveness in order to obtain reinforcers. Unlike building bridges across rivers, unlike producing computers that process data more efficiently, unlike reducing pain, teaching has had conflicting goals and uncertain outcomes. Agreed upon demands have not been placed on education to develop thorough evidence-based methods like those from other technologies.

So where do demands on evidence-based education come from? We know how demands for the behavioral services I described at the beginning of this article drove our successes. We know that the changes in the services for children with

autism and other developmental disabilities came from parents' demands. We know that changes in safety practices in industry came from the costs of injuries. We know that changes in zoo animal and pet training came from the consumers. Recently the demands on curricula in the U.S. have begun to change with the attempts to enforce standards through the No Child Left Behind Act (NCLB). The NCLB has many weaknesses. Particularly weak are those features related to the oversimplification of measurement, like standardized test performance. The use of these simple measures in a compliance model has created demands on the wrong behavior by teachers and administrators. I fear these weaknesses once again will derail attempts at data driven change. One of the more promising outcomes of NCLB, however, was the creation of the Institute of Education Sciences (IES) and IES mandates experimental validation of educational practices. For the first eight years this demand translated into using RCT as the gold standard for evidence, but IES has begun to accept single-case experimental designs (Kratochwill, Hitchcock, et al., 2010). A demand on educators in the U.S. to use RCT and single-case experiments to bolster other forms of evidence for what works in education seems to support the kinds of activities I have described here. The question remains whether behavior analysts can help meet and contribute to this demand. I think the technology of evaluation that I have discussed illustrates one behavioral technology that might help.

References

Bernstein, D. J., & Chase, P. N. (2012). Contributions of behavior analysis to higher education. In G. Madden (Ed.), *APA handbook of behavior analysis: Volume 2 Translating principles into practice* (pp. 523-543). Washington, D.C.: APA Press.

Bonardi, A., Chase, P. N., Hall, G., Lauer, E., & Nubrett-Dutra, C. (2013). *CANS training evaluation.* Technical report prepared for Interagency Service Agreement # CT EHS 8UMSCANSISA0000001CB between the University of Massachusetts Medical School and the Children's Behavioral Health Initiative of the Executive Office of Health and Human Services, Commonwealth of Massachusetts.

Branch, M. N. (1999). Statistical inference in behavior analysis: Some things significance testing does and does not do. *The Behavior Analyst, 22,* 87-92.

Cambridge Center for Behavioral Studies (October 23, 2014). *The Current Repertoire.* Retrieved from http://behavior.org/repertoire.php.

Chase, P. N. (1985). Designing courseware: Prompts from behavioral instruction. *The Behavior Analyst, 8,* 65-76.

Chase, P. N., Alai-Rosales, S., Smith, T., & Twyman, J. S. (2012). *ACE expert review*. Technical report prepared for The New England Center for Children of its Autism Curriculum Encyclopedia® (ACE).

Chase, P. N., Dickson, C., Alligood, C., Lobo, H., Walker, V., & Cook, L. (2007). *Expert review of iLearn math*. Technical Report 1 from grant #10009793-1003772R from iLearn, Inc. to West Virginia University.

Chase, P. N., Dickson, C., Alligood, C., Lobo, H., & Walker (2008). *Laboratory evaluation of iLearn math*. Technical Report 2 from grant #10009793-1003772R from iLearn, Inc. to West Virginia University.

Eichner, J., & Dullabh, P. (2007). *Accessible health information technology (health IT) for populations with limited literacy: a guide for developers and purchasers of IT.* (Prepared by the National Opinion Research Center for the National Resource Center for Health IT). AHRQ Publication No. 08-0010-EF. Rockville, MD: Agency for Healthcare Research and Quality.

Glaser, R. (1965). *Teaching Machines and programed learning II: data and directions.* Washington: D.C. National Education Association.

Horner, R. & Spaulding, S. (in press). Single-case research designs. *Encyclopedia*. Springer.

iLearn.com (October, 2014). *iLearn.* Retrieved from http://ilearn.com.

Johnson, K. R. & Ruskin, R. S. (1977). *Behavioral Instruction: An Evaluative Review*. Washington, D.C.: American Psychological Association.

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th edition). Needham Heights: Allyn & Bacon.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kulik, J. A., Kulik, C-L. C, & Cohen, P. A. (1979). A meta-analysis of outcome studies of Keller's Personalized System of Instruction. *American Psychologist, 34,* 307-318.

Lavelle, L. & Ziomek, E. (2013). *Harvard business school launching online learning initiative*. Retrieved from http://www.businessweek.com/articles/2013-10-09/harvard-business-school-launching-online-learning-initiative.

Markle, S. M. (1967). Empirical testing of programs. In P. C. Lange (Ed.), *Programmed instruction* (pp. 104–138). Chicago: University of Chicago Press.

Mechner, F. (in press). *Behavioral technology and education reform*. Beverly, MA: Cambridge Center for Behavioral Studies.

Miller, L. K. & Weaver, F. H. (1972). A multiple baseline achievement test. In G. Semb (Ed.) *Behavior analysis and education-1972* (pp. 393-399). Lawrence, KS: Support and Development Center for Follow Through, Department of Human Development, University of Kansas.

Universal Design (August 28, 2012). *The Principles of universal design.* Retrieved from http://en.wikipedia.org/wiki/Universal_design.

Walker, V. L. & Chase, P. N. (2006). *Assessment of Headsprout Basics Reading*. Technical Report for Headsprout.com.

WEBAIM, *Evaluating cognitive web accessibility with WAVE* (August 28, 2012). Retrieved from http://webaim.org/articles/evaluatingcognitive/.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1673 - 1681

Implementation of the Good Behavior Game in Classrooms for Children with Delinquent Behavior

P. Raymond Joslyn, Timothy R. Vollmer[1], & Varsovia Hernández[2]
University of Florida

Abstract

The good behavior game (GBG) is an interdependent group contingency procedure designed to reduce disruptive behavior in classroom settings. In the GBG, a class is divided into groups, simple rules are made, and contingencies are placed on the students' rule-following behavior. This procedure has been proven effective across various student ages, and its simplicity and long-term effects have contributed to its popularity in school settings. Although it has been systematically evaluated across a wide range of student ages, research on the GBG is lacking in the area of population- and setting-specific assessments. In this evaluation, the GBG was implemented in three classrooms (student ages 7-12) at a school for children with severe problem behavior; in particular, these students had already displayed behavior that might be described as "delinquent" or "predelinquent." Disruptive behavior substantially decreased across all three classrooms as a result of the intervention. This application extends the current literature by systematically replicating the results of the GBG with a different population. Implementation of the GBG, population-specific obstacles, results, and future directions are discussed.

Keywords: Good Behavior Game, Classroom, Delinquent Behavior.

Implementación del Juego del Buen Comportamiento en Aulas para Niños con Conducta Delictiva

Resumen

El Juego del Buen Comportamiento (JBC) es un procedimiento de contingencias grupales interdependientes diseñado para reducir conducta disruptiva en el salón de clases. En el JBC se divide al salón en grupos, se instauran reglas simples y se establecen contingencias sobre la conducta de seguimiento de reglas de los estudiantes. Este procedimiento ha probado ser efectivo con estudiantes de diferentes edades y su simplicidad y efectos a largo plazo han contribuido a su popularidad en ambientes escolares. Aunque ha sido evaluado sistemáticamente a través de un amplio rango de edades, hace falta investigación del JBC en el área de evaluación de poblaciones y en situaciones específicas. En este trabajo, se implementó el JBC en tres aulas (estudiantes de 7 a 12 años de edad) de una escuela para niños con problemas severos de conducta; en particular, estos estudiantes habían mostrado previamente conducta que podría describirse como "delincuente o "pre-delincuente". La conducta disruptiva disminuyó sustancialmente en los tres salones como resultado de la intervención. Esta aplicación extiende la literatura actual al replicar sistemáticamente los resultados del JBC en una población diferente. Se discute la implementación del JBC, los obstáculos específicos de la población, los resultados y las futuras direcciones.

Palabras Clave: Manejo del Aula, Juego del Buen Comportamiento, Conducta Delictiva

Original recibido / Original received: 28/05/2012 Aceptado / Accepted: 16/08/2014

[1] Correspondence: Timothy R. Vollmer. Department of Psychology, University of Florida, Gainesville, Florida 32611. E-mail: vollmera@ufl.edu

[2] The contribution of Varsovia Hernandez was supported by Grant 207587 from CONACyT

First introduced by Barrish, Saunders, and Wolf (1969), the Good Behavior Game (GBG) is now a commonly used interdependent group contingency procedure designed to reduce disruptive behavior in classroom settings. Research has demonstrated that this procedure is highly effective at reducing disruptive behavior in students of various ages (Donaldson, Vollmer, Krous, Downs, & Berard, 2011; Flower, McKenna, Bunuan, Muething, & Vega, 2014; Tingstrom, Sterling-Turner, & Wilczynski, 2006). In addition to the immediate effects produced by the GBG, longitudinal studies have shown that exposing children to the contingencies of the GBG at a young age can have lasting positive effects on students' behavior (Kellam et al., 2008). For example, Kellam et al. showed that students who were exposed to the GBG in elementary school were less likely to engage in problem behavior such as aggression and drug abuse, and were less likely to require psychiatric services in adulthood. The GBG is currently considered a best-practice behavioral classroom management procedure, and its simplicity and effectiveness has led some people to refer to it as a "behavioral vaccine" (Embry, 2002). A behavioral vaccine is defined as a repeated simple behavior that reduces morbidity or mortality and increases wellbeing (Embry, 2011).

The GBG is an interdependent group contingency in which a class is divided into two teams, simple rules are made, and contingencies are placed on students' rule following behavior. Although it has been replicated with students of various ages, research on the GBG is lacking in population- and setting-specific assessments. A replication of previous research on the GBG in a population of children who engage in delinquent behavior would provide valuable information on its efficacy in an underrepresented population as well as extend its findings into the realm of students with severe problem behavior. For the purpose of this study, delinquent behavior is defined as illegal behavior that the perpetrator is not legally responsible for due to their age. Examples of delinquent behaviors are theft and violence toward students and teachers. Students in this population are at risk for negative outcomes later in life such as drug abuse, incarceration, violence, dropping out of school, and unemployment, among others (Farrington, 1989; Fergusson & Horwood, 1998; Fergusson & Lynskey, 1998). As indicated by the longitudinal research, the GBG may have the potential to reduce or even reverse these negative outcomes (Embry, 2002; Kellam et al., 2008). Thus, the purpose of the current study was to evaluate the effectiveness of the GBG on disruptive classroom behavior at a school for children with delinquent behavior.

Method

Participants and Setting

Students from three classrooms at a school for children who engage in severe problem behavior were chosen as the subjects of this study. The students were referred to the school for severe behavioral problems that prevented them from learning and distracted other students from engaging in academic activity. The students' problem behavior can be so severe that some precautions and security measures beyond that of a typical school are in place. The classrooms are all equipped with state regulated time-out rooms, separate from the main

classroom. Each classroom is also equipped with an emergency call button that is used when one or more students display dangerous behavior and help is needed to restrain them. The school is patrolled by several uniformed police officers who provide assistance when students are dangerous, and arrest students when they commit crimes at school. Although there are some additional safeguards in place, most of the staff and teachers have minimal additional training in working with these students. This presents problems when dealing with students' disruptive and sometimes dangerous behavior. Classroom 1 was a group of 2nd and 3rd grade students (aged 7-9), Classroom 2 was a group of 4th and 5th grade students (aged 9-11), and Classroom 3 was a group of 5th and 6th grade students (aged 10-12). Each classroom had approximately 6-10 students, although the student population fluctuated due to absences, truancy, and transfers to and from other schools. Overall, classrooms were approximately 80% male and 72% African American, 24% Caucasian, and 4% biracial. Sessions occurred in the classroom during group instruction and silent work time. During these times the students were expected to remain seated at their desks, and to raise their hand if they wanted to speak. Sessions were conducted 3-5 times per week and ranged from 30 to 60 minutes in duration.

Target Responses, Data Collection, and Interobserver Agreement

Target responses were based on previous research, observations of problem behavior, and recommendations from teachers. Student behavior was scored as out of seat if they were not seated at the chair attached to their assigned desk and had not received permission to be out of their seat. Out of seat was recorded as a duration measure because students frequently stayed out of their seats for extended periods of time. Time spent out of seat began when a student got out of their seat without permission and ended when they sat back down. Duration of out of seat behavior was recorded as long as at least one student was out of their seat. Student behavior was scored as talking out of turn if they made any vocalization without first raising their hand and being called upon or addressed by the teacher. Student behavior was scored as touching another student if their hands or feet made contact with another student. In order to measure treatment integrity, experimenter tally scoring was recorded whenever the experimenter notified a student that they broke a rule and placed a tally on the board. During baseline and intervention phases, an independent observer collected data on each of the target responses for all individuals in the class using handheld computers programmed with real-time data collection software.

Interobserver agreement (IOA) was calculated using the proportional agreement method for 42% of sessions in Classroom 1, 33% of sessions in Classroom 2, and 22% in Classroom 3. The total observation time for each session was divided into 10 s intervals, records for two observers were compared within each interval and the smaller number of events (discrete responses for frequency measures and seconds for duration measures) was divided by the larger number of events recorded by an observer (agreement was considered to be 1 if no events were recorded in an interval for both observers), and these numbers were averaged across the entire session. Mean agreement for talking out was 94.8%

(range, 72.6% to 100%), 86.9% (range, 68.4% to 97.7%), and 88.6% (range, 71.9% to 98.3%) for Classrooms 1, 2, and 3, respectively. Mean agreement for out of seat was 98.1% (range, 91.6% to 100%), 97.5% (range, 94.8% to 99.4%), and 93.7% (range, 73.4% to 100%) for Classrooms 1, 2, and 3, respectively. Data for touching other students were not included in this analysis because this rarely happened. Mean agreement for experimenter behavior (tally scoring) was 98.3% (range, 90% to 100%), 98.5% (range, 96.8% to 99.9%), and 97.1% (range, 95.4% to 98.9%) for Classrooms 1, 2, and 3, respectively.

Treatment integrity was calculated by counting the total number of instances of problem behavior and the total number of tallies given for the class per session, and dividing the smaller number by the larger number. Average treatment integrity was 77.3% (range: 26.7% to 100%), 84.4% (range: 62.5% to 100%), and 74.8% (range: 38.4% to 100%) for Classroom 1, 2, and 3, respectively. Although these integrity scores are low, previous research has shown that if roughly half of the instances of problem behavior are scored, the procedure is effective (Donaldson et al., 2011).

Design and Procedure

The effects of the GBG were evaluated using a non concurrent multiple baseline across classrooms with an embedded reversal in one classroom.

Baseline. During baseline, the teachers instructed their class as they would normally. During class instruction and silent work time, standard class rules indicated that students were expected to remain seated unless they got permission to get up, remain quiet unless they had been addressed by the teacher or raised their hand and been called upon, and they were not allowed to make physical contact with other students. The teachers enforced class rules with verbal reprimand sporadically, if at all. Observations were not announced to the class and observers did not interact with students.

Good Behavior Game. Each teacher divided their class into two teams. In each class, the teacher distributed students with particularly problematic behavior evenly between the two teams to make both teams equally likely to win the game. Before beginning the game, the teams were announced to the class and written on the board. Each team and the initials of its students were written on the left side of the board with a space to the right for tally marks. The rules were also written on the board underneath the scoreboard. Prior to each session, students were reminded of the rules of the game and what they could win. The rules of the GBG were that each student must remain seated unless they have permission from a teacher to get up, students could not talk unless they had been given permission or addressed by a teacher, and students were not allowed to touch each other. A team won the game by having a lower score than the other team at the end of the session. However, both teams won the game if they met a criterion that was at least an 80% reduction in the average frequency of disruptive behavior from baseline. Rewards were selected by polling the class on what types of reinforcers they would like to work for as long as they followed school rules. These included snacks (e.g., chips, fruit snacks, sugar-free candy), tangible items (e.g., pencils, erasers, stickers), or escape from academic demands (e.g., extra free time,

computer time). Students on winning teams were allowed to choose from a selection of rewards at the end of each session.

An experimenter implemented the GBG while the teacher engaged in classroom instruction. When a rule was broken, the experimenter stated the rule (e.g., “John, you need to raise your hand to talk”), and made a hatch mark next to the corresponding team on the board. The points were counted at the end of the session, winners of the game were announced, and rewards distributed.

Results

Figure 1 shows the rate (responses per minute) of talking out of turn and the percent of session that one or more students were out of their seats without permission. Talking is graphed as rate on the left y-axis, and out of seat behavior is graphed as percent of session on the right y-axis. The data for touching another student are not presented because it happened infrequently, even in baseline. In baseline, all classrooms had high rates of talking out of turn (*M*s = 3.3, 6, and 3.9 for Classroom 1, 2, and 3, respectively) and a high percent of session out of seat (*M*s = 19.2, 16.5, and 18.6 for Classroom 1, 2 and 3, respectively). After implementation of the GBG, all classrooms saw decreased levels of talking out of turn (*M*s = 0.2, 0.3, and 0.8 for Classroom 1, 2, and 3, respectively) and percent of sessions out of seat (*M*s = 0.8, 0.4, and 1.5 for Classroom 1, 2, and 3, respectively). During treatment in Classroom 2, the teacher inadvertently added a competing reinforcer to the environment. She had begun handing out tokens to students for completing their work that could be redeemed for edible items. After this was corrected, the targeted problem behavior returned to low levels. For Classroom 1, during baseline 2 there was an increase in the rate of talking (*M*s = 1.9) and percent of session out of seat (*M*s = 9.9), compared to the previous condition. Reintroduction of the GBG in this classroom produced decreases in the rate of talking (*M*s = 0.1) and percent of session out of seat (*M*s = 0.6). Reduction in disruptive behavior for all three classrooms was well below the 80% reduction required to win the game; therefore, more often than not both teams “won” on a given day. As a measure of social validity, before sessions the teachers were asked if they wanted to implement the GBG. The answer was yes 100% of the time for all classrooms.

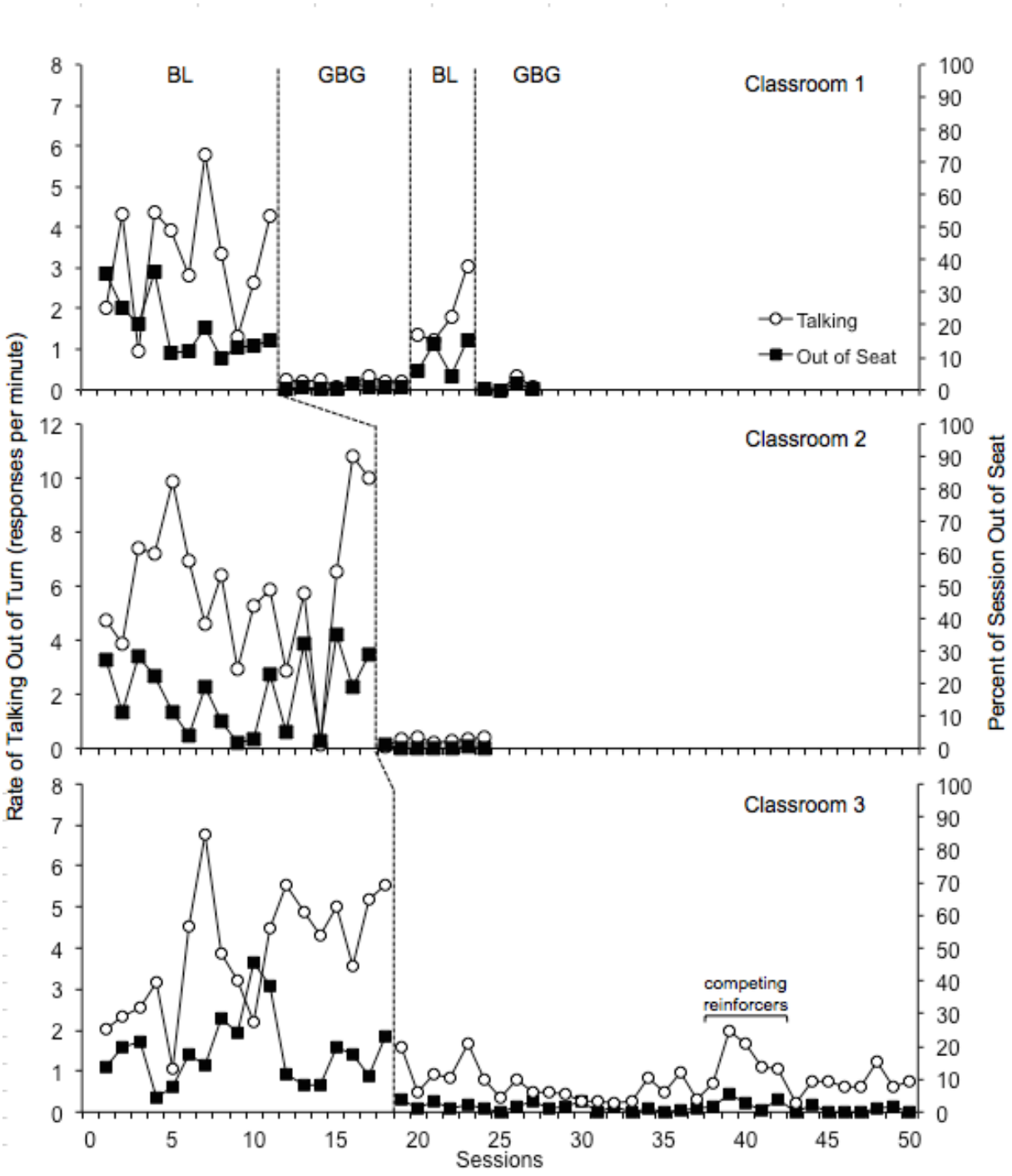


Figure 1. Responses per minute of talking out of turn and percent of session out of seat across baseline and treatment sessions in Classrooms 1, 2, and 3. BL = Baseline; GBG= Good Behavior Game.

Discussion

The findings of this study are consistent with past research on the GBG. We found that the procedures produced a marked decrease in all targeted problem behaviors. These findings indicate that the GBG is generalizable to classroom settings in which students previously displayed very severe problem behavior. Students with severe problem behavior are frequently educated at alternative learning schools (Van Acker, 2007) and it is important to note that the GBG appeared to be as effective in this setting as other, more conventional, classroom settings. All of the teachers at this school who have had experience with the GBG have spoken highly of it. They have said that it makes their classroom less stressful and that the students are less distracted. Many of students also reported that they enjoy the game and look forward to playing it.

The implementation of the GBG in this population and setting presents particular obstacles. Initially, there were some problems with reinforcer selection. Due to restricted diets and school rules, edible reinforcers were limited. Students would also save their snacks to eat later at inappropriate times and some teachers were reluctant to provide free time for students as a reinforcer because they thought it might cause disruption. To make reinforcer consumption less problematic, in some classrooms we made a rule that the students had to consume their edible reinforcers at a certain time. For the free time, we asked teachers to try giving their students the extra privilege and monitor their behavior. If the free time were to become disruptive, only edibles would be used. However, students (somewhat surprisingly) engaged in little to no problem behavior during free time. One other concern involved the behavior of individual students. On some days, certain students' behavior would be particularly difficult and students seemed to deliberately break rules to make their team gain points. To address this issue, we placed individuals having problems like this on their own team, apart from the original two teams, so that their behavior only affected themselves. These students would occasionally have tantrums so severe that they were removed from the class by resource officers. However, this problem with behavior was seen in both baseline and treatment and, therefore, was not a negative side effect of the GBG. The fact that implementation of the GBG produced marked decreases in the levels of disruptive behavior under particularly challenging conditions provides further evidence of its effectiveness in a wide range of classroom types.

There were some limitations to this application of the GBG. Some of the teachers have claimed that the game would be too difficult to implement on their own because it would disrupt their teaching. Many of the classes at the school are taught in smaller groups away from the chalkboard and it could be disruptive for the teacher to get up to mark each point. The inability to implement the GBG could be a good reason to request more resources in the classroom, or better training and use of teachers' aides. Also, other methods of scorekeeping that are less demanding on the teacher should be explored. It is worth mentioning that disruptions in teaching by implementing the GBG may be high at the beginning of the implementation but low once levels of problem behavior have decreased. Also, it is possible that the time teachers spend keeping order in a class with high levels of problem behavior could be more than the time spent implementing the GBG.

Several implications for future research can be taken from this study. First, variations on scorekeeping for the game should be explored. One way to make it easier for the teacher to implement would be to have students mark their own points on the board. This could be successful as long as the students cooperated. Another approach would be to have the teacher use a smaller board to carry around the room while teaching away from the main board, or the teacher could carry a remote that would add points to a score on a smart board, projector screen, or other electronic device. Points could also be recorded covertly, so the students were unaware of how many points they have. Teacher training could also be implemented prior to the start of the school year, which would enable the teachers to include the GBG in planning their standard classroom procedures. The GBG could also be combined with other procedures designed to increase appropriate class participation by students.

In addition to variations on scoring, other dependent variables could be evaluated. Time sampling measures could be used to observe students' on task behavior. This would provide data on problem behavior as well as what students are doing when they are not engaging in problem behavior. The evaluation of academic performance would also be an important extension for the GBG; to date, very few if any studies have correlated academic performance to instruction following during the GBG. Demonstrating that the GBG not only reduces problem behavior but also increases academic performance would provide more evidence supporting the use of the procedure in schools. One way to do this could be to examine grades and standardized test scores before and after implementation of the GBG. Another way to look at academics could be to collect data on the amount of time spent teaching and time spent reprimanding students to see if there is an increase in teaching time with the implementation of the GBG. It would also be interesting to record the frequency of disciplinary referrals and other documentations of problem behavior outside of the classroom. Demonstrating that the GBG can reduce students' problem behavior during the game as well as throughout the day in other settings would be strong evidence supporting its use in school settings. The GBG was effective at reducing targeted high frequency, low intensity behaviors, but future studies could also focus on low frequency, high intensity behaviors such as physical violence and major property destruction.

Although there is research on the GBG that has proven its effectiveness in different cultural and linguistic settings (e.g., Leflot, van Lier, Onghena, & Colpin, 2013; Ruiz-Olivares, Pino, & Herruzo, 2010); the application and research of the GBG seems to be scarce in Latin American countries with very few empirical studies conducted (Pérez, Fernández, Rodríguez, & De la Barra, 2004; Pérez, Rodríguez, De la Barra, & Fernández, 2005). In order to internationally disseminate the efficacy and utility of the GBG, a large-scale replication of these results would be very useful. In 2002, Embry wrote that large-scale replications could be key in persuading the scientific community as well as educators and policy makers to encourage the widespread use of the GBG as a behavioral vaccine. Inclusion of delinquent students in these large-scale replications would be a potentially valuable demonstration of the efficacy and generalizability of the procedure.

References

Barrish, H. H., Saunders, M. S., & Wolf, M. M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*, 119–124. doi: 10.1901/jaba.1969.2-119

Donaldson, J. M., Vollmer, T. R., Krous, T., Downs, S., & Berard, K. P. (2011). An evaluation of the good behavior game in kindergarten classrooms. *Journal of Applied Behavior Analysis, 44*, 605-609. doi: 10.1901/jaba.2011.44-605

Embry, D. D. (2002). The good behavior game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*, 273–297.

Embry, D. D. (2011). Behavioral vaccines and evidence based kernels: Non-pharmaceutical approaches for the prevention of mental, emotional and behavioral disorders. *Psychiatric Clinics of North America, 34*, 1–34. doi:10.1016/j.psc.2010.11.003.

Farrington, D. P. (1989). Early predictors of adolescent aggression and adult violence. *Violence and Victims, 4*, 79-100.

Fergusson, D. M., & Horwood, L. J. (1998). Early conduct problems and later life opportunities. *Journal of Child Psychology and Psychiatry, 39*, 1097-1108. doi: 10.1111/1469-7610.00414

Fergusson, D. M., & Lynskey, M. T. (1998). Conduct problems in childhood and psychological outcomes in young adulthood: A prospective study. *Journal of Emotional and Behavioral Disorders, 6*, 2-18. doi: 10.1177/106342669800600101

Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega, R. (2014). Effects of the good behavior game on challenging behaviors in school settings. *Review of Educational Research, 20,* 1-26. doi: 10.3102/0034654314536781.

Kellam, S. G., Brown, C. H., Poduska, J. M., Lalongo, N. S., Wang, W., Toyinbo, P., et al. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95S*, S5–S28. doi: 10.1016/j.drugalcdep.2008.01.004

Leflot, G., van Lier, P. A., Onghena, P., & Colpin, H. (2013). The role of children's on-task behavior in the prevention of aggressive behavior development and peer rejection: A randomized controlled study of the good behavior game in Belgian elementary classrooms. *Journal of School Psychology, 51,* 187-199. doi: 10.1016/j.jsp.2012.12.006

Pérez, V., Fernández, A. M., Rodríguez, J., & De La Barra, F. (2004). Evaluación del efecto de una intervención conductual en la agresividad escolar. *Terapia Psicológica, 23,* 91-98.

Pérez, V., Rodríguez, J., De La Barra, F., & Fernández, A.M. (2005). Efectividad de una estrategia conductual para el manejo de la agresividad en escolares de enseñanza básica. *Psykhe, 14,* 55-62.

Ruiz-Olivares, R. Pino, M. J., & Herruzo, J. (2010). Reduction of disruptive behavior using an intervention based on the good behavior game and the say-do-report correspondence. *Psychology in the Schools, 47,* 1046-1058. doi: 10.1002/pits.20523

Tingstrom, D. H., Sterling-Turner, H. E., Wilczynski, S. M. (2006). The good behavior game: 1969–2002. *Behavior Modification, 30,* 225–253. doi: 10.1016/B0-12-343010-0/00107-0

Van Acker, R. (2007). Antisocial, aggressive, and violent behavior in children and adolescents within alternative education settings: prevention and intervention. *Preventing School Failure: Alternative Education for Children and Youth, 51*, 5-12. doi: 10.3200/PSFL.51.2.5-12

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1682 - 1699

Increasing the Social Validity of Function-Based Treatments for Problem Behavior

Jessica P. Álvarez[1], Nathan A. Call & Joanna E. Lomas Mevers
Marcus Autism Center and Emory University School of Medicine

Abstract

Functional assessment, and function-based treatments, are the gold standard for the treatment of problem behavior. Historically, these assessment and treatment evaluations have been conducted in austere clinical settings to increase internal validity. While prioritizing internal validity is critical in the initial stages of a treatment evaluation, if there is not an eventual shift to prioritizing the external or social validity of the treatment it may inevitably fail in the natural environment. The purpose of this case example is to outline a socially valid approach to the assessment and treatment of problem behavior that ensures individuals' and their families' lives benefit in meaningful ways. More specifically, this case-example will outline a method of prioritizing social validity to identify treatment goals, conduct functional analysis, evaluate and generalize treatment, and implement caregiver training.

Keywords: Problem Behavior, Behavioral Treatment, Behavior Analysis, Social Validity

Aumentando la Validez Social de los Tratamientos Funcionales para la Conducta Problemática

Resumen

La evaluación funcional y los tratamientos funcionales son el estándar de oro para el tratamiento de la conducta problemática. Históricamente, esas evaluaciones y tratamientos se han conducido en escenarios clínicos austeros para aumentar la validez interna. Si bien el priorizar la validez interna es crítico en las etapas iniciales de la evaluación de un tratamiento, si eventualmente no hay un cambio para priorizar la validez externa o social del tratamiento, éste puede fallar en un escenario natural. El propósito del ejemplo de caso que se presenta en este trabajo es mostrar una aproximación válida para la evaluación y tratamiento de conducta problemática que asegura que las vidas de los individuos y de sus familias se beneficien de forma significativa. Más específicamente, el ejemplo de caso que se presenta mostrará un método para identificar las metas del tratamiento priorizando la validez social, para conducir un análisis funcional, evaluar y generalizar el tratamiento y entrenar al cuidador.

Palabras Clave: Conducta Problemática, Tratamiento Conductual, Análisis de la Conducta, Validez Social

Original recibido / Original received: 14/06/2014 Aceptado / Accepted: 29/09/2014

[1] Correspondence: Jessica P. Alvarez. The Marcus Autism Center. 1920 Briarcliff Rd., Atlanta, GA 30329. Phone: 404-785-9448, E-mail: jessica.alvarez@choa.org

Compared to typically developing peers, children with developmental or intellectual disabilities are at an increased risk for engaging in problem behaviors such as aggression, self-injury, property destruction, and pica (Dekker, Koot, van der Ende, & Verhulst, 2002). Estimates of the prevalence of such problem behaviors in this population vary widely, from between 10-15% (Emerson et al., 2001) to as many as 73% of individuals (Rojahn, Matson, Lott, Esbensen, & Smalls, 2001). However, what is unequivocal is that the presence of problem behaviors such as these is associated with negative outcomes for both the child engaging in the behavior and their caregivers. For example, self-injurious behaviors like head-banging, hitting/biting oneself, or skin picking can cause permanent tissue damage or permanent loss of sight or hearing, as well as increase the risk of secondary harm from infections (Minshawi et al., 2014). Wandering or running away (i.e., elopement) puts these children at risk for abduction, injury, or even death due to drowning or being struck by an automobile (Anderson et al., 2012). Such behaviors also frequently result in children being placed in more restrictive educational settings. As a result they benefit less from general educational instruction and spend less time with typically developing peers (Gresham et al., 2004). Similarly, problem behavior can require the child to be placed in restrictive residential settings, further limiting their access to important social and community interactions and supports (Friedman, Kalichman, & Council on Children with Disabilities, 2014).

Problem behaviors such as these also have profound negative effects for caregivers. Research has shown a strong correlation between having a child with problem behavior and an increased level of parental stress (Dykens, 2000; Neece, Green, & Baker, 2012) and poorer overall family functioning (Davis & Gavidia-Payne, 2009). In fact, the presence of problem behavior is more predictive of parental stress than deficits in adaptive behavior or the severity of the symptoms of the developmental disability itself (Baker, Blacher, Crnic, & Edelbrock, 2002; Estes et al., 2013; Hastings & Beck, 2004; Lecavalier, Leone, & Wiltz, 2006). Behavior problems have also been found to negatively impact marital satisfaction (Hartley, Barker, Baker, Seltzer, & Greenberg, 2012), are correlated with higher rates of problem behavior in siblings (Hastings & Beck, 2004), and parents of children with disabilities who engage in problem behavior find it harder to maintain employment (Hall, Bouldin, Andresen, & Ali, 2012). Moreover, caregivers who experience high levels of stress are less able to implement interventions for their child's problem behavior (Osbourne, McHugh, Saunders, & Reed, 2008). Thus, well-designed interventions for problem behaviors exhibited by children with developmental disabilities have the potential to produce significant benefit not only for the child but for their caregivers as well.

Effective treatments of problem behavior in this population have historically been based upon the scientific approach to the study of behavior known as behavior analysis. Within this literature, problem behavior was hypothesized by Ferster (1961) to be learned and therefore under the control of its environmental consequences. Some of the first researchers to apply this approach to the treatment of problem behavior began evaluating the effectiveness of behavior analytic principles (e.g., reinforcement, extinction, punishment, etc.) for the

reduction of self-injurious behavior. Early evidence revealed that behavior analytic approaches could effectively reduce the self-injurious behavior of individuals with developmental and intellectual disabilities (Corte, Wolf, & Locke, 1971; Lovaas & Simmons, 1969). However, these researchers noted that not all self-injurious behaviors were responsive to the same strategies. For some individuals self-injurious behavior decreased only when other non-self-injurious behavior(s) were reinforced (Corte et al., 1971). For others, self-injurious behavior decreased only when punishment was implemented contingent upon self-injurious behavior (Corte et al., 1971; Lovaas & Simmons, 1969). Still others only showed improvements when all forms of positive reinforcement were restricted following episodes of self-injurious behavior (Corte et al., 1971; Lovaas & Simmons, 1969). Carr (1977) summarized much of the research on the treatment of self-injurious behavior at the time and hypothesized that the behaviors could be maintained by extrinsic (e.g., socially mediated) or intrinsic (e.g., automatic) forms of reinforcement. For example, when attention from a caregiver is a sufficiently potent form of positive reinforcement, the child may learn that self-injury is likely to produce attention from caregivers in the form of reprimands, soothing statements, etc. Carr (1977) also suggested that treatments might not be effective at reducing self-injurious behavior if they are not correctly matched to the type of consequences maintaining the self-injury (i.e., its function). In the example above, restricting attention following self-injurious behavior is likely to produce a gradual reduction due to extinction. However, for a child for whom removal of demands serves as a negative reinforcer that maintains self-injury, restricting attention is unlikely to have a reductive effect. Thus, Carr (1977) raised the possibility that an assessment of function could be integral as a guide for selecting treatments for self-injurious behavior.

Although other researchers had developed methods for evaluating whether a particular type of reinforcer maintained an individual's self-injury (e.g., Carr, Newsom, & Binkoff, 1976), Iwata, Dorsey, Slifer, Bauman, & Richman, (1982/1994) were the first to develop a methodology that systematically evaluated several of the most likely functions within a single analysis. Three test conditions, each of which was designed to evaluate a different hypothesis about function, and one control condition were conducted with nine participants. The attention condition served as an analog for those situations in which a caregiver might deliver attention, such as a reprimand, in response to self-injury. As such, it assessed whether self-injury was maintained by social-positive reinforcement in the form of attention. The escape condition served as an analog for situations in which a caregiver discontinues tasks in an effort to stop their child's self-injury. Thus, this condition assessed whether self-injury was maintained by social-negative reinforcement in the form of escape from demands. The alone condition assessed whether self-injury was maintained by automatic reinforcement. That is, if problem behavior produced some intrinsic reinforcement, such as sensory stimulation, it would be expected to be insensitive to social consequences and persist even when the child was alone in a relatively austere environment. Finally, the toy play condition served as a control for the attention and escape test conditions because problem behavior maintained by social reinforcement would not be expected to occur in the absence of demands and free access to attention. For six of the nine participants,

self-injurious behavior occurred at high rates during only one of the test conditions, and treatments based upon results of these functional analyses (FAs) effectively reduced self-injurious behavior. These results provided support to the theories of Ferster (1961) and Carr (1977) that self-injurious behavior likely served specific behavioral function(s) and that function was a key variable in the development of effective treatments.

Since the seminal study by Iwata et al. (1982/1994), FA methodology has been extended to a wide range of problem behaviors, including (but not limited to) aggression (e.g., Baker, Hanley, & Mathews, 2006), pica (e.g., Piazza, Hanley, & Fisher, 1996), elopement (e.g., Piazza et al., 1997), stereotypy (e.g., Mace, Browder, & Lin, 1987), noncompliance (e.g., Reimers et al., 1993; Rodriguez, Thompson, & Baynham, 2013; Wilder, Harris, Reagan, & Rasey, 2007), tantrums, (e.g., Vollmer, Northup, Ringdahl, LeBlanc, & Chauvin, 1996), property destruction (e.g., Fisher, Lindauer, Alterson, & Thompson, 1998), disruptive behavior (e.g., Broussard & Northup, 1995, 1997), and inappropriate sexual behavior (e.g., Fyffe, Kahng, Fittro, & Russel, 2004). Functional analysis of severe problem behavior continues to be widely implemented, and research has continued to expand and adapt the Iwata et al. (1982/1994) methodology. The original methods have also been adapted to evaluate a wide range of additional types of reinforcers, including access to preferred edible or leisure items (e.g., Hagopian, Fisher, Thibault Sullivan, Acquisto, & LeBlanc, 1998), avoidance of social interactions (e.g., Hagopian, Wilson, & Wilder, 2001) and avoidance of transitions (e.g., McCord, Thomson, & Iwata, 2001), to name just a few.

Functional analysis test conditions are designed to evaluate the effects of antecedent and consequence variables, while controlling for extraneous variables present in the environment that make it difficult to identify the function of problem behaviors in less controlled circumstances. Thus, great care is often exerted to ensure such potential confounds are eliminated. It is perhaps not surprising then that reviews of the literature on FAs show that the majority of the assessments are conducted in hospital or institutional settings where significant resources can be allocated (Beavers, Iwata, & Lerman, 2013; Hanley, Iwata, & McCord, 2003). Within these settings, FAs are frequently conducted in specialized session rooms that are devoid of any materials not directly related to the variables being evaluated. Although these analog settings allow for extremely controlled and precise FAs, a disadvantage is that they may not be representative of the natural environment in which the problem behavior occurs. This limitation can be problematic when implementing and evaluating function-based treatments because the goal of most treatments is to reduce problem behavior in the individual's natural environment.

A review of behavioral interventions revealed that function-based treatments based on a FA resulted in significantly better outcomes compared to treatments that were not based on a FA (Campbell, 2003; Heyvaert, Saenen, Campbell, Maes, & Onghena, 2014). However, many of the studies summarized in that review limit their outcomes to demonstrations that problem behavior improved in those same well-controlled settings in which the FA was conducted. In contrast, it is important for clinicians to not only consider whether a particular treatment approach holds the

possibility of achieving a positive outcome, but whether treatment and outcome are socially valid (i.e., acceptable to the individual, their caregivers, and society; Wolf, 1978). It can be argued that function-based treatments possess some social validity by their very nature because they provide an opportunity for the individual to continue to access reinforcement that is meaningful to them. Conversely, even if a non-function-based treatment effectively reduces problem behavior it will still possess less social validity if the individual is no longer able to access the reinforcer(s) that previously maintained their problem behavior. Thus, it is not surprising that there is evidence that individuals who engage in problem behavior prefer function-based treatments to those that are not based upon the results of a FA (Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997). However, access to reinforcement is only one component of effective treatments for problem behavior that clinicians and researchers must consider to maximize social validity of treatments for problem behavior.

Designing function-based treatments that are socially valid also requires striking the appropriate balance between internal and external validity. Maximizing internal validity through the use of highly controlled sessions and specialized rooms allows for causal statements to be made regarding the function of problem behavior and any decrease observed as a result of treatment. However, emphasizing such approaches while neglecting the importance of evaluating treatment effectiveness in more naturalistic environments (i.e., external validity) can limit the chances of long-term positive outcomes. On the other hand, emphasizing external validity by conducting sessions in a less controlled manner and more naturalistic conditions can make treatment effects more likely to generalize to the natural environment, but conclusions regarding the active components of behavior change may not be definitive. Therefore, a socially valid treatment plan begins by prioritizing internal validity during the initial treatment evaluation, with a shift towards prioritizing external validity once the treatment has already been shown effective. However, historically there has been less emphasis in the literature on the inclusion of naturalistic stimuli in analog sessions or the modification of treatment components to replicate more naturally existing contingencies.

One way to enhance the probability that treatments will strike the appropriate balance between internal and external validity is to focus attention on the manner in which treatment goals are established. Frequently treatment goals are limited to a description of a desired reduction in targeted problem behaviors. However, such an approach to setting goals is structural rather than functional: this type of treatment goal merely states that the treatment will result in some degree of behavior reduction but does not specify the desired benefit for the individual or caregivers. In contrast, consideration of context or function in establishing goals for treatment is more likely to ensure clinicians develop treatments that will actually address caregivers' concerns. For example, if a treatment goal states only that "aggression will be reduced by 80%", clinicians may be prone to determining that treatment has been successful even if aggression has only been reduced within the tightly controlled circumstances of a session room. In contrast, describing the treatment goal as "reduce aggression by 80% when the child is working on difficult

demands at school", compels clinicians to incorporate academic tasks into the FA and treatment analysis, and also increases the probability of evaluating the generalization of treatment to academic settings.

Although there is ample evidence that treatments for problem behavior that are based upon results of an FA can produce significant improvements, there is little published data to show maintenance of treatment effects in the natural environment. A review of behavioral treatments in the research literature reported that only 27.4% of studies attempted to generalize treatment to a new setting or situation (Campbell, 2003). This finding suggests that treatments are commonly evaluated within a single environment, most likely those well-controlled environments in which treatments are typically first developed. This treatment approach raises concerns regarding the maintenance of treatment gains in less structured situations that are inevitable in the natural environment.

Although social validity has not always been emphasized in the literature on function-based treatment of problem behavior, the purpose of the current case study is to outline a socially valid approach to function based treatment of severe problem behavior exhibited by individuals with developmental disorders. More specifically the current case study describes how social validity can guide goal development, behavioral assessment, treatment implementation, and generalization.

Method

Participant and Setting

At the time of treatment, Stephany was a twelve-year-old female with Autism Spectrum Disorder and obesity. She communicated using single words and required moderate assistance with daily living tasks such as getting dressed. Stephany was referred to an intensive day treatment clinic for treatment of aggression, disruption, self-injurious behavior, spitting, and pica. During the months leading up to her admission, Stephany regularly punched holes in the walls of her home and had broken several windows. Her mother expressed concerns that they would be evicted from their apartment if the landlord became aware of the damage caused by Stephany's problem behavior. Her mother also reported that several respite workers hired to assist in caring for Stephany had quit due to her problem behavior. Her mother reported that Stephany most commonly engaged in problem behavior in a few situations: when preferred food was restricted, at doctor's appointments, or if item(s) in her home were rearranged and Stephany was prevented from returning them to their original location. Therefore, Stephany's mother no longer allowed certain preferred foods in their home and avoided restricting food. As a result of having provided relatively free access to food, Stephany had gained a significant amount of weight, which in turn made it harder for others to physically manage her problem behavior. Stephany's problem behavior at doctor's appointments had resulted in her not receiving necessary medical care for over two years. Lastly, prior to Stephany's admission to the day treatment program, her mother was seeking out an alternate residential placement.

At the time of the initial admission meeting, the clinician worked with Stephany's mother to identify several treatment goals that, if achieved, would most improve her and Stephany's lives. Stephany's treatment goals were to tolerate a) her morning routine (i.e., awaking, bathing, brushing her teeth, dressing, and eating breakfast); b) the restriction of food and limited meal portions; c) routine doctor's appointment; and d) community outings without engaging in problem behavior.

The first day of Stephany's admission was spent in her home in which therapists took a tour, observed Stephany and her mother in their daily routine, and instructed Stephany's mother to replicate those situations that she had reported typically resulted in problem behavior. Following the home visit, Stephany began attending the day treatment program for six hours per day, five days a week. Assessment and treatment sessions were initially conducted by therapists in a padded session room with video cameras and a one-way mirror separating an adjoining observation room. Because this setting allowed for strict control over variables hypothesized to evoke or maintain problem behavior, it also maximized internal validity. External validity was emphasized once an effective treatment had been identified, by conducting sessions in more naturalistic settings around the treatment facility (e.g., playroom, classroom, waiting area near a physician's office), and finally in Stephany's home and school.

Measurement

When attempting to strike the right balance between internal and external validity, it is important to establish a rigorous data collection system. For Stephany, data were collected separately for each topography of her problem behavior. Aggression was defined as any instance or attempt to hit, kick, head-butt, scratch, bite, push, choke, pull hair, or throw an object at another person. Each of these behaviors was operationally defined in great detail for data collection purposes. For example, pushing was defined as any instance in which Stephany's hand(s) contacted another person and applied force that altered the original standing position of that person. Other topographies of problem behavior that were similarly defined included disruption, which included throwing, hitting, or destroying objects; self-injurious behavior, which included head-banging and self-scratching; pica and spitting. Data were collected on customized data collection software via laptop computers during the FA. During stages of the treatment evaluation that were conducted in more naturalistic settings, data were collected via paper and pencil.

Functional Analysis

Prior to conducting the FA, it was important to empirically identify stimuli that were most likely to exert some influence over Stephany's problem behavior. However, rather than select items or demands arbitrarily, as is commonly done, these were initially identified based on home/school observations and interviews with her mother. As has become customary, the items identified via home/school observation and caregiver report were included in a paired-stimulus preference assessment (Fisher et al., 1992) to identify a hierarchy of preference for edible and

leisure items. Similarly, a demand assessment (Call, Pabico, & Lomas, 2009) identified demands that were likely to be aversive for Stephany. Results of these two assessments were used to select leisure and/or edible items for inclusion in toy play, attention, and tangible conditions, as well as demands to be included in the escape condition.

A modified FA based on the procedures described by Iwata et al. (1982/1994) was conducted using a variety of conditions, each of which were conducted in repeated 10 min sessions. During sessions of the *Toy play* condition Stephany was provided with continuous access to a highly preferred leisure items (i.e., Legos™ and Playdoh™) and edible item (i.e., chips or fruit snacks). The therapist provided attention no less than every 30 s, refrained from touching any of Stephany's leisure or edible items, and did not make any requests or place demands on her. There were no scheduled consequences for problem behavior during these sessions. The purpose of the toy play condition was to serve as a control for the test conditions.

Prior to the start of the *Tangible (leisure)* condition, Stephany was given 2 minutes continuous access to the highly preferred leisure items. Once the session began, the therapist restricted access to the items but provided 30 s access contingent upon problem behavior. Following the 30 s reinforcement interval, the therapist again restricted access to the item. No attention or demands were delivered during these sessions.

Prior to conducting the *Attention* condition, the therapist provided Stephany with 2 minutes of continuous high-quality attention. Once session began, the therapist restricted their attention by telling Stephany they had to do some work. Stephany had continuous access to an item identified as low preferred in the prior preference assessment (bubbles). Contingent upon problem behavior the therapist delivered a brief verbal reprimand.

During the *Escape* condition, the therapist repeatedly prompted Stephany to complete the task identified in the prior demand assessment (a number identification task) using a 3-step progressive prompting procedure: the therapist first delivered a verbal prompt to complete the task and allowed 3-5 s for compliance. If Stephany did not comply with the task, the therapist provided a model prompt. If Stephany did not comply within 3-5 s following the model prompt, the therapist physically guided the correct response. Contingent upon problem behavior, the therapist provided a 30 s break from the demand and removed all the demand materials from the table. No additional attention was provided beyond prompting the completion of the demands.

Based on caregiver reports, and observations conducted in Stephany's home and school, two additional test conditions were included in the FA. The purpose of these conditions was to evaluate additional hypotheses regarding possible antecedent and consequence variables evoking and maintaining Stephany's problem behavior. That is, Stephany's problem behavior was hypothesized to be maintained by access to food and by terminating or avoiding the interruption or manipulation of items in her environment. Again, the addition of these conditions enhanced the external and social validity of the FA by evaluating

hypotheses about potential individualized reinforcers that were maintaining her problem behavior.

Prior to the start of sessions from the *Tangible (edible*) condition, Stephany was allowed to choose between several preferred edible items. Once Stephany selected an edible item, she had 2 minutes of continuous access to the item. Tangible (*edible*) sessions were identical to those in the Tangible (*leisure*) condition with the exception that the edible item she had selected prior to session was restricted and delivered contingent upon problem behavior.

Prior to the *Interruption* condition, Stephany received 2 minutes of access to the preferred leisure items during which the therapist did not interact with the items. Once the session began, the therapist manipulated Stephany's leisure items every 5 s (e.g., rearranged the Legos™). Contingent upon problem behavior, the therapist refrained from manipulating or touching the leisure items for 30 s. During the interruption sessions, the therapist provided Stephany with continuous attention even when they were not manipulating her leisure items.

Results

Figure 1 depicts the results of the FA. Stephany consistently engaged in high rates of problem behavior during the Tangible (*leisure*), Tangible (*edible*), Interrupt, and Escape conditions compared to the Attention or Toyplay conditions. This differentiation between the rates of problem behavior in the control (i.e., Toyplay) and specific test conditions indicated that her problem behavior served multiple functions, including access to social-positive reinforcement (in the form of preferred leisure and edible items) and social-negative reinforcement (in the form of escape from demands and escape from others manipulating the items with which she was engaged).

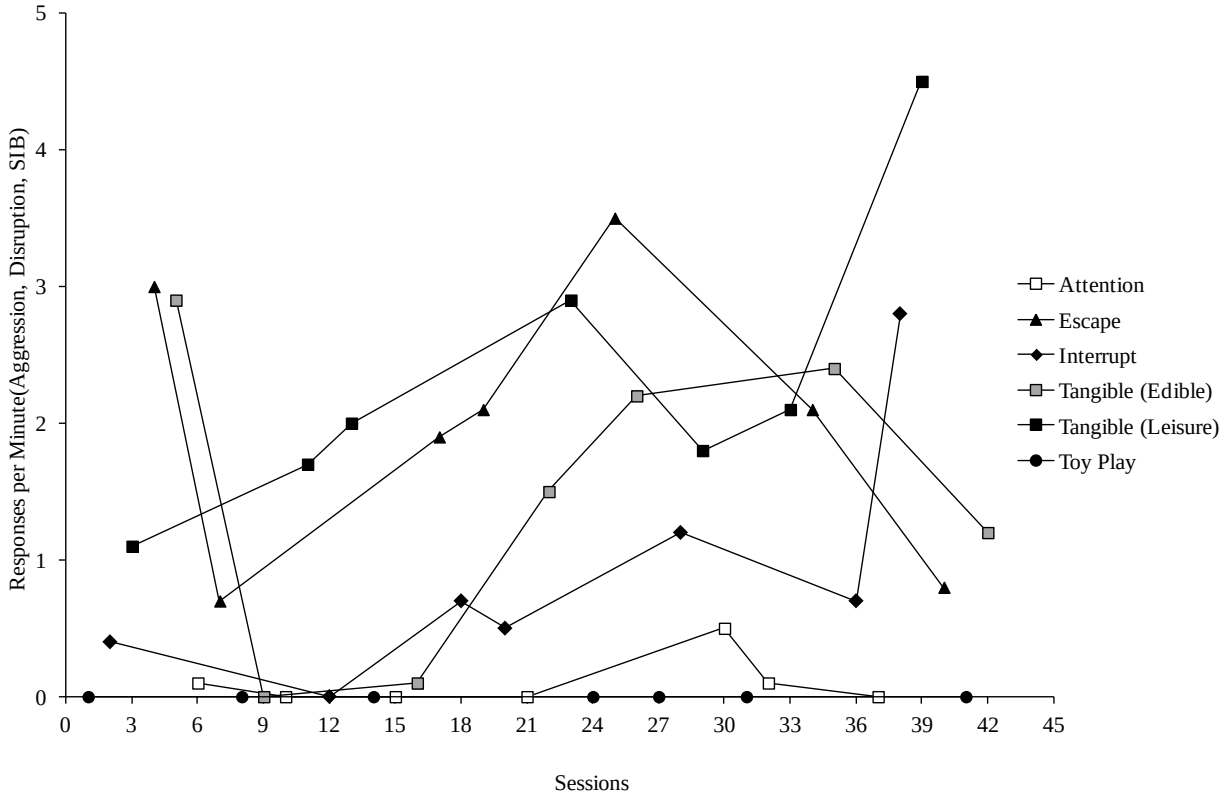


Figure 1. Functional analysis of aggression, disruption, and self-injurious behavior (SIB)

Treatment Evaluation

The clinical team used the results from Stephany's FA, in combination with her caregiver's goals, to design an intervention in which problem behavior was replaced with appropriate communicative behavior. Stephany's mother conveyed being able to restrict access to edible items was her highest priority. The initial treatment package was then modified to address subsequent treatment goals such as requiring Stephany to tolerate her morning routine tasks, academic work, leisure activities and the manipulation of items in her environment. All of these permutations of the treatment were made within the context of restricting Stephany's access to preferred edible items.

Functional Communication Training

Stephany was taught a vocal request for each of her preferred edible items using a progressive time delay prompting procedure (Charlop, Schreibman, & Garrison Thibodeau, 1985). Contingent upon mastery criteria for all vocal requests, a Differential Reinforcement for Alternative Behavior (DRA) intervention was evaluated. During these sessions, Stephany's requests were reinforced with 30 s access to an edible item, and all problem behavior was placed on EXT. A reversal design compared rates of problem behavior under DRA, and baseline conditions

that replicated the Tangible (Edible) condition of the FA (see Figure 2). Stephany's problem behavior was eliminated within three sessions of the DRA treatment.

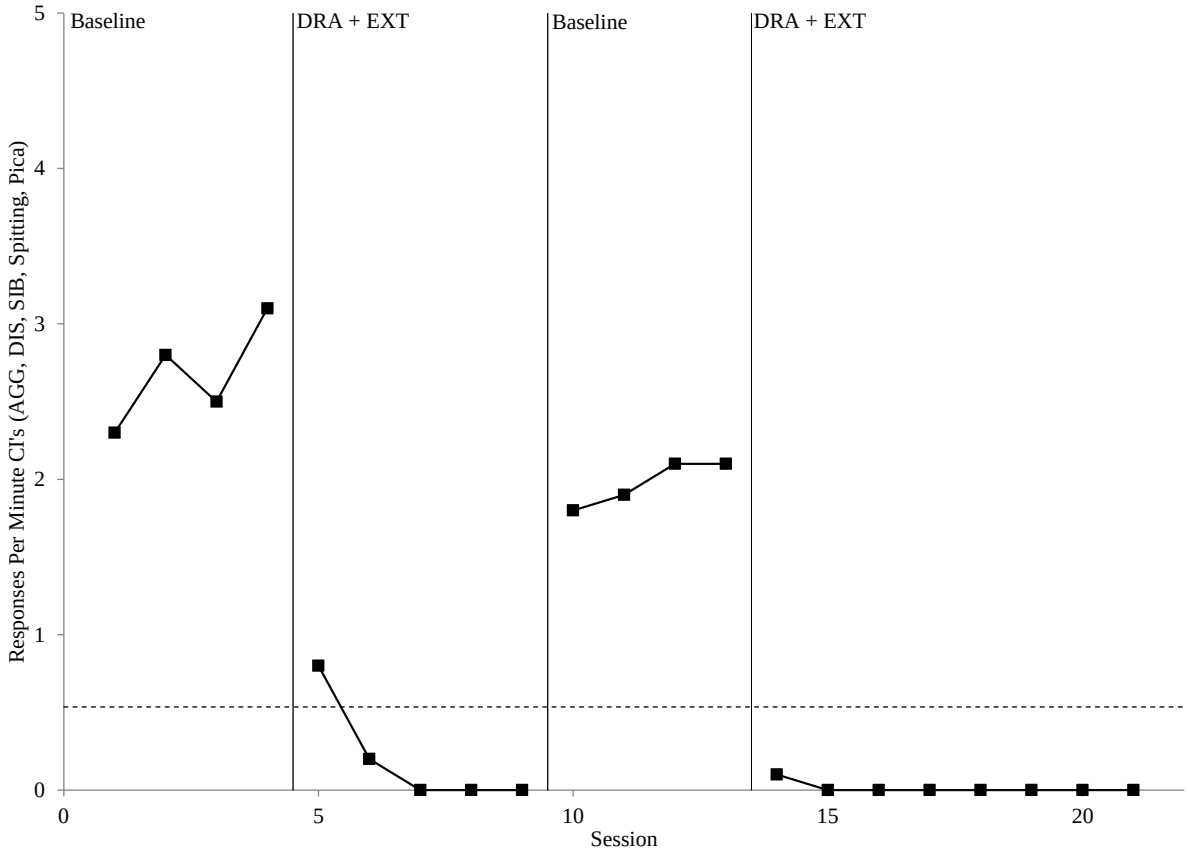


Figure 2. Treatment evaluation of differential reinforcement of alternative behavior (DRA) with extinction EXT) for aggression, disruption, and self-injurious behavior (SIB).

Treatment Evaluation

As described above, the results of treatments for problem behavior reported in the literature generally reflect the effects of interventions conducted in well controlled environments. In essence, this is what Stephany's treatment had achieved to this point. However, it was not yet a socially valid treatment that would achieve the goals established by her mother. That is, it was not reasonable to expect that Stephany's mother would be able to reinforce every appropriate request for preferred edible items. In addition, due to health concerns regarding Stephany's weight, her mother had established a goal to restrict all access to food during certain periods of the day. Therefore, the clinical team designed a treatment in which there were periods of time during which Stephany could appropriately request food, but there were also periods during which food was

unavailable even if she appropriately requested. This treatment has been conceptualized as a multiple schedule (Fisher, Kuhn, & Thompson, 1998; Hanley, Iwata, & Thompson, 2001), with the aim of bringing requests under the control of discrete stimuli.

During the multiple schedule treatment, a yellow card signaled to Stephany when appropriate requests would be reinforced, whereas a purple card signaled when appropriate requests would not be reinforced. Initially the yellow card interval was set at 45 seconds, whereas the purple card interval was set at 15 seconds. Within sessions, the therapist rotated between the yellow and purple card intervals. Therefore, within one session, Stephany experienced periods of time in which a therapist reinforced her requests for preferred edible items and periods of time in which her preferred edible items were unavailable, even if she appropriately requested them. Following the introduction of treatment there was an initial increase in problem behavior, however three consecutive sessions without problem behavior were established within 21 sessions (see Figure 3).

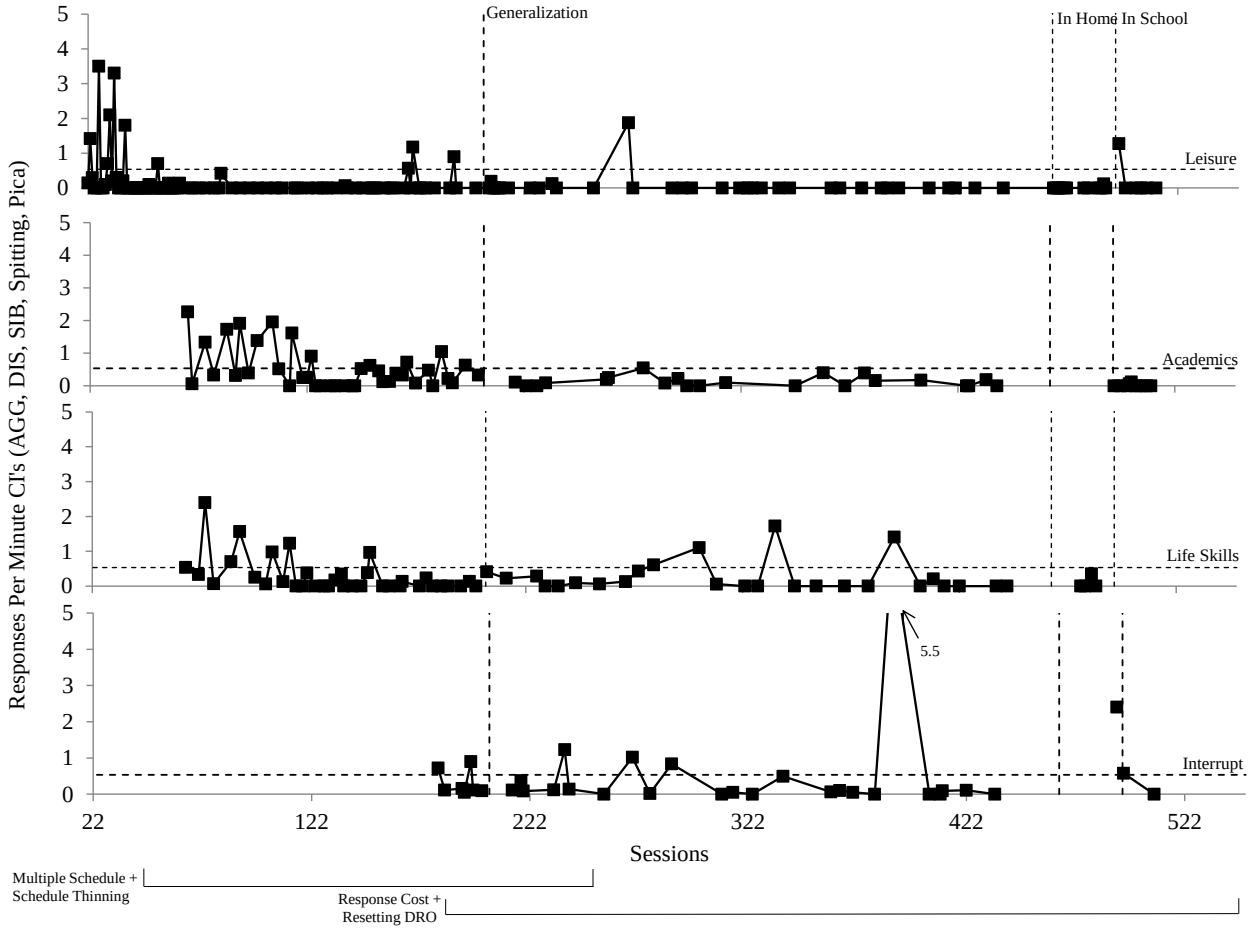


Figure 3. Treatment evaluation across leisure, academic, life skills, and interrupt conditions for aggression, disruption, and self-injurious behavior (SIB)

To further increase the social validity of the treatment, and accomplish the treatment goal of being able to restrict Stephany's access to food throughout the day, it was critical to increase the duration of the purple card interval during which edible items were unavailable. Therefore, the purple card interval was gradually increased from 15 s to the terminal goal of 15 min. Once the terminal goal had been achieved, the treatment protocol was adapted to address the additional situations Stephany's mother had identified as treatment goals. For example, the treatment was modified to address times in which Stephany was required to complete her morning routine, daily living tasks, academic tasks, leisure time with family, and to allow Stephany's family to re-arrange items in their home.

To increase Stephany's compliance with completing demands associated with her morning routine without engaging in problem behavior, a life skills treatment was implemented. For life skills sessions, the purple card interval indicated time in which Stephany was required to complete life skills tasks such as folding and putting away clothes or brushing her teeth. To address the goal that Stephany complete academic tasks without engaging in problem behavior the treatment was modified to include academic sessions. During academic sessions, Stephany was required to complete an academic task during the purple card interval. To replicate times in which Stephany and her family members were engaged in their own preferred leisure activities and preferred food was restricted, the treatment was modified to incorporate Stephany's preferred leisure activities. During purple card intervals for leisure activities, Stephany was prompted to engage in less preferred leisure activities while edible items remained restricted. Lastly, to address times in which Stephany's leisure items or other non-leisure items in her household were manipulated, the treatment was modified to require Stephany to tolerate periods of time in which another person manipulated her items without her engaging in problem behavior. During interruption sessions, a therapist manipulated Stephany's leisure items during the purple card interval. Each of these permutations of the treatment package was addressed separately (see Figure 3).

Due to an increase in problem behavior across the life skills, academic, and interruption sessions, response cost and resetting differential reinforcement for other behavior (DRO) components were implemented across all conditions. The response cost component was implemented during yellow card intervals and consisted of immediately changing the yellow card to the purple card contingent upon problem behavior. In other words, Stephany lost the opportunity to request preferred edible items contingent upon problem behavior. The DRO component was implemented during purple card intervals and consisted of the resetting of the purple card interval contingent upon problem behavior. For example, if the purple card interval was set at 5 min and Stephany engaged in problem behavior during the first minute of the interval, she would have to go an additional 5 min without engaging in problem behavior to gain access to the yellow card interval. It is important to note that had the treatment evaluation ended prior to modifying the treatment to address multiple treatment goals, Stephany would have been discharged with a treatment that had been evaluated only in a highly internally valid manner. However, through an emphasis on external and social validity, further

treatment evaluation resulted in modifications to maintain low rates of problem behavior in more naturalistic settings and achieve her caregiver's goals.

Treatment Generalization and Caregiver Training

As previously discussed, evaluating the effectiveness of treatments beyond an austere session room is a critical, and often overlooked, component of treatment evaluations. Therefore all of Stephany's treatment sessions were generalized to more naturalistic settings within the clinic and subsequently to her own home and school settings. For example, the leisure and interrupt sessions were conducted in clinic space that was designed to resemble a family's home. It contained a living room area with a couch and television; a dining area with a table and chairs; a kitchen area with a fridge, cabinets, counters, and a dishwasher; and lastly a play area with shelves containing a wide variety of leisure items. Academic sessions were generalized to a similar clinical space that was designed to resemble a classroom that contained several desks, chairs, and tables. Life skills sessions were generalized to both the classroom setting and a bathroom. In addition, other clients and therapists who were part of the day treatment clinic were present in the playroom and classroom, which replicated situations in Stephany's home and school in which other family members or peers were present.

In addition to evaluating treatment effectiveness in more naturalistic settings, it is also important to ensure that the treatments are designed to incorporate naturally existing reinforcement contingencies when possible. By doing so (e.g., delivering reinforcers after all of the clothing items are folded versus delivering reinforcement after 15 min of folding clothes), treatments can become easier for caregivers and teachers to implement and less stigmatizing for the client. Therefore, treatment for academics and life skills were both modified so that reinforcement was contingent upon compliance or task completion. Stephany engaged in low rates of problem behavior across the leisure, life skills, interrupt, and academic sessions during this generalization phase (see Figure 3).

Lastly, Stephany's caregivers completed intensive training on the treatment protocol until they were implementing all treatment components consistently (i.e., 80% or higher treatment fidelity). This training made it possible to conduct the final phase of treatment: generalization to the natural environment (i.e., Stephany's home and school). During the last two weeks of Stephany's admission, all sessions were conducted in her own home and school. Stephany's problem remained low across leisure, life skills, interrupt, and academic sessions when treatment was generalized to these settings (see Figure 3).

Discussion

Overall, Stephany's case serves as a useful example of how to ensure that FAs and function-based treatments are socially valid. Important steps that were part of Stephany's admission, but are not always discussed in the research literature, included identifying treatment goals that were meaningful to the individual and their caregivers, designing an individualized FA, adapting treatments

to address multiple goals, training caregivers to implement the treatment, and generalizing treatments to the natural environment. Stephany's treatment goals emphasized social validity in that they were nominated by her mother and targeted specific situations that were problematic in their lives. Functional analysis conditions emphasized social validity in that they were individualized to assess the specific antecedents and consequences that were problematic for Stephany (e.g., evaluating interruption and manipulation of items, restricted access to preferred edible items). When implementing function-based treatments for problem behavior, service providers should strive to develop treatments that are feasible for caregivers to implement in the natural environment. Methods of increasing the feasibility of treatments can range from incorporating naturally occurring contingencies for appropriate behavior to modifying a single treatment package to address multiple treatment goals. A final component of ensuring that a treatment is socially valid is to evaluate the treatment in increasingly naturalistic environments. Ultimately, a successful treatment means that caregivers are able to implement a treatment within their daily lives and maintain the treatment effects that are meaningful to them. Through a more socially valid approach to the assessment and treatment of problem behavior, service providers can produce more meaningful changes in individuals' and their caregivers' lives.

References

Anderson, C., Law, K. J., Daniels, A., Rice, C., Mandell, D. S., Hagopian, L., & Law, P. A. (2012). Occurrence and family impact of elopement in children with autism spectrum disorders. *Pediatrics, 130,* 87-877. doi:10.1542/peds.2012-0762

Baker, B. L., Blacher, J., Crnic, K. A., & Edelbrock, C. (2002). Behavior problems and parenting stress in families of three-year-old children with and without developmental delays. *American Journal on Mental Retardation, 107,* 433-444. doi:10.1352/0895-8017(2002)

Baker, J. C., Hanley, G. P., & Mathews, R. M. (2006). Staff-administered functional analysis and treatment of aggression by an elder with dementia. *Journal of Applied Behavior Analysis, 39,* 469-474. doi:10.1901/jaba.2006.80-05

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46,* 1-21. doi:10.1002/jaba.**30**

Broussard, C. D., & Northup, J. (1995). An approach to functional assessment and analysis of disruptive behavior in regular education classrooms. *School Psychology Quarterly, 10,* 151-164. doi:10.1037/h0088301

Broussard, C. D., & Northup, J. (1997). The use of functional analysis to develop peer interventions for disruptive classroom behavior. *School Psychology Quarterly, 12, 65-76*. doi:10.1037/h0088948

Call, N. A., Pabico, R. S., & Lomas, J. E. (2009). Use of latency to problem behavior to evaluate demands for inclusion in functional analysis. *Journal of Applied Behavior Analysis, 42,* 723-728. doi: 10.1901/jaba.2009.42-723

Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 24,* 120-138. doi:10.1016/S0891-4222(03)00014-3

Carr, E. G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84,* 800-816. doi:10.1037/0033-2909.84.4.800

Carr, E. G., Newsom, C. D., & Binkoff, J. A. (1976). Stimulus control of self-destructive behavior in a psychotic child. *Journal of Abnormal Child Psychology, 4,* 139-153. doi: 10.1007/BF00916518

Charlop, M. H., Schreibman, L., & Garrison Thibodeau, M. (1985). Increasing spontaneous verbal responding in autistic children using a time delay procedure. *Journal of Applied Behavior Analysis, 18,* 155-166. doi:10.1901/jaba.1985.18-155

Corte, H. E., Wolf, M. M., & Locke, B. J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis, 4,* 201-213. doi:10.1901/jaba.1971.4-201

Davis, K., & Gavidia-Payne, S. (2009). The impact of child, family, and professional support characteristics on the quality of life in families of young children with disabilities. *Journal of Intellectual and Developmental Disability, 34,* 153-162. doi:10.1080/13668250902874608

Dekker, M. C., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2002). Emotional and behavioral problems in children and adolescents with and without intellectual disability. *Journal of Child Psychology and Psychiatry, 43,* 1087-1098. doi:10.1111/1469-7610.00235

Dykens, E. M. (2000). Annotation: Psychopathology in children with intellectual disability. *Journal of Child Psychology and Psychiatry, 41,* 407-417. doi:10.1111/1469-7610.00626

Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R., Mason, L., & Hatton, C. (2001). The prevalence of challenging behaviors: A total population study. *Research in Developmental Disabilities, 22,* 77-93. doi:10.1016/S0891-4222(00)00061-5

Estes, A., Olson, E., Sullivan, K., Greenson, J., Winter, J., Dawson, G., & Munson, J. (2013). Parenting-related stress and psychological distress in mothers of toddlers with autism spectrum disorders. *Brain & Development, 35,*133-138. doi:10.1016/j.braindev.2012.10.004

Ferster, C. B. (1961). Positive reinforcement and behavioral deficits of autistic children. *Journal of Child Development, 32,* 437-456. doi:10.1177/108835768900400501

Fisher, W. W., Kuhn, D. E., & Thompson, R. H. (1998). Establishing discriminative control of responding using functional and alternative reinforcers during functional communication training. *Journal of Applied Behavior Analysis, 31,* 543-560. doi:10.1901/jaba.1998.31-543

Fisher, W. W., Lindauer, S. E., Alterson, C. J., & Thompon, R. H. (1998). Assessment and treatment of destructive behavior maintained by stereotypic object manipulation. *Journal of Applied Behavior Analysis, 31,* 513-527. doi:10.1901/jaba.1998.31-513

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491-498. doi: 10.1901/jaba.1992.25-491

Friedman, S. L., Kalichman, M. A., & Council on Children with Disabilities (2014). Out-of-home placement for children and adolescents with disabilities. *Pediatrics, 134,* 836-846. doi:10.1542/peds.2014-2279

Fyffe, C. E., Kahng S. W., Fittro, E., & Russell, D. (2004). Functional analysis and treatment of inappropriate sexual behavior. *Journal of Applied Behavior Analysis, 37,* 401-404. doi:10.1901/jaba.2004.37-401

Gresham, F. M., McIntyre, L. L., Olson-Tinker, H., Dolstra, L., McLaughlin, V., & Van, M. (2004). Relevance of functional behavioral assessment research for school-based

interventions and positive behavioral support. *Research in Developmental Disabilities, 25,* 19-37. doi:10.1016/j.ridd.2003.04.003

Hagopian, L. P., Fisher, W. W., Thibault Sullivan, M., Acquisto, J., & Leblanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis, 31,* 211-235. doi:10.1901/jaba.1998.31-211

Hagopian, L. P., Wilson, D. M., & Wilder, D. A. (2001). Assessment and treatment of problem behavior maintained by escape from attention and access to tangible items. *Journal of Applied Behavior Analysis, 34,* 229-232. doi:10.1901/jaba.2001.34-229

Hall, A. G., Bouldin, E. D., Andresen, E. M., & Ali, A. K. (2012). Maintaining employment among caregivers of individuals on a medicaid waitlist for services. *Journal of Disability Policy Studies, 23,* 121-128. doi:10.1177/1044207311420567

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36,* 147-185. doi:10.1901/jaba.2003.36-147

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis, 34,* 17-38. doi:10.1901/jaba.2001.34-17

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis, 30,* 459-473. doi:10.1901/jaba.1997.30-459

Hartley, S. L., Barker, E. T., Baker, J. K., Seltzer, M. M., & Greenberg, J. S. (2012). Marital satisfaction and life circumstances of grown children with autism across 7 years. *Journal of Family Psychology, 26,* 688-697.doi:10.1037/a0029354

Hastings, R. P., & Beck, A. (2004). Practitioner review: Stress intervention for parents of children with intellectual disabilities. *Journal of Child Psychology & Psychiatry, 45,* 1338-1349. doi:10.1111/j.1469-7610.2004.00841

Heyvaert, M., Saenen, L., Campbell, J. M., Maes, B., & Onghena, P. (2014). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: An updated quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 35,* 2463-2476. doi:10.1016/j.ridd.2014.06.017

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197–209. doi: 10.1901/jaba.1994.27-197 (Reprinted from *Analysis and Intervention in Developmental Disabilities*, *2*, 3–20, 1982).

Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of beaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability, 50,* 172-183. doi:10.1111/j.1365-2788.2005.00732.x

Lovaas, O. I., & Simmons, J. Q. (1969). Manipulation of self-destruction in three retarded children. *Journal of Applied Behavior Analysis, 2,* 143-157. doi:10.1901/jaba.1969.2-143

Mace, F. C., Browder, D. M., & Lin, Y. (1987). Analysis of demand conditions associated with stereotypy. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 25-31. doi:10.1016/0005-7916(87)90068-1

McCord, B. E., Thomson, R. J., & Iwata, B. A. (2001). Functional analysis and treatment of self-injury associated with transitions. *Journal of Applied Behavior Analysis, 34,* 195-210. doi:10.1901/jaba.2001.34-195

Minshawi, N. F., Hurwitz, S., Fodstad, J. C., Biebl, S., Morriss, D. H., & McDougle, C. J. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology Research & Behavior Management, 7,* 125-136. doi: 10.2147/PRBM.S44635

Neece, C. L., Green, S. A., & Baker, B. L. (2012). Parenting stress and child behavior problems: A transactional relationship across time. *American Journal on Intellectual and Developmental Disabilities, 117,* 48-66. doi:10.1352/1944-7558-117.1.48

Osborne, L. A., McHugh, L., Saunders, J., & Reed, P. (2008). Parenting stress reduces the effectiveness of early teaching interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 38,* 1092-1103. doi:10.1007/s10803-007-0497-7

Piazza, C. C., Hanley, G. P., Bowman, L. G., Ruyter, J. M., Lindauer, S. E., & Saiontz, D. M. (1997). Functional analysis and treatment of elopement. *Journal of Applied Behavior Analysis, 30,*653-672. doi:10.1901/jaba.1997.30-653

Piazza, C. C., Hanley, G. P., & Fisher, W. W. (1996). Functional analysis and treatment of cigarette pica. *Journal of Applied Behavior Analysis, 29,* 437-450. doi:10.1901/jaba.1996.29-437

Reimers, T. M., Wacker, D. P., Cooper, L. J., Sasso, G. M., Berg, W. K., & Steege, M. W. (1993). Assessing the functional properties of noncompliant behavior in an outpatient setting. *Child and Family Behavior Therapy, 15,* 1-15. doi:10.1300/J019v15n03_01

Rodriguez, N. M., Thompson, R. H., & Baynham, T. Y. (2013). Assessment of the relative effects of attention and escape on noncompliance. *Journal of Applied Behavior Analysis, 43,* 143-147. doi:10.1901/jaba.2010.43-143

Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The behavior problems inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders, 31,* 577-588. doi: 0162-3257/01/1200-0577/0

Vollmer, T. R., Northup, J., Ringdahl, J. E., LeBlanc, L. A., & Chauvin, T. M. (1996). Functional analysis of severe tantrums displayed by children with language delays: an out clinic assessment. *Behavior Modification, 20,* 97-115. doi:10.1177/01454455960201005

Wilder, D. A., Harris, C., Reagan, R., & Rasey, A. (2007). Functional analysis and treatment of noncompliance by preschool children. *Journal of Applied Behavior Analysis, 40,* 173-177. doi:10.1901/jaba.2007.44-06

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11,* 203-214. doi:10.1901/jaba.1978.11-203

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1700- 1715

Prácticas de Crianza Asociadas a la Reducción de los Problemas de Conducta Infantil: Una Aportación a la Salud Pública

Silvia Morales Chainé & Fernando Vázquez Pineda
Universidad Nacional Autónoma de México

Resumen

El entrenamiento conductual a padres resulta en cambios conductuales que están asociados con la reducción de la conducta problemática de los niños. El objetivo del trabajo fue identificar los componentes de un programa de entrenamiento conductual a padres para cambiar la conducta parental y la de sus hijos. Participaron 84 padres mexicanos de niños con problemas de conducta con edades de entre dos y 12 años. Los padres participaron en un programa de crianza positiva, que consistió en un juego de roles y que los entrenó a responder de forma positiva ante las diferentes conductas del niño con el fin de disminuir su conducta problemática. La conducta de los niños se evaluó mediante cuestionarios de auto-reporte que respondieron los padres. Los resultados mostraron que después del entrenamiento se observó que los padres modificaron su conducta, lo cual redundó en que reportaran que la conducta de sus hijos también se modificó. Los componentes más efectivos del programa fueron la corrección del comportamiento, el elogio, las instrucciones claras, el establecimiento de reglas, la solución de problemas, la interacción social y la reducción en el uso del castigo.

Palabras Clave: Conducta Infantil, Crianza Positiva, Prevención.

Child Raising Practices Associated to the Reduction of Problematic Child Behavior: A Public Health Approach

Abstract

Parent behavioral training results in behavioral changes that are associated with the reduction of a child's problematic behaviors. The purpose of the present study was to identify the most effective components of a parent behavioral training program on both, the observed parents' behavior and their report on the reduction of their child's problematic behavior. Participants were 84 Mexican parents of two to twelve years old children with behavioral problems. Parents attended a positive child raising program based in role playing that trained them to respond in a positive manner to their children's behaviors with the ultimate goal of reducing problematic behaviors. The children's behaviors were assessed using self-report questionnaires answered by the parents. After the behavioral training, a significant change in the parents' behavior was observed. Parents also reported a significant reduction of their child's problematic behavior. The most effective components of the intervention program were behavior correction, praise, giving clear instructions, rule establishment, problem solving, social interaction, and the reduction in punishment.

Keywords: Child Behavior, Positive Child Raising, Prevention.

Original recibido / Original received: 13/05/2014 Aceptado / Accepted: 10/09/2014

De acuerdo con la Organización Mundial de la Salud (OMS, 2014a) en América Latina entre el 3% y 4% de los niños y adolescentes padecen trastornos de conducta que requieren tratamientos especializados. Específicamente, el rango de la prevalencia del desorden por conducta negativista desafiante va del 1% al 11% con una prevalencia mundial promedio del 3.3%, mientras que el déficit de inatención con hiperactividad ocurre en el 5% de los niños en el mundo (Asociación Psiquiátrica Americana, APA, 2013). En México, los problemas más frecuentes en niños entre los 4 y los 7 años de edad han sido la conducta negativista desafiante, la desobediencia, la conducta agresiva y la hiperactividad (Medina-Mora et al., 2003).

Los problemas de conducta infantil se han definido como el grupo de "…conductas que violan los derechos de otros (p. ej., agresión o violación de la propiedad privada) y/o que promueven que el individuo se involucre en conflicto con las normas sociales o las figuras de autoridad" (APA, 2013, p. 461). Es común referirse a dichos problemas como el desorden negativista desafiante, la conducta agresiva e incluso el déficit de atención con o sin hiperactividad. Sin una intervención apropiada, es usual observar que los problemas de conducta se transforman en comportamiento antisocial o de consumo de drogas (Frick & White, 2008). Por lo que resulta importante interrumpir la progresión hacia el comportamiento antisocial a partir de intervenciones exitosas.

Existe evidencia de que favorecer el que los padres utilicen un estilo de crianza basado en apoyar a los niños y reforzarlos por sus logros es eficaz para la prevención de la violencia (OMS, 2014b). Boardman (1962) sugirió la necesidad de cambiar la conducta de los padres como una forma efectiva para cambiar la conducta de los niños. El entrenamiento conductual para padres se ha reconocido como la estrategia líder de intervención de las conductas disruptivas (Forehand, Jones, & Parent, 2013). El entrenamiento conductual se ha dirigido principalmente a las conductas de oposición, desobediencia y agresión infantil (Forgatch & Patterson, 2010). Chorpita et al. (2011) revisaron la efectividad de 23 tratamientos para modificar la conducta disruptiva de niños de entre dos y 18 años y encontraron que el entrenamiento conductual a padres mostró ser la forma más efectiva de intervención. Este hallazgo fue independiente de la edad de los niños, de su grupo étnico, de la modalidad de la intervención (familiar, grupal, individual o auto-administrado) y del escenario de aplicación (la clínica, el hogar, el ámbito hospitalario o la escuela).

La meta del entrenamiento conductual para padres es reducir su propia conducta coercitiva o negativa, así como el incremento de su conducta positiva, con el fin de lograr eliminar o disminuir las conductas disruptivas del niño. Las metas específicas del entrenamiento a padres han sido incrementar la atención a conducta apropiada, ejercer control a través de límites consistentes en escenarios e ignorar o usar tiempo fuera para la conducta inapropiada (McMahon, Wells, & Kotler, 2006). Cartwright-Hatton et al. (2011) utilizaron procedimientos típicos del entrenamiento conductual para padres (juego dirigido al niño, elogiar, ignorar y tiempo fuera) para modificar conductas de ansiedad de los niños. Lo importante, entonces, es que el padre sea el agente del cambio de las conductas problema del

niño, independientemente de si son conductas observables directamente o "internas".

Forehand et al. (2013) reportaron que pocos estudios han examinado las conductas de los padres y no han evaluado el efecto de su cambio conductual sobre los problemas de conducta infantil. Con el fin de describir qué prácticas de crianza promueven o previenen el desarrollo o mantenimiento de problemas de conducta en niños, Cornell y Frick (2007) analizaron la interacción de 87 diadas madre-hijo con niños entre los tres y los cinco años de edad. Los resultados indicaron que la disciplina basada en el establecimiento claro y consistente de reglas es más efectiva para que los niños aprendan a reaccionar favorablemente a las normas establecidas por los adultos. Concluyeron que las prácticas de crianza basadas en estrategias de obediencia, que evitan el uso de castigo corporal, fueron efectivas para reducir los problemas de conducta. Kochanskay y Murray (2000) también observaron que el incremento en la interacción positiva entre padres e hijos basada en la cooperación, el apego y la mutua reciprocidad generó emisión de conducta pro-social en los niños. A pesar de estos resultados, Cornell y Frick (2007) subrayaron que una limitación de sus hallazgos, así como el de otros estudios, fue que utilizaron únicamente pruebas psicométricas cuya confiabilidad y validez podría cuestionarse.

Eyberg, Nelson, y Boggs (2008) resaltaron la importancia de evaluar y reportar datos observacionales del comportamiento de los padres durante las intervenciones conductuales. Analizaron 34 estudios en los que se empleó una intervención para modificar la conducta disruptiva de niños. Encontraron que en 24% de dichos estudios no se reportaron datos de los padres y de los estudios que los reportaron, en el 20% no utilizaron datos observacionales. Forehand et al. (2013) señalaron que con sistemas de observación directa del comportamiento tanto de padres como de niños se podría identificar qué conductas de los padres y qué tipo de disciplina favorecen cambios en el comportamiento infantil.

En México, Morales, Félix, Rosas, López, y Nieto (en prensa) evaluaron la asociación entre las prácticas de crianza empleadas por 300 padres y la conducta negativista desafiante y de agresión de sus hijos. Emplearon tanto un sistema de observación directa de la conducta de los padres, como instrumentos psicométricos de evaluación. Encontraron que las conductas positivas en la interacción social y de seguimiento instruccional por los padres para la promoción de la obediencia infantil se asociaron con un reporte de bajo grado de conducta negativista desafiante y agresiva. Sin embargo, aún es necesario realizar estudios relacionados con la evaluación de las estrategias de crianza que podrían estar asociadas con la disminución de problemas de conducta. Particularmente, habría que averiguar el efecto de la interacción social positiva, el uso de la técnica de ignorar, el uso de las instrucciones claras, la solución de problemas, el establecimiento de reglas y la interacción académica positiva (Cornell & Frick, 2007; Morales et al., en prensa; Rakow et al., 2011).

En resumen, debido a que existe un número pequeño de estudios en los que se ha explorado cómo las prácticas de crianza fomentan la aparición de problemas de conducta, a que la edad de los niños participantes ha variado en los diferentes estudios y a que los hallazgos se han basado en auto-reportes de los

padres y no en la observación de su conducta y la de sus hijos, no se puede afirmar cuál práctica de crianza fomenta la emisión de conducta problemática y cuál de conducta pro-social (Forehand et al., 2013; Kendall, Settipani, & Cummings 2012). El manejo de los problemas de conducta infantil deben enfocarse en la evaluación del cambio en las conductas de los padres que permitan probar si tal modificación genera un cambio en el reporte de las conductas disruptivas de los niños (Forehand et al., 2013), especialmente, en un contexto de la salud pública donde es necesario promover la práctica basada en la evidencia (Chorpita et al., 2011; Morales, 2012). En consecuencia, el objetivo de este trabajo consistió en identificar cuáles componentes de un entrenamiento a padres son efectivos para promover el cambio en la conducta parental y consecuentemente en el reporte de la conducta infantil en niños mexicanos a través de un estudio pre-experimental en instituciones de salud pública.

Método

Participantes

Participaron 84 padres de ocho entidades de la República Mexicana, seleccionados usando un muestreo por cuotas de las instituciones de salud pública para la atención primaria a las adicciones existentes en cada entidad. Los padres fueron convocados a participar en el programa de intervención de crianza positiva por algún problema de conducta con alguno de sus hijos, entre los dos y los 12 años de edad. Ni el tamaño ni el nivel académico de los grupos en cada entidad fue homogéneo. El 6% de los participantes provenía del estado de Campeche, el 4% de Chihuahua, el 12% del Distrito Federal, el 5% Durango, el 4% de Guerrero, el 11% de Puebla, el 4% de San Luis Potosí y el 54% de Veracruz.

El promedio de edad de los padres participantes fue de 36 años, con un rango entre los 22 a los 71. El 93% fueron mujeres, el 12% eran solteros, el 54% casados, el 32% vivían en unión libre y el 2% estaban separado o divorciados. El 5% de los participantes no tenía estudios, el 20% estudiaron la primaria, el 44% secundaria, el 24% preparatoria y el 7% eran profesionales. El 71% se dedicaba al hogar, el 26% era empleado y el 3% comerciante. El 31% de los niños de los padres participantes tenía entre dos y seis años de edad y 69% tenía entre siete y 12 años. El 71% de los niños fueron varones y el 29% niñas.

Los participantes firmaron un consentimiento informado donde se estableció que la duración de su participación sería de ocho sesiones. Establecía que los padres aceptaban que se utilizaran los resultados del estudio para investigación epidemiológica y difusión de resultados. Se indicó que se cuidaría plenamente su identidad y se guardaría la confidencialidad de la información utilizando promedios grupales. También se especificó que tenían derecho a declinar el uso de su información y participación en cualquier momento del estudio sin perjudicar su intervención en el plan de tratamiento. El estudio no otorgó ningún tipo de incentivo a los participantes pero se les explicó el beneficio social de su participación en la implementación de estrategias efectivas para la atención

psicológica de su problemática social. Finalmente se les otorgó información del contacto para recibir información adicional.

Instrumentos

Se utilizaron cuestionarios psicométricos que respondieron los padres para obtener un reporte de su propia conducta y de la de sus hijos. También se empleó un sistema de observación directa del comportamiento de los padres. Los cuestionarios psicométricos que se emplearon fueron el Inventario de Prácticas de Crianza, el Cuestionario de Habilidades de Manejo Infantil, el Inventario de Conducta Infantil y el Cuestionario de Validez Social.

El Inventario de Prácticas de Crianza *(*IPC; López, 2013) es un cuestionario auto-aplicable de lápiz y papel de 20 minutos de aplicación aproximadamente. Consta de 40 preguntas cerradas, que se responden en una escala de siete opciones, que van de nunca (0) hasta siempre (6), que evalúan las conductas de los padres con respecto a la disciplina y a la promoción del afecto de sus hijos. El IPC fue validado con una muestra de 260 participantes y se obtuvo un nivel de confiabilidad de .92 (a través del análisis de consistencia interna por alfa de Cronbach). Mediante un análisis factorial exploratorio se encontraron seis factores (castigo*,* ganancias materiales, interacción social, ganancias sociales y la dimensión límites) que explicaron el 64% de la varianza (para la descripción de las escalas empleadas en este trabajo véase Morales et al., en prensa).

El Cuestionario de Habilidades de Manejo Infantil está basado en situaciones hipotéticas de crianza (CHAMI; Morales & Vázquez, 2011). Es un cuestionario auto-aplicable de lápiz y papel con 10 viñetas de evaluación sobre las habilidades de manejo de conducta infantil. Son situaciones hipotéticas de interacción problemática con el niño donde los padres responden, de manera abierta, qué harían ante dicha situación. Cada viñeta es calificada con base en tres posibilidades: 0 si el padre no describe la habilidad; 1si la describe parcialmente; o 2 si la describe completamente. El instrumento fue validado con 294 participantes de distintos estados del país, obteniendo una consistencia interna por alfa de Cronbach de .62 y una varianza explicada del 55%, a través del análisis factorial exploratorio que arrojó cuatro escalas: ignorar como técnica para promover conducta adecuada (ITCA), elogio*,* instrucciones claras, solución de problemas y establecimiento de reglas (ICSE) y la dimensión de interacción social-académica (ISA). La concordancia entre evaluadores fue del 80%.

El Inventario de Conducta Infantil (ICI; Morales & Martínez, 2013) es un instrumento auto-aplicable de lápiz y papel de 30 reactivos que puede resolverse aproximadamente en 30 minutos. La consistencia interna del instrumento fue de .93 (por alfa de Cronbach). Un análisis factorial exploratorio mostró la existencia de cuatro factores: comportamiento oposicionista desafiante, comportamiento agresivo, inatención y la dimensión de hiperactividad. Los cuatro factores explicaron el 53% de la varianza. Cada pregunta se responde empleando una escala tipo Likert de cinco puntos (0 = nunca; 4 = siempre; Morales et al., en prensa).

El Cuestionario de Validez Social (Morales & Martínez, 2013) es un cuestionario auto-aplicable de lápiz y papel con 22 reactivos que evalúan la satisfacción de los participantes con el entrenamiento a padres. El instrumento obtuvo una consistencia interna

por alfa de Cronbach de .91 y una varianza explicada del 71%, a través del análisis factorial exploratorio que arrojó tres factores: metas (evalúa la congruencia entre las metas del programa y las del participante); procedimientos (evalúa la complejidad que percibe el participante con los procedimientos del entrenamiento a padres); y la dimensión resultados (evalúa la satisfacción con los resultados logrados con el entrenamiento).

El Sistema de Observación Directa (Morales & Martínez, 2013) está constituido por tres listas cotejables, cuatro registros de evento y tres registros de intervalo parcial de tiempo. Las tres listas cotejables evaluaron la corrección simple del comportamiento, el establecimiento de reglas y la solución de problemas. Los cuatro registros de evento evaluaron el elogio de conducta académica, el seguimiento instruccional (e.g., obtiene atención, da instrucción clara, espera, elogia) y las conductas de interacción social (e.g., compartir, elogiar, risa provocada, mirar, sonreír, reír, tocar y peticiones verbales) e interacción académica.

Los tres registros de intervalo parcial de tiempo de 10 segundos durante 5 minutos evaluaron las mismas conductas de interacción social, interacción académica y enseñanza incidental, pero en estos se obtuvo el porcentaje de intervalos donde ocurrieron las conductas. A partir de las listas cotejables, de los registros de evento y de los registros de intervalo parcial se obtuvieron los porcentajes de la conducta meta o de los intervalos registrados por conducta y promedios globales de interacción, así como porcentaje de padres que puntuaron la ocurrencia de conductas observadas a partir del cuartil más alto del total de conductas ejecutadas por otros padres durante las sesiones de entrenamiento. Para revisión de las definiciones de las conductas consulte Morales y Martínez (2013).

En todos los registros, se obtuvo la concordancia entre dos observadores independientes y sólo se consideraron aceptables aquellos registros cuya concordancia fue mayor al 80%. La concordancia se obtuvo a partir del cálculo de los acuerdos [(acuerdos/acuerdos + desacuerdos) x 100].

Procedimiento

Se utilizó un estudio pre-experimental. Tanto la evaluación previa como la posterior al entrenamiento consistieron en dos sesiones de evaluación (dos escritas con formato grupal y dos individuales) con duración de 120 minutos cada una y el entrenamiento a padres consistió en cuatro sesiones más. En la sesión de evaluación escrita grupal, los participantes recibieron los cuestionarios psicométricos descritos en el apartado de instrumentos y de forma grupal se dieron las siguientes instrucciones:

“En esta sesión se realizarán una serie de cuestionarios que nos permitirán conocer las habilidades con las que ustedes cuentan para corregir a sus hijos en este momento y la frecuencia con la que se observan ciertas conductas en ellos. El llenado de los cuestionarios es individual, ¿Tienen alguna pregunta? Podemos comenzar”.

En la sesión individual de 120 minutos se llevó a cabo una evaluación de situaciones simuladas a través de ensayos conductuales entre el profesional de la salud (ejecutando el papel del niño) y cada padre, a lo largo de ocho grupos de

estímulos relacionados con las habilidades y conductas de interacción entre ambos. El primer grupo de estímulos que presentó el profesional a los padres estuvo constituido por 10 situaciones de evaluación sobre corrección del comportamiento infantil (ante la entrega de reportes escolares, la hora de la comida, en el supermercado y la pelea entre hermanos) con duración máxima de 15 segundos por cada ensayo. El segundo grupo de estímulos estuvo constituido por la presentación de obediencia a 10 instrucciones académicas para la evaluación del reforzamiento positivo (elogio) ante la obediencia en esta situación. El tercer grupo de estímulos estuvo constituido por la oportunidad de ocurrencia de conductas de interacción social (el profesional, en su papel de niño, repetía la misma conducta observada en el padre de manera contingente a su emisión (e.g., sonreír). El cuarto grupo de estímulos estuvo constituido por la oportunidad de obediencia a 10 instrucciones que se solicitaba al padre dar (el profesional en su papel de niño mostraba obediencia si el padre seguía el formato de instrucciones claras o desobedecía si el padre omitía alguno de los pasos de la instrucción clara). El quinto grupo de estímulos estuvo constituido por el establecimiento de reglas en una situación simulada a la hora de la comida. El sexto grupo de estímulos estuvo constituido por la evaluación de habilidades a la hora de la tarea (el profesional realizaba correctamente tres sumas y se equivocaba en otras dos durante el ensayo conductual). El séptimo grupo de estímulos estuvo constituido por la evaluación de la situación de enseñanza incidental durante la preparación del agua de limón (Morales, 2001). El octavo grupo de estímulos estuvo constituido por la evaluación de la situación de solución de problemas para mantener la habitación del niño arreglada (para mayor detalle ver Morales & Martínez, 2013).

Para la realización de los ensayos conductuales se dieron las siguientes instrucciones:

“A continuación realizaremos una serie de situaciones de evaluación en las que yo jugaré el papel de un niño. Sé que es un poco inusual para usted este tipo de evaluación, pero lo importante de estos ensayos consiste en brindarle la oportunidad de mostrar las herramientas con que cuenta actualmente para resolver la conducta de su hijo. Por eso, por favor, es importante que procure imaginar que yo soy un niño (su hijo, si es posible) y actúe y diga lo que considere necesario para resolver las situaciones que se le presenten. ¿Tiene alguna duda? Comencemos.”

La fase de entrenamiento a padres consistió en la aplicación de los procedimientos derivados de los principios básicos del comportamiento tales como el reforzamiento positivo, el castigo negativo, la extinción de la conducta mantenida por reforzamiento positivo o negativo y el control de estímulos para la programación de la generalización del comportamiento entre escenarios, participantes, o a lo largo del tiempo. En las cuatro sesiones sesiones se utilizaron como estrategias de entrenamiento conductual: la instrucción verbal, el modelamiento de habilidades, los ensayos conductuales y la retroalimentación de la ejecución de habilidades en situaciones simuladas, en ese orden.

Particularmente, en la primera sesión del entrenamiento se trabajaron las estrategias relacionadas con el análisis funcional del comportamiento infantil a

través de la identificación del contexto asociado a la conducta meta y las consecuencias inmediatas a la misma (CCC) y el reforzamiento diferencial del comportamiento alternativo (RDA). Durante esta sesión se llevaron a cabo tres ensayos conductuales para la identificación del CCC y se asignaron dos tareas: CCC y RDA. En la segunda sesión se revisaron las tareas, verificando el número de comportamientos identificados y el tipo de consecuencia otorgada (reforzamiento o corrección), se procedió al entrenamiento de conductas parentales para la interacción positiva (IP) y de seguimiento de instrucciones (SI). Durante esta sesión se llevaron a cabo dos ensayos conductuales, uno para interacción y otro para seguimiento instruccional y se pidieron cuatro tareas: CCC, RDA, IP y SI. En la tercera sesión se revisaron las tareas (verificando la cantidad de conductas identificadas por cada padre y el porcentaje de reforzamiento y correcciones otorgadas a éstas) y se procedió al entrenamiento de estrategias para la corrección (C) y extinción (E) de comportamiento meta. Durante esta sesión se llevaron a cabo dos ensayos conductuales de interacción académica y dos ejercicios de identificación del CCC sobre berrinche. Se dejaron seis tareas: CCC, RDA, IP, SI, C y E. En la sesión cuatro se revisaron las tareas y se entrenaron las habilidades para el control de estímulos: organización y manejo del tiempo, identificación de situaciones de riesgo al comportamiento meta infantil, establecimiento de reglas, aplicación de reprimendas, pérdida de privilegios y solución de problemas en familia. Se llevó a cabo la identificación de problemas de conducta en el hogar (mañana, tarde y noche) y en la comunidad (visitas, salidas, viajar, compras y separación) y se llevaron a cabo ensayos conductuales de enseñanza incidental, establecimiento de reglas y solución de problemas y un ejercicio de identificación de técnicas de corrección (ver Morales & Martínez, 2013). Se concluyó con el cierre de la sesión y la programación de la evaluación.

En la evaluación final, se aplicaron todos los instrumentos de la pre-evaluación escrita grupal y del sistema de observación directa en ensayos conductuales, con las mismas instrucciones y se agregó el cuestionario de validez social.

Análisis estadísticos

Para el análisis de los datos se llevaron a cabo los análisis descriptivos (promedios y las desviaciones estándar, como medidas de tendencia central) de cada variable medida psicométricamente y a través de la observación directa. Para estimar el efecto de una variable sobre otra se utilizó el modelo de regresión lineal múltiple (conducta del padre sobre el reporte de conducta infantil) y se aplicó la prueba *t* para muestras relacionadas. Todos los análisis se realizaron a través del paquete estadístico SPSS® versión 15.0 para Windows®. Se estableció un nivel de significancia menor a .003.

Resultados

Primero se muestra el porcentaje promedio de las conductas de los padres y el reporte de conducta infantil en las evaluaciones previas y posteriores al entrenamiento conductual a padres. En seguida se muestran los porcentajes de adherencia a los procedimientos durante los ensayos conductuales y actividades de las sesiones del entrenamiento a padres y finalmente se presenta la relación predictiva significativa entre las mediciones post de las conductas de los padres con el reporte de conducta infantil.

En la Tabla 1 se muestran los porcentajes promedio de los puntajes en las escalas psicométricas y obtenidos mediante el sistema de observación directa antes y después del entrenamiento a padres y los coeficientes *t*. Como se puede observar, en la escala IPC, los padres reportaron un promedio de castigo antes del entrenamiento mayor (*M* = 71.64, *D.E.* = 16.65) que después del mismo (*M* = 52.1, *D.E.* = 22.16*)*. Hubo incrementos estadísticamente significativos en los porcentajes promedio de todas las escalas del CHAMI. Los incrementos fueron del 34% para el ITCA, del 35% para el elogio e ICSE y del 25% para ISA.

En la misma Tabla 1, se observa que resultaron significativos todos los incrementos en los porcentajes promedio de las conductas del sistema de observación directa. Los incrementos fueron del 31% en las conductas de los padres de corrección simple de la conducta infantil, del 23% en las conductas de elogio, del 14% en las conductas de interacción, del 24% en las conductas de seguimiento instruccional, del 26% en las conductas de establecimiento de reglas, del 6% en las conductas de interacción académica, del 11% en la enseñanza incidental y del 27% en la solución de problemas.

En cuanto al reporte de conducta infantil (Tabla 1), se puede observar que hubo decrementos significativos en los trastornos del comportamiento infantil. En el caso del comportamiento negativista desafiante hubo un decremento del 25.13% al 20.11%. El reporte de conducta agresiva disminuyó significativamente del 60.11% al 51.41%. El reporte de conducta de inatención decrementó del 50.57% al 19.26%. El promedio de comportamiento de hiperactividad se redujo del 50.30% al 19.39%.

Tabla 1
Promedio (*M*) y desviación estándar (*DE*) de los puntajes en las escalas psicométricas y en el sistema de observación directa de los participantes pre y post entrenamiento, coeficientes *t* y su significancia.

Escalas	Evaluación Promedio Pre *M (DE)*	Post *M (DE)*	*t* (82)	*p*
IPC				
Castigo	71.64 (16.65)	52.1 (22.16)	8.296	.001
Ganancias materiales	52.54 (21.92)	50.50 (23.13)	.609	0.544
Interacción	72.4 (17.85)	70.68 (23.62)	.502	0.617
Normas	77.17 (21.34)	79.38 (26.67)	-.574	0.568
Ganancias sociales	76.30 (21.16)	78.23 (27.30)	-.508	0.613
Limites	70.08 (23.55)	76.24 (27.77)	-1.561	0.122
CHAMI				
ITCA	27.51 (27.84)	61.44 (26.91)	-10.169	.001
ELOGIO	40.96 (25.17)	75.90 (22.24)	-10.196	.001
ICSE	30.27 (27.96)	65.06 (18.90)	-9.754	.001
ISA	36.75 (27.97)	61.75 (30.07)	-6.563	.001
OBS				
Corrección simple	14.66 (17.2)	45.27 (25.22)	-9.331	.001
Elogio	33.27 (20.23)	56.55 (22.31)	-11.404	.001
Interacción social	17.19 (11.16)	31.35 (17.63)	-8.533	.001
Seguimiento instruccional	46.68 (25.58)	70.32 (17.11)	-9.737	.001
Establecimiento de reglas	40.15 (23.95)	66.53 (20.96)	-9.705	.001
Interacción académica	12.45 (8.09)	18.85 (11.88)	-6.392	.001
Enseñanza incidental	8.04 (7.07)	19.31 (16.46)	-8.064	.001
Solución de Problemas	42.08 (21.18)	69.49 (18.88)	-11.226	.001
ICI				
ODD	25.13 (6.64)	20.11 (5.13)	8.138	.001
AGRESIÓN	60.11 (2.5)	51.41 (1.41)	26.534	.001
Inatención	50.57 (7.6)	19.26 (5.76)	36.252	.001
Hiperactividad	50.30 (10.06)	19.39 (5.85)	25.721	.001

En la Tabla 2 se muestra el porcentaje de padres o del promedio de conductas paternas observadas durante los ensayos conductuales del entrenamiento. El porcentaje de padres que llenaron correctamente el CCC fue del 66.09% en la Sesión 1 y del 72.90% durante la Sesión 3. Los padres mostraron un incremento en el porcentaje promedio de comportamiento de corrección de la Sesión 2 (23.69%) a la 3 (57.10%) y 4 (50%) en los auto-reportes del CCC, así como porcentajes relativamente altos en los ensayos conductuales durante el entrenamiento: instrucción clara (70.94%, Sesión 2), establecimiento de reglas (66%) y solución de problemas (57.40%, Sesión 4). El 27.10% de padres registró conductas de interacción positiva por arriba de 22 conductas durante la Sesión 2;

el 32.40% de padres conductas de interacción académica por arriba de 13 conductas en la Sesión 3; y el 32% conductas de enseñanza incidental por arriba de 15 en la Sesión 4. En cuanto a las listas cotejables, se puede observar que los padres reportaron un porcentaje promedio del 11.43% de problemas en el hogar siempre y del 3% en la comunidad (después de tres sesiones de trabajo), para ser abordados en la cuarta sesión (porque seguían considerándose por los padres como situaciones a resolver, por ejemplo a la hora de la comida o al irse a dormir). Finalmente, el 65.30% de los padres identificó las correcciones durante un procedimiento de establecimiento de reglas durante la cuarta sesión y el 58.30% reportó la ejecución de los procedimientos de actividades planeadas durante la hora de la comida, en la última sesión.

Tabla 2

Porcentaje promedio de conductas correctas del padre y porcentaje de padres superiores al cuartil 75 durante los ensayos conductuales o los ejercicios de práctica de las cuatro sesiones del entrenamiento a padres.

Padres o conductas	Porcentaje
Sesión 1	
Padres que identificaron correctamente el contexto-conducta.consecuencias de la conducta infantil	66.09%
Sesión 2	
Corrección por pérdidas del comportamiento problema	23.69%
Padres en el cuartil más alto de conductas de interacción social positiva (22 conductas o más)	27.10%
Instrucción clara	70.94%
Sesión 3	
Corrección por pérdidas del comportamiento problema	57.10%
Padres en el cuartil más alto de conductas de interacción académica (13 conductas o más)	32.40%
Padres que llenaron correcto del contexto-conducta.consecuencias del berrinche	72.90%
Sesión 4	
Corrección por pérdidas del comportamiento problema	50%
Problemas en el hogar siempre	11.43%
Problemas en la comunidad siempre	3%
Padres en el cuartil más alto de conductas de enseñanza incidental (15 conductas o más)	32.00%
Padres que identificaron correcciones deseadas durante el establecimiento de reglas	65.30%
Establecimiento de reglas a la hora de la comida	66.00%
Actividades planeadas	58.30%
Solución de problemas	57.40%

La Tabla 3 representa los reportes y comportamiento de los padres que predicen el reporte de comportamiento infantil. Un análisis de regresión lineal múltiple mostró un nivel predictivo tanto del reporte, como de las conductas de los padres (R^2 = .44) sobre el reporte de comportamiento oposicionista desafiante F(1, 81) = 4.04, p = .001; el del comportamiento agresivo R^2=.36; F(1, 81) = 2.89, p = .002; del comportamiento de inatención R^2 = 43; F(1, 81) = 4.02, p = .001; y del comportamiento hiperactivo R^2 = .31; F(1, 81) = 2.39, p = .010. El análisis del cuestionario de validez social mostró que los participantes reportaron una satisfacción del 81% con las metas del programa, del 68% con los procedimientos y del 80% con los resultados del mismo.

Tabla 3
Reporte de conducta infantil y comportamiento de los padres que predicen dicho reporte de comportamiento infantil

	Coeficientes de regresión		
ESCALAS/OBS	*B*	*t*	*p*
Conducta Negativista Desafiante			
Castigo	0.487	3.613	.001
ICSE	0.303	2.21	.03
Corrección del comportamiento	-.282	-2.284	.026
Elogio	.345	2.205	.031
Interacción social	-.353	-2.019	.047
Agresión			
Corrección del comportamiento	.330	2.502	.015
Elogio	-.383	-2.294	.025
Inatención			
Castigo	.426	3.157	.002
ICSE	.326	2.381	.020
Corrección del comportamiento	-.334	-2.701	.009
Hiperactividad			
Castigo	.376	2.530	.014
Corrección del comportamiento	-.311	-2.287	.025

Discusión

El objetivo de este trabajo fue identificar cuáles componentes de un entrenamiento a padres son efectivos para promover el cambio en la conducta parental y consecuentemente en el reporte de la conducta infantil en niños mexicanos, de una muestra en instituciones de salud pública. Con esto se intentó verificar la aplicación de procedimientos derivados de los principios básicos, como el reforzamiento y la extinción, en el diseño de intervenciones preventivas que ponen particular atención en la conducta de los padres y reducen el reporte de conducta problemática infantil.

Los resultados del presente estudio parecen indicar que las estrategias de la instrucción verbal (sobre el análisis funcional, la interacción social positiva, seguimiento de instrucciones, interacción académica, corrección del comportamiento y la planeación de actividades), el modelamiento de habilidades (de interacción y seguimiento de instrucciones), los ensayos conductuales (del análisis funcional, de interacción social, académica, enseñanza incidental y seguimiento de instrucciones) y la retroalimentación de todas estas ejecuciones (como del análisis funcional en situaciones reales y de los problemas de conducta en el hogar y la comunidad) pudieran influir tanto en la adquisición de conductas parentales de crianza positiva como en la reducción del reporte de comportamiento negativista desafiante, agresión, inatención y de hiperactividad en los niños. Estos hallazgos son consistentes con la literatura previa sobre el entrenamiento a padres (Forehand et al., 2013; McMahon et al., 2006).

Los resultados concuerdan con las afirmaciones previas de que el entrenamiento conductual a padres promueve su adquisición de conductas relacionadas con la corrección apropiada del comportamiento infantil y la promoción de conducta pro-social en sus hijos (Chorpita et al., 2011). En particular, los resultados fueron congruentes con lo señalado por Cartwright-Hatton et al. (2011) quienes reportaron que el uso del reforzamiento positivo (elogio) y la extinción de la conducta mantenida por reforzamiento positivo o negativo (el ignorar conducta inadecuada) favorecen el éxito en la reducción de problemas de conducta en niños. En el presente estudio, dichos procedimientos, aunados al castigo negativo y al control de estímulos (Morales & Vázquez, 2011) favorecieron la reducción en el reporte de conductas problemáticas.

Parece ser que las estrategias como la interacción positiva entre padres e hijos constituyen una estrategia exitosa dentro del entrenamiento a padres, tal como señalaron Krochanska y Murray (2000) y que el seguimiento de instrucciones favorece el control efectivo de los padres del comportamiento infantil (McMahon et al., 2006). Tal como lo indicaron Chorpita et al. (2011) y Forehand et al. (2013), durante el entrenamiento a padres se promovió la ejecución de habilidades para llevar a cabo el análisis funcional del comportamiento, el RDA, estrategias para la corrección simple, castigo negativo, extinción de comportamiento meta y el manejo apropiado de contingencias. En el presente trabajo, la estrategia de ignorar, el elogio, el seguimiento de instrucciones, el establecimiento de reglas, la interacción social y académica y la corrección simple del comportamiento se asociaron con la reducción del comportamiento negativista

desafiante, la agresión, la inatención y la hiperactividad, lo que coincide con los resultados de Cornell y Frick (2007) y McMahon et al. (2006).

Es posible que la asociación entre las conductas observadas en los padres y el reporte de conducta infantil indique la presencia de las primeras asociada al reporte de las segundas como sugiere Forehand et al. (2013). Si éste fuera el caso, entonces los resultados del presente estudio representarían un esfuerzo de la medición precedente en tiempo de la conducta de los padres como predictor del comportamiento infantil como señala Kendall et al. (2012). En este estudio, el entrenamiento conductual resultó en un esfuerzo efectivo para que el comportamiento de los padres representara un papel importante en el reporte de reducción de la conducta problemática infantil (Granic & Patterson, 2006). Específicamente, promover la interacción positiva entre padres e hijos y estrategias de obediencia por medio del seguimiento de instrucciones, podría estar reduciendo la interacción coercitiva entre la diada padre hijo y por lo tanto reducir el comportamiento problemático infantil (Burke Pardini, & Loeber, 2008). Incrementar la interacción durante el juego, el elogio e ignorar conducta no deseada como plantearon Cartwright-Hatton et al. (2011) podría promover una incompatibilidad con la ocurrencia de interacción coercitiva y por lo tanto una mayor probabilidad de reducción del comportamiento problema del niño. Sin embargo, y a partir del uso de ensayos conductuales, es necesario que nuevos estudios se dediquen a observar directamente el cambio en el comportamiento infantil, y analizar su concordancia con el reporte de conducta por parte de los padres. Así mismo, estudios adicionales podrían abordar el efecto de la adquisición de conducta particular en los padres sobre la conducta de los niños, con comparación de grupos o evaluando problemas de conducta más severos o con trastorno dual en los niños y su reactividad al entrenamiento a padres. También, en los nuevos estudios se debe considerar, el nivel de involucramiento óptimo de los padres con los niños en el hogar y en la comunidad (Lavigne et al., 2008).

Una aportación del presente estudio fue evaluar, a través de ensayos conductuales, el comportamiento de los padres a través de un sistema de observación directa. Los hallazgos de la efectividad del programa de entrenamiento a padres, observados en este estudio, promueven su adopción en las instituciones de salud pública, en tanto son innovaciones prácticas como intervenciones breves, útiles por la cobertura a la demanda de servicio. Es decir, el contexto donde se obtuvieron los hallazgos fue en escenarios de salud pública, observándose evidencia del papel del comportamiento parental que podría estar asociado al de conducta infantil (Eyberg et al., 2008; Rakow et al., 2011).

A partir de las aportaciones del presente estudio surge la necesidad de considerar los elementos que permitan entender los cambios en la conducta de los padres y los del niños, considerando que estas prácticas basadas en la evidencia han formado parte de una política de salud pública reciente en México. Por ejemplo, estudios adicionales podrían considerar la varianza del cambio alcanzada por tratamientos dirigido exclusivamente a los niños y mostrar también si el tratamiento para el niño y el padre es más eficaz que el tratamiento sólo para el niño. Forehand et al. (2013) sugirieron evaluar además los diferentes papeles o

funciones del padre, es decir, como agente de cambio o como modelo de conductas saludables. Es común observar que los padres modelan y muestran cómo ejecutar la solución de problemas o la interacción positiva, por lo que resulta primordial evaluar el grado en que el niño las imita y con su adopción funge como predictor del cambio en la conducta problemática.

La contribución del presente estudio a la psicología aplicada, consiste en conocer cómo funcionan los programas conductuales de intervención efectivos, que derivan de los principios básicos del comportamiento y que explican la reducción del comportamiento infantil, en escenarios clínicos de alta demanda. La identificación de las prácticas de crianza asociadas a la reducción del comportamiento infantil permitirán diseñar procedimientos de intervención de bajo costo y alta efectividad, aceptados también socialmente por las instituciones que los adoptarán, y que van dirigidos a los niños denominados como de temperamento difícil buscando en un futuro detener la progresión del comportamiento hacia actos delictivos, de violación de normas y de la propiedad privada en las comunidades (Frick & White, 2008).

Referencias

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition. Arlington, VA, E.U: Autor

Boardman, W. K. R. (1962). A brief behavior disorder. *Journal of Consulting Psychology, 26*, 293- 297.

Burke, J. D., Pardini, D. A.& Loeber, R. (2008). Reciprocal relationships between parenting behavior and disruptive psychopathology from childhood through adolescence. *Journal of Abnormal Child Psychology, 26,* 679-692.

Cartwright-Hatton, S., McNally, D. Field, A. P., Rust, S., Laskey, B., &Woodham, A. (2011). A new parenting-based group intervention for young anxious children: Results of a randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 50,* 242-251.

Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., Phillips, L., Ward, A., Lynch, R., Trent, L., Smith, R. L., Okamura, K., & Satarace, N. (2011). Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice, 18*, 155-172.

Cornell, A. H., & Frick, P. J. (2007).The moderating effects of parenting styles in the association between behavioral inhibition and parent-reported guilt and empathy in preschool children. *Journal of Clinical Child and Adolescent Psychology*, *36*, 305–318. doi: 10.1080/15374410701444181

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology, 37,* 215-237.

Forehand, R., Jones, D. J. & Parent, B. A. J. (2013). Behavioral parenting interventions for child disruptive behaviors and anxiety: What´s different and what´s the same? *Clinical Psychology Review*, 33, 133-145. doi:10.1016/j.cpr.2012.10.010.

Forgatch, M. S., & Patterson, G. R. (2010). Parent management training – Oregon model: An intervention for antisocial behaviors in children and adolescents. En: Weisz, J. R., Kazdin, A. E., (Eds), *Evidenced-based psychotherapies for children and adolescents 2*, p. 159-178. Nueva York, E.U.: Guilford.

Frick, P. J. & White, S. F. (2008). Research review: The importance of callous-unemotional traits for development models of aggressive and antisocial behavior. *The Journal of Child Psychology and Psychiatry, 49*, 359-375. doi: 10.1111/j.1469-7610.2007.01862.x

Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 11,* 101-131.

Kendall, P. C., Settipani, C. A., Cummings, C. M. (2012). No need to worry: The promising future of child anxiety research. *Journal of Clinical Child and Adolescent Psychology, 41,* 103-115.

Kochanska, G., & Murray K. (2000). Mother–child mutually responsive orientation and conscience development: From toddler to early school age. *Child Development*, *71*, 417–431. doi: 10.1111/1467-8624.00154

Lavigne, J., LeBaillly, S., Gouze, K. R., Cicchetti, C., Jessup, B. W., Arend, R., & Binns, H. (2008). Predictor and moderator effects in the treatment of oppositional defiant disorder in pediatric primary care. *Journal of Pediatric Psychology, 33,* 462-472.

López, C. F. (2013). Inventario de prácticas de crianza. En C. S. Morales, & R. M. J. Martínez (Eds), *Prevención de las conductas adictivas a través de la atención del comportamiento infantil para la crianza positiva. Manual del Terapeuta* (pp. 14-19). México: CENADIC-SSA.

McMahon, R. J., Wells, K. C. & Kotler, J. S. (2006). Conduct problems. En Mash, E. J., Barkley, R. A. (Eds), *Treatment of childhood disorders,* p. 137-268. Nueva York, E. U.: Guilford Press.

Medina-Mora, M. E., Borges, G., Lara, C., Benjet, C., Blanco, J., Fleiz, C., y Aguilar-Gaxiola, S. (2003).Prevalencia de trastornos mentales y uso de servicios: Resultados de la encuesta nacional de epidemiología psiquiátrica en México. *Salud Mental*, *4*, 1-16.

Morales, S. (2001). Programa de entrenamiento conductual a padres. En H. Ayala, C. L. Chaparro, J. M. Fulgencio, C. Pedroza, S. Morales, T. Pacheco, G. Mendoza, A. Ortiz, S. Vargas, & N. Barragán (Eds), Tratamiento de agresión infantil: Desarrollo y evaluación de programas de intervención conductual multi-agente. *Revista Mexicana de Análisis de la Conducta, 27*,1-118.

Morales, C. S. (2012). Las habilidades metodológicas y conceptuales en el quehacer cotidiano de la ciencia del comportamiento. *Revista Mexicana de Investigación en Psicología, 4*, 126-129.

Morales, C. S., Félix, R. V., Rosas, P. M., López, C. F., & Nieto, G. J. (en prensa). Prácticas de crianza asociadas al comportamiento infantil. *Avances en Psicología Latinoamericana*.

Morales, C. S. & Martínez, R. M. J. (2013). *Prevención de las conductas adictivas a través de la atención del comportamiento infantil para la crianza positiva. Manual del Terapeuta*. México: CENADIC-SSA.

Morales, C. S., & Vázquez, P. F. (2011). Evaluación de conocimientos sobre habilidades de manejo conductual infantil en profesionales de la salud. *Acta de Investigación Psicológica, 1*, 428-440.

Organización Mundial de la Salud (2014a). Revisado el 26 de Julio de 2014 en http://www.paho.org/bol/index.php?option=com_content&view=article&id=221:la-mayoria-menores-problemas-salud-mental-no-recibe-tratamiento&catid=667:notas-de-prensa.

Organización Mundial de la Salud (2014b). La prevención de la violencia: Evaluación de los resultados de programas de educación para padres. Revisado el 26 de Julio de 2014 en http://apps.who.int/iris/bitstream/10665/103610/1/9789243505954_spa.pdf?ua=1&ua=1.

Rakow, A., Forehand, R., McKee, L., Roberts, L., Haker, K. & Compas, B. (2011). Parent guilt induction and child internalizing problems when a caregiver has a history of depression. *Journal of Family Psychology, 25,* 147-151.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1716 - 1745

Actions vs. Words: How We Can Learn Both

Claire S. Cahill[1] & R. Douglas Greer
Graduate School of Arts and Sciences and Teachers College Columbia University

Abstract

In three experiments we investigated the relation between observing responses and incidental language acquisition by children ages 3 to 5 with and without disabilities. In Experiment I, participants heard the name of an object while observing an accompanying action with the object. The participants consistently acquired the actions associated with the objects, but learned few names. Experiment II compare responses to stimuli presented with and without actions, with the results indicating that the presence of an action hindered rather than facilitated incidental acquisition of names. In Experiment III, we selected participants who acquired listener responses when actions were present, but did not readily acquire the speaker responses. Following a multiple exemplar intervention, participants acquired both speaker and listener responses along with the action responses for novel stimuli. The findings suggest that when children are provided with a specific instructional history, they can acquire multiple benefits from a single language exposure experience.

Keywords: Observing Responses, Stimulus Control, Conditioned Reinforcement, Sensory Dominance, Language Acquisition

Acciones vs. Palabras: Cómo Podemos Aprender Ambas

Resumen

En tres experimentos se investigó la relación entre respuestas de observación y la adquisición de lenguaje incidental por niños de 3 a 5 años con y sin discapacidad. En el Experimento I, los participantes escucharon el nombre de un objeto mientras observaban una acción que acompañó al objeto. Los participantes consistentemente adquirieron las acciones asociadas con los objetos, pero aprendieron pocos nombres. El Experimento II comparó las respuestas ante estímulos presentes con y sin acciones. Los resultados indicaron que la presencia de una acción dificultó en lugar de facilitar la adquisición incidental de los nombres. En el Experimento III, se seleccionaron participantes que adquirieron respuestas de oyente cuando las acciones estaban presentes, pero que no habían adquirido las respuestas de hablante. Después de una intervención múltiple ejemplificada, los participantes adquirieron tanto las respuestas de oyente como las de hablante conjuntamente con las respuestas de acción para estímulos novedosos. Los resultados sugieren que cuando se provee a los niños con una historia instruccional específica adquieren beneficios múltiples de una sola exposición de experiencia con el lenguaje.

Palabras Clave: Respuestas de Observación, Control de Estímulos, Reforzamiento Condicionado, Dominancia Sensorial, Lenguaje, Adquisición

Original recibido / Original received: 03/06/2014 Aceptado / Accepted: 05/10/2014

[1] E-mail: Claire S. Cahill, claire.cahill@gmail.com; R. Douglas Greer, rdg13@columbia.edu
Graduate School of Arts and Sciences and Teachers College Columbia University, Box 76, 525 West 120th Street, New York, New York 10027

In our everyday experiences, each of our senses is simultaneously bombarded by a variety of stimuli. In order to function, humans have developed a capability to selectively attend to some aspects of the environment and filter out others. Although we are immersed in constant stimulation, we only contact a select few stimuli. Two individuals in the same setting can have entirely different experiences. Both are presented with the same information, but their attention is turned in different directions. This is the same phenomenon by which we "suddenly notice" something. Although it has been present in our environment, it does not catch our attention until it becomes relevant (Keohane, Luke, & Greer, 2008; Skinner, 1974).

As young children contact environmental experiences, they encounter objects and actions that they do not yet know the names of. At the same time, they are only selectively aware of limited environmental stimuli in the vast array of available stimuli. As language develops, these objects and actions become connected to the arbitrarily applicable words for things that have evolved in a given culture. Learning actions, and words for actions and things, develops as a function of which of the available environmental stimuli attract the child's attention. While phylogeny contributes a great deal to the process (i.e., visual acuity, auditory acuity, and neurophysiology), environmental experiences play a key role, especially at the cultural level and in the development of language (Christiansen & Chater, 2008; Kenneally, 2007; Tomasello, 2008). Different disciplines approach the contributions of experience to this phenomenon from different perspectives. We believe that combining findings from different disciplinary approaches to development can lead to a more complete understanding of learning and development. To that end, when a child is drawn to a movement, the object moving, and the word for that object, the discipline of the behavioral analysis of language or verbal behavior uses the term stimulus control (Catania, 2003; Dinsmoor 1983, 1985, 1995; Skinner, 1957). Stimulus control develops from a history of positive and negative experiences and contributes to how we individually contact our world (Keohane, Luke, & Greer, 2008; Skinner, 1974).

In the behavior analytic literature on language development (Greer & Ross, 2008; Novak & Pelaez, 2004), the acts of noticing are referred to as observing responses. Observing responses incorporate the afferent sensory pathways with which we attend to the stimulus (Wykoff, 1952). Different stimuli will select out our observing responses depending, in part, on prior experiences. Our history of prior experiences contributes to what we observe (Keohane et al., 2008). When an individual encounters a multi-sensory event, some evidence suggests that we are either listening or looking; humans rarely devote equal attention to both experiences (Sinnett, Soto-Faraco, & Spence, 2008). Although we respond to stimuli with multiple senses, the dominance of vision over the other senses has been consistently replicated. In a frequently cited experiment, Colavita (1974) reported that participants consistently attended to a visual rather than an auditory stimulus when both were presented simultaneously, and this finding has been consistently replicated in the four decades since the initial publication (See Spence, 2009 for a summary). The implications of these findings are far reaching, especially for the development of language, which involves auditory stimuli as

children acquire the capability to learn words for things incidentally. The incidental learning of language requires observing auditory and visual stimuli, or other sensory stimuli, simultaneously. Thus, how does the dominance of vision affect learning words for things?

The co-occurrence of multiple stimuli is referred to as multisensory perception, requiring "integration of the information" presented to the different senses and as multiple stimulus control in the analysis of the development of verbal behavior (Greer & Ross, 2008; Novak & Pelaez, 2004). Research suggests that multisensory interaction can either facilitate responses, or hinder responses or learning (Sinnett et al., 2008). Although it seems impossible that the presentation of multiple stimuli can be both beneficial and detrimental at the same time, Sinnett et al. suggested that the nature of the task is involved. The researchers found that when presented with auditory and visual stimuli simultaneously, the accuracy and rate of participant responses was affected by the complexity of the required response. In the more difficult stimulus discrimination task, visual stimuli were dominant over auditory. Task demands determine whether multisensory stimuli compete to hinder or are joined to facilitate responses. In the case of multisensory stimuli, there is clearly a predisposition to attend to the visual aspects of a stimulus, but that alone does not determine how the individual will respond to the stimulus.

Some researchers found a beneficial relationship between gesture and speech to facilitate comprehension. Kelly, Ozyurek, and Maris (2010) found that pairing gestures with speech influenced speech comprehension, such that when gestures and speech convey the same information, comprehension and response rates are improved. Others found that gestures hindered learning of novel words and impeded comprehension (Hirata & Kelly, 2010). In the case of the Kelly et al. study, gestures were part of the verbal or language function of a previously learned communicative repertoire, while in the Hirata and Kelly study learning was involved. Perhaps one difference in whether or not multisensory stimuli hinder or facilitate language effects on a listener concerns whether one is learning a language function or using previously learned language.

There are multiple variables affecting the relationship between gesture and language in learning components of language. Kelly and Lee (2012) compared the acquisition of simple and complex Japanese word pairs taught simultaneously with gestures for English speaking adults. Participants learned "easy" words when they were taught with gestures, while the presence of gesture inhibited the acquisition of the "hard" words. These findings mirror earlier research that found gestures facilitate vocabulary acquisition in a second language only when the phoneme constructions of the words are similar to the learner's native language (Kelly, McDevitt, & Esch, 2009; Sueyoshi & Hardison, 2005). Kelly and Lee suggest that when gesture is paired with more difficult words, it is possible that the added visual information interfered with the comprehension of the newly learned words. The researchers pose an explanation that adding gestures to speech sounds creates a visual distraction that interferes with comprehension.

Distraction also describes an observing response that is under the control of a stimulus, and that stimulus control is at least partially a result of a cumulative history of consequences. From this perspective, distraction refers to an occasion in

which multiple stimuli are present, but the individual's observing responses are selected out by certain stimuli over others. Having redefined distraction, the experimenter can then present multiple stimuli to the participant, and systematically measure which of the stimuli select out his or her observing responses. When contradictory visual and auditory stimuli are presented simultaneously, Choi (2012) found that variations in responding were a function of observing responses determined by instructional history. The researcher simultaneously demonstrated an action (e.g., touching his nose) while giving a vocal direction (e.g., to jump), without specifying which of the two antecedent stimuli, visual or auditory, the participant should respond to. Prior to intervention, the participants overwhelmingly attended to the visual antecedent and imitated the experimenter's actions without regard for the vocal direction. But following intensive auditory discrimination training, the vocal directions selected out participants' observing responses and they responded to the directions without imitating the demonstrated actions. This finding underscores the role of experiences in establishing particular observing responses. Establishing a history of reinforcement experiences for auditory responses increases the likelihood that an individual will respond to an auditory stimulus. But it is interesting to note that the default observing response prior to intervention was visual, again supporting the Colavita effect.

In a study most relevant to the studies presented herein, Hahn (2005) found that when children between 18 and 40 months old were taught either arbitrary object names or object actions, they demonstrated more object actions when compared to object names. With respect to object names, the participants had more correct listener responses, when compared to speaker responses. In follow-up series of three experiments, Hahn and Gershkoff-Stowe (2010) found that when 2 and 3-year old participants were presented with object names and object actions, object names were first learned receptively, (i.e., responding as a listener) then productively (i.e., responding as a speaker). Actions on the other hand, were acquired predominantly as production responses, in which the participants imitated the actions they had observed the experimenter perform with the objects. Overall, the participants produced few object names, but were able to produce nearly all of the actions. The researchers conducted a subsequent experiment with four and 5 year-old participants, in which actions and object names were taught simultaneously, again finding that the actions were learned at a higher rate as compared to the names as production responses. The names were learned as listener responses (i.e., receptive), but not as speaker responses (i.e., productive responses), such that the participants could select the specified object when it was named, but did not produce the name of the object. Replication of this experiment with adults yielded comparable results. These results suggest that the processes involved in learning names and actions for objects do not drastically change with age and development, without direct intervention (Hahn & Gershkoff-Stowe, 2010).

Childers and Tomasello (2002), compared the numbers of exposures needed by 2.5 year old children to learn nouns, verbs, and actions for novel objects. Listener responses requiring the selection of the named stimulus were consistent across nouns, verbs, and actions, but significant differences were found for speaker responses in which the participants were required to produce the

names. Children consistently produced the actions. But the children had few correct responses for the production of the name of the object or name of the action. They examined the number of exposures to acquire the nouns, verbs, and actions, and found that the children learned the actions after fewer exposures, while learning the nouns and verbs required multiple exposures over multiple sessions. Childers and Tomasello also found that when adults and children were taught novel names and arbitrary actions for unfamiliar objects, all of the participants consistently acquired the actions before learning the object names. Clearly the observing responses are selected out by actions more so than object names. This is not to say that actions are acquired rather than names, these findings reflect more on the rate of acquisition of these responses, which has important implications for the incidental learning of language. Incidental learning is the capability that allows an individual to learn from his or her environmental experiences or simple exposure, rather than from direct instruction (Greer & Ross, 2008; Greer & Speckman, 2009; Hart & Risley, 1995; Horne & Lowe, 1996; Rodriguez & Tamis-LeMonda, 2011). When presented with multi-sensory stimuli, we appear to have a phylogenetic predisposition to readily acquire actions and slowly acquire language.

The central theme to all of this research is the role of incidental learning. When individuals encounter multisensory stimuli, the elements that are acquired are learned simply through contact. We are not directly taught the names and functions of most things in our environment (Hart & Risley, 1995; McGuiness, 2004), rather we observe and learn incidentally. Much of the previously described research focused on the human tendency to observe the environment through visual rather than auditory observing responses. But clearly this tendency does not prevent incidental language acquisition: it only affects the rate with which it is acquired.

The mechanisms by which children come to learn the names of things incidentally comprises another, and we think complementary, line of research in language, referred to as verbal behavior development, where the term verbal refers to communicative functions regardless of topography. Similar to the social pragmatic analysis (Tomasello, 2008; Tomasello & Farrar, 1986), this discipline analyzes the effects of experience on the development of language (Greer & Longano, 2010; Greer & Ross, 2008; Greer & Speckman, 2009). However, verbal behavior development supplements the social pragmatic account by experimental analyses of the learning experiences , specifically the history of experience that culminates in developmental capabilities. The analysis of the development of verbal behavior focuses on how children come to learn language through the incidental language learning capability or ILLC. Greer and Ross (2008) describe the ILLC as a learned capability by which an individual simply hears a word or phrase while observing an object in any of the senses and can then produce the word or phrase as a speaker or respond as a listener for the object at a later time without direct instruction. Research in verbal behavior development identified typically developing children, and children with autism or other language delays, who lacked ILLC and provided interventions that established ILLC (Greer, 2008; Greer & Keohane, 2005; Greer & Speckman, 2009). Before the children had ILLC

they could not acquire language incidentally but once they did, they learned language through incidental exposure (Fiorile & Greer, 2007; Gilic, 2005; Greer, 2008; Greer & Keohane, 2005; Greer, Nirgudkar, & Park, 2003; Greer, Stolfi, Chavez-Brown, & Rivera-Valdes, 2005; Greer, Stolfi, & Pistoljevic, 2007; Helou-Care, 2008; Longano, 2008; Pistoljevic, 2008) similar to the exposures described in Childers and Tomasello (2002). These findings supplement the social pragmatic research by suggesting how experiences come to establish language functions.

A great deal of evidence supports the importance of children's capability to contact name-learning opportunities from simple exposure (Childers & Tomasello, 2002; Crystal, 2005; Hart & Risley, 1995, 1999). Some evidence also suggests that this language learning capability is itself learned from experiences (Fiorile & Greer, Gilic & Greer, 2011; Greer & Longano, 2010; Greer & Speckman, 2009; Greer, Stolfi, & Pistoljevic, 2007). Yet, evidence also supports the superiority of visual stimulus control over the auditory stimuli (Colavita, 1974; Hahn, 2005; Hahan & Gereskhoff-Stowe, 2010; Spence, 2009). We address two questions in the following experiments. First, given the simultaneous presentation of actions and names, are visual stimuli dominant over auditory in tests of incidental language learning? Second, does experience make it possible for children to simultaneously learn both actions and names?

Method

Participants

Participants were recruited from a publicly funded private preschool, serving 200 students with and without disabilities from ages 16 months to 5 yrs old. They were recruited from classrooms that included both typically developing students and students with language delays. The participants were 16 preschool students ranging in age from 3.1 to 5.0 years old, with a mean age of 4.2. Thirteen of the participants were diagnosed as preschoolers with speech and language delays, and three were typically developing. These participants were selected based on their verbal behavior developmental cusps and capabilities that are empirically identifiable behaviors critical to development (Greer & Ross, 2008), with each participant having the prerequisite repertoires of generalized imitation, generalized visual identity matching, tacts (i.e., declaratives), and the listener component of naming. The listener component of naming means that they can learn the names of stimuli as a listener but not produce the names productively. The presence or absence of these repertoires was established through administration of the criterion referenced *CABAS International Curriculum and Inventory of Repertoires for Children from Pre-School through Kindergarten* (C-PIRK) (Greer & McCorkle, 2009; Waddington & Reed, 2009) as well as the *Verbal Behavior Developmental Assessment* (Greer & Ross, 2008).

Setting and Materials

All sessions were conducted in a classroom at a time when no other students were present to minimize distractions from competing stimuli. The sessions took place at a child-sized table with the participant seated in a child-sized chair. The experimenter was seated directly across from the participant so that the experimenter's movements were easily viewed throughout the session.

The materials used for both the dependent and independent variables consisted of stimuli sets of three target stimuli, objects that were novel to the participants. They were three-dimensional objects, obscure tools, hardware items, household objects, and kitchen utensils, listed in Table 1. Two identical exemplars of each target stimulus were included in the set. The objects were each assigned a contrived name and grouped into sets of three stimuli. Actions were assigned to the stimuli sets, and were rotated within the sets across participants. Actions were assigned to the stimuli sets, rather than the objects, such that the actions paired with stimuli were interchangeable within each set. In order to eliminate the possibility that the participant could infer the action based on the form of the objects, the actions were arbitrarily assigned and not dictated by object structure. The novel verbal labels and nonverbal actions are listed in Table 1. Twelve of the novel labels and nonverbal actions are the same ones used by Hahn and Gershkoff-Stowe (2010). To create additional stimuli sets, six novel names and actions were created in addition to those developed by Hahn and Gershkoff-Stowe. Stimuli that were known to any participant in either name or function were removed from the sets prior to the experiment.

Table 1

List of Stimuli Sets with Objects, Names, and Actions for All Experiments

Set #	Names	Objects	Action
1	Bek	Cookie cutter	With one hand, swoop object through air in circles
	Tata	Dog toy	With object on table, tap with one hand
	Peeb	Wood tone block	Make object jump vertically
2	Mup	Napkin ring	Hold object in front of mouth and blow on it
	Tam	Strainer	Place object on head
	Pimmel	Silicone poacher	Bat object back and forth between two hands
3	Deet	Drink clip	Touch object to nose
	Mig	Wood spinner	Walk object forward and back on table
	Ibby	Loofah	Rotate object in air using two hands
4	Ziz	Strainer	Touch object to table
	Lupa	Note holder	Hide object behind back
	Dop	Jar opener	Rub on stomach
5	Tay	Juicer	Roll between hands
	Niff	Thimble	Drive on table top in a figure 8
	Gugi	Brillo	Slide on arm from hand to shoulder
6	Dow	Reusable ice cubes	Hold against ear
	Oot	Tube roller	Balance on palm with arm extended
	Booma	Wheels	Move horizontally in air back and forth

Note. Stimuli sets consist of contrived names, actual objects, and actions. Each object is assigned a specific name, while the actions associated with the objects are rotated and counterbalanced across participants.

Design

Each participant received two sessions of the ILLC opportunity experiences, which consisted of a visual match to sample instruction (MTS) while hearing the experimenter say the names for the stimulus with demonstration of actions. This was followed by measures of the dependent variable, consisting of correct responses to no-feedback probe trials for action selection, action demonstration, and listener and speaker responses to the stimuli.

The results were analyzed using a repeated measure ANOVA with two within subject factors: Condition (Action, Name) and Test (Receptive, Productive). The Action Condition was comprised of action demonstration and action selection, and the Name condition included listener and speaker responses to the stimuli. The Receptive Test consisted of correct responses to the selection trials for action selection and listener responses, while the Productive Test was measured as the number of correct response for action demonstration and speaker responses to the stimuli.

Procedure

Incidental language learning experience: Match to sample with action demonstration. During the ILLC experience, each participant received instructional trials for visual identity matching by selecting identical visual versions of each target stimulus while hearing the experimenter name the stimulus and simultaneously demonstrating its function. The instructional trials consisted of the experimenter obtaining the participant's attention, demonstrating the action, giving the direction to match, and providing feedback for the participant's response. Although the response topography consisted of visual identity matching, the critical component of the ILLC experience for the participant was visually attending to the stimulus while hearing the experimenter say its name. The visual match-to-sample instruction simultaneously with hearing the word spoken functioned as a context in which the participant received opportunities to observe both visual and auditory aspects of the stimulus. This constituted a name learning exposure or incidental language learning experience. Inclusion of the match to sample response topography ensured that the participant visually attended to the stimulus by requiring the selection response.

The experimenter placed one exemplar of each stimulus in the set on the table in front of the participant, and obtained the participant's attention. The experimenter demonstrated an action with an identical visual version of one of the target stimuli, and presented the direction, "Find _____." The direction was intentionally non-specific, such that the participant's response did not require a demonstration of action but allowed he or she to pick up the stimulus and manipulate it. Correct responses were recorded if the participant pointed to or picked up the stimulus from the field of three stimuli. The experimenter provided reinforcement in the form of praise and tokens contingent on correct responses. In the case of an incorrect response, the experimenter delivered a correction procedure in which the action demonstration and direction were re-presented and the correct response was prompted but not reinforced. Data were collected for the

numbers of correct and incorrect responses to instructional trials for the MTS instruction.

Criterion for mastery of the MTS instruction was two consecutive sessions with 100% accuracy, which we determined to be adequate exposures for ILL. One session of match instruction consisted of six instructional trials for matching each of the three stimuli, with a total of 18 instructional trials per session. The trials were rotated such that the same target stimulus was not presented for two consecutive probe trials. Sessions for this experiment were presented across consecutive days, with no more than one session of match to sample instruction presented per day.

Dependent variable

Following mastery of MTS in the ILL experience, the experimenter allowed a minimum of one hour and maximum of two hours to elapse and presented probe trials without feedback for measures of the dependent variable. For each measure, two probe trials were presented for each of the stimuli for a total of six probe trials per measure. The trials were rotated such that the same target stimulus was not presented for two consecutive probe trials.

Correct and incorrect responses to action demonstration, action selection, listener, and intraverbal speaker responses were recorded as measures of the dependent variable. The instructions and responses for each of the four measures of the dependent variable are summarized in Table 2.

Table 2

List of Dependent Measures with Experimenter Presentation and Participant Response for All Experiments

Response	**Experimenter Presentation**	**Target Participant Response**
Action Selection	Demonstrate action without stimulus and ask, "Which one does this?"	Select stimulus associated with the demonstrated action
Action Demonstration	Give participant the stimulus and ask, "Show me what this does."	Demonstrate action associated with stimulus
Joining Action to Object Name	Ask, "Show me what a _____ does."	Select named stimulus and demonstrate the action associated with the stimulus
Listener	Ask, "Find ____."	Select named stimulus
Speaker: Tact	Present stimulus without a verbal antecedent	Name stimulus
Speaker: Intraverbal	Present stimulus and ask, "What is this?"	Name stimulus

Interobserver Agreement and Procedural Fidelity

Throughout the experiment, interobserver agreement (IOA) was collected using a second observer simultaneously recording data during the matching responses during the ILL experiences and probe trials. The second observer was previously trained and calibrated in observing both fidelity of the experimenter presentations and accuracy in recording participants' responses. The percentage of IOA was calculated by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100%. IOA was calculated for 38% of the match to sample instruction, with 100% agreement, and for 69% of the measures of the dependent variables, with 99% agreement.

Results and Discussion

A repeated measure ANOVA was used to analyze two within subject factors: Condition (Name, Action) and Test (Receptive, Productive). These results are summarized in Figure 1. The results showed that the main effect of Condition (Name, Action) was significant, $F(1, 15) = 24.61$, $p < .001$. Participants acquired all of the actions ($M = 6.00$, $SD = 0.00$), but fewer names ($M = 4.78$, $SE = .25$). The main effects of Test (Receptive or listener response, Productive or speaker response) $F(1, 15) = 20.35$, $p < .001$ was also significant. The participants had more correct receptive responses ($M = 5.94$, $SE = .04$) in comparison to the productive responses ($M = 4.84$, $SE = .24$). The interaction between Condition and Test was significant $F(1, 15) = 20.35$, $p < .001$. The participants acquired the names as a receptive response ($M = 5.88$, $SD = .34$) more readily in comparison to the names as a productive response ($M = 3.69$, $SD = 1.92$). No difference was found between the receptive and productive responses to the actions ($M = 6.00$, $SD = 0.00$).

Across all of the participants, the actions associated with the objects were readily acquired, as both a selection and production response. Consistent with the findings of Hahn and Gershkoff-Stowe (2010), the actions selected out the observing responses of these participants. In this case, the stimulus control was exerted by the action of the objects rather than the name. The stimuli consisted of the physical object, its actions, and its name. All of these aspects were available, but particular aspects of the stimulus selected out the observing responses of the individual participants.

All of the participants selected and produced actions with 100% accuracy, indicating that actions select out attention. At the same time, the participants consistently acquired the names for the stimuli as a listener with 98% accuracy. Given the name of an object, the participants were able to select the corresponding object from a field. But, this did not extend to the speaker response, and when asked to independently produce the name of an object, participants responded with 61% accuracy. In fact, it is clear that a sharp distinction existed between the listener and speaker responses to the stimuli. The concurrent lack of speaker responses indicates that the speaker and listener repertoires were not joined. The

developmental independence of the listener/receptive and speaker/productive responses is consistent also with a large body of research in the behavioral analysis of development (Greer & Ross, 2008; Greer & Speckman, 2009; Rosales-Ruiz & Baer, 1997).

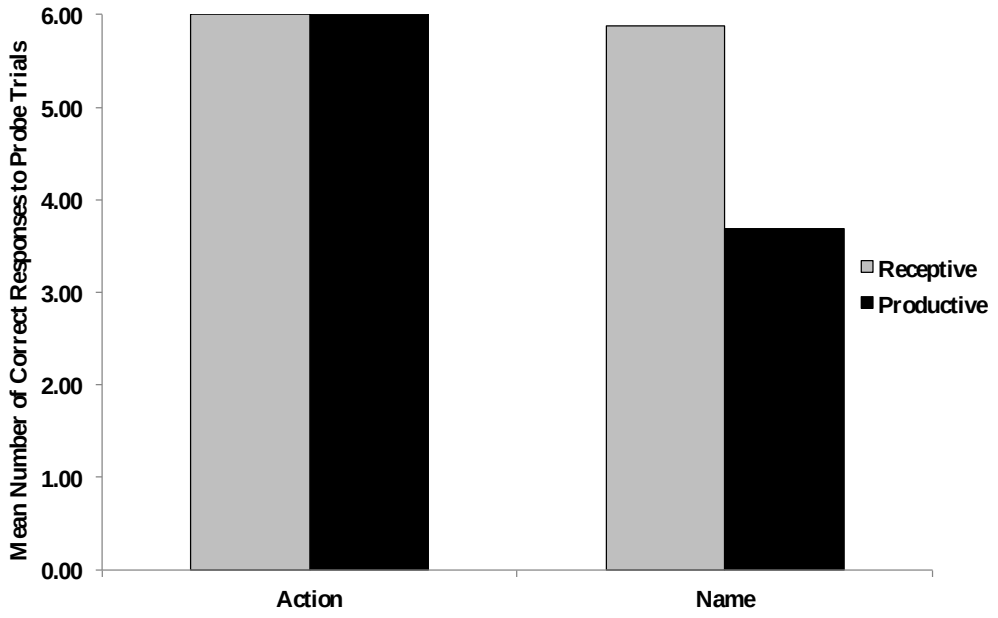


Figure 1. Responses to Condition (Action, Name) and Test (Receptive, Productive) for Experiment I

Also, according to current theory and findings in behavior analysis (Greer & Ross, 2008; Hayes, Barnes-Holmes, & Roche, 2001; Rehfeldt, Barnes-Holmes, & Hayes, 2009) when these initially developmentally independent repertoires join as a result of certain experiences, or direct instruction, one becomes capable of incidental language learning of listener and speaker responses simultaneously. Simply hearing a word, on one or more occasions, as the child attends to the stimuli along with the caregiver, provides the incidental language learning experience(s), resulting in both listener and speaker responses. This is the ILLC/Naming verbal behavior developmental capability that is one of, or the source of, the acceleration of language development in children.

When viewed in reference to the ILLC capability, our findings together with those of Hahn (2005), Hahn and Gershkoff-Stowe (2010), and Childers and Tomasello (2002), raises questions about the relationship between observing responses and the corresponding stimulus control of objects, names, and actions in language acquisition. If these participants were provided with ILLC experiences for the same sets of stimuli, without the presence of actions, would the responses differ significantly when compared to those presented with actions? Will the participants readily acquire the names of objects as a speaker without the presence of actions in the ILLC experience?

One of the primary benefits of single-case design used in behavior analysis is that the results provide an opportunity to view individual differences and variations that are not apparent in a group design. Since the question of interest focuses on the responses of the same individual to differing stimulus conditions, a single subject design with alternating conditions within each participant was used for the Experiment II.

EXPERIMENT II

Method

Participants

The participants in this study were seven preschool students ranging in age from 3.10 to 5.5 years old. Three of the participants were diagnosed as preschoolers with speech and language delays, and four were typically developing. The participants were selected from the same setting as Experiment I, and participants were selected based on the same criteria. A description of the participants is presented in Table 3.

Table 3
Participant Characteristics for Experiments II and III

Participant	Gender/ Age	Verbal Capabilities	Diagnosis
1a	Female/ 5.0	Listener ILLC Conversational exchanges	Typically Developing
1b	Female/ 3.9	Listener ILLC Conversational exchanges	Typically Developing
2a	Female/ 5.0	Listener and Speaker ILLC Conversational exchanges	Speech and Language Delay
2b	Male/5.5	Listener and Speaker ILLC Conversational exchanges	Speech and Language Delay
3a	Female/ 5.0	Listener and Speaker ILLC Conversational exchanges	Typically Developing
3b	Female/ 4.3	Listener and Speaker ILLC Conversational exchanges	Typically Developing
4b	Female/ 3.10	Listener and Speaker ILLC Conversational exchanges	Speech and Language Delay
5	Female/ 4.0	Listener and Speaker ILLC Conversational exchanges	Typically Developing
6	Male/ 3.1	Listener and Speaker ILLC	Speech and Language Delay
7	Male/ 4.1	Listener ILLC	Speech and Language Delay
8	Male/ 4.5	Listener and Speaker ILLC	Typically Developing

Note. The above listed verbal capabilities are in addition to the prerequisite capabilities of generalized imitation, listener, and speaker repertoires required for participant selection criteria.

Design

For each participant, experimental action conditions and no-action control conditions were alternated for a total of six phases. Participant responses under the two conditions were compared using single case experimental design with alternating treatments counterbalanced across matched pairs. Each participant completed six phases, with the phases alternated in a counterbalanced fashion across participants (e.g., ABABAB or BABABA).

Participants were paired based on capabilities and levels of verbal behavior, and the conditions were counterbalanced such that one participant in the pair received the no action condition for a set and the paired participant received the action condition for the same set. The sequencing of the stimuli sets was counterbalanced across pairs. It should be noted that Participant 4a was unable to complete the experiment, and is not included in the results.

Procedure

Action condition: ILLC experience with match to sample and demonstration of function. The ILLC experience for the action condition was identical to Experiment I.

Action condition: Dependent variables. Following mastery of match to sample instruction in the ILLC experience, probe trials were conducted for the dependent measures of demonstration of actions; listener responses; and intraverbal speaker responses ("What is this?").

Procedures were identical to those in Experiment I. The action selection response was omitted, due to the redundancy of the responses for action selection and action demonstration in Experiment I. Additional dependent measures were conducted for actions emitted during the ILLC experience, joining an action to the object name, and tact speaker responses. The tact speaker response differs from the intraverbal speaker response in that there is no verbal direction or question from the experimenter. For the intraverbal speaker response, the experimenter asks, "What's this?" but for the tact speaker response the experimenter simply visually displays the item in order to elicit "spontaneous" speech.

The sequence in which the dependent variables were measured was: 1) actions imitated during the ILLC experience; 2) action demonstration; 3) listener; 4) tact speaker; 5) intraverbal speaker; and 6) joining an action to the object name. The antecedents and responses for each of the six measures are summarized in Table 2. The additional measure of action demonstration imitation during the ILLC experience is described as follows.

Action demonstration imitation during the ILLC experience. During match-to-sample instruction while hearing the word for the object alone or object/action, the experimenter recorded whether the participant imitated the actions demonstrated with the objects. The required response during match instruction was the selection of the identical visual version of the stimulus presented by the experimenter. Action demonstration was not a required response and therefore was not corrected or reinforced; however, experimenters recorded

whether the participant imitated the action demonstration at any point during the instruction. The number of actions demonstrated was recorded as the number of occurrences out of the total number of action opportunities, which in this case was the total number of ILLC experience matching instructional trials presented.

No action condition: ILLC experience with match to sample. During the ILLC experience, each participant received instructional trials for match to sample responses while hearing the experimenter say the name of the stimulus without the action demonstration. Otherwise, for the no action condition, the responses were recorded and provided with feedback identical to those in the action condition described in Experiment I.

No action condition: Dependent variables*.* Following mastery of match to sample instruction in the ILLC experience, measures of the dependent variables 1) listener, 2) tact, and 3) intraverbal speaker responses were conducted using the same procedures as the action condition. Since there were no actions associated with the stimuli in this condition, the measures for action demonstration, joining an action to an object name, and occurrences of actions during the ILLC experience were not included. The antecedents and responses for each of the three measures are summarized in Table 2.

Interobserver Agreement and Procedural Fidelity

The methods for collecting and calculating interobserver agreement (IOA) for the ILLC experience and measures of the dependent variables were identical to those used in Experiment I. IOA was calculated for 51% of the match to sample instruction, with 100% agreement, and for 60% of the measures of the dependent variables, with 100% agreement.

Results and Discussion

For the ILLC experience match instruction, all of the participants in the experimental condition met the criterion within two sessions. It was unlikely that the participants would have made errors, since the required response of matching was a prerequisite repertoire for all participants. All of the dependent variables responses summed across participants and conditions are presented in Figure 2. It is clear that in the action condition, the participants accurately produced the actions during the probe trials, with 96% correct responses. In comparing the listener and speaker responses in both conditions, there were more correct listener responses than speaker responses, regardless of the condition. When analyzing responses across the two conditions, there were more correct responses for the listener and speaker responses (98% and 79%, respectively) in the no action condition compared to the action condition (90% and 62%, respectively). The findings are discussed in greater detail as follows.

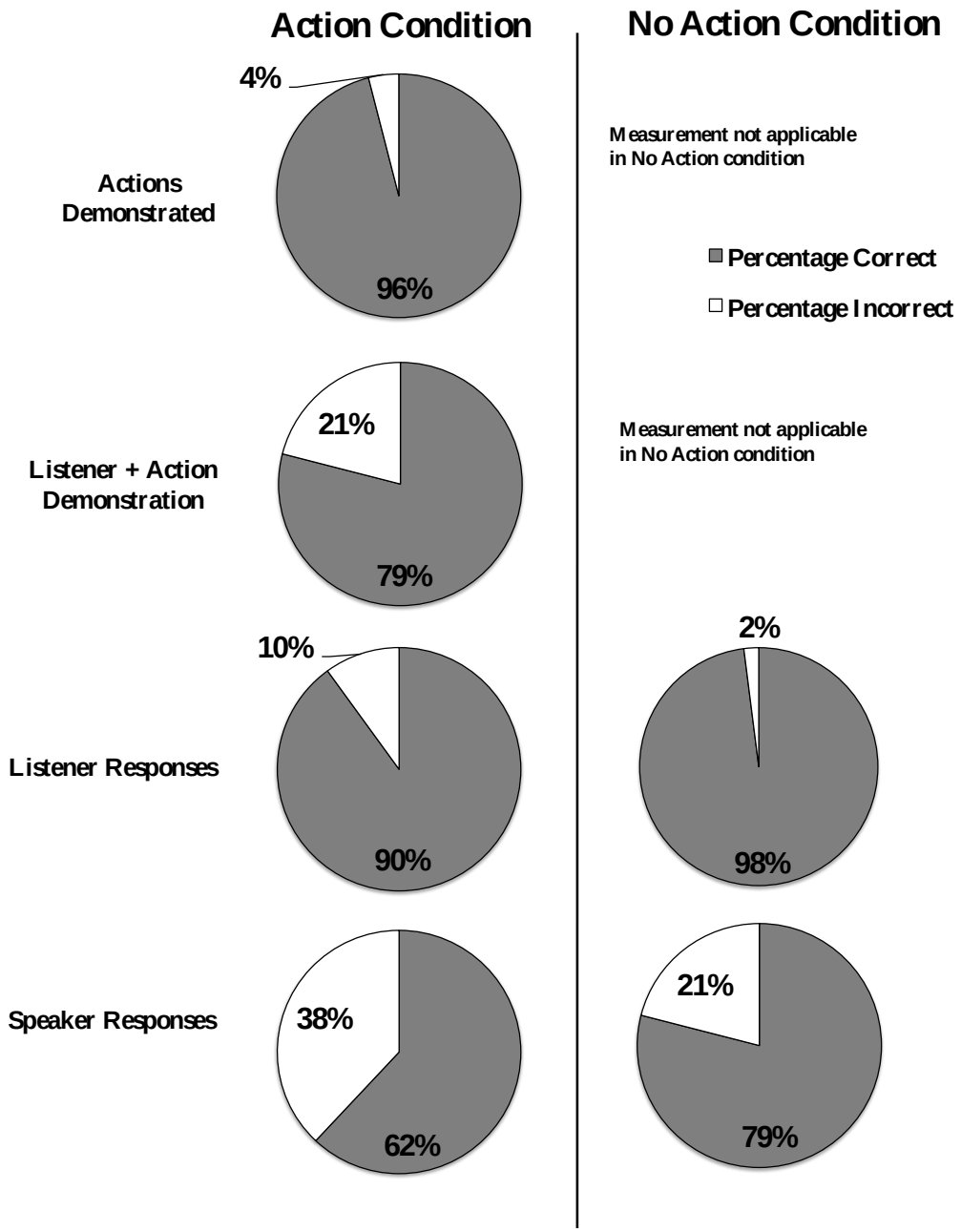


Figure 2. All correct responses to probe trials, summarized across participants by conditionfor Experiment II.

In the probe trials for demonstration of actions**,** Participants 1a, 1b, 2a, 2b, 3a, and 4b responded with 100% accuracy for all object demonstration trials. Participant 3b responded with 72% accuracy. Overall, the participants responded with the correct action demonstration with 96% accuracy across all of the probe trials.

For the probe trials for joining object names to actions, Participants 3a and 4b responded with 100% accuracy across the three stimuli sets. Participants 1a,

1b, 2a, 2b, and 3b had similar response patterns, such that the initial probe trials for the first sets of stimuli had a lower number of correct responses followed increases in both or one of the second and third sets. The increases in correct responses indicate that the participants learned from the initial set what responses would be required for future stimuli sets. It is likely that the initial set resulted in a shift of stimulus control and subsequent observing responses, such that the participant attended to different aspects of the stimulus during the next instructional sessions based on prior experience. In this case, the probe trials may have evoked an observing response, resulting in the participants "noticing objects one may be asked about" (Skinner, 1957, p. 415).

The probe trials for ILLC were conducted across both experimental and control stimuli sets, and included the listener and speaker responses to the stimuli and are summarized by action and no action conditions in Figures 3 and 4. In general, the participants acquired the listener responses consistently across both the action and no action conditions. In this experiment, the listener responses were acquired with relative ease across both conditions. The listener responses are displayed in a pie chart in Figure 3. An effect can be observed for Participants 1a, 1b, 2b, and 3b in which there was a greater number of correct listener responses for the no action condition. Both Participants 2a and 3a showed no difference in listener responses across the two conditions while Participant 4b had fewer correct listener responses in the no action condition.

Across all of the participants, the number of correct speaker responses for the stimuli were consistently the same as or less than the number of correct listener responses for both conditions. These data show that regardless of condition, the listener response was acquired at the same rate or more readily than the speaker responses. These results are consistent with findings from the ILLC research discussed previously, in which the listener responses are acquired prior to the speaker responses. In comparing the responses across the conditions, six of the seven participants had a greater number of correct responses to the speaker probe trials for the control, or no action, condition. These results are displayed in a pie chart in Figure 4. The results for these participants are consistent with those of Childers and Tomasello (2002), Hahn (2005), Hahn and Gershkoff-Stowe (2010), who also found that when actions, objects, and names were presented simultaneously, the participants effortlessly produced the actions, and that the listener responses were acquired more often than the speaker responses.

During the match instruction, all of the participants imitated the actions with the stimuli as demonstrated by the experimenter. These responses were not required and were not provided feedback. Although variability was observed, all of the participants imitated the actions with the objects, although no directions were given to do so and the participants were not reinforced for emitting the response. In terms of stimulus control, it appears that actions demonstrated with objects select the attention of participants. These actions warranted an immediate visual observing response, while the auditory observing response for the names of the objects did not. This is not to say that the participants cannot learn the names of the objects, on the contrary, the participants were able to select named objects as a listener. But a dramatic difference was observed when participants were

required to produce those names as a speaker. Based on phylogenic or ontogenetic factors, these participants selectively acquired the see-do response of action demonstration.

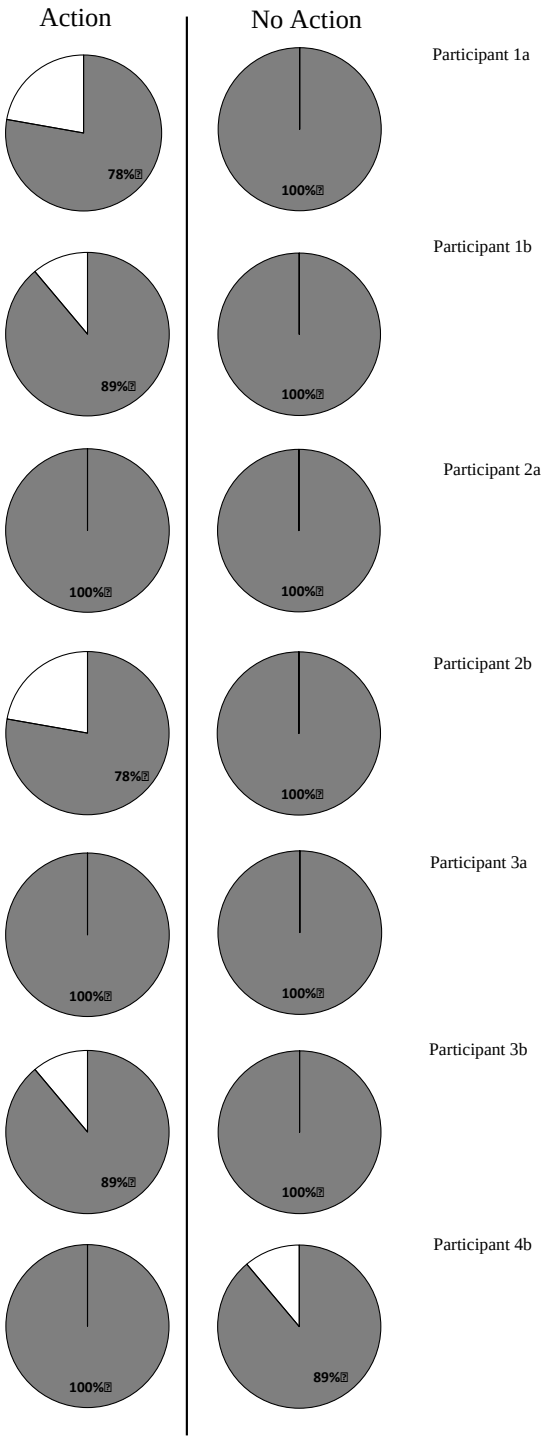


Figure 3. Listener responses to probe trials summarized for all participants, with the responses summarized across conditions for Experiment II.

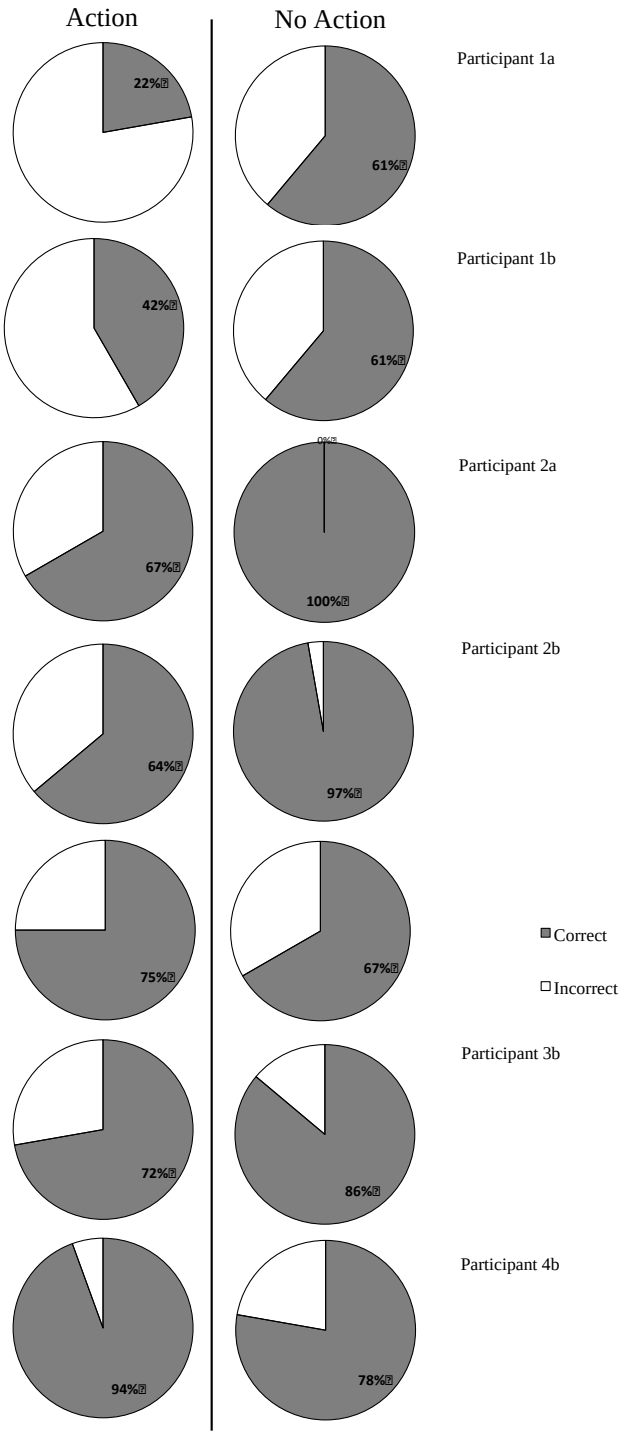


Figure 4. Speaker responses for all participants, summarized across conditions for Experiment II.

One of the primary benefits of single-case design is that the results provide an opportunity to view individual differences and variations that are not apparent in a group design. Based on the results of Experiment II, it is clear that the participants' observing responses were selected out by particular stimuli. Although there was an overall tendency to attend to the actions of the object, there were participant variations in stimulus control that can be attributed to, at least in part, the collective experiential history of reinforcement for that individual. In order to better address variations in participant observing responses, in the third experiment, participants were selected based on their responses to multiple stimuli for one object. Specifically, participants were selected who imitated actions and responded as a listener to the stimuli, but emitted fewer speaker responses. By selecting participants whose observing responses were selected out by actions rather than names, the third experiment sought to create a test of whether a common history of reinforcement could establish multiple stimulus control for observing both actions and names. The purpose of Experiment III was to determine if a history of reinforcement experiences could extend the scope of observing responses to include both actions and names simultaneously, such that the participant consistently acquired multiple responses following contact with the multiple stimuli.

EXPERIMENT III

Method

Participants

The participants in this study were four preschool students ranging in age from 3.1 to 4.5 years old. Two of the participants were diagnosed with language delays, and two were typically developing. The participants were selected from the same setting as Experiments I and II, and participants were selected based on the same criteria. A description of Participants 5, 6, 7, and 8 is presented in Table 3.

Setting and Materials

The setting was identical to those in Experiments I and II. The materials from Experiments I and II were used for both the dependent and independent variables.

Design

The experimental design was a non-concurrent multiple probe design across participants to isolate the role of experience on the establishment of the capability to action and language under incidental learning conditions. The dependent measures were participants' responses to no-feedback probe trials for: (a) selection of objects associated with actions, (b) action demonstration, (c) ILLC responses, and (d) joining of an action to the object name for novel stimuli. The independent variable was Multiple Exemplar Instruction (MEI) across actions, name learning, and the joining of name learning with different sets of stimuli. Different stimuli sets were used for each phase, such that four to six sets were used for each participant. The sequencing of the stimuli sets was counterbalanced across participants.

The sequence of the experiment began with the ILLC experience (visual match-to-sample instruction with the opportunity to hear the name of the stimulus and action), followed by probe trials for the dependent variables conducted for one set of stimuli at the

outset of the experiment. This was repeated with a second set of stimuli immediately prior to the implementation of the independent variable of MEI for each participant respectively. MEI was conducted with a new set of stimuli, until criterion was met for all responses. After mastery of MEI, the ILLC experience with match to sample instruction was repeated with a new set of stimuli, followed by probe trials for the dependent variables. The alternation between MEI and measures of the dependent variables were rotated until criterion of 100% accuracy was achieved for all of the dependent variables. The sequencing of the experiment is summarized in Figure 5.

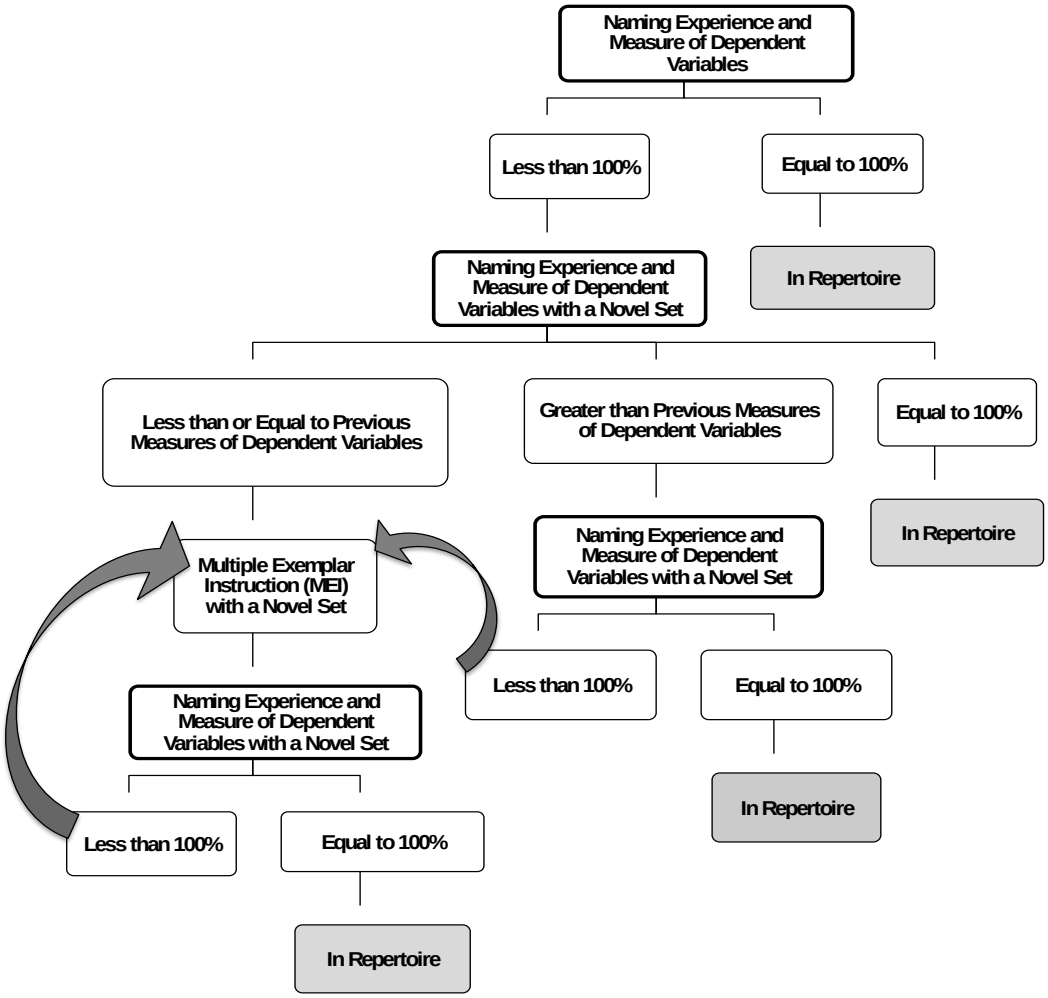


Figure 5. The experimental sequence for measures of the dependent variables and Multiple Exemplar Instruction (MEI) for Experiment III.

Procedure

ILLC experience: Match to sample with demonstration of function. The procedures for the match to sample instruction were identical to those in Experiment I and the action condition of Experiment II.

Dependent variables. Following mastery of match to sample instruction in the ILLC experience, procedures for the measures of the dependent variables were identical to those used in Experiment I. Unlike Experiment II, imitation was not recorded during the match to sample instruction because the data did not show a clear relation to the condition or the other responses. The dependent measures were 1) probe trials for action selection, 2) demonstration of actions, 3) listener responses, 4) tact speaker responses, 5) intraverbal speaker responses, and 6) joining an action to the object name. The probe trials were conducted using the same procedures as in Experiment I, and the action condition of Experiment II. The antecedents and responses for each of the six measures are summarized in Table 2.

Pre-experimental screening. The probe trials described in the preceding dependent variables section also served a dual purpose as a pre-experimental screening for participants. Experiment III required that all participants had similar responses to the stimuli, when the stimuli were comprised of objects, actions, and names. Participants were selected who imitated actions, responded as a listener to the stimuli, but emitted few speaker responses. The responses indicated that the participants' observing responses were selected out by actions more so than names. Participants whose responses to the probe trials differed from the selection criteria were not included in the experiment.

Initially, probe trials for the dependent measures were conducted for each of the participants at the outset of the experiment. Prior to implementing the MEI intervention, these measures were repeated using a new set of stimuli. Repetition of the dependent measures prior to the intervention was used to control for maturation or other variables that may have affected participant responding. If there was an increase in the number of correct responses to the dependent measures in the second set, match instruction and probe trials for the dependent measures were conducted for additional sets of stimuli until stable responding or a descending trend was observed prior to implementing the intervention. Provided that the dependent measures were consistent across the first and second stimuli sets, or there were fewer correct responses for the second set, the independent variable of MEI was implemented. After completion of MEI, the post-experimental ILLC experience with match to sample instruction was presented followed by probe trials for the dependent variables.

Independent variable

Multiple exemplar instruction with demonstration of function. After obtaining the pre-experimental measures of the dependent variables, the experimenter implemented the independent variable of multiple exemplar instruction (MEI). Using a new set of stimuli, the experimenter presented instructional trials for four different responses to each stimulus: 1) imitating actions, 2) listener, 3) tact speaker, and 4) intraverbal speaker responses. All responses were immediately followed by experimenter delivered reinforcement for correct responses and corrections for incorrect responses. The antecedents and responses for each of the MEI instructional trials are summarized in Table 4.

Table 4
List of Experimenter Antecedents and Participant Responses for Multiple Exemplar Instruction for Experiment III

Response	Experimenter Presentation	Target Participant Response
Action Imitation	Demonstrates action with a stimulus, and asks, "Do this."	Imitates demonstrated action with identical visual version of the stimulus
Listener	Asks, "Find ____."	Selects named object from field of 3 stimuli
Speaker: Intraverbal with Action	Presents stimulus while demonstrating the action and asks, "What's this?"	Names stimulus
Speaker: Tact	Presents stimulus without a verbal direction	Names stimulus

The field of three stimuli remained on the table in front of the participant throughout all of the responses. The participant received reinforcement in the form of praise, social attention, or tokens for emitting correct responses to instructional trials. For incorrect responses, the experimenter modeled the correct response for the participant to imitate or echo, but did not reinforce the correction. Correct and incorrect responses were recorded for all of the response topographies for each stimulus.

The action imitation, listener, intraverbal speaker, and tact speaker instructional trials were rotated across all three of the stimuli, such that consecutive instructional trials did not consist of responses to the same stimulus. The instructional trials were rotated across the stimuli and response forms until all of the responses were mastered concurrently. A session consisted of 24 instructional trials, comprised of six instructional trials per response form for the action imitation, listener, intraverbal speaker, and tact speaker responses. Criterion was set at 100% accuracy for one session.

Post MEI measures of the dependent variables. Following mastery of the multiple exemplar instruction, the ILLC experience with match to sample instruction was followed by no-feedback probe trials for the dependent variables *with a new set of stimuli*. These were identical to those presented prior to the MEI intervention. Criterion for mastery of the dependent measures was set at 100% accuracy across the six response topographies. If the participant met criterion with the novel set following MEI, then it was determined that multiple stimulus control was acquired along with the necessary observing responses to learn multiple responses from a single experience. On the other hand, if criterion was not met for the post MEI measures of the dependent variables, the participant repeated MEI with a new set of stimuli, until criterion was achieved. Again, the post MEI

measures of the dependent variables were repeated. If criterion was achieved, the participant was considered to have acquired multiple stimulus control as described above. Otherwise, this sequence in which MEI was rotated with measures of the dependent variables was repeated until criterion was met.

Interobserver Agreement and Procedural Fidelity. Interobserver agreement (IOA) was collected and calculated using the same methods as the previous experiments. IOA was calculated for 42% of the match to sample sessions, with 100% agreement; for 74% of the probe trials, with 99% agreement; and for 40% of the MEI instruction, with 99% agreement.

Results and Discussion

For the pre-experimental match to sample instruction, across all of the participants in the experimental condition, criterion was met within two sessions. In probe trials for both selection of objects associated with actions and for action demonstration, all of the participants responded with 100% accuracy. In the probe trials for joining an action to the object name, participants responded with 100% accuracy across both stimuli sets, with two exceptions. Participant 6 responded with 100% and 83% accuracy. Participant 8 responded with 100% and 67% accuracy.

In the probe trials for ILLC, the probe trials for listener and speaker responses were repeated with two stimuli sets for each participant prior to the MEI intervention, and the results are summarized in Figure 6. Participants 5, 7, and 8 responded with 100% accuracy to all probe trials for the listener responses for both sets of stimuli prior to the MEI intervention. Participant 6 responded with 100% accuracy for the first set and 83% accuracy for the second set. For the speaker response responses, Participant 5 responded with 67% and 33% accuracy. Participant 6 responded with 33% accuracy to both stimuli sets. Participant 7 responded with 67% and 50% accuracy. Participant 8 responded with 67% and 33% accuracy. For the intraverbal speaker responses, Participant 5 responded with 83% and 33% accuracy. Participant 6 responded with 33% accuracy to both stimuli sets. Participant 7 responded with 83% and 33% accuracy. Participant 8 responded with 67% and 33% accuracy. Each participant received multiple sessions of MEI until the criterion was met with 100% accuracy across all responses. The number of sessions required to meet criterion varied across participants, although all of the participants only required MEI for one set of stimuli. Participants 5 and 6 required six sessions, Participant 7 required five, and Participant 8 required three.

Following the MEI intervention, match instruction was conducted with novel sets of stimuli and across all of the participants in the experimental condition, criterion was met within two sessions. After mastery of the match to sample instruction for a novel set of stimuli, probe trials were presented for the six dependent measures. All of the participants responded with 100% accuracy across all of the probe trials for the six dependent measures following MEI instruction, and the results are summarized in Figure 6. Prior to the MEI intervention, the responses were not only below criterion level for mastery, but also

indicated a descending trend in correct responses across stimuli sets. Since each participant met criterion with the novel set following MEI, it was determined that the participant had acquired multiple stimulus control and the necessary observing responses to learn multiple responses from a single experience.

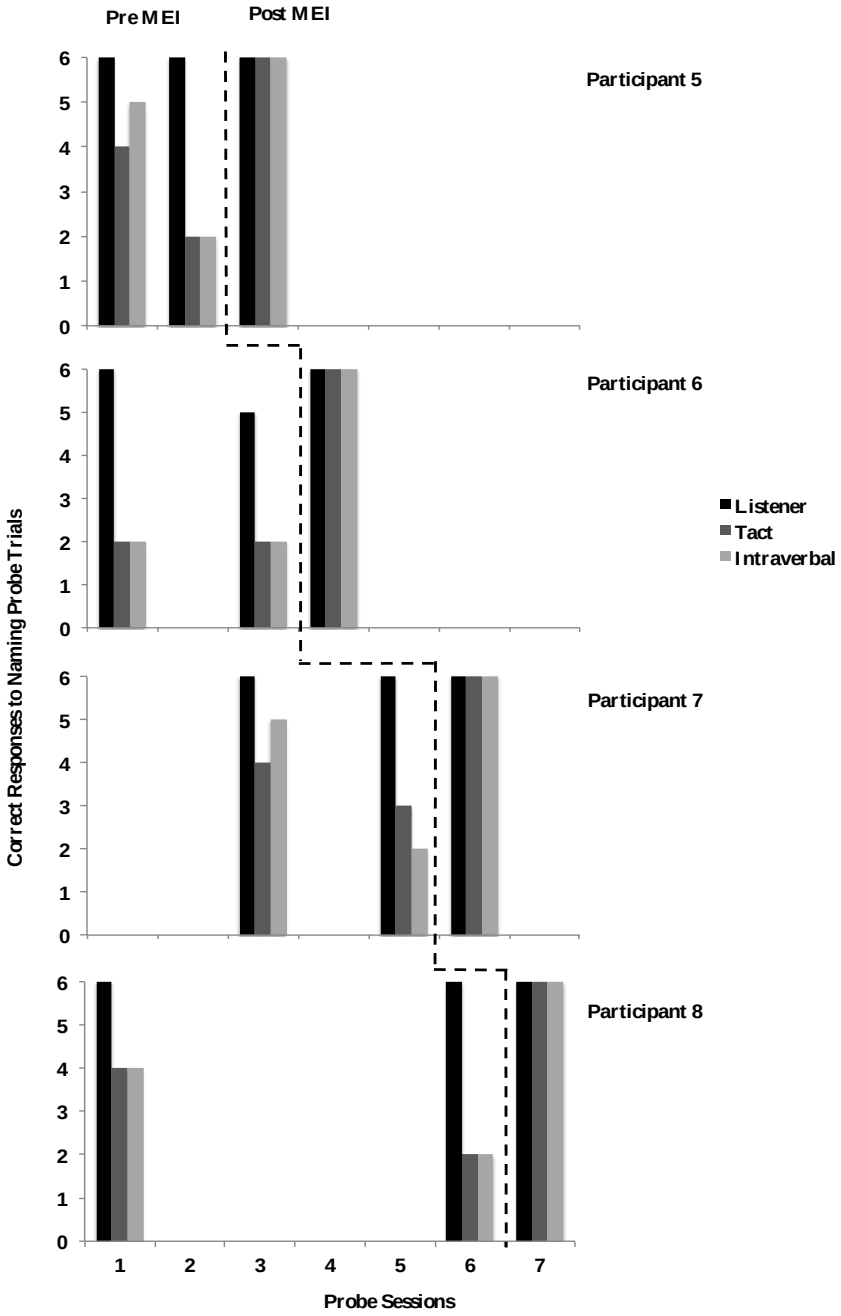


Figure 6. Listener and speaker responses prior to and following the MEI intervention for Experiment III.

For this experiment, participant selection required that each participant readily acquire actions and listener responses to the stimuli, but acquire fewer speaker responses. Prior to and following the MEI intervention, the participants selected objects associated with an action, demonstrated an action, joined an action to an object name, and acquired the listener responses. These responses are consistent with the results from the previous two experiments, which indicated that actions and listener responses are acquired with relative ease. Acquisition of the names as a listener did not extend to the accuracy of the speaker responses, which consistently had fewer correct responses prior to the MEI intervention. Based on the responses prior to MEI, it is clear that actions and names as a listener selected out the observing responses of the participants.

Each of the four participants received varied numbers of MEI sessions, dependent on the individual rate of acquisition. But following mastery of MEI, all of the participants responded to all of the probe trials for the dependent measures with 100% accuracy. Instructionally, MEI provided rotated opportunities for multiple responses to the same stimuli in the presence of reinforcement. Following this cumulative history of reinforcement for multiple responses to stimuli, the participants acquired multiple responses to probe trials for the novel set of stimuli. Most notably, the participants acquired speaker responses to the stimuli as a result of exposure to the ILLC experience. The increased speaker responses following MEI indicated that observing responses and stimulus control shifted as a result of the intervention.

As an instructional intervention, MEI pairs reinforcement with the rotated opportunities for multiple response topographies for a stimulus. When the procedure of MEI and the capability of ILLC are reduced to the underlying principles of behavior, it becomes apparent that reinforcement underlies both the intervention and the capability. It is a history of reinforcement that shapes observing responses and stimulus control, and MEI creates a history of reinforcement for multiple responses. In this case, MEI creates a history of reinforcement for actions, listener, and speaker responses, which results in a shift of stimulus control such that both names and actions select out the observing responses of the individual. In reference to the Colavita effect (Colavita, 1974; Spence, 2009), visual stimuli such as actions, select out observing responses over auditory stimuli, such as names. Multiple exemplar experiences or direct instruction establishes a history of reinforcement that overrides this general tendency, allowing the individual to simultaneously acquire names and actions of objects.

General Discussion

Taken as a whole, the results of the three experiments give a clearer picture of the relationship between actions and object names in language acquisition. It has been suggested that the presence of an action can hinder, or in some cases facilitate the acquisition of names. These experiments dissected the relations among object, name, and action, to reveal a complex interaction of conditioned reinforcement and observing responses unique to the individual. In relation to the

"Colavita effect" (Spence, 2009), these experiments establish that actions, as a visual stimulus, consistently select out observing responses, while observing responses for the listener, less difficult, and the speaker, more difficult, components of names varies widely. Reduced to the basic principles, stimulus control for objects or actions is established through a cumulative history of reinforcement, determining which stimuli select out observing responses. Observing responses then determine which aspects of multi-sensory stimuli are available to the individual.

The focus of this series of experiments was the acquisition of multiple responses to a single stimulus through incidental contact. This ties closely to ILLC/Naming, which allows one to observe a stimulus, hear its name, and subsequently acquire the name-object relation as both a speaker and a listener. ILLC/Naming is thought to account for the rapid expansion of vocabulary in young children, and is critical to language development (Greer & Longano, 2010). In most typically developing children, the capability emerges effortlessly, but for some children an intensive intervention is required to induce the capability. Various interventions have been successful for inducing ILLC, but underlying all of these interventions is the pairing of reinforcement with the visual and auditory observing responses necessary to acquire language. Knowing how ILLC is induced experimentally also sheds some light on its development in children without intervention. An experientially learned reinforcer (i.e., conditioned reinforcer) must be present such that during the ILLC experience, the observing responses of the individual selects out visual and auditory stimuli, which in turn results in the acquisition of names for objects (Longano & Greer, in press).

When evaluating the source of reinforcement in ILLC, Longano and Greer (in press) tested the role of conditioned reinforcement for observing visual and auditory stimuli. For participants without ILLC, the researchers systematically paired reinforcement with observing responses for non-preferred visual and auditory stimuli on a computer screen. The stimuli were then combined, such that an animated visual stimulus was presented while the recorded auditory stimulus (object name) was spoken for four stimuli in a set. No prosthetic reinforcement (i.e., reinforcement not a natural outcome of the response) was provided while the participants observed the simultaneous stimulus presentation, and after multiple observations of the paired stimuli, the participants acquired the names of the stimuli. Additional probe trials with novel sets of stimuli confirmed that the participants acquired the capability of ILLC as a result of this procedure. The researchers suggest that the ILLC requires the joining of visual and auditory stimuli as conditioned reinforcers. In this case, establishing a history of conditioned reinforcement for observing multiple aspects of a stimulus was sufficient to induce the capability for incidental language acquisition. These findings closely parallel the findings from the present series of experiments. Establishing a history of reinforcement for observing visual and auditory stimuli resulted in acquisition of multiple responses from a single experience. These interventions allowed children to learn from incidental environmental exposures, which provides exponentially more learning opportunities.

The findings from these three basic science experiments have translational value by contributing to a better understanding of interactions between development and teaching. It becomes apparent that presenting multiple pieces of information does not necessarily benefit the learner. In fact, the different aspects are more likely to compete for attention than to facilitate multiple responses. The findings from the third experiment have the greatest development by teaching implications. By establishing which aspect of the stimulus that the student is attending to, the teachers or psychologists can then identify which aspect has acquired stimulus control, and more importantly, which one has not. The MEI procedure from Experiment III was successfully used to extend stimulus control such that the participants attended to multiple stimuli simultaneously and subsequently acquired multiple responses. Rather than replacing one observing response with another, the MEI intervention multiple simultaneous observing responses. Both visual and auditory stimuli selected out observing responses after the intervention.

Effective instruction requires attention to language development. Although a teacher may demonstrate a math problem or science experiment while describing the steps, the students may only attend to the visual presentation or auditory aspects. Greer, Corwin, and Buttigieg (2011) found that students without the capability for ILLC did not benefit from the common teaching practices. Successful learning in the typical classroom setting requires that students observe and learn from teacher demonstrations. These students lacked the capability for Naming, necessary for incidental language learning. By implementing an MEI intervention, the researcher found that pairing reinforcement with multiple responses to a stimulus induced Naming. This capability not only allowed for incidental language acquisition, but the ability to learn from teacher demonstration. Essentially, this developmental intervention provided students with the observing responses that are critical to learning in the classroom setting (Greer et al., 2011). This capability might ultimately be the deciding factor for success or failure in school.

There is an implicit assumption in most classrooms that when the teacher presents a lesson with demonstration, modeling, and description, that the students should learn through observation. The accumulation of recent research suggests otherwise; there are critical prerequisite repertoires required for learning from teacher presentations. Optimally, teachers should approach learners as individuals and evaluate what methods are successful. If students are attending selectively to portions of the instruction and are not learning, then interventions can be implemented for those students who cannot learn from traditional methods. Educational research has afforded us with tools that can help not only to prevent student failure, but also accelerate learning. Initial assessment and intervention are crucial to student success.

It is our history of reinforcement that determines which stimuli are salient and will select out our observing responses. Each individual has his or her own accumulation of experiences that shape observing responses. But the present research demonstrates that stimulus control and observing responses are not static, rather they can be shifted through an experimentally manipulated history of reinforcement. Despite predispositions, consistent pairing of reinforcement with

observing responses allows a child to contact new stimuli, and acquire new responses. This implies that educational interventions should focus not on teaching repertoires, but instead on changing conditioned reinforcers for students which will in turn allow them to learn in new ways that were not possible before.

References

Catania, A. C. (2003). Why behavior should matter to linguists. *Behavioral and Brain Sciences 26*, 670-672.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*, 967–978.

Choi, J. (2012) *Effects of mastery of auditory match-to-sample instruction on echoics, emergence of advanced listener literacy, and speaker as own listener cusps by elementary school students with ASD and ADHD.* Unpublished Ph.D Dissertation, Columbia University.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31,* 489-558.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics, 16,* 409-412.

Crystal, D. (2005). *How language works*. New York, NY: Overlook Press.

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, 6, 693–728.

Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior, 43,* 365-381.

Dinsmoor, J. A. (1995). Stimulus control: Part I. *The Behavior Analyst, 18,* 51-68.

Fiorile, C. A., & Greer, R. D. (2007). The induction of naming in children with no echoic- to-tact responses as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior, 23, 71-88*.

Gilic, L. (2005). *Development of naming in two-year-old children.* Unpublished Ph.D Dissertation, Columbia University.

Greer, R. D. (2008). The ontogenetic selection of verbal capabilities: Contributions of Skinner's verbal behavior theory to a more comprehensive understanding of language. *International Journal of Psychology and Psychological Therapy, 8*, 363-386.

Greer, R. D., Corwin, A., & Buttigieg, S. (2011). The effects of the verbal developmental capability of naming on how children can be taught. *Acta de Investigación Psicológica, 1, 23-54.*

Greer, R. D., & Keohane, D. D. (2005). The evolution of verbal behavior in children. *Behavioral Development Bulletin, 1,* 31-47.

Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn to do it. *The Analysis of Verbal Behavior, 26*, 73-106.

Greer, R. D., & McCorkle, N. (2009). *CABAS® curriculum and inventory of repertoires for children from pre-school through kindergarten,* 4th edition. Yonkers, NY: CABAS®/Fred S. Keller School.

Greer, R. D., Nirgudkar, A., & Park, H. (2003). The effect of multiple exemplar instruction on the transformation of mand and tact functions. Paper presented at the annual international conference of the Association for Behavior Analysis, San Francisco, CA.

Greer, R. D., & Ross, D. E., (2008). *Verbal behavior analysis: Inducing and expanding complex communication in children severe language delays*. Boston: Allyn & Bacon.

Greer, R. D., & Speckman, J. M. (2009). The integration of speaker and listener responses: A theory of verbal development. *Psychological Record, 59,* 449-488.

Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdez, C. (2005). The emergence of the listener to speaker component of Naming in children as a function of multiple exemplar instruction*. The Analysis of Verbal Behavior, 21,* 123-134.

Greer, R. D., Stolfi, L., & Pistoljevic, N. (2007). Emergence of Naming in preschoolers: A comparison of multiple and single exemplar instruction. *European Journal of Behavior Analysis*. *8,* 119-131.

Hahn, E. R. (2005). *Faster mapping: How children and adults learn labels and actions for objects.* Unpublished PhD Dissertation, Carnegie Mellon University.

Hahn, E. R., & Gershkoff-Stowe, L. (2010). Children and adults learn actions for objects more readily than labels. *Language Learning and Development, 6,* 283-308.

Hart B. M., & Risley, T. R. (1995). *Meaningful differences in the everyday life of America's children*. Baltimore, MD: Paul Brookes.

Hart, B., & Risley, T. R. (1999). *The social world of children learning to talk*. Baltimore: Paul H. Brookes Publishing Co.

Hayes, S., Barnes-Homes, D., & Roche, B. (2001). *Relational frame theory: A post Skinnerian account of human language and cognition.* New York: Kluwer/Academic Plenum.

Helou-Care, Y. (2008). *The effects of the acquisition of naming on reading comprehension with academically delayed middle school students with behavioral disorders.* Unpublished PhD Dissertation, Columbia University.

Hirata, Y., & Kelly, S. D. (2010). The effects of lips and hands on auditory learning of second language speech sounds. *Journal of Speech, Language and Hearing Research, 53,* 298 -310.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 185-241.

Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phoneme demands are high. *Language and Cognitive Processes, 27*, 693-807.

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Neural correlates of learning Japanese words with and without iconic gestures. *Language and Cognitive Processes, 24*, 313-324.

Kelly, S. D., Ozyurek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science, 21*, 260 -267.

Kenneally, C. (2007). *The first words: The search for the origins of language.* New York: Viking.

Keohane, D. D., Luke, N., & Greer, R. D. (2008). The things we care to see: The effects of rotated protocol immersion on the emergence of early observing responses. *Journal of Early and Intensive Behavioral Interventions, 5,* 23-39.

Longano, J. M. (2008). *The effects of echoic behavior and a second order classical conditioning procedure as the reinforcement history of emergent naming.* Unpublished Ph.D Dissertation, Columbia University.

Longano, J. M., & Greer, R. D. (in press). The effects of echoic behavior and a second order classical conditioning procedure as the reinforcement history of emergent Naming. *The Analysis of Verbal Behavior.*

Novak, G., & Pelaez, M. (2004). *Child and adolescent development: A behavioral systems approach.* Thousand Oaks, CA: Sage.

McGuiness, D. (2004). *Early reading instruction: What science really tells us about how to teach reading*. Cambridge, MA: MIT Press.

Pistoljevic, N. (2008). *The effects of multiple exemplar and intensive tact instruction on the acquisition of naming in preschoolers diagnosed with autism and other language delays.* Unpublished Ph.D Dissertation, Columbia University.

Rehfeldt, R. A., Barnes-Holmes, Y., & Hayes, S. (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities.* Oakland, CA: New Harbinger Publications.

Rodrigues, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Development, 82,* 1058-1075.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis, 30(3),* 533-544.

Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica, 128,* 153-161.

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing Group and the B. F. Skinner Foundation.

Skinner, B.F. (1974). *About behaviorism.* New York: Knopf

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in Brain Research 176,* 245–258.

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning, 55*, 661 -699.

Tomasello, M. (2008). Origins of human communication. MIT Press: Cambridge, MA

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development, 57,* 1454–1463.

Waddington, E. M., & Reed, P. (2009). The impact of using the "Preschool Inventory of Repertoires for Kindergarten" (PIRK) on school outcomes of children with Autism Spectrum Disorders. *Research in Autism Spectrum Disorders, 3*, 809-827.

Wyckoff, L. B. (1952). The role of observing responses in discrimination learning. Part 1. *Psychological Review, 66,* 68-78.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1746- 1756

Using Computer-Based Programmed Instruction to Train Goal-Directed Systems Design

Nicholas L. Weatherly*[1], Aubrey Daniels International* & Richard W. Malott**
*Atlanta, GA, ** Western Michigan University

Abstract

The purpose of this study was to evaluate two versions of a programmed instruction training program designed to teach undergraduate college students a goal-directed systems approach to analyzing organizational systems (Malott & Garcia, 1987). The first version was a paper-based programmed instruction module that had previously been shown to be effective at training the basic knowledge of the concepts, however was ineffective at training the application of these concepts. A computer-based programmed instruction (CBPI) version was created to improve the application of these concepts, which was tested through a series of three open-ended posttests with increasingly explicit prompts for each successive test. The results of the study showed higher performance results for the CBPI versions across all three dependent variables. The results of a nonparametric global test showed a standardized effect size of .86 and a p-value of < .001.

Keywords: Computer-based Instruction, E-Learning, Instructional design, Performance management, Programmed instruction, Student training.

Usando Instrucción Programada por Computadora para el Entrenamiento de Diseños de Sistema Orientados a Metas

Resumen

El propósito de este estudio fue evaluar dos versiones de un programa de instrucción programada diseñado para enseñar a estudiantes no graduados una aproximación al diseño de sistemas dirigidos a metas para analizar sistemas organizacionales (Malott & Garcia, 1987). La primera versión consistió en un modulo de un programa de instrucción personalizada de papel que anteriormente había probado ser efectivo para enseñar conocimientos básicos de conceptos, pero que no obstante era ineficiente para entrenar en la aplicación de dichos conceptos. Se creó una versión computarizada de instrucción programada (CBPI) para mejorar la aplicación de los conceptos, la cual se probó a través de una serie de tres post-tests con respuestas libres que incluyeron pistas cada vez más explícitas en cada prueba sucesiva. Los resultados del estudio mostraron una mejor ejecución en las tres variables dependiente cuando se usaron las versiones del CBPI. Los resultados de una prueba no paramétrica global mostraron un tamaño del efecto estandarizado de .86 y un valor p < .001.

Palabras clave: Instrucción Computarizada, E-aprendizaje, Diseño Instruccional, Administración de la Ejecución, Instrucción Programada, Entrenamiento a Estudiantes

Original recibido / Original received: 14/06/2014 Aceptado / Accepted: 21/09/2014

[1] Nicholas L. Weatherly, Aubrey Daniels International; Richard W. Malott, Department of Psychology, Western Michigan University. The authors thank Dr. Brad Huitema of Western Michigan University for his statistical advice and for administering the nonparametric global test (O'Brien, 1984). Address correspondence to Nicholas L. Weatherly, Aubrey Daniels International, 3344 Peachtree Road, Suite 1050, Atlanta, GA 30326 (e-mail: nweatherly@aubreydaniels.com)

Since 2000, there have been numerous attempts at improving educational systems through legislative programs such as the No Child Left Behind Act of 2001, school finance reform, and increased accountability for teacher performance (Superfine, 2014). Unfortunately, the quality and efficacy of these programs is often interpreted from lagging indicators of success such as graduation rates, or an overgeneralization of standardized test scores. Omitted are the key behaviors and leading indicators of success that are critical when tracking progress and systematically improving performance deficits. Efforts to improve educational practices and outcomes are often based on assumptions that money, curriculum materials, facilities, and regulation cause learning without an understanding of the variables responsible for change (Cohen, Raudenbush, & Ball, 2003). However, resources alone have proven ineffective at maximizing learning and performance without an understanding of the contingencies surrounding learning and an application of sound behavior-analytic techniques.

In 1968, Skinner's seminal text The Technology of Teaching described the need for a behavioral analysis of educational practices and the importance of applying the principles of behavior to improve educational systems. Over the years, decades of research and practice have built an empirically proven behavior-analytic technology that can be used to impact all levels of these systems, from the students through the teachers and administrators. One of the greatest contributions behavior analysis has offered in this area has been programmed instruction, a behavior-analytic technology that is the cornerstone needed when blending the principles of behavior into the steady advancements of computer-based educational technologies.

Programmed instruction (PI) is a teaching method based on behavior-analytic research and principles of behavior such as feedback, prompting, shaping, and discrimination training (Jaehnig & Miller, 2007). Programming is referred to by Skinner (1963) as "the construction of carefully arranged sequences of contingencies leading to the terminal performances which are the object of education" (p. 183). Computer-based instruction (CBI) provides a continuously evolving technology that offers an efficient way of achieving the behavior-analytic standards of programmed instruction. Computer-based or web-based training can offer a wide array of stimulus-presentation modes and programming options that can improve instruction by making careful and specific programming possible (Clark, 1983, 1985, 1994; Kozma, 1994; Tudor & Bostow, 1991). Keys to success include interactivity, a central component of CBI, that involves the user of an instructional program making a response in order to advance through the program (Chou, 2003; Kritch & Bostow, 1998), and feedback provided contingent upon an overt response made during the training (Bates, Holton, & Seyler, 1996; Fredrick & Hummel, 2004; Jaehnig & Miller, 2007). With the exponential advancements seen in the area of computer technology and the widespread application of electronic resources into educational systems, computers are being used more than ever as a viable educational tool that can be effective at all levels of training.

Computer-based instruction has been proven effective across a number of training areas directly relevant to educational systems including training classroom teachers to interact with parents (Ingvarsson & Hanley, 2006) and training college

students on computer software use (Karlsson & Chase, 1996). This technology has also been effectively used to train the concepts and principles of behavior analysis (Miller & Malott, 1997; Munson & Crosbie, 1998; Tudor, 1995). Given the importance of integrating behavior-analytic tools into educational methods, it is of value to staff to understand the principles of behavior so they can properly apply these principles to their teaching techniques.

When compared with traditional lecture-based instruction, programmed instruction has been reported to be more effective (Chatterjee & Basu, 1987; Daniel & Murdoch, 1968; Fernald & Jordan, 1991; Kulik, Cohen, & Ebeling, 1980) and produce more rapid acquisition (Fernald & Jordan, 1991; Hughes & McNamara, 1961; Jamison, Suppes, & Wells, 1974; Kulik, Kulik, & Cohen, 1980) than traditional instruction. However, other comparisons of programmed instruction and traditional forms of instruction have produced mixed results (Bhushan & Sharma, 1975; Kulik, Schwalb, & Kulik, 1982), possibly as a result of discrepancies in the design of the instruction (Jaehnig & Miller, 2007). Several meta-analyses on the effectiveness of programmed instruction have indicated that programmed instruction continues to improve due to persistent advancements in programming technology (Hartley, 1978; Kulik, Cohen, & Ebeling, 1980; Kulik, Schwalb, & Kulik, 1982), but there is a continued need for additional data in this area. The purpose of this study was to assess the efficacy of programmed instruction modules by experimentally evaluating two types of programmed instruction modules, paper-based programmed instruction and computer-based programmed instruction (CBPI).

Method

Participants and setting

This study involved three groups of undergraduate college students enrolled in a behavior analysis course over a three-semester period. The first group contained 19 students, the second group contained 32 students, and the final group contained 45 students. The primary researcher provided all posttests and any instructional materials that were not provided as part of the department's course materials the students purchased at the beginning of the course. The student participants completed the instructional materials on their own time, in accordance with the course syllabus for the semester.

Goal-Directed Systems Design

The current programmed instruction module used in the course was a paper-based module that was previously created, evaluated, and revised to teach the concept of *goal-directed systems design* (Malott & Garcia, 1987). Goal-directed systems design is a way of analyzing the structure of an organization in terms of the inputs, processes, and desired outputs that can be viewed at all levels within the system. Training college students to effectively analyze and improve the respective system they will end up working in was a curriculum goal identified by the university. Historically, a quiz was given to students after they finished the paper-based programmed instruction, using questions almost identical to those

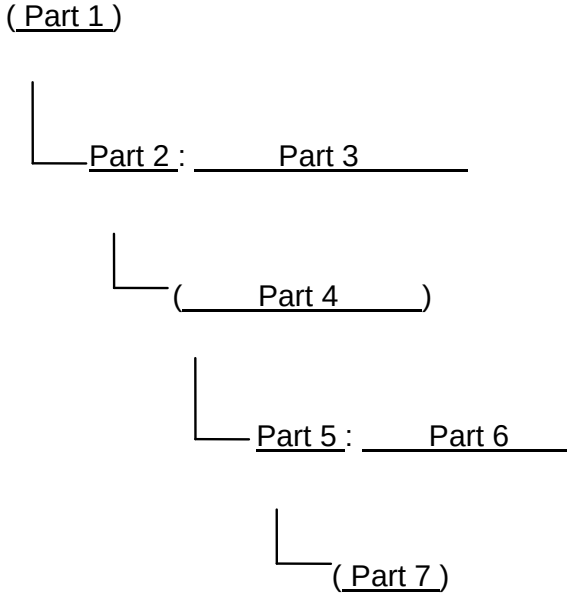
provided in the instructional module. The quiz consisted of one fill-in-the-blank question and 11 multiple-choice questions. Although these quiz data were reportedly high (approximately 92% accuracy across a six-year period), instructors also reported that the students consistently had difficulty applying the concepts once the training module was completed. For example, class discussions and the final course project centered around giving the students a sample organizational scenario with the expectations that the students could complete an Input-Process-Output model as instructed in the goal-directed systems design paper-based programmed instruction. There are limited data on generalization and response induction in the areas of programmed instruction and computer-based instruction, including the ability to vocalize and apply concepts taught textually (Ingvarsson & Hanley, 2006; Tudor & Bostow, 1991). For these reasons, evaluating the efficacy of paper-based vs. computer-based programmed instruction on the generalized application of the principles learned in the programmed instruction modules was prioritized for this study.

Data Collection – Dependent Variables

For the purpose of this study, three posttests were given to assess the application of the concepts taught in the training modules. These tests provided the opportunity to apply goal-directed-systems design concepts to the same problem across all three tests, with increasingly explicit prompts for each successive test. Given the previous success of the paper-based programmed instruction in training the basic knowledge of the goal-direct systems design concepts, the primary dependent variables for the current study were the percentage of questions answered correctly on these three posttests designed to test the application of the concepts. Each student received these three posttests in succession after completing their respective programmed-instruction module.

*Posttest scoring criteria***.** In order to properly test the validity of the first, second, and third posttests, grading criteria were created to score those specific tests. There were four main areas selected by the researchers and instructors to be evaluated in order to properly assess the applied knowledge the students acquired by using the programmed instruction. These areas of evaluation were as follows: 1. Correct bracketing of the Input-Process-Output diagram; 2. Correct sequence of information provided in the diagram; 3. Proper exclusion of unnecessary information as part of the diagram (e.g., improperly placing “customers” as a main resource in the diagramming); 4. Correct information provided within the diagram.

Posttest #1 offered an opportunity for responding to all of the target areas by providing a written example of a process within an organization (a recycling plant), along with a blank space for the student to provide their own Input-Process-Output diagram. For each test, one point was awarded for each component of the diagram that was correctly provided by the participant, allowing for a total of seven points for properly providing all components (see Figure 1). One point was awarded for correct bracketing of the entire diagram and one point for having the correct sequence of information throughout the diagram. One point was also awarded for proper exclusion of unnecessary information from the diagram.



Total Points Possible:
*7 Points: Correct content in each of the 7 parts of the diagram
1 Point: Correct bracketing
1 Point: Correct sequence of information
1 Point: Proper exclusion of unnecessary information
**Note: Given the written prompts in Posttest #3, only 4 points are possible for correct content.*

Figure 1. Posttest Scoring Criteria

Independent Variables

*Paper-based programmed instruction***.** The first training condition consisted of a phase in which the undergraduate students were only provided with the paper-based programmed instruction. Prior to the current study, this instruction was the primary training tool used to train the goal-directed systems design unit. The 54-page paper-based programmed instruction included written descriptions of concepts and applied examples, 20 practice activities located throughout the instruction that contained questions pertaining to these concepts, and one set of review questions at the end of the instruction. Each question provided the user with an opportunity to write a response (multiple-choice answer, provide a diagram, fill-in-the-blank), which could then be verified by turning to the answer page immediately following each activity. In addition, 12 of the 20 activities provided detailed feedback along with the correct answer for the questions. During this training condition, students were assigned the paper-based programmed instruction and given the three posttests during the next seminar. The primary researcher administered all tests.

*Computer-based programmed instruction***.** The improvement conditions consisted of providing students with the computer-based programmed instruction (CBPI) that was created as a possible replacement for the paper-based instruction previously used. The initial CBPI version contained the same examples and covered all content areas as identified in the paper-based version. Just as with the

paper-based programmed instruction condition, students were assigned the CBPI during the goal-directed systems design unit and given the three posttests during the next seminar. The primary researcher administered all tests.

The training components that are essential to effective CBPI, such as interactivity and contingent feedback, were included in the components of the program. The program was created in Microsoft® PowerPoint® using text, graphics, and animations, and included an interactive response-requirement from the user, which involved opportunities for overt responses with feedback. Feedback was presented contingent on the response, which has been shown to be superior as compared to programs in which the feedback information can be viewed either non-contingently or after an arbitrary response (Anderson, Kulhavy, & Andre, 1971, 1972). These interactive components involved multiple-choice questions, although the CBPI also required constructed responses directly produced by the students that were turned in to the course instructor during the seminar following the completion of the program. The participants received feedback on their constructed responses during seminar. The CBPI also contained navigational aids that allowed participants to access definitions, tables and diagrams, and any sections that the participants wished to review throughout the instruction.

The first version of CBPI contained 348 total frames (slides) that included training and activity components taken directly from the paper-based version. It was designed with a table of contents that allowed for navigation to various practice activities, training sections, and review aids such as definitions, charts, and diagrams. Each user was required to make an overt response on their computer in order to progress through the instruction. During each practice activity, the user was not offered the opportunity to advance to a subsequent question/section until the correct response was made.

After the first CBPI implementation, error analyses were conducted to assess questions in need of improvement, just as they had been when developing the paper-based version of the programmed instruction prior to the current study. Based on the data from the error analyses and feedback received from the users, a number of improvements were made prior to the next CBPI implementation. Detailed feedback was added to all of the practice questions that had limited feedback (an approximately 36% increase in the amount of feedback) and a review section was added at the end of the instruction. A revised applied activity was also added at the end of the program, along with a number of cosmetic improvements including the addition of more color and animation. The final CBPI version contained a total of 438 frames, which included 138 instructional frames, 260 practice activity frames, and 34 review frames along with the table of contents.

Procedures and Design

The three programmed instruction modules were administered separately across three semesters. During the first semester, the group of students (N = 19) received the computer-based programmed instruction. The group in the second semester (N = 32) received the previously developed paper-based programmed instruction and the group in the third semester (N = 45) received the final version of

the CBPI. A posttest only design was used for the three posttests due to open-ended nature of these tests, which was designed to assess proper application of the goal-directed systems design concepts. The posttests were administered during the first class that followed the completion of the respective training program. Once an instructional module was assigned to a semester, all students enrolled in the course completed only that assigned module. The three posttests were used across all three semesters.

Results

When compared with the paper-based programmed instruction, CBPI produced greater performance results across all three posttests, with performance increasing as the explicit prompts were added for each successive posttest (see Figure 2). The first version of the computer-based programmed instruction (N = 19) resulted in a posttest #1 mean score of 51%, a posttest #2 mean score of 63%, and a posttest #3 mean score of 87%. During the subsequent semester, students (N = 32) were given the paper-based programmed instruction resulting in a posttest #1 mean score of 38%, a posttest #2 mean score of 55%, and a posttest #3 mean score of 78%. The final version of computer-based instruction was administered to students (N = 45) the next semester, resulting in a posttest #1 mean score of 51%, a posttest #2 mean score of 68%, and a posttest #3 mean score of 88%.

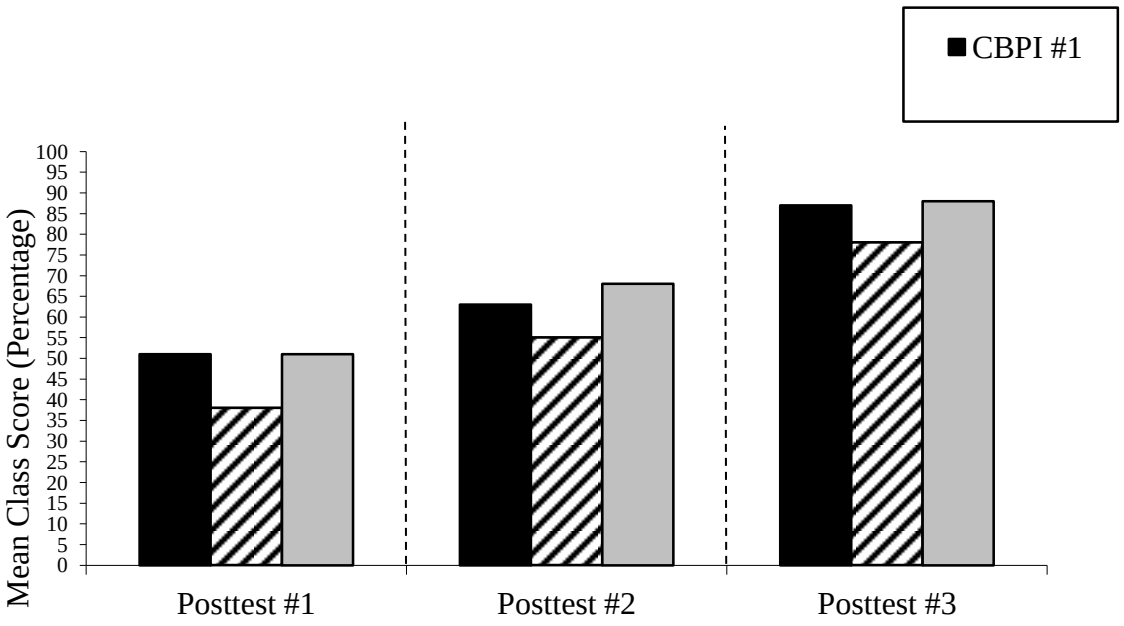


Figure 2. Mean differences in posttest scores across the three different programmed-instruction conditions. Both CBPI versions produced greater performance results across all three posttests when compared with the paper-based programmed instruction, with performance increasing as the explicit prompts were added for each successive posttest

Our primary objectives for this study were to continue improving the training of the goal-direct systems design concepts and to place a larger emphasis on the application of these concepts as tested through posttests #1, #2, and #3. Given this goal, a nonparametric global test (O'Brien, 1984) was used to compare the three dependent variables (i.e., posttest #1, #2, and #3) between all semesters that used the computer-based versions of the programmed instruction and the semester that used the paper-based version. The results of the global test showed a larger, more statistically significant effect in the mean scores of the students in the computer-based programmed instruction groups as compared to the paper-based group. The standardized effect size was .86 and the *p*-value was $< .001$. Interobserver agreement (IOA) was assessed using the point-by-point agreement formula {# of agreements / (# of agreements + disagreements) x 100%} for all dependent variables. IOA was evaluated for 53% of all posttest evaluations across the three separate implementations, resulting in 92% agreement. It should be noted that the quizzes used to assess basic knowledge of the material as part of the course requirements maintained at approximately 92% accuracy across the duration of the study, with the final version of the CBPI resulting in 93% performance accuracy.

Discussion

The purpose of the current study was to evaluate the effectiveness of a computer-based programmed instruction (CBPI) version of a previously developed paper-based programmed instruction that trained undergraduate students in the concept of goal-directed systems design. A 54-page paper-based programmed instruction module was used to create a CBPI version that involved 438 frames and provided approximately 90 minutes of training. The results of the study showed that converting a well researched and designed paper-based programmed instruction module into CBPI produced a more effective training program that can be easily accessed in an increasingly popular computer-based medium. The study also showed that CBPI can be designed to appropriately impact the generalized application of the concepts taught in the instructional program, an important element of any training program.

Comments and survey data from students showed preference for computer-based instructional methods, particularly commenting on the feedback, entertainment, and educational value provided with CBPI, and the stepwise delivery of the content in both the paper-based and computer-based versions of programmed instruction. Based on user feedback, future versions of the instruction should assess less lengthy versions or effective ways to break-up the training components to allow for breaks in the training. Future research could also assess the amount of transfer to applied settings produced with CBPI. Given the limited data in this area and the high social validity for this type of behavior-analytic technology this could be a valuable contribution to the field. A constraint found within the PowerPoint® program was the limited data-collection capabilities for this programming tool. Future CBPI programs should consider a platform that can accommodate tools that automatically record user responses, such as SCORM (Sharable Content Object Reference Model) or Tin Can API. Using an authoring

tool that incorporates this feature, along with the other important features of programmed instruction, could provide an efficient training platform for computer-based instruction.

With the continued advancements in computer technologies and authoring tools, computer-based instruction can greatly benefit settings with limited resources and large receiving systems. Computer-based instruction provides an ideal platform for behavior-analytic standards, allowing for proper manipulation of instructional antecedents, high rates of responding, contingent consequences, and an array of data collection and performance tracking options. The empirically-supported benefits demonstrated by the improved student performance resulting from the CBPI used in the current study shows CBPI to be a very efficient and effective training option over traditional training methods. Given the trends in web-based learning and the ever-present need for effective training solutions, CBPI can be a valuable tool that can impact any system and all areas of training and should continue to be examined and improved upon for future applications.

References

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology, 62,* 148-156.

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology, 63,* 186-188.

Bates, R. A., Holton, E. F., III, & Seyler, D. L. (1996). Principles of CBI design and the adult learner: The need for further research. *Performance Improvement Quarterly, 9,* 3-24.

Bhushan, A., & Sharma, R. D. (1975). Effect of three instructional strategies on the performance of B.Ed. student-teachers of different intelligence levels. *Indian Educational Review, 10,* 24-29.

Chatterjee, S., & Basu, M. K. (1987). Effectiveness of a paradigm of programmed instruction. *Indian Psychology Review, 32,* 10-14.

Chou, C. (2003). Interactivity and interactive functions in web-based learning systems: A technical framework for designers. *British Journal of Educational Technology, 34,* 265-279.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53,* 445-459.

Clark, R. E. (1985). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. *Educational Communication and Technology Journal, 33,* 249-262.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development, 42,* 21-29.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instructions, and research. *Educational Evaluation and Policy Analysis, 25,* 119-142.

Daniel, W. J., & Murdoch, P. (1968). Effectiveness of learning from a programmed text covering the same material. *Journal of Educational Psychology, 59,* 425-431.

Fernald, P. S., & Jordan, E. A. (1991). Programmed instruction versus standard text in introductory psychology. *Teaching of Psychology, 18,* 205-211.

Fredrick, L. D., & Hummel, J. H. (2004). Reviewing the outcomes and principles of effective instruction. In D. J. Moran, & R. W. Malott (Eds.), *Evidence-based educational methods* (p. 9-22). San Diego: Elsevier Academic Press.

Hartley, S. S. (1978). Meta-analysis of the effects of individually-paced instruction in mathematics. *Dissertation Abstracts International, 38,* 4003.

Hughes, J. L., & McNamara, W. (1961). A comparative study of programmed and conventional instruction in industry. *Journal of Applied Psychology, 45,* 225-231.

Ingvarsson, E. T., & Hanley, G. P. (2006). An evaluation of computer-based programmed instruction for promoting teachers' greetings of parents by name. *Journal of Applied Behavior Analysis, 39,* 203-214.

Jamison, D., Suppes, P., & Wells, S. (1974). The effectiveness of alternative instructional media: A survey. *Review of Educational Research, 44,* 1-67.

Jaehnig, W., & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record, 57,* 219-232.

Karlsson, T., & Chase, P. N. (1996). A comparison of three prompting methods for training software use. *Journal of Organizational Behavior Management, 16,* 27-44.

Kritch, K. M., & Bostow, D. E. (1998). Degree of constructed-response interaction in computer-based programmed instruction. *Journal of Applied Behavior Analysis, 31,* 387-398.

Kulik, J. A., Cohen, P. A., & Ebeling, B. J. (1980). Effectiveness of programmed instruction in higher education: A meta-analysis of findings. *Educational Evaluation and Policy Analysis, 2,* 51-64.

Kulik, C. L. C., Kulik, J. A., & Cohen, P. A. (1980). Instructional technology and college teaching. *Teaching of Psychology, 7,* 199-205.

Kulik, C. C., Schwalb, B. J., & Kulik, J. A., (1982). Programmed instruction in secondary education: A meta-analysis of findings. *Journal of Educational Research, 75,* 133-138.

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development, 42,* 7-19.

Malott, R. W., & Garcia, M. E. (1987). A goal-directed model for the design of human performance systems. *Journal of Organizational Behavior Management, 9,* 125-129.

Miller, M. L., & Malott, R. W. (1997). The importance of overt responding in programmed instruction even with added incentives for learning. *Journal of Behavioral Education, 7,* 497-503.

Munson, K. J., & Crosbie, J. (1998). Effects of response cost in computerized programmed instruction. *The Psychological Record, 48,* 233-250.

O'Brien, P. C. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics, 40,* 1079-1087.

Skinner, B. F. (1963). Reflections on a decade of teaching machines. *Teachers College Record, 65,* 168-177.

Skinner, B. F. (1968). *The technology of teaching.* New York: Appleton-Century-Crofts.

Superfine, B. M. (2014). The promises and pitfalls of teacher evaluation and accountability reform. *Richmond Journal of Law and the Public Interest*, *17*, 591-622.

Tudor, R. M. (1995). Isolating the effects of active responding in computer-based instruction. *Journal of Applied Behavior Analysis, 28,* 343-344.

Tudor, R. M., & Bostow, D. E. (1991). Computer-programmed instruction: The relation of required interaction to practical applications. *Journal of Applied Behavior Analysis, 24,* 361-368.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1757 - 1771

Effects of Positive and Negative Reinforcement in a Concurrent Operants Arrangement on Compliance and Problem Behavior[1]

Breanne Byiers, Adele Dimian, Jennifer J. McComas, & Frank J. Symons
The University of Minnesota

Abstract

Functional analysis (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994) is a robust approach to identifying function-based interventions for problem behavior, including self-injury, aggression, and destruction. Such interventions, however, may be difficult for untrained caregivers to implement with fidelity in natural environments. Further research is needed to identify simple antecedent strategies for promoting appropriate behavior among children with significant problem behavior. The purpose of the current study was to utilize a concurrent schedules arrangement to identify conditions under which two children with autism spectrum disorder (ASD) and developmental delays who engaged in problem behaviors would choose to complete academic tasks to earn access to preferred items. In both cases, problem behaviors were shown to be sensitive to reinforcement in the forms of escape from task demands and access to preferred items. A concurrent operant arrangement in which the participants could choose to complete work tasks to earn access to preferred activities, or to take a break without demands or preferred items, was implemented. The schedule requirements in the demand component were systematically increased across opportunities, while the amount and type of reinforcement was kept constant. The results show, at the lowest levels of task demands, both participants allocated more opportunities to the work option. At higher levels, however, both participants allocated a majority of their choices to the break option. Despite the absence of preferred items in the break component, no instances of problem behavior were observed following selection of the break option. This indicates that this type of analysis could be used to identify conditions for compliance among individuals who engage in escape- or multiply-maintained problem behaviors, without the need to provoke or reinforce problem behavior. Limitations of the current study and recommendations for future research are discussed.
Keywords: Problem Behavior, Compliance, Positive and Negative Reinforcement.

Efectos del Reforzamiento Positivo y Negativo en un Arreglo de Operantes Concurrentes sobre la Obediencia y la Conducta Problemática

Resumen

El análisis funcional (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994) es una aproximación robusta para identificar intervenciones basadas en la función para la conducta problemática, incluyendo la autolesión, agresión y destrucción. Tales intervenciones, no obstante, pueden ser difíciles de implementar fidedignamente en escenarios naturales por cuidadores no entrenados. Se requiere más investigación para identificar estrategias antecedentes simples para promover la conducta apropiada entre niños con problemas significativos de conducta. El propósito del presente estudio fue utilizar un arreglo de programas concurrentes para identificar las condiciones bajo las cuales dos niños con trastorno del espectro autista (ASD, por sus siglas en inglés) y retraso en el desarrollo que emitían conducta problemática escogerían completar tareas académicas para ganar acceso a ítems preferidos. En ambos casos, las conductas problemáticas fueron sensibles al reforzamiento consistente en escape de las demandas de la tarea y en acceso a ítems preferidos. Se implementó un arreglo de programas concurrentes en el que los participantes podían escoger entre completar una tarea académica para ganar acceso a ítems preferidos o tomar un descanso sin demandas y sin ítems preferidos. Los requisitos del programa en el componente de demanda fueron incrementados sistemáticamente a través de las oportunidades de elección, mientras que el tipo y cantidad de reforzamiento se mantuvo constante. Los resultados mostraron que en el nivel más bajo de demandas, ambos participantes prefirieron la opción de trabajo. A niveles de demanda más altos, no obstante, ambos participantes eligieron la opción de tomar un descanso. A pesar de la ausencia de ítems preferidos en el componente de descanso, no se observaron instancias de conducta problemática después de esta opción. Esto indicó que este tipo de análisis puede ser usado para identificar condiciones que conducen a la obediencia entre individuos cuya conducta problemática se mantiene por escape o bien por múltiples reforzadores, sin la necesidad de provocar o reforzar la conducta problemática. Se discuten las limitaciones del presente estudio y se ofrecen recomendaciones para futura investigación.
Keywords: Conducta Problemática, Obediencia, Reforzamiento Positivo y Negativo.

Original recibido / Original received: 06/06/2014 Aceptado / Accepted: 19/09/2014

[1] This research was supported by NICHD Grant No. 44763. Jennifer J. McComas, Ph.D., University of Minnesota, 347 Education Sciences Building, 56 River Rd. East Minneapolis, MN 55455

Functional analysis (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994) permits identification of functional relations between problem behavior, including self-injury, aggression, and destruction, and its consequences. Determining the function of problem behavior, in turn, facilitates altering the relevant reinforcement contingencies to decrease problem behavior and increase appropriate behavior (Steege, Wacker, Berg, Cigrand, & Cooper, 1989). For example, treatment for negatively reinforced behavior often includes escape extinction, reinforcement of alternative, appropriate behavior (contingent on alternative behavior or noncontingent), or a combination of both (e.g., DRA, DRO; Vollmer, Marcus, & Ringdahl, 1995; Wacker et al., 1990). With any intervention, its effectiveness depends on the fidelity with which it is implemented. Consequence-based interventions such as extinction or differential reinforcement may be particularly difficult for parents, teachers, or other caregivers to implement with adequate fidelity, especially if the target behavior is frequent or of such severity that it is challenging or impossible to ignore. When the problem behavior includes aggression or self-injury, procedures that might produce a side effect such as an extinction burst may be inappropriate. However, research has shown also that positive reinforcement in the form of access to preferred items or activities, can be effective for improving problem behavior that is maintained by negative reinforcement or multiple functions (i.e., positive and negative reinforcement) (Payne & Dozier, 2013). Importantly, the effectiveness of positive reinforcement for reducing negatively reinforced and increasing appropriate behavior has been demonstrated even in the absence of extinction (Lalli et al., 1999).

One strategy that shows promise in assessing the effects of access to preferred items or activities on behaviors maintained by negative reinforcement is the use of concurrent schedules. Several studies have used concurrent schedules to examine the effects of manipulations of both positive and negative reinforcement on problem behavior and task completion. For example, Golonka, Wacker, Berg, Derby, Harding, and Peck (2000) provided two participants with negatively reinforced problem behavior to choose between continued work or taking a break during demanding tasks. In an alternating treatments design, the effects of providing access to preferred items during the break (combined negative and positive reinforcement condition) were compared to the effects of a break without preferred items (negative reinforcement alone). The combined condition resulted in greater reductions in problem behavior and increases in appropriate requesting. Nevertheless, escape extinction was necessary to increase completion of task demands without problem behavior for both participants.

Similarly, Piazza, Fisher, Hanley, Remick, Contrucci, and Tammera (1997) compared the effects of negative reinforcement with combined positive and negative reinforcement, with and without extinction. Three participants with multiply maintained problem behavior participated. They demonstrated that, for two participants, providing breaks with preferred items contingent on appropriate behavior was effective for decreasing problem behavior and increasing compliance without escape extinction. For the final participant, however, escape extinction was necessary. When the schedule of reinforcement for appropriate behavior was

faded, escape extinction and access to multiple reinforcers for appropriate behaviors were necessary for optimal results for all participants.

Finally, Hoch, McComas, Thompson, and Paone (2002) used a concurrent schedules arrangement to evaluate the effects of positive and negative reinforcement without extinction on the behavior of three children with autism whose problem behavior was maintained at least in part by negative reinforcement. They demonstrated that problem behavior was eliminated and task completion increased when problem behavior produced a break from task demands and task completion produced a break with access to preferred activities. These results were maintained even when the response requirement was increased and the schedule of reinforcement was thinned.

Overall, these results indicate that combining positive and negative reinforcement may be more effective than either form alone for decreasing problem behavior and increasing compliance. In many cases, however, escape extinction was necessary to achieve optimal results. One possible reason for this pattern is that participants in these studies were required to complete a certain amount of a difficult task in order to get access to an opportunity to escape from the task. One alternative strategy could be to provide opportunities for individuals to avoid the task entirely by presenting choice opportunities prior to presentation of task demands. In this case, escape or avoidance are always available for appropriate behavior (choice making), which reduces the likelihood of problem behavior. On the other hand, by manipulating the quantity or difficulty of the work presented, or the quantity or quality of the reinforcement available for task completion, it should be possible to bias the individuals' responding away from escape/avoidance and toward task completion.

In the current study, we evaluated the effects of positive reinforcement on the amount of work completed by two children with autism spectrum disorder (ASD) and developmental delays who engaged in problem behaviors maintained by escape from demands and access to tangible items (i.e., negative and positive reinforcement). We created a concurrent operant arrangement in which two response options were presented prior to the initiation of any difficult task demands: (a) negative reinforcement in the form of escape contingent on a request for a break, and (b) positive reinforcement in the form of access to a highly preferred edible item contingent on completing a pre-determined and signaled amount of work. Across trials, the amount and type of reinforcement available remained constant, as did the alternative option (break contingent on a request). A progressive-ratio schedule was implemented in which the schedule requirements for the positive reinforcer increased after each session in order to identify the highest number of work tasks that each participant would choose to complete in order to gain access to the preferred items.

Method

Participants and setting. Two individuals with ASD and developmental delay participated in this study. Both participants were referred for a functional assessment of severe problem behavior in the form of self-injurious behavior (SIB)

and/or aggression and property destruction. Ian was an 8 -year old Caucasian boy. Due to the severity of his behavior, Ian lived in a group home for adolescents with developmental disabilities and behavioral problems. He had age-typical gross and fine motor skills and some delays in the area of communication. Ian spoke in 2-3 word utterances, usually to request access to preferred items or to avoid/escape from non-preferred situations. He also engaged in echolalia. Ian had some basic self-help skills, including toileting and dressing with minimal prompting, but needed prompts to begin these tasks, and required help in most other areas of daily living. Ian exhibited occasional SIB, which typically occurred following episodes of aggression and property destruction. Aggression and property destruction occurred several times a week and included hitting, biting, pinching and throwing objects at people and had resulted in changes in residential placement.

The second participant was a 10- year old African- American boy named Gavin. He lived at home and attended a center-based behavioral treatment program that specialized in addressing the needs of children with ASD for 40 hours per week. Gavin received speech and language services and occupational therapy at the center. He had age typical gross and fine motor skills and used gestures to communicate. Gavin could produce word approximations with prompting. Gavin had limited self-help skills and needed help with all aspects of daily living. Gavin had a history of severe problem behavior including fecal smearing. His primary target behavior for the purposes of this analysis was self-injury that occurred several times per day and included hitting his chin and head and biting and pinching himself.

Sessions for Ian were conducted at his group home in the common eating area. The room was approximately 10 by 10 feet and contained a table with chairs and was adjacent to the kitchen, living room, and Ian's bedroom, which contained preferred items such as a television and toys. Sessions for Gavin were conducted at the day treatment center in an approximately 14 by 14 feet assessment room with a table and two chairs. All sessions were video recorded by the research team.

Dependent variables, response measurement, and interobserver agreement. Four dependent variables were coded: problem behavior during the functional analysis and choice analysis, item chosen during the preference assessment, response option chosen (break or work) during each trial of the choice analysis, and the number of work tasks successfully completed during trials in which the work option was selected in the choice analysis.

Trained research assistants collected direct observation data. Frequency counts were used to record instances of problem behavior during the functional analysis and choice analysis. For Ian, yelling and screaming nearly always preceded aggression and property destruction, which were severe at times, resulting in significant injury to others or damage to the environment. Thus, for safety reasons, yelling or screaming (e.g., any instance of a verbal noise or utterance at a volume louder than a typical speaking voice) served as the target behavior. For Gavin, SIB was operationally defined as any instance of chin hitting, open and closed hand head hitting, banging his head against objects, biting or

pinching himself. The remaining variables were coded by marking the item or response option chosen during the preference assessment and choice analysis and by tallying the number of tasks successfully completed during the choice analysis. A choice was defined as: a verbal response (i.e., saying “work” or “break”), a manual sign for work or break (Gavin only), touching or picking up the picture icon or token board associated with the choice, or starting the work trial (Ian only).

Inter-observer agreement (IOA) data were collected during 100% of the preference assessment trials, and approximately 30% of sessions across the functional analysis and the choice analysis sessions, respectively. IOA was calculating by dividing the number of agreements by the sum of the number of agreements and disagreements and then multiplying by 100%. For Ian’s sessions, IOA for all behaviors was 100%. For Gavin’s sessions, average IOA for all behaviors was 91% (*r* = 81-100%).

Procedure

Functional analysis. Analog functional analyses were conducted using multi-element designs to evaluate the influence of social reinforcement on problem behavior for each participant. The conditions implemented are described below and are based on procedures described in Iwata et al. (1982/1994) with the addition of a tangible condition. Sessions were 5 min long for Ian and were implemented by the group home manager with coaching by the research team. Gavin’s sessions were 10 min each and were conducted by his lead therapist with coaching from a research team. The order of the sessions was randomized and the analog conditions were designed based on descriptive assessments and functional assessment interviews for each participant.

Free Play: This condition was designed as a control condition. The participant and staff person were seated at a table with a variety of preferred activities available. The staff person provided verbal praise for appropriate engagement, commented about the activity every 10-15 s, and honored requests whenever possible. No programmed consequences for problem behavior were provided.

Attention (positive reinforcement): This condition was designed to assess the influence of contingent attention on problem behavior. The staff person instructed the participant to go play independently. All staff and other adults moved at least 10 feet away from the participant. The staff person ignored all social approaches, including verbal requests, and physical contact. Contingent on problem behavior, the staff person provided a brief period of attention in the form of verbal redirection (e.g., “No, you don’t need to yell”).

Escape from demands (negative reinforcement): This condition was designed to assess the influence of negative reinforcement, in the form of escape from demands, on problem behavior. The staff person instructed the participant to complete tasks identified as non-preferred by the staff (i.e., discrete- trial academic tasks, wiping the table, sweeping the floor). The staff member provided verbal prompts to continue the activity every 15-20 seconds, and physical prompts if

necessary. Contingent on problem behavior, the staff person said, “OK, you can take a break,” removed all materials and staff moved at least 5 feet away for 10-15 s.

Tangible (positive reinforcement): This condition was designed to assess the influence of positive reinforcement, in the form of access to preferred edibles or items, on problem behavior. Preferred edibles were selected based on reports from treatment staff. The staff person and participant were seated across from one another at a table. The edible was placed within sight, but out of the participant’s reach and the staff person told the participant that he had to wait for the edible. If the participant made an appropriate request (e.g., “toast please” or signed for candy), the staff person told him to wait. Contingent on problem behavior, the staff member gave the participant a small piece of the edible.

ABLA. The Assessment of Basic Learning Abilities (ABLA; Stubbings & Martin, 1995) is a hierarchical assessment in which standard prompting and reinforcement procedures are used to assess the ease or difficulty with which an individual is able to learn novel imitation and two-choice discrimination tasks. The ABLA was conducted with both participants in order to confirm that each had sufficient 2-item discrimination skills to complete the paired-choice preference assessment and the choice analysis.

Preference assessment. A paired-choice preference assessment (Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992) was conducted to identify preferred activities/items. The stimuli presented for each participant were selected based on staff report. Eight edibles were evaluated for Ian and 6 items/activities, including edibles, were presented for Gavin. The participants sampled each item for 30 s before the assessment began. Pairs of items were randomly presented 6” apart and 6” on a table in front of the participants. Each pair of items was presented twice, with the left-right position of each pair counterbalanced across presentations to reveal any location bias.

Concurrent operant. A concurrent operant analysis was conducted to examine the schedule arrangements in which each participant would choose either (a) to complete work tasks to earn access to preferred edible items or (b) to take a break from task demands. For each participant, academic tasks served as the task demands. The same academic tasks that were used in the functional analysis (Gavin) and that the participant could complete independently and accurately were used instead of the housekeeping tasks (Ian) that were used in the functional analysis because the academic tasks had a clear discrete trial format and allowed experimenters to systematically manipulate the response requirements presented across the choice trials. The academic task selected for Ian was 2-item non-identity matching tasks including colors, numbers, and letters. For Gavin, one-step tasks, including identifying pictures, gross motor and verbal imitation tasks were presented in random order.

Prior to beginning each trial, the staff person arranged the choice of work and break options with 2” picture icons symbolizing the work (i.e., a picture of a child sitting at a desk and writing) and break (a picture of a child sitting in a bean bag chair) options 6” apart on the table in front of the participant. In addition, either the full set of items to be matched (Ian), or the token board with the number of

tokens indicating the number of work trials to be completed (Gavin) was placed behind the “work” symbol with the rewards available for task completion.

Before presenting the first choice trial, the staff person exposed the participants to the consequences associated with each choice option by providing physical prompts to select each option and then followed through with the consequences of each choice. No prompts were delivered on subsequent trials. To start each trial, the participant was brought to the table and asked “Do you want to work or take a break?”

During trials in which the work option was selected, the staff person offered the participant a choice between two edible rewards identified as highly preferred via the paired choice preference assessment. After the participant selected the edible, the staff person instructed him to complete the task. If there was a delay of more than 3 s between responses, the staff person verbally and/or physically prompted the participant to continue working. Incorrect responses resulted in neutral verbal responses (e.g., “Ok, nice try”), and prompts to restart the incorrect task. If the participant engaged in problem behavior, the staff person physically prompted the participant to complete the current work task. After successful completion of one work task with prompting, the staff person asked the participant whether he wanted to continue working for access to the preferred edible, or if he wanted to take a break. If the participant said “break” or pointed to the break card, the trial was terminated and the participant was allowed to take a break away from the work table for 2 min, and the next trial was presented as usual. Conversely, if the participant said “work”, the name of the preferred edible, or continued working without problem behavior, the work trial continued. When all of the items were correctly matched (Ian) or all of the tokens had been removed from the token board (Gavin), the staff person gave the participant the selected preferred item, and provided him with up to 2 min to consume the item. Requests for more of the preferred item, or for continued work resulted in termination of the break and presentation of the next choice trial.

On trials in which the participant selected the break option, he was told to go play independently, and a timer was set for two min. Verbal requests for attention were honored and no demands were placed on participants during break times. During the break times, if a participant requested work or access to preferred edibles, the break was terminated and a choice new trial was initiated.

The number of tasks to be completed was increased incrementally across trials to determine the point at which each participant chose break instead of work. Trials were increased by 5 for Ian and by 2 for Gavin throughout the choice assessment. The goal was to find the maximum amount of work the participant would choose to complete rather than choose the break option. When Ian chose ‘break’ in 50% or more of the trials, the number of tasks was reduced and then increased again in order to replicate the effect. The number of tasks required of Gavin continued to increase until he reached 32 tasks, at which time experimenters and staff agreed that 32 was a sufficiently high number of tasks and to reduce the number of tasks required to avoid the risk of Gavin having an aversive experience of exposure to long ratios (Dardano, 1973). For both participants, a changing

criterion design was used to demonstrate experimental control (Gast & Ledford, 2014).

Results

Functional analysis. The results of the functional analysis for Ian (top panel) and Gavin (bottom panel) are depicted in Figure 1. For Ian, a high frequency of yelling and screaming was observed during the tangible and escape conditions of the functional analysis. These results suggest that his problem behavior was maintained by access to positive reinforcement (preferred edibles), and negative reinforcement (escape from task demands). For Gavin, SIB was elevated in both the tangible and escape conditions, whereas it only occurred in one session of the control condition. These results suggest that Gavin's problem behavior was maintained by access to positive reinforcement (food or preferred items), and negative reinforcement (escape from tasks). In both cases, these functions were consistent with the observations of the research staff and the reports of caregivers with regard to the antecedents (task demands, denied access to preferred items) that frequently preceded instances of problem behavior for both participants, as well as the consequences (escape from demands, access to preferred items) that were frequently provided in order to calm the participants during or after episodes of problem behavior.

ABLA. Both Ian and Gavin successfully completed the first four of six levels of the ABLA, which involves simple imitation, position discrimination, visual discrimination, and non-identity match-to-sample tasks. Neither participant was able to complete the final level, which involves a two-choice auditory-visual discrimination.

Preference assessment. Figure 2 shows the percentage of paired choice trials in which each stimulus was selected by Ian (top panel) and Gavin (bottom panel). Chocolate candies and animal crackers, and candy and play dough® were the highest preferred for Ian and Gavin, respectively. These items were used as the preferred items that the participants could earn for task completion during the subsequent choice analysis.

Concurrent operant. Figure 3 shows the result of the concurrent operant analyses for Ian (top panel) and Gavin (bottom panel). The results of the analysis for Ian indicate that problem behavior was relatively rare across the analysis, occurring on only 4/72 of trials overall. In every case, problem behavior occurred when Ian had selected the work option and had begun the task, and in all cases, he opted to complete the work task when given the option to take a break following the problem behavior. Looking specifically at the trials in which 15 or fewer tasks were required to earn access to the preferred items, Ian selected the break option only once, and successfully completed all of the required tasks without problem behavior during 98% of the trials. When the number of work tasks was increased beyond 15, Ian selected the break option and engaged in problem behavior somewhat more frequently but it was not until the work requirement reached 25 that he stopped consistently choosing the work option. These results were

replicated after reducing the work requirement back to 20 and then 15 and then increased again in increments of five.

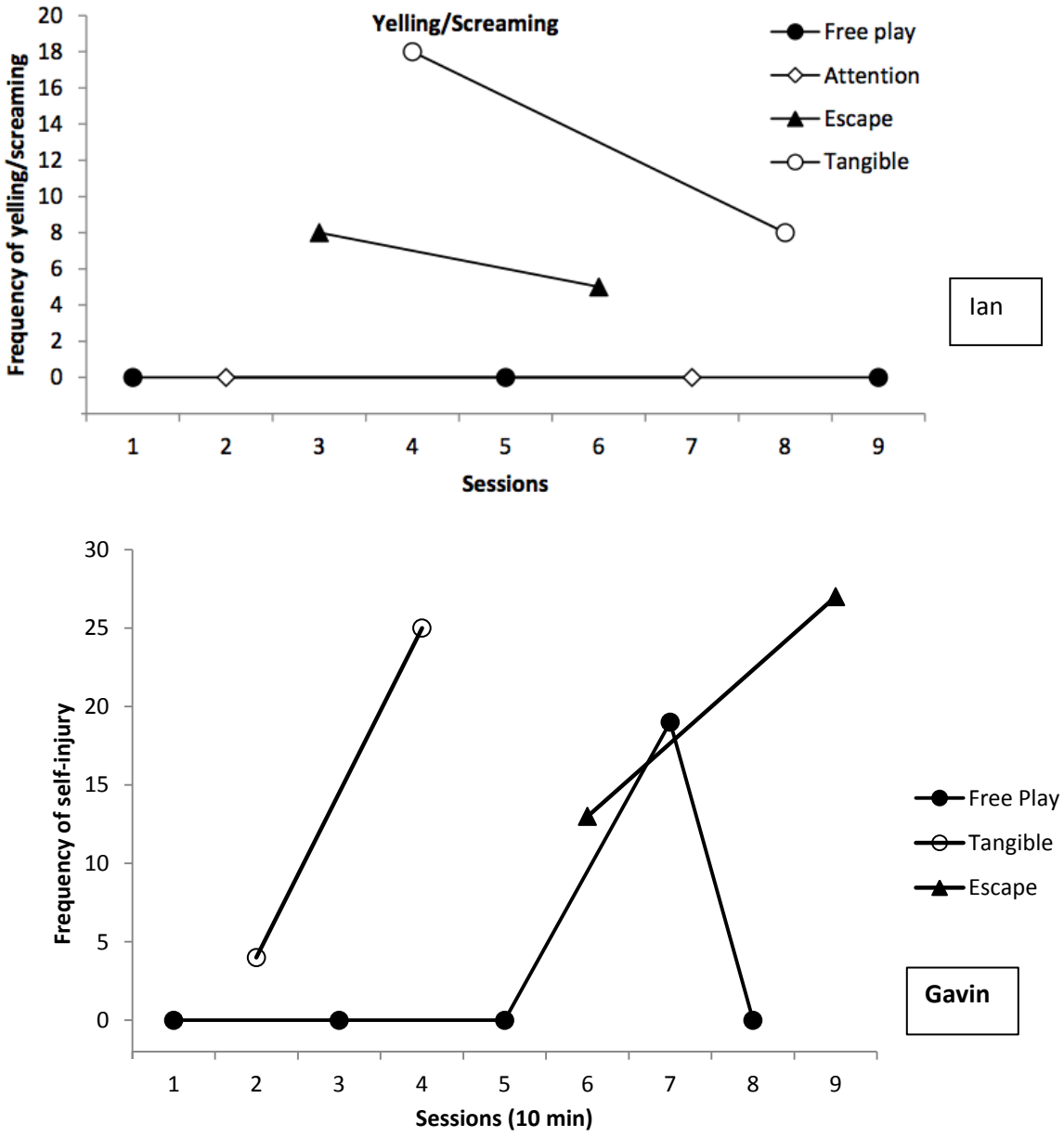


Figure 1. Frequency of challenging behavior across the functional analysis conditions for Ian (top panel) and Gavin (bottom panel).

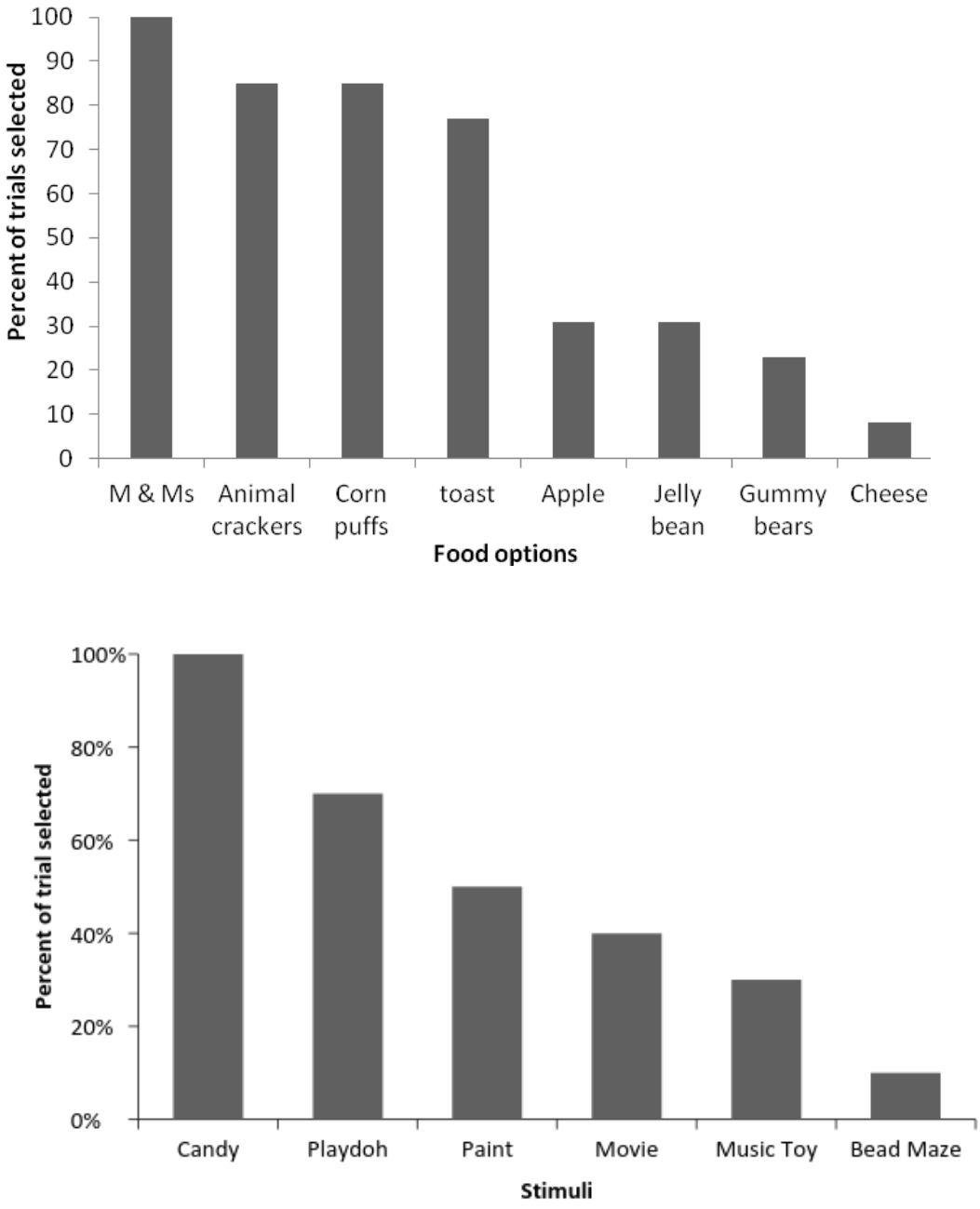


Figure 2. Percent of trials in which each item was selected during the forced choice paired preference for Ian (top panel), and Gavin (bottom panel).

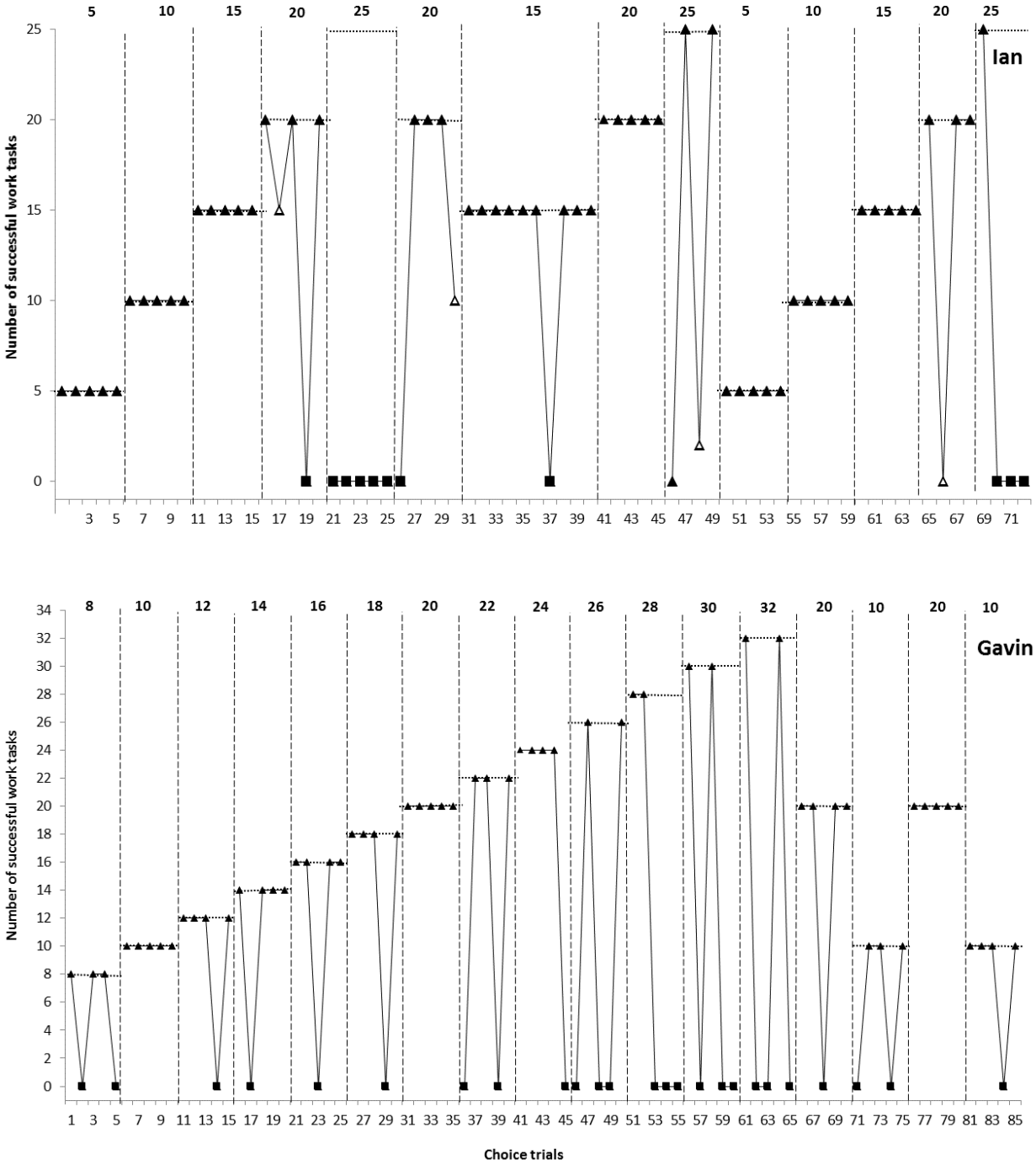


Figure 3. The number of work tasks completed prior to break requests, problem behavior, or task completion across different task demand levels for Ian (top panel) and Gavin (bottom panel).
Note: The numbers above the graphs, as well as the horizontal dashed lines represent the number of work tasks required to earn access to the preferred items. Closed shapes indicate that no challenging behavior occurred during the trial, whereas open shapes indicate that challenging behavior occurred at some point during the trial. Triangles indicate that the work option was chosen at the beginning of the trial, and squares indicate that the break option was chosen. Vertical dashed lines indicate changes to the work task criterion.

The results of Gavin's concurrent operant analysis were somewhat more variable. Across all 85 trials, Gavin selected the break option on 25, or less than 1/3, of the trials and successfully completed the work tasks on the remaining more than 70% of the trials. At the beginning of the analysis, when eight work tasks were required in the trial, Gavin selected the break option on 2/5 (40%) of the trials. However, in subsequent trials, he became overall less likely to select the break option. When 10 work tasks were required, Gavin selected the break option in only 3/15 trials. For trials in which 12, 14, 16, 18, and 20 tasks were required, he chose break no more than once per trial. For 22 and 24 tasks, he selected the break option on 2/5 (40%) and 1/5 (20%) trials, respectively. As the demands increased beyond 24 tasks, Gavin increasingly selected break, with 3/5 (60) trials being break choices in each of the 26, 28, 30, and 32 work task conditions. However, no level was identified at which Gavin switched entirely to picking the break option. Gavin did not engage in problem behavior during any of the trials throughout the analysis.

Discussion

The purpose of the current study was to utilize a concurrent operants arrangement as a tool for identifying conditions under which two children with autism spectrum disorder (ASD) and developmental delays who engaged in problem behaviors maintained by positive and negative reinforcement would choose to complete academic tasks to earn access to preferred items. The results show that as the schedule requirement to access the preferred items was systematically increased, there was a point at which both participants were less likely to choose to work over taking a break. Both participants showed some variability in their choice allocations to the work and break options, especially at the levels with highest task demands. Whereas Ian was extremely consistent in his selection of the work option at the lowest levels of task demands, Gavin's responding was more variable, selecting the break option on some trials even when task demands were very low. The reasons for these differences in choice allocation are unclear, but may be due to differences in the potency of the rewards selected for each participant, or in the level of difficulty or preference for the work tasks presented.

An important finding of the study was the extremely low levels of problem behaviors that occurred throughout the study for both participants. Given their long histories of both negative and positive reinforcement for problem behavior, it was possible that participants would engage in problem behavior rather than select the break option, because no preferred edibles were available during the break time. Neither of the participants, however, engaged in problem behaviors on trials in which the break option was selected even though preferred items were not available. Although Ian did show some problem behaviors during a small proportion of the sessions, the behaviors were less severe than those typically reported by his caregiver, and in each case when the choice between work and break was represented following an instance of problem behavior, Ian chose to complete the tasks, and no additional instances of problem behavior occurred.

Overall, these results suggest that this type of concurrent operant arrangement presented prior to the initiation of task demands may lead to increases in task completion without the need for escape extinction. Therefore, this antecedent approach to intervention could be an important tool for parents, teachers, and other care providers who work with individuals with IDD who engage in escape- or multiply-maintained problem behaviors but who are unable or unwilling to implement escape extinction. This analysis provided specific information regarding the conditions under which the participants would choose to complete tasks, and could be expanded to further parametric manipulations of task difficulty, or duration, as well as parameters of reinforcer quality. The information gained from this type of analysis (i.e., the number of trials that a participant is likely to choose to complete given the opportunity to earn access to a specific item or activity) is likely to be easily understood by individuals who are not well-versed in behavioral principles, potentially leading to better treatment integrity over the long-term.

Because several parameters were manipulated concurrently in the present study, it is unclear which specific elements of the design were responsible for the results. Specifically, studies have demonstrated that simply providing individuals with choices regarding activities may lead to decreases in problem behavior (Shogren, Faggella-Luby, Bae, & Wehmeyer, 2004). Therefore it is possible that simply allowing participants to choose between working or taking a break may have resulted in similar effects. However, considering the results of the other concurrent schedule arrangements in which escape extinction was necessary to establish compliance with task demands (e.g., Golonka et al., 2000; Piazza et al., 1997), this seems like an unlikely explanation. Other important components may have included the use of highly preferred items, and visual signals indicating the duration of the task to be completed. Because a component assessment was not conducted, it is currently unclear which components are necessary and/or sufficient for the observed reductions in problem behavior. Future research should systematically vary the presence or absence of these components in order to assess their independent effects.

The current study has several additional limitations that should be noted. First, no baseline data were collected to determine whether the participants would have engaged in problem behavior when presented with the specific work tasks used in the work option of the concurrent operant analysis, in the absence of the choice opportunity and access to preferred items. It is noteworthy, however, that a baseline session was attempted with Ian, but severe problem behavior (aggression and property destruction) immediately followed the instruction to complete the work task and led to the termination of the session. Considering the relatively high levels of problem behavior observed with both participants during the functional analysis sessions, it seems likely that both participants would have engaged in problem behavior without some or all of the intervention components, but future research should address this issue directly.

Second, the starting points for the number of work tasks presented were selected somewhat arbitrarily based on reports from treatment/group home staff,

and clinical judgment. The design could be strengthened by using a data-based selection of the number of work tasks required in the initial phase.

Third, although the choice analysis was conducted in the participants' natural environments, it was implemented by members of the research team, rather than by treatment staff or other caregivers in the natural environment. It is possible that having familiar caregivers, with whom there is likely to be a history of reinforcement for problem behaviors, implement the assessment would have affected the results, and future research should address this possibility.

Finally, no data on the generalization or maintenance of the findings were collected in this study. Knowing whether the participants' allocation to the work and break options was consistent over time could have important implications for using the information gathered in this type of assessment to inform the manner in which tasks presented. In addition, knowing how allocation might differ with different types of tasks and with different types and quantities of reinforcement would be valuable. For example, a participant might require more or higher quality reinforcement for a daily living task versus an academic task. Future research could examine the effects of varying the quantity or quality of reinforcement and/or task type on choice allocation.

Finally, these results could be viewed through a behavior economics lens, where the unit price for the positive reinforcer increased, consumption of the reinforcer decreased (see Madden, Bickel, & Jacobs 2000, Prediction 1). Future investigators might consider finding the break point for one reinforcer and then assessing additional potential reinforcers to determine whether it is possible to identify a reinforcer that has a higher break point. Findings of such a study might suggest which reinforcers to use under what conditions.

In conclusion, concurrent schedules arrangements with parametric manipulations of task or reinforcement quantity and quality show promise as a strategy for identifying the conditions under which children with autism and severe problem behavior will choose to comply with task demands, even in the absence of escape extinction.

References

Dardano, J. F. (1973). Self-imposed timeouts under increasing response requirements. *Journal of the Experimental Analysis of Behavior, 19*, 269-287.

Fisher, W., Piazza, C., Bowman, L.G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*, 491-498.

Gast, D. L. & Ledford, J. R. (2014). *Single case research methodology: Applications in special education and behavioral sciences*. New York: Routledge

Golonka, Z., Wacker, D., Berg, W., Derby, K. M., Harding, J., & Peck, S. (2000). Effects of escape to alone versus escape to enriched environments on adaptive and aberrant behavior. *Journal of Applied Behavior Analysis, 33,* 243-246.

Hoch, H., McComas, J. J., Thompson, A. L., & Paone, D. (2002). Concurrent reinforcement schedules: Behavior change and maintenance without extinction. *Journal of Applied Behavior Analysis, 35,* 155-169.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman C. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis,* 27, 197-209. (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2*, 3-20, 1982.

Lalli, J. S., Vollmer, T. R., Progar, P. R.,Wright, C., Borrero, J., Daniel, D., … May, W. (1999). Competition between positive and negative reinforcement in the treatment of escape behavior. *Journal of Applied Behavior Analysis, 32*, 285–296.

Madden, G. J., Bickel, W. K., & Jacobs, E. A. (2000). Three predictions of the economic concept of unit price in a choice context. *Journal of the Experimental Analysis of Behavior, 73*, 45-64.

Payne, S. W. & Dozier, C. L. (2013). Positive reinforcement as treatment for problem behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis, 46*, 699-703,

Piazza, C. C., Fisher, W. W., Hanley, G. P., Remick, M. L., Contrucci, S. A., & Tammera, L. A. (1997). The use of positive and negative reinforcement in the treatment of escape-maintained destructive behavior. *Journal of Applied Behavior Analysis, 30*, 279- 298.

Shogren, K. A., Faggella- Luby, M. N., Bae, S. J., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior: A meta-analysis. *Journal of Positive BehaviorInterventions, 6*, 228-237.

Steege, M.W., Wacker, D.P., Berg, W.K., Cigrand, K.K., & Cooper, L.J. (1989). The use of behavioral assessment to prescribe and evaluate treatments for severely handicapped children. *Journal of Applied Behavior Analysis, 22*, 23-33.

Stubbings, V., & Martin, G. (1995). The assessment of basic learning abilities test for predicting performance of developmentally disabled persons on prevocational training tasks. *International Journal of Practical Approaches to Disability*, *19*, 12-17.

Vollmer, T. R., Marcus, B. A., & Ringdahl, J. E. (1995). Noncontingent escape as treatment for self-injurious behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis, 28,* 15-26.

Wacker, D. P., Steege, M. W., Northup, J., Sasso, G., Berg, W., Reimers, T., Cooper, L., Cigrand, K., & Donn, L. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Analysis, 23*, 417-429.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1772 - 1792

The Sciences of Learning, Instruction, and Assessment as Underpinnings of the Morningside Model of Generative Instruction

Elizabeth M. Street & Kent Johnson*
Central Washington University, *Morningside Academy

Abstract

This paper focuses on a subset of the practices that have created the powerful learning technology developed and disseminated by Morningside Academy in Seattle, Washington, U.S.A. We briefly describe this technology, known as the Morningside Model of Generative Instruction, and tell how it builds on the selectionist approach of B. F. Skinner and the pragmatic approach of John Dewey. We also describe the critical role Precision Teaching plays at Morningside Academy and its dependence on findings from the science of learning and the science of instruction, including placement of learners, task analysis, content analysis, instructional protocols, and principles of instructional design. Last, we acknowledge the symbiotic relation between effective Direct Instruction programs that teach skills to accuracy levels and Precision Teaching, which takes these accurate repertoires and systematically turns them into high frequency performances that take on the character of *fluent* repertoires. Over time, using Precision Teaching across multiple and successive repertoires also creates more agile learners.

Keywords: Assessment, Direct Instruction, Instruction, Learning, Precision Teaching.

Las Ciencias de Aprendizaje, Instrucción y Evaluación como Cimientos de la Instrucción Generativa del Modelo Morningside

Resumen

Este trabajo se enfoca en una serie de prácticas que han creado la poderosa tecnología de aprendizaje desarrollada y diseminada por la Academia Morningside en Seattle, Washington, E. U. Se describe brevemente dicha tecnología, conocida como el Modelo Generativo de Instrucción Morningside y se menciona cómo se construyó bajo la aproximación seleccionista de B. F. Skinner y la aproximación pragmática de John Dewey. También se describe el rol crítico que la Instrucción de Precisión juega en la Academia Morningside y su dependencia en hallazgos de la ciencia del aprendizaje y en la ciencia de la instrucción, incluyendo el papel de los aprendices, el análisis de tareas, el análisis de contenido, los protocolos instruccionales y los principios del diseño instruccional. Finalmente, se reconoce la relación simbiótica entre los programas de Instrucción Directa efectiva, que enseñan habilidades para lograr niveles de precisión y la Enseñanza de Precisión, que considera dichos repertorios precisos y sistemáticamente los convierte en ejecuciones de alta frecuencia que tienen el carácter de repertorios fluidos. Con el paso del tiempo, usar la Enseñanza de Precisión a través de múltiples repertorios sucesivos también crea aprendices más ágiles.

Palabras Clave: Evaluación, Instrucción Directa, Instrucción, Aprendizaje, Enseñanza de Precisión.

Original recibido / Original received: 14/07/2014 Aceptado / Accepted: 30/09/2014

Practices derived from the learning sciences and the philosophical underpinnings that guide them are combined to create a powerful learning technology at Morningside Academy in Seattle, Washington, U.S.A. This technology, known as the Morningside Model of Generative Instruction (MMGI), has resulted in changed learning trajectories for over 1,000 learners at the Academy and over 30,000 students in over 130 schools and agencies in the United States and Canada through the Morningside Teachers' Academy, the Morningside Summer School Institute, and countless presentations at the annual conference of the Association for Behavior Analysis International and other similar conferences.

Morningside Academy is well known as a Precision Teaching (PT) school. However, PT technology works only to the degree that it is faithful to the philosophical underpinnings of the learning sciences, analysis of the material to be taught, and elegant instruction that constitute the MMGI. The MMGI builds on five important streams of research: (a) generativity and contingency adduction; (b) content analysis, instructional design, and implementation; (c) program placement and modification based on continuous measurement; (d) classroom organization and management; and (e) critical thinking, reasoning, problem solving, and self-regulated decision-making (Johnson & Street, 2004b).

Writing for the Society for the Teaching of Psychology, Benassi, Overson, and Hakala (2014) describe "the interplay between the science of learning, the science of instruction, and the science of assessment" (p. 3) in the learning sciences, an excellent description of the work that is done at Morningside Academy. Morningside Academy is—to borrow from a taxonomy described by Jim Johnston (1996)—a third level research institution. That is, it is primarily service-oriented and rarely conducts basic research. Still, the Morningside leadership team and faculty are guided by a well-honed understanding and application of the basic and applied research related to these three sciences. In this article, we describe the ways in which these three sciences inform practices at Morningside, but we begin with the philosophical tenets that undergird them.

Philosophical Underpinnings

Morningside Academy's team is conversant with the science of learning and its philosophical roots. The Morningside Model of Generative Instruction blends the selectionist approach of B. F. Skinner with the pragmatic approach of John Dewey. Skinner (1969) "draws a parallel between the emergence of complex behavioral repertoires and the emergence of more complex and variably functional forms in evolutionary biology. The environment selects simple forms, and a more complex entity gradually emerges" (Johnson & Street, 2004b, p. 20). At Morningside, we see the selectionist principle as establishing the repertoire of the learner that forms the building blocks for more complex repertoires. As these building blocks become fluent, that is, "accurate, speedy, durable, smooth, and useful" (Johnson & Street, 2004b, p. 20), the selectionist principle builds more complex intellectual skills by combining the building blocks in ways that meet learners' learning needs.

We also draw heavily on the pragmatic functional approach espoused by John Dewey (1896, 1976, 1981, 1986). One of several Dewey tenets that underpin

MMGI is his emphasis on "natural influences over learning, taken from the student's current activity, goals, and value systems" (Johnson & Street, 2004b, p. 21). Morningside's process becomes organic as a student moves up the curriculum ladders. After mastering a core foundation of basic skills, when areas of interest that are important to one or more learners emerge for which their basic repertoires are not yet fully seeded or when learners are asking why a particular basic skill is important, the teacher may analyze the area(s) of interest or next steps in the curriculum to determine ways that learners' current repertoires overlap with them and provide prompts that enable learners to engage successfully in the activities. For more on the organic approach, see Johnson and Street (2013).

In addition, based on the requirement of The National Council for Accreditation of Teacher Education's (NCATE, 2008) for U.S. schools of education to specify the *conceptual framework* that underlies their educational preparation programs, two pseudo-philosophical positions have become prominent: constructivism and instructivism. Instructivism is a molecular approach to education, while constructivism is a molar approach. At Morningside Academy, we find the molecular approach which instructivism promotes to be advantageous for teaching new behaviors. However, we also find that looking at educational programming through the lens of the constructivists keeps Dewey's concern for natural influences over learning in the mix. We use instructivist practices to seed the learner's repertoire and thus prepare them to participate in constructivist practices such as Project-Based Learning. In other words, we attempt to turn the upside-down constructivist world that begins with composite, real-world activities right side up by seeding repertoires with component skills so that learners are competent to participate fully in the composite constructivist world.

The philosophies that underpin educational practice are thought-provoking, however, the focus at Morningside Academy is on ensuring that learners who begin the program lagging behind same-age peers are provided instruction that brings them to, or ahead of, the level of their peers, makes them good and productive scholars, and improves their scores on standardized tests. In fact, parents of learners enrolled at Morningside are offered money-back guarantees for their children's tuition if their children do not gain at least two years on standardized tests for one year of participation in their area of greatest weakness. It is evidence of the effectiveness of its teaching technologies that, during the 34-year history of this pledge, Morningside has returned less than one percent of tuition. To ensure that the promised gains occur, Morningside's team relies on relevant findings from the science of learning, the science of instruction, and the science of assessment.

Principles Derived From the Science of Learning

Findings from the science of learning are routinely incorporated in Morningside's classrooms. First, teachers draw on the power of reinforcement, specifically, and feedback, generally. For example, Morningside's Daily Support Card (Johnson & Street, 2004b), the conduit for distributing points for good performance, serves as a daily form of communication among learner, parent, and

teacher. Teachers provide a pre-determined maximum number of points based on each of four categories of behavior—academic, learning skills, organization, and citizenship. Each teacher defines and exemplifies rules related to each category early in the year and awards points immediately when desirable behavior is evident. Learners take their support cards home each day, and their parents have the opportunity to reward their hard work as well. When parents provide reinforcement from their own menu at home for work well done, it further strengthens the behaviors that will ultimately recruit reinforcement from others.

Second, observing a lesson at Morningside reveals that teachers apply findings related to the selection of effective prompts that can be withdrawn systematically and easily (MacDuff, Krantz, & McClannahan, 2001). Teachers are also conversant with the applied research on shaping (Pryor, 1999), discrimination and generalization (Tiemann & Markle, 1990), errorless learning (Terrace, 1963), stimulus control (Mayer, Sulzer-Azaroff, & Wallace, 2012), establishing and motivating operations (Laraway, Snycerski, Michael, & Poling, 2003; Michael, 1982), and schedules of reinforcement (Vargas, 2013).

A very important principle derived from the science of learning is the delayed prompting procedure. It underpins Morningside's reading and other comprehension procedures where learners need to make sense of what they have read or heard and apply it elsewhere. Based on the work of Touchette and Howard (1984), by delaying prompts for six seconds, learners are provided the least amount of prompting needed to respond correctly to a question. This reduces prompt dependence and the need to fade prompts later. It also provides an opportunity for learners to "show what they know" before help is provided.

Morningside teachers also focus on teaching students how to learn on their own. Many parents who brought their children to Morningside only to catch them up find that a more important result occurred: their children became effective learners. Effective learners demonstrate *generativity***.** In a generative process, behaviors learned under prior conditions or circumstances are *recruited* by new, very different conditions to form new combinations or blends that serve a new or different function or outcome in a new context and in the absence of instruction. Generative Instruction involves arranging conditions that produce novel and complex behaviors, in new circumstances, without directly teaching them. (See examples that appear later in the article.)

To promote generativity, Morningside teachers apply strategy and problem solving research from the science of learning by arranging contingencies that recruit current relevant repertoires learned under one set of conditions for new purposes. (See, for example, Andronis, Layng, & Goldiamond, 1997; Epstein, 1991). In some circumstances, there is an obvious connection between what has been learned and what is now required, which improves the likelihood of successful recruiting. To promote more distant generative connections, Morningside has been influenced by the work of Whimbey (1975) and Whimbey and Lochhead, (1991) on reasoning and problem solving. Morningside's principal has adapted Whimbey and Lochhead's Think Aloud Pair Problem Solving approach for learners at Morningside, who learn to recruit current relevant

repertoires for figuring out how to solve a problem and complete novel tasks in the absence of instruction (Robbins, 2011, 2014).

Principles Derived From the Science of Instruction

There are at least five aspects of the science of instruction that play a prominent role in Morningside practices. They include a) *placement of learners* in groups for instruction; b) *task analysis*; c) *content analysis*; d) *instructional protocols* and e) *principles of instructional design*.

We briefly describe each in turn though, in practice, they are much more organically applied.

Learner Placement

The Joplin Plan, which was originally developed to facilitate gains in reading (Wahlberg, Reynolds, & Wang, 2004) is used at Morningside for placement of learners in all academic areas (Kulik, 2004). It is an ability grouping approach in which learners are placed with those whose skill levels are similar to their own. The Joplin Plan also facilitates another important aspect of the Morningside approach, peer coaching, which we describe later.

Task Analysis

In behavioral circles, task analysis began as a systematic way to dissect a specific task into the skills needed to perform it and the order in which the skills should be performed for maximum efficiency. Mayer et al. (2012) define task analysis as "breaking down a complex skill, job, or behavior chain into its component behaviors, sub-skills, or subtasks." (p. 710). The Morningside team conducts this kind of task analysis when the situation calls for it, but it specializes in content-area level analyses using an approach described by Eric Haughton (1972).

Haughton, who worked with severely mentally handicapped adults, found it most effective to identify three sets of skills his learners needed to function in their environment: tool skills, component skills, and composite skills. *Tool skills* are the basic skills in a field, those which are necessary to acquire higher-level skills. Haughton specifically compiled evidence that there were 12 self-help tool skills—he called them the "big 6 plus 6" (DesJardins, 1980). Haughton's second-level skill set are *component skills*—skills which depend on one or more tool skills. *Composite skills* are "authentic, higher-level performances that socially validate a learner's mastery of a content area" (Johnson & Street, 2013, p. 41; also see Johnson & Street (2013) for our analyses of reading, writing, and arithmetic.)

For example, in teaching reading, a tool skill might be accurately saying the sound(s) of each letter presented individually and in combinations. A component skill might be phonetically reading regular one to three syllable words. A related composite skill might be reading passages with expression. The categorization of

an objective as tool, component, or composite depends not only on the content being analyzed but also on the incoming skill of the learner.

Haughton (1980) also introduced the concept of learning channels. The learning channel describes the way in which the learner comes in contact with a stimulus (an input) and the way in which the response is to be composed (an output). Haughton identified seven potential inputs including, among others, taste, see, and hear, and eleven potential outputs including, among others, mark, match, say, do, and write. Thus a learner might see (input) and then say (output) the names of letters of the alphabet (abbreviated "see/say" names of letters of the alphabet) or hear/say words composed of sounds presented one at a time. Haughton believed, based on evidence he had compiled, that a learner isn't automatically able to transfer across learning channels. That is, because they can accurately see/write math facts doesn't necessarily mean they can hear/say math facts.

Content Analysis

Content analysis categorizes the skills that have been identified in a task analysis into different types that are best served by differing instructional and practice procedures. The two educators who are credited with first providing content analysis taxonomies are Bloom (1956a, 1956b) and Gagné (1965). Their work was followed by that of Engelmann and Carnine (1982) and Tiemann and Markle (1978, 1990). The Morningside team finds Tiemann and Markle's approach to be the most user-friendly of these four approaches. Tiemann and Markle posit nine types of learning and provide the reader with specific steps and sample programs for encouraging learning of each. The three umbrella terms in their model are psychomotor learning, simple cognitive learning, and complex cognitive learning. Psychomotor learning is made up of single responses, response chains, and kinesthetic repertoires; simple cognitive learning consists of associations, (verbal) sequences, and verbal repertoires; and complex cognitive learning is made up of concepts, principle applying, and strategizing.

Instructional Protocols

Instructional protocols is a generic name for the manner in which concrete tasks that have pre-specified outcomes are presented to learners. Gilbert (1962a, 1962b) established the four-step protocol known as *mathetics,* which continues to inform instructional design today. In Gilbert's protocol, the teacher first *demonstrates* the skill. Second, the teacher *guides* the learner through the use of prompts. Gilbert called the third step *release*, in which the teacher provides the learner with an opportunity to perform the skill on his or her own. The last step, *delayed release* or *spontaneous completion,* occurs after either time or other items are interposed with the target before returning to it. Gilbert's mathetics was recursive; that is, the teacher would demonstrate, then move to the guide stage when the learner appeared to be ready to do the skill with prompts. However, if the learner was unsuccessful, the teacher would immediately return to the

demonstration stage with more examples. Similarly, if the teacher had moved ahead to the release phase and the learner made errors, Gilbert's model called for reverting to prompted examples before introducing another release trial. Continuing the basic protocol, an unsuccessful response to the stimulus after a delay would return the learner to the release phase until the learner appeared ready to try again with a delayed release trial. In Engelmann and Carnine's (1982, 1991) Direct Instruction scripts, *model, lead, test,* and *delayed test* are equivalent to Gilbert's demonstrate, guide, release, and delayed release phases. Archer's (See Archer & Hughes, 2011) *I do it. We do it. You do it. You do it again* are similar equivalents. Like Gilbert, both Engelmann and Carnine and Archer and Hughes instruct teachers to use the process recursively. An important aspect of the protocol is that it works for all the different kinds of learning where a pre-specified answer is called for—for example, in teaching the steps in long division, as well as in teaching concepts, such as identifying examples and non-examples of Romantic music, classifying plant phyla, or distinguishing fair from unfair social relations. If the learner is unsuccessful at any of the stages, the teacher drops back to the previous stage, seeking the point where the learner is successful and then moves forward again in an iterative process.

Two additional protocols that are most often attributed to Engelman and Carnine (1991) significantly improve teaching scripts: signaling and faultless communication. Signaling—for example, the teacher tapping his pen on the whiteboard or snapping his fingers—cues learners when responding is required and appropriate. The skilled teacher hears when a learner is struggling with the task and can do a quick error-correction procedure until the learner is responding correctly and on signal with others in the group. Teachers also strive for *faultless* or unambiguous—*communication* as described by Engelmann and Carnine (1982). Both Adams and Engelmann (1996) and Engelmann and Colvin (2006) describe features of official Engelmann Direct Instruction programs,

Instructional Design

When there currently exist no workable protocols for teaching instructional objectives that are important to creating well-rounded learners, the Morningside faculty and leadership team develop their own materials, using their adaptation of a "System of Instruction" model that was developed by Markle and Tiemann (1967). This adaptation (See Figure 1), along with the influence of Markle (1990), Gilbert (1962a, 1962b), and Engelmann and Carnine (1991) forms the basis for the instructional design work that is done at Morningside.

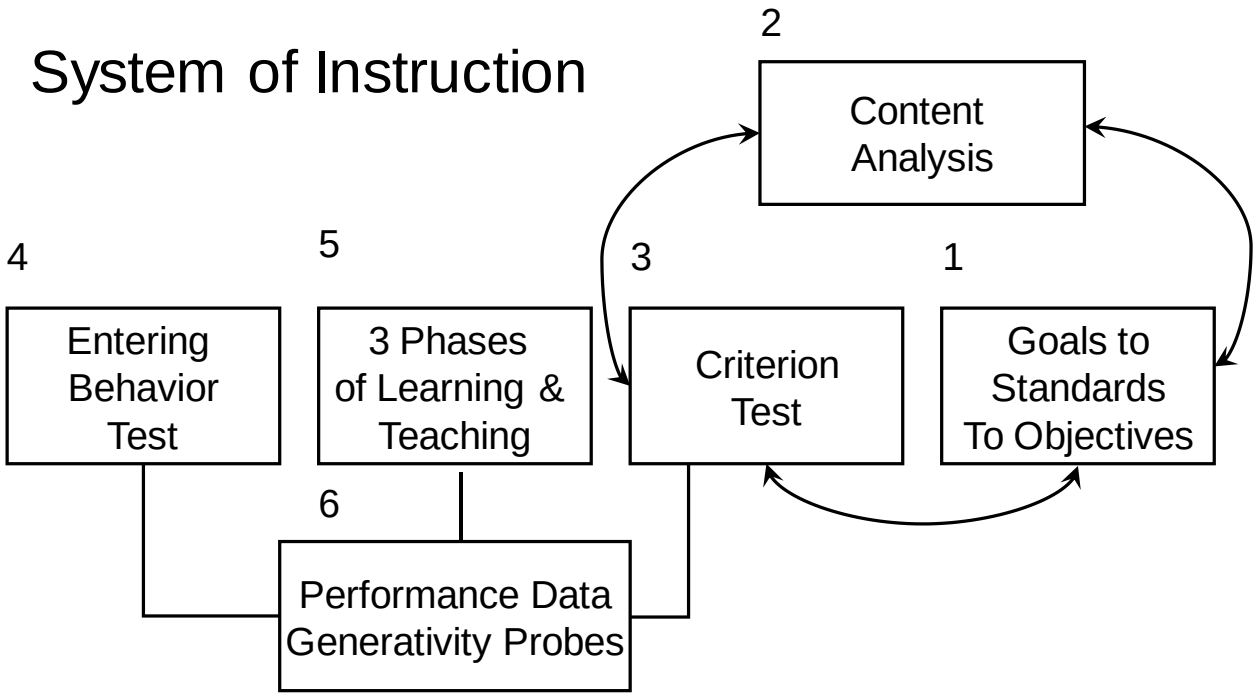


Figure 1.System of Instruction, adapted from Markle and Tiemann (1967).

The system of instruction model works equally well with a curricular strand within a field of study or for the entire field. Thus, it can be applied to a curricular strand such as phonemic awareness within the field of reading or to reading as a whole. The critical aspects of the work include a thoroughgoing analysis of the content area or curricular strand; selecting and using one of the learning typologies that we've discussed earlier, determining the correct ordering of elements in the curriculum so that learners' progress is seamless, finely tuning the instructional protocols, and ensuring that data are collected that provide evidence that the design has been learner verified when a substantial percentage of learners achieve mastery.

In addition, Morningside's programmers review new and promising materials that come on the market and—after obtaining appropriate permissions—use or modify them to expand its bank of programs. For example, Morningside programmers designed a direct instruction script and practice worksheets based on *Word Workout* (Lewkowicz, 1994), a program designed to teach learners to decode complex multi-syllable words. Sometimes individual teachers complete less formal adaptations when current materials aren't achieving the desired level of mastery. For example, one Morningside teacher adapted the vocabulary development work of Beck and her colleagues (Beck, McKeown, & Kucan, 2002) to improve her students' mastery of vocabulary.

Morningside also creates programs *de novo* when there are none available that meet the Academy's standards. In these cases, they begin with Morningside's "system of instruction" as the basis of the programs, conduct a

component/composite analysis, clarify the types of learning involved and appropriate learning channels for the objectives, and write scripts that use mathetics, signals, and faultless communication. Two examples include their recently available program related to computation (Johnson & Melroe, 2014) and a soon-to-be-available word problems program (Johnson, Isbell, Delgado, & Leon, 2015). Available from Morningside Press, these programs include a direct instruction script as well as practice sheets appropriate for Precision Teaching practice.

Principles Derived from the Science of Assessment

According to Malmquist (2004) "a hallmark of Morningside's procedure is the continuous interplay between instruction and assessment" (p. 52). Malmquist proceeds to describe three levels of assessment used at Morningside: micro-level, meta-level, and macro-level.

Three Levels of Assessment

The Micro Level: Precision Teaching (PT) serves as the micro-level assessment at Morningside. Originating from the work of Lindsley and his students at the University of Kansas in the 1960s (Johnson & Street, 2014), PT provides a mechanism through which changes in performance frequency can be tracked. Frequency—the number of performances of a tool or component skill over time—provides a reliable mechanism to determine the fluency of the skill.

Lindsley chose frequency as the best indicator of fluency because frequency measurement is much closer to direct observation of behavior than percent correct, percent of intervals, or time samples of behavior and is a true measure of behavior in time. Frequency also very accurately represents the probability of future action. Thus, Lindsley believed that building behaviors to high frequencies would make their future performance more likely (Pennypacker, Gutierrez, & Lindsley, 2003).

Fluency, as a qualitative concept, has been described as performance that is "flowing, flexible, errorless, automatic, confident, second nature, . . .masterful" (Johnson & Street, 2013, p. 21). Although most people recognize a fluent performance when they see it, they would be hard pressed to say the frequency required to achieve that end. That's why, over time, fluency has been defined by its by-products, of which five have emerged to date: (a) The behavior is at a *frequency* where it is *maintained* and thus is easily executed when needed (Haughton, 1972; 1980); (b) it has ***e**ndurance* necessary to stay in play for as long as real-world contingencies require (Binder, 1985); (c) it has ***s**tability* in the face of distraction (Johnson & Layng, 1992, 1994, 1996); (d) it is available for real-world *applications* that require it (Haughton, 1972, 1980); and (e) it results in *generativity* (Johnson & Layng, 1992, 1994, 1996; Johnson & Street, 2013) and thus "is easily combined with other performances as necessary to solve novel problems" (Johnson & Street, 2013, p. 28). A mnemonic--"Get the MESsAGe!—helps novices remember these by-products. Morningside's team has identified frequency ranges in reading,

writing, and arithmetic (See Johnson & Street, 2013) that correlate well with these by-products.

Practice in a PT classroom is far from the “drill and kill” approach, in which practice was an end in itself. In the Precision Teaching approach to practice, learner performance is timed, most typically in one-minute intervals. The learner and his teacher or a peer coach then review his performance with respect to a goal based on the previous day’s performance and the learner’s ultimate aim, the frequency that “gets the MESsAGe.” Typically learners will practice a skill several times a day within the time set aside for practice and most will meet their daily goal. When a learner fails to meet his goal, the teacher will review his chart and may talk with him and his coach to determine the reason for the challenges he is facing. Two of many possible courses of action include slicing back to an earlier piece of the curriculum or isolating items which were causing particular trouble onto a new practice sheet (Johnson & Street, 2004b, 2013). The teacher assigns these new sheets as the next day’s practice. Practice at Morningside is daily, highly structured, and individualized. Progress toward frequency goals is charted every day.

The Standard Celeration Chart (SCC; see Figure 2) is the vehicle through which changes in frequency are tracked. Johnson and Street (2013) report that the SCC shares the following charting conventions with some other charts: It is “(1) standardized, for easy communication, and chart and program comparison; (2) calendar-based, not session-based, to show the effects on performance of programs when they are in place and when they are not; (3) focused upon frequencies, not percent correct; and (4) focused on learning, not performance” (p. 30).

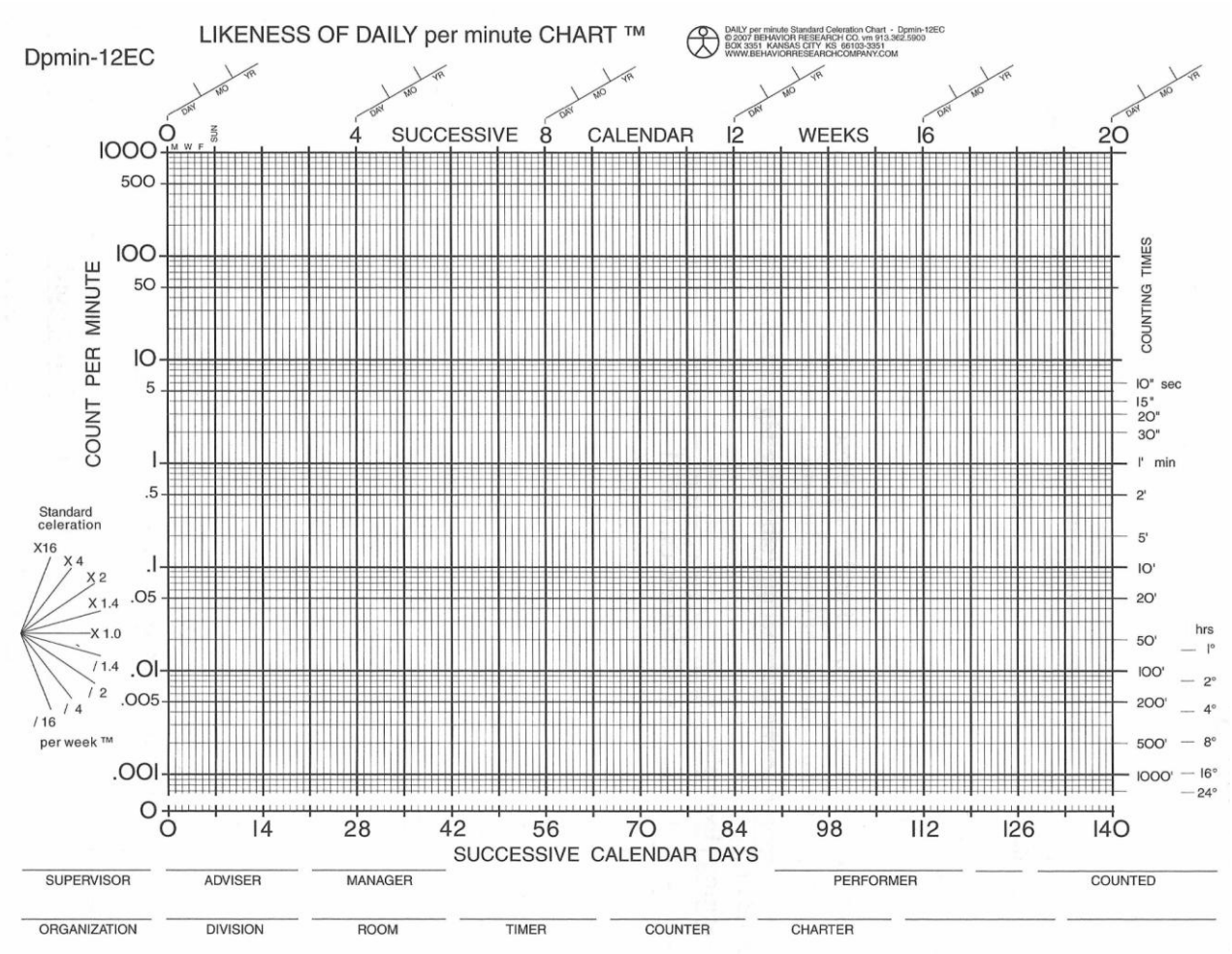


Figure 2. Likeness of a Daily per minute Standard Celeration Chart. Standard Celeration Charts are available at Behavior Research Company, Box 3351, Kansas City, KS 66103-3351. VM 913-362-5900, www.behaviorresearchcompany.com

The Standard Celeration Chart differs from other instruments in that it plots ratios of frequencies, not raw performance frequencies, over time. “The growth that learners make is proportional to their previous growth. Proportional growth is much more representative of the way people really learn” (Johnson & Street, 2013, p. 30). An inspection of Figure 2 reveals the ratio scale up the *y* or left axis of the chart. Rather than being equidistant from each other in a linear fashion; they are arranged by multiples of 10, more like charts or graphs that one sees in the physical sciences than it is like those used in education. Lindsley was drawn to the ratio scale because he believed that, just like other things in nature, behavior changed in relation to where it was when one started charting it (White & Haring, 1980). The chart also accommodates virtually any behavior since the range of possible frequencies is from .001 per minute to 1,000 per minute.

The chart was named the Standard *Celeration* Chart because Lindsley (1992) was more concerned about the rate of growth over time in performance than he was about performance at any point in time. He coined the word *celeration* to refer to the rate of growth (**ac**celeration) or deterioration (decelertion) in learning. Because celeration measures how much time it took for a learner to reach a frequency aim, Lindsley (2001b) thought of celeration as synonymous with learning. The chart is designed in such a way that celeration is easily determined by drawing a line from the first frequency the learner posts to the prescribed frequency, when it is achieved, and comparing the slope of the line with the *Standard celeration per week*^{TM} legend on the left side of each chart. (See Figure 2.)

As evidence has emerged that higher frequencies appear to be characteristic of "expert" performance, precision teachers have attempted to find ways to increase learners' celerations. Two benefits have emerged from encouraging higher frequency performance on tool and component skills. The first is what the staff at Morningside call *curriculum leaps*--learners require little if any instruction or practice to acquire next steps in a curriculum series when the previous steps are at prescribed frequencies. For example, a learner may acquire long division with minimal practice if both math facts and estimation are at high enough frequencies. The Morningside team estimates that approximately 33 percent of the curriculum is acquired in this manner. The second benefit is that new learning channels emerge with minimal practice when other channels are at prescribed frequencies. For example, a learner who is fluent with a "see/say" also is fluent with a "see/write" or a "hear/say" related to the same content with no or only minimal additional practice.

Later in his life, Lindley (2001a) talked about a relation between celeration and agility as akin to the relation between frequency and fluency. An agile learner is one who is mentally quick and resourceful, able to adjust quickly to unfolding events in learning something new. The Standard Celeration Chart shows growth in agility as steeper and steeper slopes across time and across performances. Lindsley thought it was possible and even likely that speedier celeration on several sets of related behaviors would improve the ability to acquire other related behaviors more speedily (Lindsley, 2001a). We have seen some compelling evidence of this phenomenon at Morningside with some of our more advanced learners. Although we have not consistently documented agility patterns in our students' data, others are beginning to do so. (See, for example, the work of Meyer, Newsome, & Newsome, 2013).

The Meta Level: Meta-level assessments occur less frequently than micro-level ones, but more frequently than macro-level assessments. Morningside has adapted curriculum-based measurement (CBM) procedures (Deno, 1985, 1989; Shinn, 1989) to track growth on important curriculum indicators in reading, writing, and mathematics. These adapted CBM measures both validate the results that learners are charting on their SCCs—their mini-level assessments—and suggest how learners are likely to perform on the macro-level assessments at the end of the year. To do this, using the previous year's data, the team conducts a simple linear regression between the scores at a particular point in time on a standardized

meta-assessment and scores on the macro assessment. This regression line then allows them to determine what the current year's learner needs to achieve on each meta-level assessment to achieve the promised two year gain (Gire, Testa, & Johnson, 2010). Typically, Morningside collects meta-level assessment monthly or bi-monthly and, when learners aren't on track to make the gains that parents have been promised, faculty and the leadership team huddle to determine programmatic changes that are likely to increase the learner's growth to be consistent with expectations. (For more on the history of Precision Teaching or Lindsley's legacy, see Binder, 1996; Johnson & Street, 2014; and Potts, Eshelman, & Cooper, 1993)

The Macro Level: Macro-level assessments utilize published criterion- and norm- referenced tests to compare the performance of learners from the beginning to the end of the year in relation to a designated peer group. This is the "show me the money" part of the assessment process for two reasons: 1) for those enrolled in the laboratory school in Seattle, these tests determine whether the school or the parents get to keep the learner's tuition; and 2) for partner schools who participate through the Morningside Teachers' Academy, pre- to post-score gains on these tests determine eligibility for federal funding. It is also how many of them determine whether or not to renew their contracts with Morningside Teacher's Academy.

The standardized tests we use for the pre- to post-test comparisons are the state-approved tests in states where participating programs are located. They change periodically; however, Morningside stays current with the state's selection so that comparisons with other schools in the state are possible.

Putting it All Together

We build our instructional programs using our adaptation of Markle and Tiemann's (1967) System of Instruction (Figure 1). One very important piece is in box 5: The three phases of teaching. These three phases are instruction, practice, and application. It is this three-stage model that is at the heart of the Morningside Model of Generative Instruction. Learners typically begin new content with the first phase—*instruction*—during which the instructional protocols we described earlier are evident. In this phase, we *establish a new repertoire*; that is, the learner *acquires* a performance that she could not perform previously. The format of the lesson is determined by the learning channel and learning outcome it is designed to teach.

Students and teacher engage in a highly interactive lesson that focuses on only one performance or skill at a time and they are then combined as accuracy emerges. During this phase, learners are dependent on prompts, make errors early on, and are distracted by extraneous stimuli. It is also in this phase that response topographies are shaped and discriminations among and stimulus control by novel and familiar stimuli is assured. As Johnson and Street (2004b) note, "Student performance comes under the control of the parameters that define acceptable variability of stimuli and acceptable latitude for responses" (p. 99). Instructional lessons are characterized by increasingly higher rate volleys with the teacher providing continuous feedback about the correctness of the response. As learners

become more and more confident and their responses are very nearly always correct, they move to the second phase: *practice*.

At Morningside, students spend as much as 40% of their school day practicing in highly structured and timed activities. Practice is goal-oriented and continuously monitored. Practice activities exist on paper, on computer, and/or with flashcards for each major tool and component foundation skill in the curriculum. Continuous monitoring, which is critical to achieving efficiency, occurs as a function of Morningside's well-oiled peer coaching system (Johnson & Street, 2013). Learners use Lindsey's Timings Chart (See Figure 3) and his Daily per minute Standard Celeration Chart (Figure 2) to track performance and to suggest and verify that the learner is improving—**ac**celerating—at the prescribed rate. Performance aims are established to tell the student how many of a skill they should be able to do in the timing period, based on the celeration aim for the task. The learner and his peer coach use the Timings Chart, which accommodates up to 10 practice sessions per day, to track the learner's daily performance and celerations and to ensure that he stays on the prescribed trajectory for the skill. At the end of the daily lesson, the teacher, peer coach, or learner charts the learner's best performance on her Daily per minute Standard Celeration Chart. Based on the learner's performance and her celeration, the teacher—following discussion with the peer coach—may recommend an alternate form of the current day's practice, recommend that the learner move on to the next practice sheet in the sequence, or suggest that a new practice sheet be created that includes a subset of items on the sheet on which the learner virtually always stumbled during the day's practice session.

The third phase of teaching is *application* and *generativity*. Application, strictly speaking, refers to the learner's ability to use a newly acquired skill in real world situations that are similar but not identical to those that were practiced. For example, a good application of see/say words in isolation on a practice sheet is correctly reading them on a bus schedule and a good application of hear/write numbers is to write correctly on one's hand one's friend's phone number. Too ensure these important characteristics of learning, Morningside's teachers provide explicit compound and composite tasks including simulations, games, and real-world applications to encourage generalization of what has been learned in the world of practice.

Generativity—also called contingency adduction—is different from application in that it is the recombination of previously acquired skills to solve a *novel* or unfamiliar problem (Epstein, 1991, 1993; Epstein, Kirshnit, Lanza, & Rubin, 1984). New environmental contingencies *recruit* behaviors learned under different contingencies to solve a novel problem. Morningside students have successfully solved many problems that were slated for explicit instruction without it. For example, students have (a) sounded out new words that are re-combinations of taught words; (b) solved fraction word problems by applying the algebraic equation procedures taught for whole number problems, and using fractions computation skills instead, (c) made a prediction at a certain point in reading a selection, after learning how to draw a conclusion, (d) identified an author's bias after learning how to identify an author's point of view, and (e) written sentences

with appositives, (i.e., The candidate, a surly and arrogant man, lost the election.) after learning how to modify nouns with adjectives in the standard way (i.e., a surly and arrogant candidate lost the election.).

Both types of skill extensions—application and generativity—are critical for learners to be efficient. They also account for what we described earlier: curriculum leaps. We noted this earlier when we said that once tool and component skills are learned to levels that promise the by-products of fluency, some learners are able—without further instruction or practice—to achieve frequencies on other skills in the curriculum on their first opportunity.

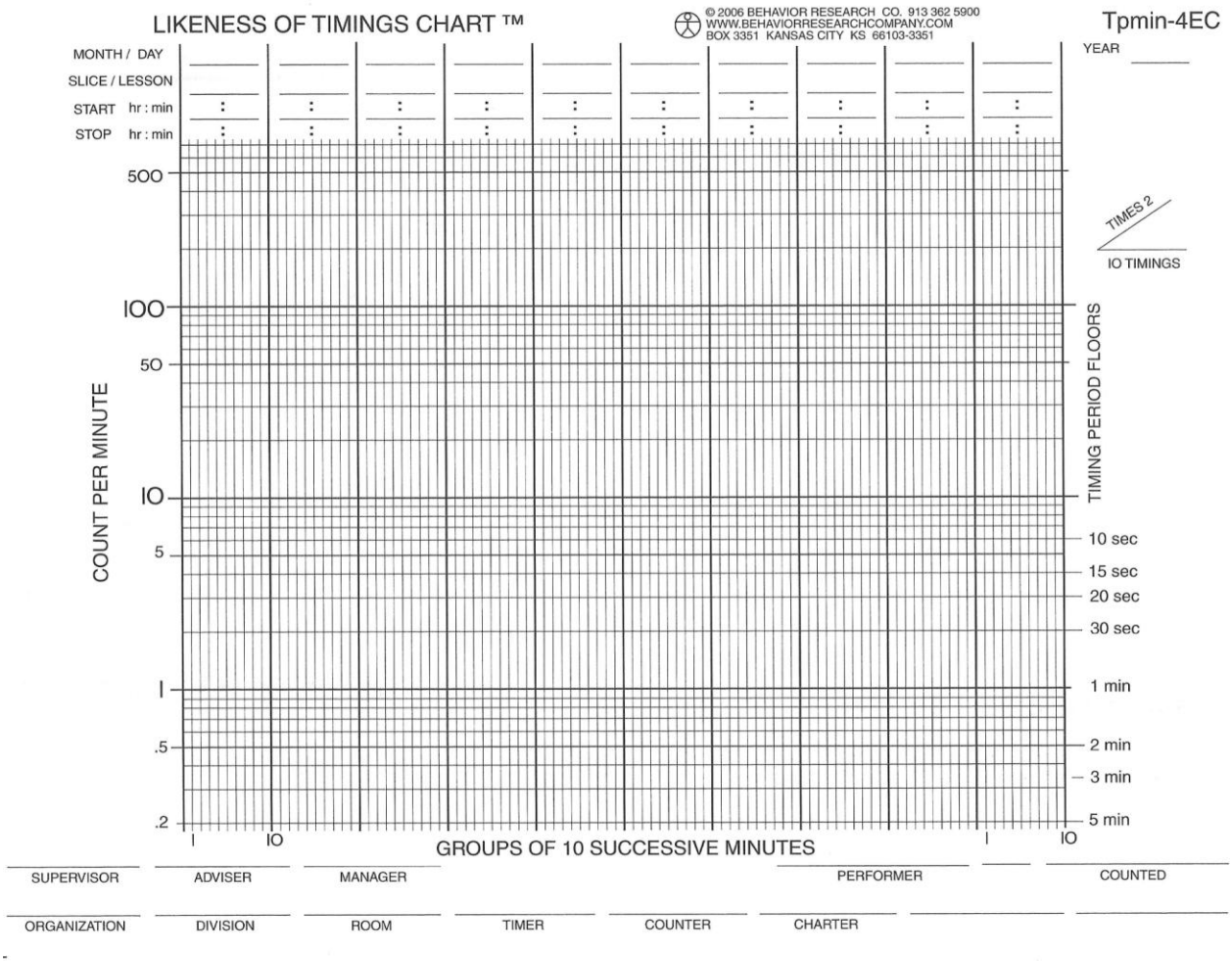


Figure 3. Likeness of a Timings Standard Celeration Chart. Standard Celeration Charts are available at Behavior Research Company, Box 3351, Kansas City, KS 66103-3351. VM 913-362-5900, www.behaviorresearchcompany.com

Results

Morningside has consistently produced results in learners who attend the program in Seattle, Washington that far exceed their historic performance. In fact, pre- to post-test scores on nationally standardized tests reveal average growth of two grade levels for each year of instruction in reading and mathematics for the past five years. Although the gains are not as great at schools which have contracted for services from Morningside Teachers' Academy, they too are impressive. For example, at an early implementation of the Morningside Model of Generative Instruction in a First Nation school—Ft. Frasier—in British Columbia, Canada, learners whose reading scores on the *Canadian Test of Basic Skills* (King-Shaw, 1995) at pretest were in the 20th percentile earned scores on end-of-year posttests at the 50th percentile within two years and above the 60th percentile by the end of the fourth year of implementation. In five years, students' percentile ranks in mathematics jumped from the 22nd percentile to the 74th percentile. During the five-year period, the school's ranking went from 13th in a district of 25 schools to second in math and fifth in reading.

Similar changes in growth trajectories are evident in the data (available on request) when Riverside Indian School in Anadarko, Oklahoma, the second largest of the Native American off-reservation boarding schools in the U.S., contracted with Morningside Teachers' Academy for assistance in reading.

Summary

Morningside Academy is nearing its 35th anniversary and, during that time, it has led the way through its combination of a variety of learner-verified curricula, its adoption and strengthening of practices that have been pioneered by others, and its creation of new programs. As we noted at the beginning of the article, Morningside Academy is best known as a Precision Teaching school. Precision Teaching has, over the years since Ogden Lindsley first conceived it, incorporated the findings of scientists in the fields of learning, instruction, and assessment. At Morningside, learners don't begin to chart data until they are at very close to 100% accuracy, a very high standard according to most other school systems. Morningside's teachers and leaders know that percent correct standards don't fare well in the face of evidence that supports building performance frequencies to levels that correlate with fluency and that makes all the difference.

Further, anecdotal evidence compiled over the more than 35 years of operation of Morningside Academy suggests that learners who achieve both accuracy and speed display confidence and competence not only about what they have learned, but also about how to learn new content. They recognize dysfluency in themselves and take their learning into their own hands to ameliorate the situation. Still, they and the Morningside faculty are indebted to those who have developed efficient and effective *DI* and *di* programs that ensure accuracy which is a necessary condition for achieving the frequencies which correlate with fluency. Morningside's team believes that these two parts of the work they do are in a symbiotic relation, each feeding on and being fed by the other. They also set the

stage for engineering application and generativity opportunities, which allow students to widely apply their learning in everyday circumstances and to figure out how to think and to do many things they did not learn in school—the signature of a smart, successful adult.

However, there is still more work to do. While our primary goal is to provide a service to our students, this has not kept Morningside's leadership team from posing questions for which answers derived from a rigorous program of basic research would allow its staff to further strengthen and perhaps even streamline procedures. We have described these questions elsewhere (Johnson & Street, 2004a, 2004b, 2012) and invite readers who conduct basic research to consider them as candidates for their own research agendas.

For those wishing to learn more about Direct Instruction, we recommend Engelmann and Carnine (1991) as well as Stein, Kinder, Silbert, and Carnine (2006) and Carnine, Silbert, Kame'enui, and Tarver (2009). For those wishing to learn more about Precision Teaching, we recommend White and Haring (1980), which—though dated—is the classic "how-to" book for teachers. In addition, Pennypacker et al. (2003) is the classic procedural handbook. Finally, Johnson and Street (2004b; 2013) provide details about the Morningside Model of Generative Instruction and the role Precision Teaching plays in creating its results.

References

Adams, G., & Engelmann, S. (1996). *Research on direct instruction: 25 years after DISTAR.* Seattle, WA: Educational Achievement Systems.

Andronis, P. T., Layng, T. V. J., & Goldiamond, I. (1997). Contingency adduction of symbolic aggression" by pigeons. *The Analysis of Verbal Behavior*, *14*, 5-17.

Archer, A., & Hughes, C. (2011). *Explicit instruction: Effective and efficient teaching.* New York, NY: Guilford Press.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction.* New York, NY: Guilford.

Benassi, V. A., Overson, C. E., & Hakala, C. M. (Eds) (2014). Applying science of learning in education: Infusing psychological science into the curriculum. Retrieved from http://teachpsych.org/ebooks/asle2014/index.php.

Binder, C. (1985). *The effects of explicit timing and performance duration on academic performance frequency in elementary school children.* Unpublished doctoral dissertation, Columbia Pacific University.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19,* 163-197.

Bloom B. S. (1956a). *Taxonomy of educational objectives, Handbook I: The cognitive domain.* New York, NY: David McKay.

Bloom, B. S. (1956b). *Taxonomy of educational objectives: The classification of educational goals.* New York, NY: Longman

Carnine, D. W., Silbert, J., Kame'enui, E., & Tarver, S. G. (2009). *Direct instruction: Reading* (5th Ed.). Upper Saddle River, NJ: Pearson.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232.

Deno, S. L. (1989). Curriculum-based measurement and special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 1–17). New York, NY: Guilford Press.

Desjardins, A. (1980). Letter from Ann Desjardins to Dr. Leslie Wiedenman describing techniques for frequency development of the "Big 6 Plus 6" self-help skills. Retrieved from http://www.fluency.org/Desjardins_Big6.pdf

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review, 3,* 357-370.

Dewey, J. (1976). The child and the curriculum. In J. A. Boydston (Ed)*, John Dewey: The middle works, 1899-1924* (Vol. 2: 1902-1903) (pp.271-291). Carbondale, IL and Edwardsville, IL: Southern Illinois University Press. (Original work published in 1902).

Dewey, J. (1981). Experience and nature. In J. A. Boydston (Ed), *John Dewey: The later works, 1925-1953* (Vol. 1: 1925) (pp.1-326). Carbondale, IL and Edwardsville, IL: Southern Illinois University Press. (Original work published 1925; rev ed., 1929).

Dewey, J. (1986). Logic: The theory of inquiry. In J. A. Boydston (Ed.), *John Dewey: The later works, 1925-1953* (Vol. 12: 1938) (pp.1-527). Carbondale, IL and Edwardsville, IL: Southern Illinois University Press. (Original work published 1938).

Engelmann, S., & Carnine, D. W. (1982). *Theory of instruction: Principles and applications*. Eugene, OR: ADI Press

Engelmann, S., & Carnine, D. W. (1991). *Theory of instruction: Principles and applications* (2nd Ed.). Eugene, OR: ADI Press

Engelmann, S., & Colvin, G. (2006). Rubric for identifying authentic direct instruction programs. Retrieved from http://www.zigsite.com/PDFs/rubric.pdf.

Epstein, R. (1991). Skinner, creativity and the problem of spontaneous behavior, *Psychological Science, 2,* 362-370.

Epstein, R. (1993). Generativity theory and education. *Educational Technology, 33,* 40-45.

Epstein, R., Kirshnit, R., Lanza, R., & Rubin, R. (1984). "Insight" in the pigeon: Antecedents and determinants of an intelligent performance. *Nature, 308,* 61-62.

Gagné, R. (1965). *The conditions of learning*. New York, NY: Holt, Rinehart & Winston.

Gilbert, T. (1962a). Mathetics: The technology of education. *Journal of Mathetics, 1,* 7-74.

Gilbert, T. (1962b). Mathetics II: The design of teaching exercises. *Journal of Mathetics, 1*, 7-56.

Gire, J., Testa, J., & Johnson, K. (2010). *Predicting performance on the Iowa Test of Basic Skills (ITBS) using the Scholastic Reading Inventory (SRI).* Presentation at the 36th annual Association for Behavior Analysis International Convention, San Antonio, TX. (Copies are available from Julian Gire at Morningside Academy, 901 Lenora Street, Seattle, WA 98121.)

Haughton, E. (1972). Aims: Growing and sharing. In J. B. Jordan, & L. S. Robbins (Eds.), *Let's try doing something else kind of thing* (pp. 20-39). Arlington, VA: Council for Exceptional Children.

Haughton, E. (1980) Practicing practices: Learning by activity. *Journal of Precision Teaching, 1*, 3-20.

Johnson, K., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist, 47*, 1475-1490.

Johnson, K., & Layng, T. V. J. (1994). The Morningside Model of Generative Instruction. In R. Gardner, D. Sainato, J. Cooper, T. Heron, W. Heward, J. Eshleman, & T. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 173-197). Belmont, CA: Brooks-Cole.

Johnson, K., & Layng, T.V. J. (1996). On terms and procedures: Fluency. *The Behavior Analyst, 19*, 281-288.

Johnson, K., & Melroe, K. (2014). *Teaching computation skills: A diagnostic prescriptive instructional sequence.* Seattle, WA: Morningside Press.

Johnson, K., Isbell, S., Delgado, M., & Leon, M. (2015). *Teaching standard word problem solving to fluency: An algebraic approach.* Seattle, WA: Morningside Press.

Johnson, K. J., & Street, E. M. (2004a). The Morningside Model of Generative Instruction: An integration of research-based practices. In D. J. Moran, & R. Malott, (Eds.). *Empirically supported educational methods,* pp. 247-265. St. Louis, MO: Elsevier Science/Academic Press

Johnson, K. J., & Street, E. M. (2004b). *The Morningside Model of Generative Instruction: What it means to leave no child behind.* Concord, MA: Cambridge Center for Behavioral Studies.

Johnson, K. J., & Street, E. M. (2012). From the laboratory to the field and back again: Morningside Academy's 32 years of improving students' academic performance. *The Behavior Analyst Today, 13*, 20-40.

Johnson, K., & Street, E. M. (2013). *Response to intervention and precision teaching: Creating synergy in the classroom.* New York, NY: Guilford Publications

Johnson, K., & Street, E. M. (2014). Precision teaching: The legacy of Ogden Lindsley. In F. K. McSweeney, & E. S. Murphy (Eds.) *The Wiley Blackwell handbook of operant and classical conditioning* (pp. 581-609). Hoboken, NJ: Wiley.

Johnston, J. A. (1996). Distinguishing between applied research and practice. *The Behavior Analyst, 19,* 35-47.

King-Shaw E. M. (1995). *Canadian test of basic skills.* [Test]. Don Mills, Ontario, Canada: Thomas Nelson.

Kulik, J. A. (2004). Grouping, tracking, and de-tracking: Conclusions from experimental, correlational, and ethnographic work. In H. J. Walberg, M. J. Reynolds, & M. C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 157-182). Greenwich, CT: Information Age.

Laraway, S., Snycerski, S., Michael, J., & Poling. A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36,* 407-414.

Lewkowicz, N. (1994). *Word workout.* Yellow Springs, OH: The Word Workshop.

Lindsley, O. R. (1992). Precision Teaching: Discoveries and effects. *Journal of Applied Behavior Analysis, 25,* 51-57.

Lindsley, O. R. (2001a). Celeration and agility for the 2000's [sic]. *Journal of Precision Teaching and Celeration, 17,* 107-111.

Lindsley, O. R., (2001b). Do times two, then go for four or more: Precision teaching aims for the 21st century. *Journal of Precision Teaching and Celeration, 17,* 99-102.

MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (2001). Prompts and prompt-fading strategies for people with autism. In C. Maurice, G. Green, & R. M. Foxx (Eds), *Making a difference: Behavioral intervention for autism* (pp. 37-50). Austin, TX: PRO-ED.

Malmquist, S. (2004). Using a multi-level system of assessment to inform instructional decisions and determine program effectiveness. In K. Johnson, & E. M, Street, *The Morningside Model of Generative Instruction: What it means to leave no child behind* (pp. 52-94). Concord, MA: Cambridge Center for Behavioral Studies.

Markle, S. M. (1990). *Designs for instructional designers.* Champaign, IL: Stipes.

Markle, S. M., & Tiemann, P. W. (1967). *Programming is a process.* [Film] Chicago: University of Illinois.

Mayer, G., Sulzer-Azaroff, B., & Wallace, M. (2012). *Behavior analysis for lasting change*. Cornwall-on-Hudson, NY: Sloan Publishing, LLC.

Meyer, S., Newsome, W. D., & Newsome, K. (2013). *A program-wide evaluation of learner agility and the impact on communication with parents and schools.* Symposium paper presented at the International Precision Teaching Conference, Saint Pete Beach, FL. (Copies are available from Dr. William D. Newsome, Fit Learning, 3953 South McCarran Blvd., Reno, NV. 89502).

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37,* 149-155.

National Council for Accreditation of Teacher Education (2008). *Professional standards for the accreditation of teacher preparation institutions.* Washington, D. C.: National Council for Accreditation of Teacher Education. Retrieved from the National Council for Accreditation of Teacher Education website http://www.ncate.org/Portals/0/documents/Standards/NCATE%20Standards%202008.pdf

Pennypacker, H. S., Gutierrez, A., & Lindsley, O. R. (2003). *Handbook of the Standard Celeration Chart, deluxe edition.* Concord, MA: Cambridge Center for Behavioral Studies.

Potts, L., Eshleman, J. W., & Cooper, J. O. (1993). Ogden R. Lindsley and the historical development of precision teaching, *The Behavior Analyst, 16,* 177-189.

Pryor, K. (1999). *Don't shoot the dog: The new art of teaching and training (rev. ed.).* New York, NY: Bantam Books.

Robbins, J. K. (2011). Problem solving, reasoning, and analytical thinking in a classroom environment. *The Behavior Analyst Today, 12,* 40-47.

Robbins, J. K. (2014). *Learn to reason with TAPS: A talk aloud problem solving approach.* Seattle, WA: Robbins/Layng & Associates.

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis.* New York, NY: Appleton-Century-Crofts.

Shinn, M. R. (Ed.; 1989). *Curriculum-based measurement: Assessing special children.* New York, NY: Guilford Press.

Stein, M., Kinder, D., Silbert, J., & Carnine, D. W. (2006). *Designing effective mathematics instruction: A direct instruction approach* (4th Ed.). Upper Saddle River, NJ: Pearson.

Terrace, H. S. (1963). Discrimination learning with and without "errors." *Journal of the Experimental Analysis of Behavior*, *6*, 1–27.

Tiemann, P. W., & Markle, S. M. (1978). *Analyzing instructional content: A guide to instruction and evaluation.* Champaign, IL: Stipes.

Tiemann, P. W., & Markle, S. M. (1990) *Analyzing instructional content: A guide to instruction and evaluation* (4th ed). Champaign, IL: Stipes. Currently available from Seattle, WA: Morningside Press.

Touchette, P. E., & Howard, J. S. (1984). Errorless learning, reinforcement contingencies, and stimulus control. *Journal of Applied Behavior Analysis, 17,* 175-188.

Vargas, J. (2013). *Behavior analysis for effective teaching*. New York: Routledge.

Wahlberg, H. J., Reynolds, M. J., & Wang, M. C. (Eds., 2004), *Can unlike students learn together: Grade retention, tracking, and grouping*. Greenwich, CT: Information Age.

Whimbey, A. (1975). *Intelligence can be taught.* New York, NY: E.P. Dutton & Co.

Whimbey, A., & Lochhead, J. (1991). *Problem solving and comprehension.* Hillsdale, NJ: Lawrence Erlbaum.

White, O. R., & Haring, N. G. (1980). *Exceptional teaching* (2nd Ed.). Columbus, OH: Merrill.

ACTA DE INVESTIGACIÓN PSICOLÓGICA, 2014, 4 (3), 1793 - 1808

Prácticas de Crianza Positiva: Entrenamiento a padres para reducir Bullying

Brenda Mendoza González[1], Francisco Javier Pedroza Cabrera* & Kalina Isela Martínez Martínez*
Universidad Autónoma del Estado de México, *Universidad Autónoma de Aguascalientes

Resumen

El estudio buscó evaluar la eficacia de un programa de prácticas de crianza positiva dirigido a padres para reducir el bullying y aumentar la conducta pro-social de sus hijos. Participaron ocho parejas y dos padres solteros de 10 niños identificados como bullies. La mitad de los padres conformó el grupo control y se entrenó a la otra mitad para identificar la conducta agresiva y pro-social de sus hijos, sus antecedentes y consecuentes. Durante ocho sesiones semanales se entrenó a los padres a establecer límites, reforzar la conducta pro-social y las alternas a las agresivas, a sobre-corregir, desaprobar levemente, castigar y extinguir la conducta agresiva. Padres y maestros registraron durante tres semanas de línea base y ocho de tratamiento, la frecuencia de emisión de conductas pro-sociales (hacer la tarea, ayudar con tareas domésticas y recoger sus juguetes) y de conducta agresiva física y verbal. Los resultados mostraron una reducción significativa de la conducta agresiva y un aumento de la pro-social respecto tanto a la línea base como a la frecuencia de emisión de esas conductas por los niños del grupo control. Notablemente, la conducta también cambió en la escuela. Se discute la efectividad de las intervenciones con padres para reducir la conducta de bullying en diferentes contextos.

Palabras Clave: : bullying, prácticas de crianza, modificación conductual, intervención.

Positive Child Rearing Practices: Parents training for reduce bullying

Abstract

The study aimed at assessing the effectiveness of a positive child rearing program with parents for reducing bullying and incrementing pro-social behavior of their children. Participants were eight couples and two single parents of 10 children identified as bullies. Half of the parents were assigned to a control group and the other half were trained to identify aggressive and pro-social behaviors of their children, as well as their antecedents and consequences. During eight weekly sessions parents were trained to set limits, reinforce both pro-social behavior and alternative responses to the aggressive ones, to correct, slightly disapprove, punish and extinguish aggressive behavior. Frequency of emission of specific pro-social behaviors (doing homework, helping in domestic chores and picking up toys) and of physical and verbal aggression was registered by parents and teachers during three weeks of base line and during eight weeks of treatment. Results showed a significant reduction of aggressive behavior and an increase of pro-social behavior compared both to base line and to the frequency of the same behaviors by children of the control group. Notably, behavior also changed at school. Results are discussed regarding the usefulness of interventions with parents in reducing bullying behavior by their children in different contexts

Keywords: bullying, child rearing practices, behavior modification, intervention.

Original recibido / Original received: 22/05/2014 Aceptado / Accepted: 25/09/2014

[1] Proyecto DRODEP 752905 DSA/103.5/14/7529

Las prácticas de crianza han sido un tópico ampliamente estudiado tanto en el ámbito internacional (Forehand & Long, 2010; Patterson, DeGarmo, & Forgatch, 2004) como en el nacional (Andrade & Betancourt, 2012; Ayala et al., 2001; González & Landero, 2012; Pedroza, Aguilera, Cervantes, & Martínez, 2013). El principal objetivo de los estudios ha sido la reducción del comportamiento agresivo y de la desobediencia por niños. Muchos han sido los esfuerzos para identificar la relación entre las prácticas de crianza y el desarrollo de conducta agresiva y antisocial y se ha encontrado que las prácticas de crianza positiva son un factor que reduce el riesgo del desarrollo de conducta delictiva (Patterson & Stouthamer-Loeber, 1984; Pedroza & Martínez, 2011).

Resultados de estudios en México han mostrado que uno de los factores de riesgo en el desarrollo de comportamiento agresivo en edades tempranas, es la falta de habilidades de los padres para educar a sus hijos (Ayala, Pedroza, Morales, Chaparro, & Barragán, 2002). Dicha falta de habilidades se relaciona con implementar una disciplina inconsistente, explosiva, y agresiva, que favorece el comportamiento agresivos de los niños (Patterson, 1974). Patterson llamó a esto teoría de la coerción y propuso la necesidad de implementar programas para la reducción de problemas de conducta agresiva de niños a partir de la modificación de estrategias disciplinarias de los padres (Patterson et al., 2004; Pedroza, 2006). Particularmente, con respecto al estudio del comportamiento agresivo en niños, actualmente existe un interés por el estudio del bullying o acoso escolar. Éste se caracteriza por el desbalance de poder entre el agresor y la víctima, la repetición de la agresión y su direccionalidad (Cervantes & Pedroza, 2012; Pedroza, Aguilera, et al., 2013) que ocurre únicamente en el escenario escolar y entre pares (Monks & Smith, 2006).

Se han desarrollado diferentes estrategias de prevención e intervención para la reducción de la agresión en escenarios escolares (Karna et al., 2013; Salmivalli, Poskiparta, Athola, & Hataja, 2013; Olweus & Limber, 2010). Dichos programas han estado dirigidos a los profesores y los estudiantes, con el propósito de disminuir la frecuencia del comportamiento agresivo de los niños clasificados como agresores. También se ha buscado mejorar las relaciones entre pares a partir del establecimiento de límites, el uso consistente de consecuencias no hostiles ante el comportamiento problemático, así como incrementar el interés de los profesores hacia los niños desarrollando un ambiente de cordialidad.

En algunas investigaciones sobre el acoso escolar se ha establecido la relación entre el estilo de crianza que utilizan los padres y la tendencia de un niño a desempeñar un rol de víctima, de acosador, de espectador o incluso un doble rol de víctima y acosador (Cervantes & Pedroza, 2012). Se ha encontrado que la sobreprotección se encuentra estrechamente relacionada con el desarrollo del perfil de víctima (Mendoza, 2014; Olweus, 1993). La supervisión infrecuente de los niños (Olweus, 1993), las prácticas de crianza autoritarias, negligentes, o punitivas (Mendoza, 2014; Patterson, DeBaryshe, & Ramsey, 1989) se relacionan con el desarrollo del rol de acosador.

En algunas investigaciones se ha documentado que los padres de niños con problemas de conducta carecen de habilidades para fomentar la conducta social y para controlar adecuadamente la conducta problemática de sus hijos, por

lo que emplean técnicas hostiles para tratar de eliminar esos comportamientos (Farrington, 2004; Patterson et al., 2011; Pedroza & Martínez, 2011). Por ejemplo, se ha documentado que los padres de niños que están recibiendo tratamiento por presentar conducta problemática no refuerzan el comportamiento adecuado de sus hijos y además emplean el castigo físico (Farrington, 2004). La tasa de interacciones coercitivas entre padres e hijos es tres veces mayor entre muestras clínicas que entre familias de muestras no clínicas (Patterson, DeGarmo, & Forgatch, 2004; Pedroza & Martínez, 2011).

Se ha mostrado que los programas de entrenamiento para padres son más efectivos en la reducción de las conductas problema de los niños, en comparación con otros programas en los que no participan los padres (Barkley, 1997). Los niños tratados sin que sus padres reciban intervención muestran un cambio en escenarios clínicos, sin embargo, una vez que el niño regresa a su ambiente familiar se encuentra bajo el control de estímulos que propician la ocurrencia del comportamiento agresivo (Webster-Stratton & Hancock, 1998).

En México Pedroza (2006) desarrolló un programa de entrenamiento conductual para padres de niños clasificados como agresivos, cuyo propósito fue modificar las prácticas de los padres que mantienen el comportamiento agresivo de sus hijos. El programa incluyó la identificación y el análisis funcional del comportamiento agresivo, el reforzamiento del comportamiento pro-social, el empleo del castigo y extinción del comportamiento agresivo, la interacción social positiva y la supervisión de las actividades del niño. Pedroza encontró que la conducta agresiva de niños y de padres disminuyó en un 75%.

En suma, se ha demostrado que existe una relación entre ciertas prácticas de crianza inadecuadas con las conductas agresivas de los niños y que implementar programas de intervención con los padres para que aprendan a responder adecuadamente ante la conducta agresiva de sus hijos ha sido efectivo (Forehand & Long, 2010; González & Landero, 2012). Si se considera que los padres son quienes pasan la mayor parte del tiempo con los niños y son los modelos de aprendizaje de comportamientos adecuados e inadecuados, resulta importante el desarrollo de programas de intervención para padres con el propósito de modificar sus hábitos de crianza y reducir en sus hijos los episodios de agresión en casa. También es importante evaluar si las intervenciones con padres no sólo disminuyen la conducta agresiva en el contexto familiar, sino si se generaliza al contexto escolar, disminuyendo el bullying. Algunos investigadores han sugerido que enseñar a los padres una disciplina de control no coercitiva del comportamiento de sus hijos y fomentar el comportamiento pro-social, permite que el comportamiento pro-social del niño se transfiera a otros escenarios como el escolar (Orpinas & Horne, 2006). Es posible que lo mismo sea cierto para la conducta agresiva. Por lo tanto, el propósito del presente trabajo fue evaluar la efectividad de un programa de entrenamiento para padres en prácticas de crianza positiva (Pedroza, Mendoza & Martínez, 2013), para reducir episodios de agresión en el contexto familiar e incrementar la conducta pro-social. También se buscó averiguar si la reducción de la conducta agresiva y el aumento de la conducta pro-social se generalizaría al contexto escolar.

Método

Participantes

Participaron 18 padres de familia, nueve parejas y dos padres solteros (una madre y un padre), con un promedio de edad de 36 años. También participaron sus hijos (ocho niños y dos niñas) de entre seis y once años, todos estudiantes de educación básica, cada uno de ellos inscritos en una de las diez escuelas de educación primaria participantes. El alumnado participante fue referido por los directivos de las escuelas participantes por exhibir comportamiento agresivo hacia sus pares en el contexto escolar (bullying) a un tratamiento especializado.

Criterios de inclusión y exclusión a la muestra.

El primer criterio de inclusión fue el empleado por Santoyo (2007). De acuerdo con este criterio se pide al maestro del grupo, al psicólogo escolar, al maestro de apoyo de la institución y al directivo identificar a los niños que muestran conductas agresivas. Para considerar a un niño como agresivo por lo menos tres de esas personas deben identificarlo como tal. El segundo criterio de inclusión fue que sus padres estuvieran dispuestos a participar en un programa de entrenamiento en prácticas de crianza positiva. Se excluyó a niños que aún cuando tuvieran tres nominaciones por parte del personal escolar tuvieran un diagnóstico o estuvieran en tratamiento psiquiátrico.

Diseño de la investigación

Se empleó un diseño combinado de grupo control y línea base múltiple entre conductas (Kazdin, 2000). Este diseño permite medir el impacto del programa al comparar el comportamiento observado en la línea base con el comportamiento observado en la fase de tratamiento. Además se puede comparar el resultado con un grupo de control.

Instrumentos

Contexto Familiar

Evaluación Funcional. Esta evaluación consistió en un registro diario de la frecuencia de ocurrencia de la conducta agresiva y pro-social (Pedroza, Mendoza, et al., 2013). La hoja de registro permite identificar el antecedente de cada conducta, la conducta blanco y sus consecuencias y consecuentemente se le denomina registro antecedente, conducta, consecuente (ACC). La tarea del respondiente (i.e., los padres en el presente estudio) es escribir una descripción de cada conducta agresiva y pro-social y su frecuencia de ocurrencia. Además debe señalar a qué persona en el ámbito familiar fue dirigida dicha conducta. Los padres debían así mismo describir los eventos antecedentes y consecuentes del comportamiento registrado. El comportamiento agresivo se clasificó en físico (e.g., golpear, empujar, morder, aventar objetos a otra persona), verbal (e.g., burlas, apodos, insultos, hablar con malos modales, denigrar a otros) y antisocial (e.g., robar o dañar sus pertenencias o las de otros). Los padres también registraron el

comportamiento de berrinche. La conducta pro-social se definió como el que el niño participara en actividades familiares (e.g., ayudar en la colocación en la mesa de los utensilios necesarios para comer y llevar los alimentos, recoger los juguetes inmediatamente después de usarlos y realizar las tareas escolares antes de realizar cualquier actividad lúdica).

Contexto Escolar

Los profesores registraron diariamente la conducta agresiva que el niño dirigía a los compañeros de clase o a las autoridades educativas (profesor u otros adultos). El registro permitió identificar el número de conductas agresivas que cada niño participante en el estudio emitió y dirigió a sus pares o autoridades. El comportamiento agresivo en el contexto escolar se clasificó de igual manera que en el contexto familiar, como físico y verbal. Además, los profesores también registraron el comportamiento sexual (e.g., tocar genitales de otros, intimidar con insultos con carácter sexual) y la exclusión (e.g., ignorar al otro e impedir a otros participar en la actividad académica).

Procedimiento

Línea Base

Los padres emplearon el registro ACC para anotar las conductas pro-sociales y de comportamiento agresivo, así como sus antecedentes y consecuentes de forma cotidiana durante tres semanas continuas. Durante el mismo periodo de tiempo, los profesores de cada niño registraron diariamente el número de veces, así como el tipo de conducta de agresión que el niño emitió hacia sus pares o hacia adultos. Los maestros enviaron diariamente a los padres dicho registro.

Tratamiento

Se conformaron dos grupos, un control (n = 5) y otro experimental (n = 5). Se asignó aleatoriamente a los padres de los niños a cada uno de los grupos. Se asignó a cada grupo a uno de los dos padres solteros que participaron en el estudio. El grupo control no recibió ningún tipo de tratamiento, únicamente se les solicitó que registran las conductas pro-sociales y agresivas de sus hijos y que las entregaran a los investigadores una vez por semana. El grupo experimental participó en un programa de entrenamiento a padres en prácticas de crianza positiva. El programa duró ocho sesiones semanales de 60 minutos cada una.

El programa de entrenamiento a padres de prácticas de crianza positiva (Pedroza, Mendoza, et al., 2013), comprende varios componentes de terapia conductual como el análisis funcional del comportamiento, el establecimiento de límites, el reforzamiento de la conducta pro-social, el reforzamiento de respuestas alternas a la conducta agresiva, la sobre-corrección, la desaprobación leve, el castigo y la extinción de la conducta agresiva. Todas las sesiones del programa tuvieron la siguiente estructura: a) recordatorio de la sesión anterior y análisis de los registros ACC realizados en la semana, b) análisis de la tarea, retroalimentación y reforzamiento de la ejecución, c) exposición de los contenidos

programados, d) ejercicios prácticos de los contenidos, e) resumen de lo visto en la sesión, y f) asignación de tareas para casa.

El programa se dividió en tres secciones: 1) Disciplina efectiva y selectiva: el terapeuta expuso a los padres el plan de trabajo y se les dio a conocer los contenidos del programa. Posteriormente, se definió el comportamiento agresivo y se resaltó el papel que juegan los padres en la gestación, mantenimiento y desarrollo de las conductas agresivas de sus hijos. En este componente se revisó el papel de los reforzadores en el mantenimiento de la conducta. Se explicó a los padres cómo refuerzan accidentalmente el comportamiento agresivo de sus hijos y cómo ignoran los comportamientos positivos de éstos. Se expusieron los tipos de reforzadores y su empleo principalmente a través de tablas de contingencia y economía de fichas. El terapeuta y el co-terapeuta modelaron la forma de reforzar las conductas de los niños y se dieron ejemplos de comportamientos que deben reforzar. También, los padres aprendieron el empleo de estrategias de castigo no aversivo, aunque se les hizo hincapié que antes de aplicar cualquier estrategia de castigo se deben emplear procedimientos de reforzamiento de conductas incompatibles o funcionalmente equivalentes al comportamiento agresivo pero de mayor aceptación social. Se enseñaron estrategias de castigo no aversivo, como la desaprobación social leve, tiempo fuera, costo de respuesta y sobre-corrección. El terapeuta junto con el co-terapeuta modelaron cada una de las estrategias y el padre realizó ensayos conductuales. Se enseñó a los padres a establecer reglas para guiar el comportamiento de sus hijos en casa y se les enseñó a realizar contratos conductuales con sus hijos. 2) Establecimiento de instrucciones e interacción social: El objetivo de este componente fue que los padres emitieran instrucciones claras y precisas y que interactuaran con sus hijos con calidad y calidez. Los padres aprendieron a dar instrucciones claras y precisas a sus hijos a partir del moldeamiento del comportamiento. Con respecto a la interacción, se indicó a los padres las conductas agresivas y no agresivas que ocurrieron durante la interacción con su hijo; el terapeuta modeló cada uno de los comportamientos esperados por el padre en una interacción cálida (para una descripción de los componentes véase González, Vargas, Galván, & Ayala, 1998). 3) Supervisión de las actividades del niño: Este componente tuvo como objetivo que los padres aprendieran a supervisar las actividades que sus hijos realizaban en casa y fuera de ella. El terapeuta identificó las situaciones que los padres monitoreaban con menor frecuencia, se hizo un recordatorio de la elaboración de los contratos conductuales y se expusieron las habilidades necesarias para la solución de problemas en la interacción social. Los padres aprendieron la importancia de monitorear las actividades del niño y se enseñó a realizar el registro de supervisión en la escuela. Los padres elaboraron un contrato conductual con el niño para llevar al cabo el monitoreo de actividades en la escuela y le indicaron que día con día entregarían a sus maestros una tarjeta en donde éstos evaluarían su comportamiento en el salón de clase y en el patio de recreo.

Durante el entrenamiento se capacitó a los padres para lograr que sus hijos realizaran ciertas conductas específicas. Se les entrenó a dar acceso al niño a realizar alguna actividad lúdica después de que realizaran sus tareas escolares, reforzando así la conducta de hacer la tarea antes de jugar. Los padres también

fueron entrenados a enseñarle a sus niños a participar en tareas domésticas como “poner la mesa” para comer, solicitando la participación de los niños por lo menos una vez al día. Si los niños obedecieron, los padres los elogiaron y les dieron una ficha por participación, como parte de un programa de economía de fichas, que los niños podían cambiar semanalmente por acceso a alguna actividad. Los padres también fueron entrenados a establecer reglas en su casa, como que el niño debía recoger sus juguetes después de jugar con ellos. El entrenamiento para reducir la conducta agresiva consistió en enseñar a los padres a usar técnicas de castigo no aversivas como la de sobre-corrección y el costo de respuesta, al tiempo que le enseñaron al niño la ejecución de la respuesta alterna a la agresión. Se analizó con los padres la función de la conducta de agresión de los niños, identificando los reforzadores positivos y negativos que la mantenían, por lo que se aplicaron técnicas aversivas que permitieran la reducción del comportamiento agresivo y que se reforzará las conductas alternas a la agresión. En el programa de entrenamiento los padres recibieron capacitación para identificar un episodio de berrinche, extinguirlo y enseñar una conducta que tuviera la misma función que la conducta de berrinche (e.g., comunicar asertivamente).

Resultados

En la Figura 1 se muestra el promedio semanal de la conducta de hacer la tarea antes de jugar o de realizar cualquier otra actividad de ocio de los niños del grupo experimental y del control durante la línea base y durante el tratamiento. El análisis visual de los datos muestra que el grupo experimental incrementó la conducta de hacer la tarea antes de jugar de una vez durante la línea base a alrededor de cinco ocasiones durante el tratamiento. En cambio, la frecuencia de esa conducta de los niños del grupo control se mantuvo con una frecuencia menor a uno a lo largo de las doce semanas de registro.

En la Figura 2 se muestra la frecuencia semanal de la conducta pro-social de ayudar a “poner la mesa” de los niños del grupo experimental y del control durante la línea base y el tratamiento. La frecuencia de la conducta de ayudar a “poner la mesa” aumentó de cero a cinco veces por semana para los niños del grupo experimental, mientras que se mantuvo cercana a cero para los niños del grupo control.

En la Figura 3 se muestra la frecuencia promedio semanal de la respuesta de los niños del grupo experimental y del control durante la línea base y durante el tratamiento al establecimiento de la regla familiar de recoger sus juguetes después de jugar con ellos. La inspección visual de la figura muestra que la frecuencia de la conducta aumentó de menos de una vez a la semana durante la línea base a cerca de seis veces durante el tratamiento para los niños del grupo experimental. En cambio, esta conducta permaneció con una emisión menor a uno durante las 12 semanas de registro para los niños del grupo control.

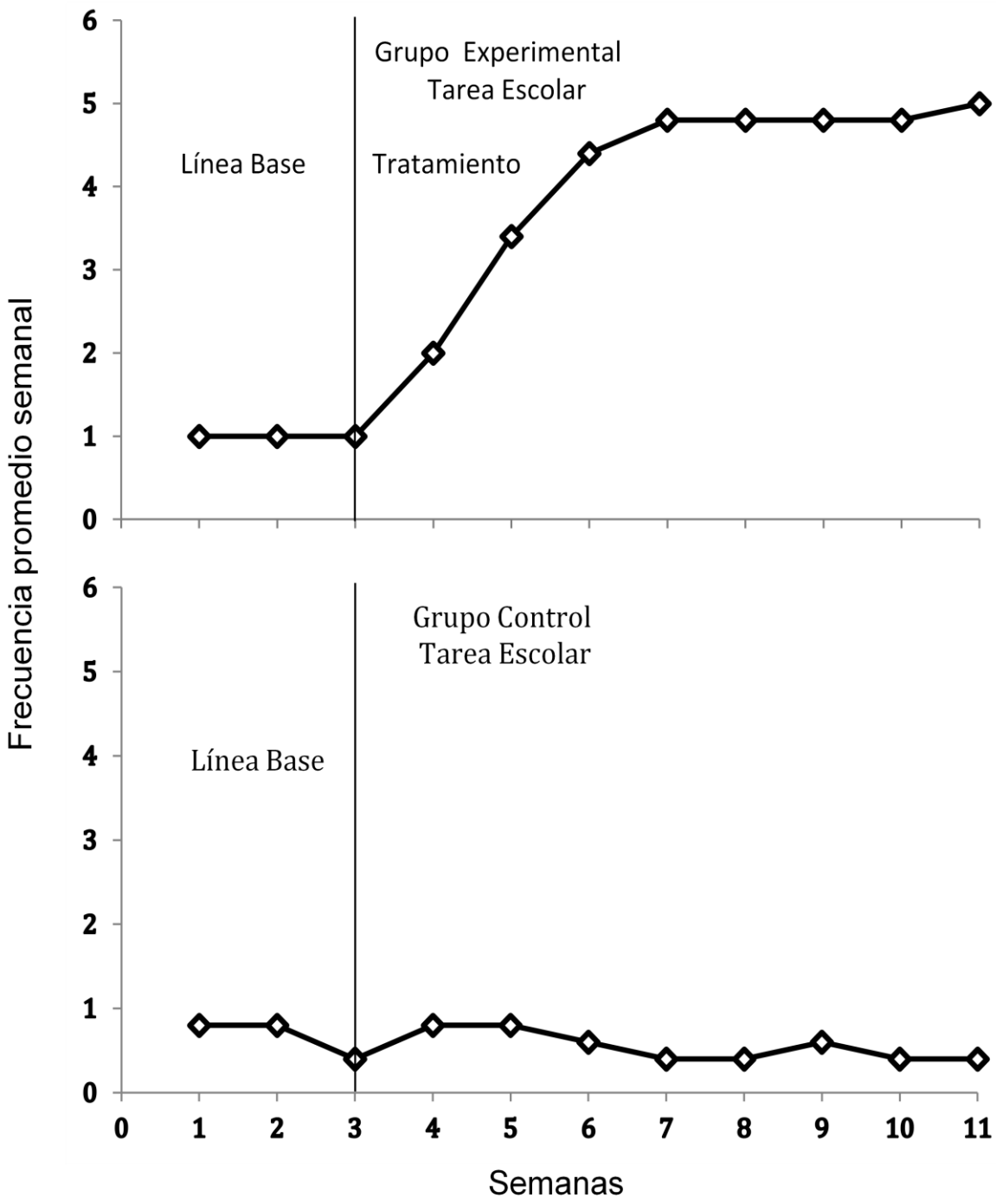


Figura 1. Promedio semanal de la frecuencia de la conducta de hacer la tarea antes de jugar. El panel superior muestra los datos correspondientes a los niños del grupo experimental y el inferior a los del grupo control.

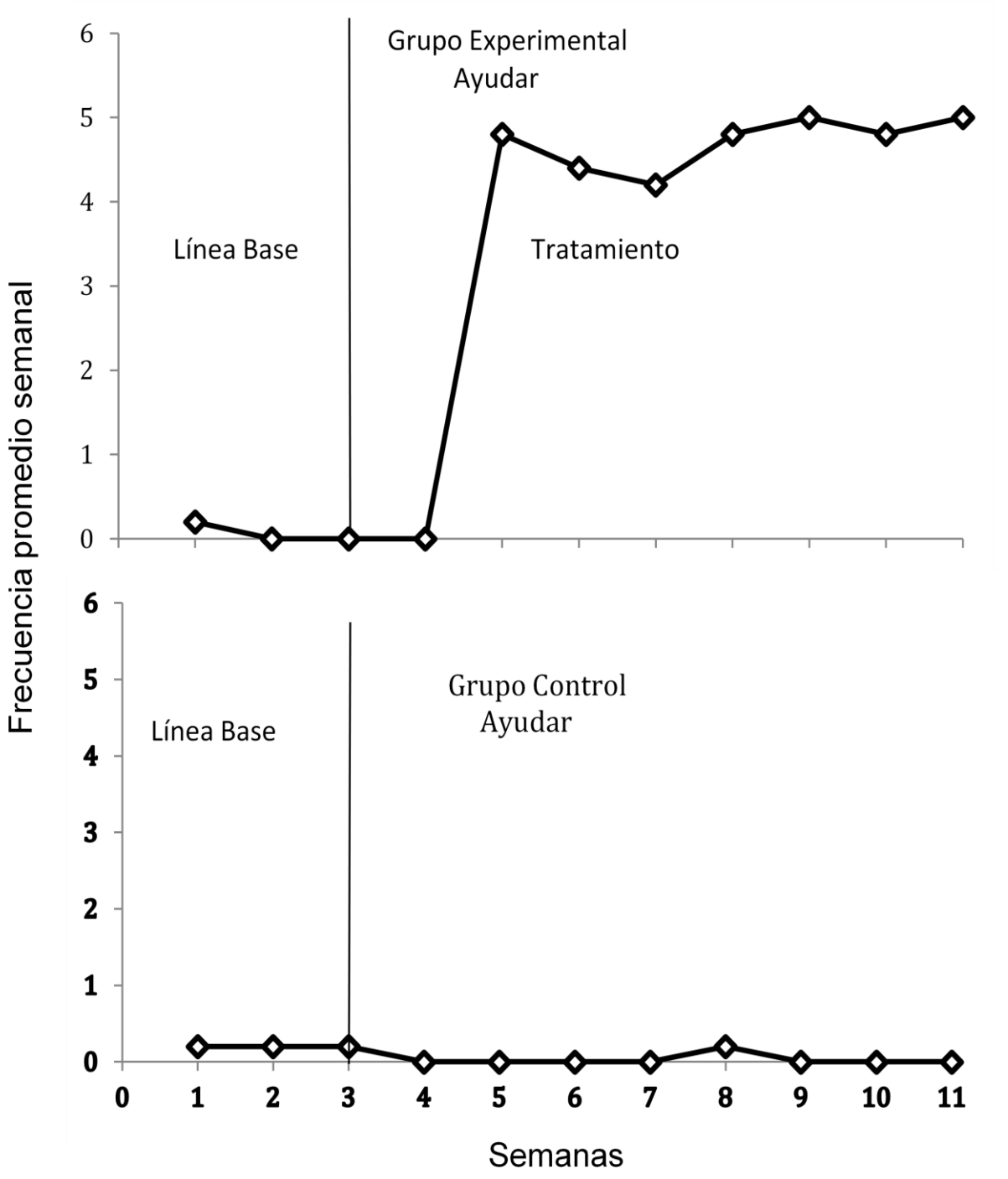


Figura 2. Promedio semanal de la frecuencia de la conducta de ayudar a “poner la mesa” emitida por los niños del grupo experimental (panel superior) y control (panel inferior).

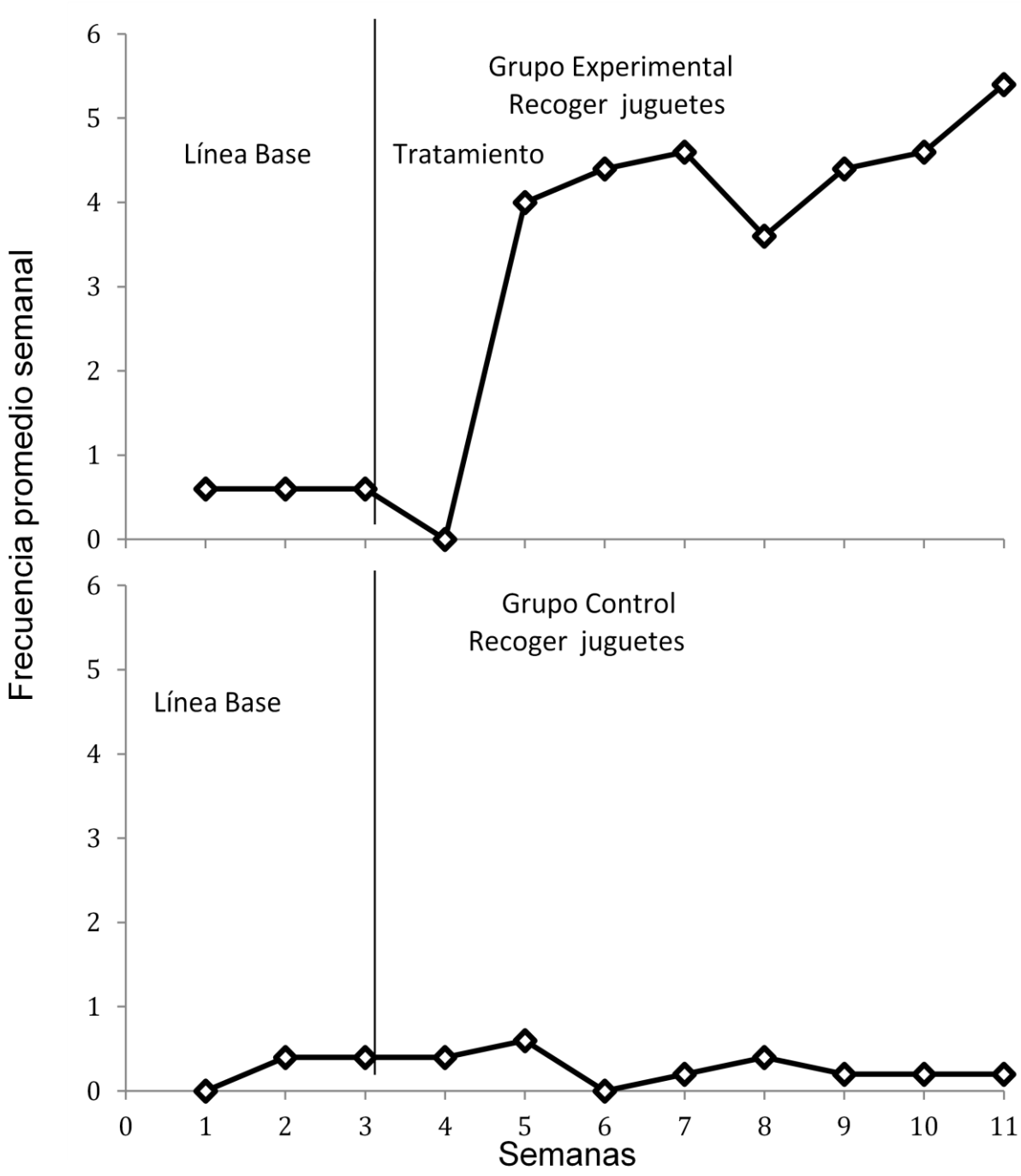


Figura 3. Promedio semanal de la frecuencia de la conducta de recoger juguetes de los niños del grupo experimental (gráfica superior) y control (gráfica inferior).

En la Figura 4 se muestra la frecuencia de emisión de conducta agresiva hacia los compañeros de la escuela. Se puede observar que el entrenamiento a padres tuvo como consecuencia la reducción de la conducta agresiva en la

escuela de los niños del grupo experimental de alrededor de cinco veces por semana durante la línea base a cero después de entrenar a sus padres. La frecuencia de la conducta agresiva de los hijos de los padres que no participaron en el programa permaneció en alrededor de cinco veces por semana durante las 12 semanas de registro.

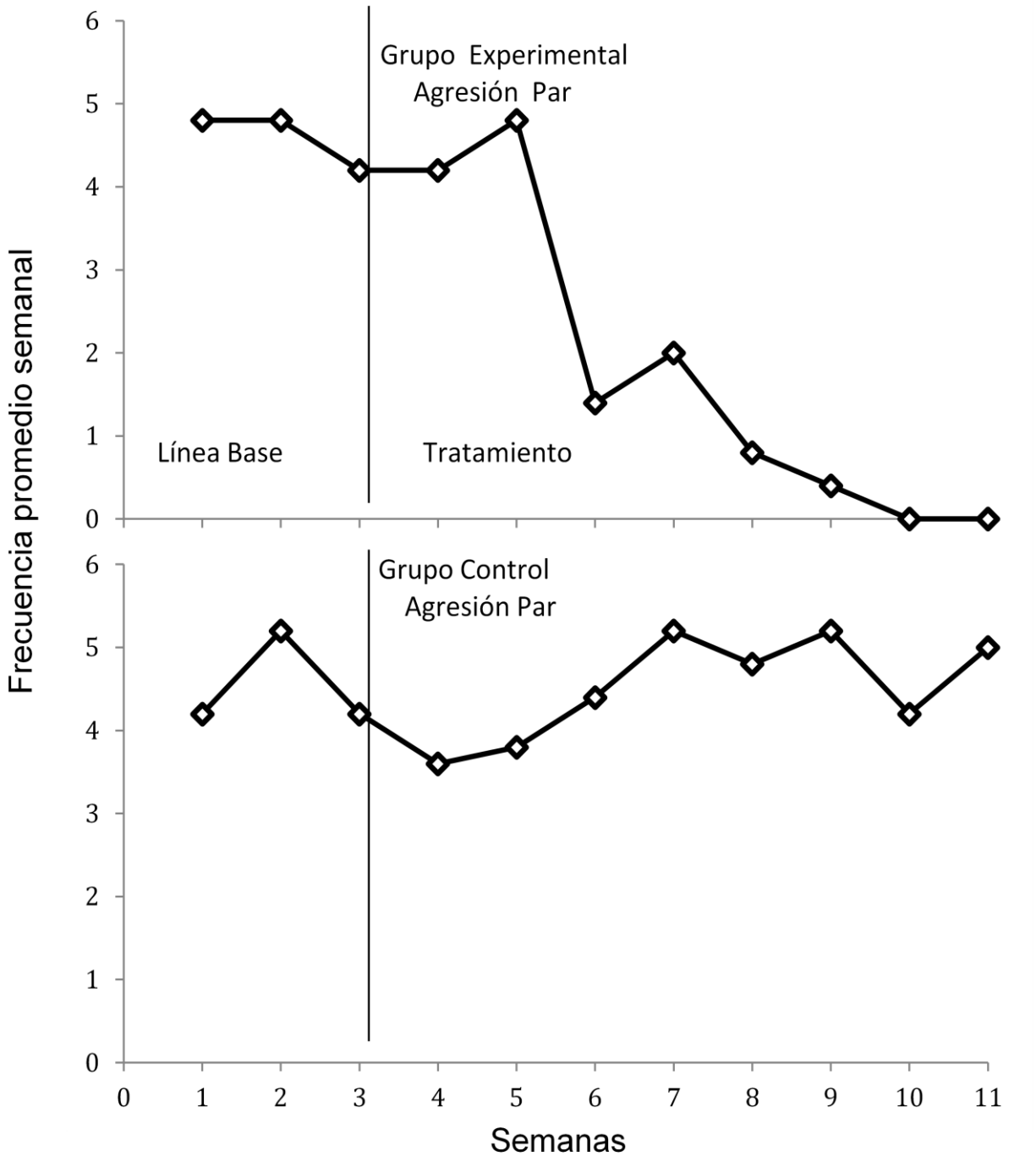


Figura 4. Promedio semanal de la frecuencia de las conductas agresivas que los niños del grupo experimental y del control dirigieron hacia un par específico.

En la Figura 5 y se muestra la frecuencia promedio semanal de la conducta agresiva que los niños del grupo experimental y del grupo control dirigida a algún miembro de la familia o a algún adulto en la escuela. Los datos de la figura muestran que la conducta agresiva dirigida a adultos disminuyó de cuatro veces por semana a cero para los niños del grupo experimental. Esa misma conducta se mantuvo en un promedio de alrededor de tres veces por semana en el caso de los niños del grupo control.

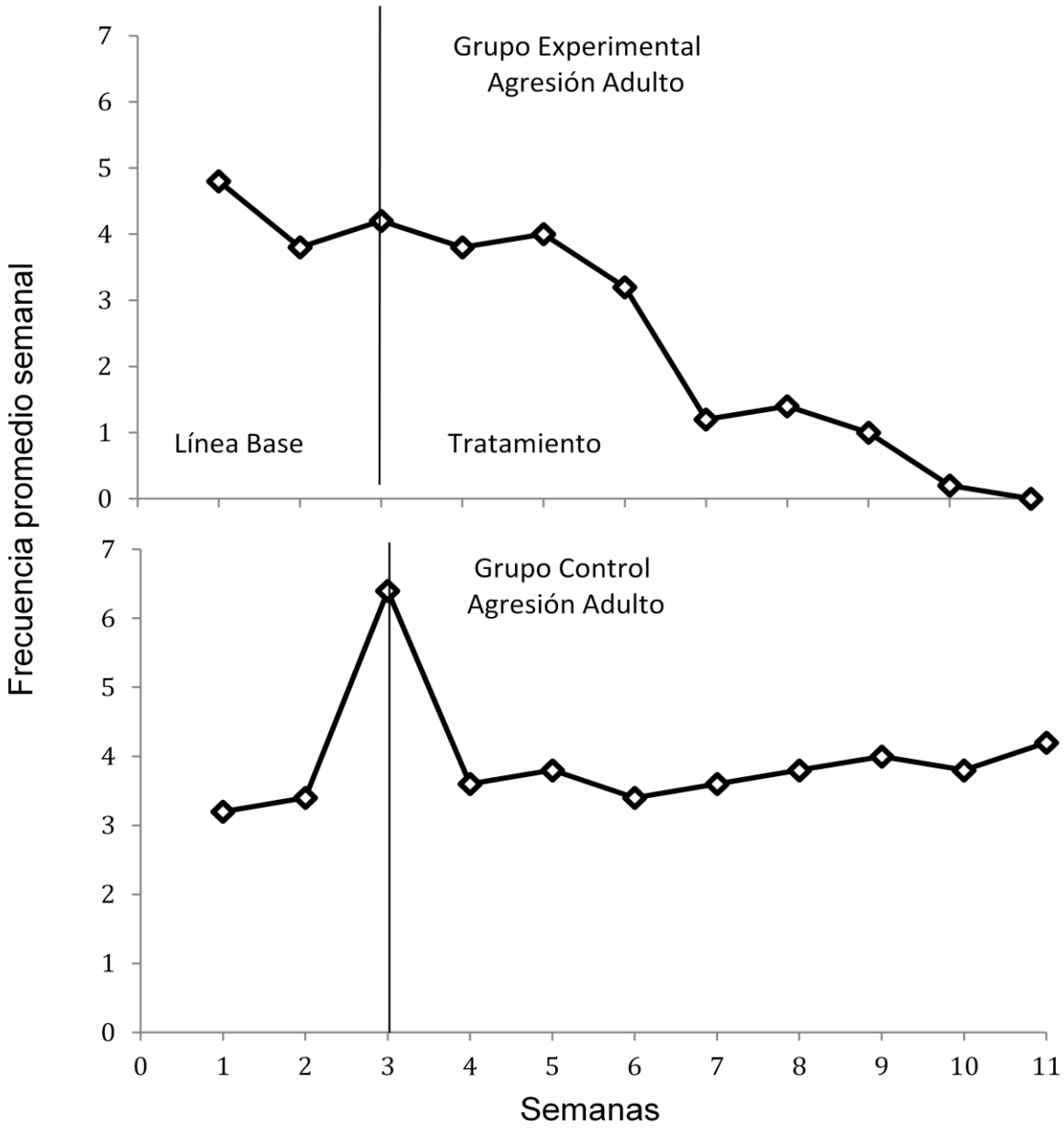


Figura 5. Promedio semanal de la frecuencia de conductas agresivas que los niños dirigieron hacia adultos en el contexto familiar y escolar. En el panel superior se muestran los datos para los niños del grupo experimental y en el inferior para los del grupo control.

En la Figura 6 se muestra el efecto del programa de entrenamiento para la disminución de los episodios de berrinche. Los niños de los padres que recibieron tratamiento disminuyeron sus berrinches de alrededor de siete semanales a uno. Los berrinches de los niños cuyos padres no recibieron tratamiento se mantuvieron en alrededor de seis por semana.

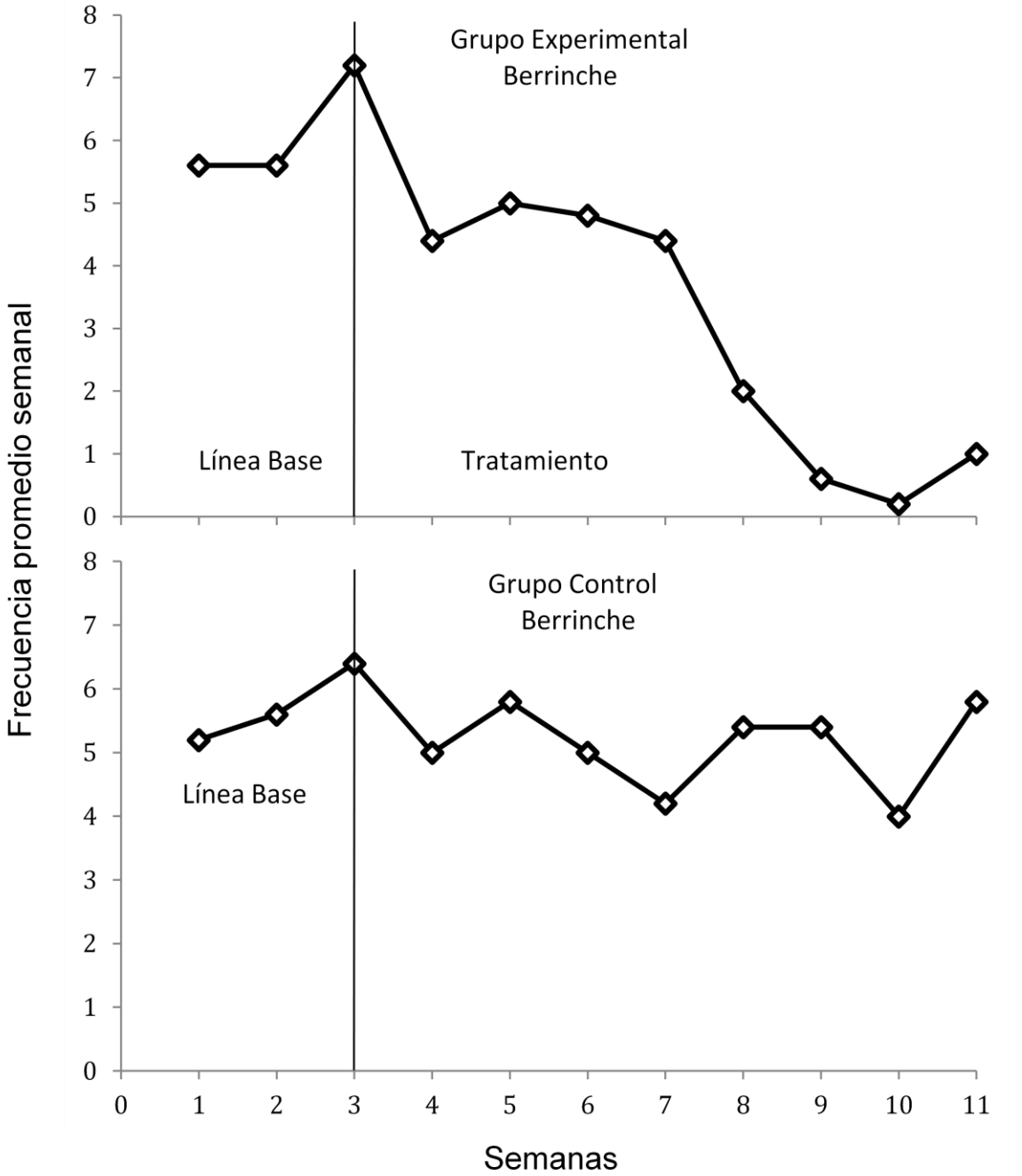


Figura 6. Promedio semanal de la frecuencia de conducta de hacer berrinches que los niños del grupo experimental (panel superior) control (panel inferior) exhibieron en el contexto familiar.

Discusión

La evaluación del cambio conductual en los programas de intervención es fundamental para determinar si una intervención es efectiva. Sin embargo, como señaló Kazdin (2000) la evaluación no es suficiente para demostrar que los cambios se deben al programa de intervención, sino que es necesario emplear diseños experimentales que permitirán corroborar que la causa del cambio conductual se debe al programa de intervención y no a otra variable no controlada. El objetivo de este estudio fue conocer la efectividad de un programa de intervención con padres sobre la frecuencia de emisión de conducta agresiva y pro-social de niños que exhiben acoso escolar. Los resultados encontrados mostraron que existió una relación funcional entre la disminución del comportamiento agresivo y el incremento de conductas pro-sociales y el entrenamiento a padres mediante un programa de prácticas de crianza positiva (Pedroza et al., 2013).

El programa de entrenamiento de prácticas de crianza positiva mostró su efectividad, ya que incrementó la frecuencia de conductas pro-sociales y disminuyó la frecuencia de conducta agresiva. Esto se logró aumentando la calidad de la interacción entre padres e hijos y entre los niños y sus compañeros de clase y la supervisión y monitoreo de los padres hacia sus hijos, que son condiciones necesarias para la disminución del comportamiento agresivo de los niños. El programa de prácticas de crianza positiva está dirigido a padres, dado que son los principales agentes de cambio para disminuir la conducta agresiva e incrementar la conducta pro-social de sus hijos. El considerar a los padres como los principales agentes de cambio de la conducta infantil fue propuesto por Patterson (1974). Este autor mostró que los padres fueron capaces de identificar los eventos consecuentes que mantienen el comportamiento agresivo y después de un entrenamiento fueron capaces de modificarlos para lograr cambios en el comportamiento de sus hijos. Al igual que lo hizo Patterson, en el presente estudio se capacitó a padres de familia para realizar un análisis funcional del comportamiento agresivo con el objetivo de que identificaran los estímulos antecedentes y consecuentes de dicha conducta y modificaran sus prácticas de crianza y consecuentemente su propia conducta y la de sus hijos.

Una contribución importante del presente estudio fue la demostración de que el cambio en la conducta se generalizó del ámbito familiar al escolar. Este hecho mostró generalidad de la variable independiente, lo cual es una característica fundamental del análisis conductual aplicado. El hecho de que el entrenamiento a padres repercuta no sólo en la conducta de sus hijos en la familia, sino que trascienda a la escuela, confirma que los padres son los principales agentes del cambio de la conducta de sus hijos. Dando entrenamiento únicamente a los padres fue posible reducir significativamente la conducta de bullying de los niños dirigida hacia alumnos específicos, que fueron identificados como víctimas tanto por el profesorado como por sus propios padres de familia. El entrenar a los padres a manejar las contingencias de reforzamiento apropiadas para cada conducta, así como enseñarles a reforzar conductas alternas al comportamiento agresivo, usando técnicas de sobre-corrección, costo de respuesta y economía de fichas fue muy efectivo para reducir la conducta de bullying. Estas técnicas han

probado su efectividad en otros programas de intervención con padres pero únicamente se había probado en el contexto familiar (Ayala et al., 2001; Ayala et al., 2002; Pedroza, 2006). En este estudio, en cambio, no sólo se confirmó su efectividad en el ámbito familiar, sino que se probó su efectividad para reducir conducta agresiva en otro ámbito diferente, el escolar.

Es importante señalar que los padres que participaron en el presente estudio estaban siendo presionados por las autoridades escolares o legales para que se redujera la conducta de bullying de sus hijos y aumentara su conducta pro-social. Este hecho pudo afectar los resultados encontrados, dada la motivación que tenían los padres por mostrar que la conducta de sus hijos cambió en un tiempo corto. Por tanto, será necesario comprobar los hallazgos utilizando el programa de entrenamiento de prácticas de crianza positiva con padres de otros niños que no tengan dicha presión para ver si resulta igualmente efectivo. Otro aspecto a resaltar fue que ocho de los diez padres participantes, asistieron con sus parejas a todas las sesiones. Este dato es importante, ya que en investigaciones mexicanas se ha identificado que existen patrones diferenciales de la influencia de las prácticas de crianza del padre y de la madre sobre el comportamiento de sus hijos e hijas (e.g., Andrade & Betancourt, 2012). Esos resultados condujeron a los investigadores a sugerir la necesidad de analizar las interacciones entre padres e hijos, considerando las prácticas de crianza de cada uno de los padres. Por lo tanto, en futuros estudios sería conveniente utilizar el programa de prácticas de crianza positiva y establecer cuál de los dos padres fue el verdadero agente de cambio o si fueron ambos.

El programa de intervención dirigido a padres empleado en el presente estudio puede ser de gran utilidad para disminuir el acoso escolar y su implementación en las instituciones educativas sería una herramienta útil no sólo para disminuir el bullying en las escuelas, sino posiblemente también otras conductas problema desde el contexto familiar. El éxito del programa para disminuir el acoso escolar redundó en que el Sistema Nacional de Desarrollo Integral de la Familia promoverá su uso y se capacitará a especialistas del DIF en todo el país (Mendoza, 2014).

Referencias

Andrade, P., & Betancourt, D. (2012). Problemas emocionales y conductuales en niños: Predictores desde la percepción de los padres y de los hijos. *Acta de Investigación Psicológica, 2*, 650–664.

Ayala, H., Chaparro, A., Fulgencio, M., Pedroza, F. J., Morales, S. Pacheco, A. Mendoza, B., Ortiz, A., Vargas, E., & Barragán, N. (2001). Tratamiento de agresión infantil: Desarrollo y evaluación de programas de intervención conductual*. Revista Mexicana de Análisis de la Conducta*, *1*, 1-34.

Ayala, H., Pedroza, F., Morales, S., Chaparro, A., & Barragán, N. (2002). Factores de riesgo, factores protectores y generalización del comportamiento agresivo en una muestra de niños en edad escolar. *Salud Mental, 25*, 27-40.

Barkley, R. A. (1997). *Defiant children. A clinician's manual for assessment and parent training.* New York, NY.: The Guilford Press.

Cervantes, A. R., & Pedroza, F. J. (2012). El bullying: Una aproximación a la delimitación operacional del concepto. *INFAD, Revista de Psicología, 1,* 451-560.

Farrington, D. P. (2004). Conduct disorder, aggression, and delinquency. In Lerner, R. M. & Steinberg, L. (Eds.), *Handbook of adolescent psychology* (pp. 627-664). New York. NY: Wiley.

Forehand, R., & Long, N. (2010). *Parenting the strong-willed child.* New York: McGraw- Hill.

González, M., & Landero, R. (2012). Diferencias en la percepción de estilos parentales entre jóvenes y adultos de las mismas familias. *Summa Psicológica, 9,* 53-64.

González, J. L., Vargas, E. E., Galván, E. & Ayala, H. E., (1998). Intervenciones para mejorar la calidad de las interacciones medres-hijos: Evaluación analítico conductual del impacto social, validez externa y validez interna. *Psicología Conductua*l, *3,* 473-499.

Karna, A., Voeten, M., Little, T. D., Alanen, E., Poskiprta, E., & Salmivalli, C. (2013). Efectiveness of the KiVa antibullying progam: Grades 1-3 and 7-9. *Journal of Educational Psychology*, *2*, 535-551.

Kazdin, A. E. (2000). *Modificación de la conducta y sus aplicaciones prácticas*. México: Manual Moderno.

Mendoza, B. (2014). *Prevención de la violencia desde el contexto familiar. Conferencia Magistral*. 5to Congreso Internacional de la Familia. DIF Nacional. México.

Monks, C. P., & Smith, P. (2006). Definitions of bullying: Age differences in understanding of the term, and of the role of experience. *British Journal of Developmental Psychological, 24,* 801-821.

Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Oxford: Blackwell Publishers Inc.

Olweus, D., & Limber, S. P. (2010). Bullying in school: Evaluation and dissemination of the Olweus bullying prevention program. *American Journal of Orthopsychiatry, 1*, 124-134.

Orpinas, P., & Horne, A. M. (2006). *Bullying prevention: Creating a positive school climate and developing social competence.* Washington, D.C.: American Psychological Association.

Patterson, G. R. (1974). A basis for identifying stimuli which control behaviors in natural settings. *Child Development*, *45*, 900-911.

Patterson, G. R., DeBaryshe, B. D., Ramsey, E. (1989). A developmental perspective on antisocial behavior. *Ameraican Psychologist, 44*, 329-335.

Patterson, G. R., DeGarmo, D., & Forgatch, M. S. (2004). Systematic changes in families following prevention trials. *Journal of Abnormal Child Psychology, 6,* 621-633.

Patterson, G. R., & Stouthamer-Loeber, M. (1984). The correlation of family management practices and delinquency. *Child Development*, *55*, 1299-1307.

Pedroza, F. J. (2006). *Entrenamiento conductual para padres de niños clasificados como agresivos. Tesis doctoral inédita*. UNAM. México.

Pedroza, F. J., Aguilera, S. J., Cervantes, A. R., & Martínez, K. I. (2013). Estabilidad y cambio en roles de agresor y víctima de episodios bullying. En F. J. Pedroza, & S. J. Aguilera (Coord.), *La construcción de identidades agresoras: El acoso escolar en México* (pp. 49-71). México: CONACULTA.

Pedroza, F. J. & Martínez, K. I. (2011). Procesos de interacción social facilitadores del establecimiento y mantenimiento de patrones de consumo de sustancias. *International Journal of Developmental and Education Psychology (INFAD), 2*, 363-372.

Pedroza, F., Mendoza, B. & Martínez, K. (2013). *¡Auxilio! ¡Mi hijo no trae manual!* México: Pax.

Salmivalli, C., Poskiparta, E., Athola, A. & Hataja, A. (2013). The implementation and effectiveness of the KiVa antibullying program in Finland. *European Psychologist, 2*, 79-88.

Santoyo, C. (2007). *Estabilidad y cambio de patrones de comportamiento en escenarios naturales: Un estudio longitudinal en Coyoacán*. México: UNAM/CONACYT.

Webster-Stratton, C., & Hancock, L. (1998). Training for parents of young children with conduct problems: Content, methods, and therapeutic processes. En J.M. Briesmeister, & C.E. Schafer (Eds.), *Handbook of parent training* (pp. 98-152). New York, NY.: Wiley.

Lineamientos para los autores

La Revista Acta de Investigación Psicológica (AIP) tiene como propósito publicar simultáneamente en papel y en forma electrónica artículos científicos originales de investigación empírica en todos los ámbitos de la psicología. El manuscrito no debe someterse al mismo tiempo a consideración de otra revista. Además, se debe garantizar que los contenidos del manuscrito no han sido publicados y que todas las personas incluidas como autores han dado su aprobación para su publicación.

Se pueden someter a la revista manuscritos describiendo investigación original en español o en inglés. En ambos casos, la primera página debe incluir el título en ambos idiomas, el título no mayor a 85 caracteres incluyendo espacios, se recomienda que sea claro, preciso y que contenga las variables del estudio, nombre(s) del(os) autor(es) completo(s) y afiliación institucional. En la parte inferior de la página se debe incluir el nombre del autor o el de la autora a quien se dirigirá cualquier correspondencia, número telefónico, correo electrónico y domicilio completo.

El manuscrito debe presentarse en un único documento escrito a doble espacio con letra Arial 12 puntos, y no debe exceder de 25 páginas, incluyendo tablas y figuras. El formato del texto debe apegarse estrictamente al Manual de Estilo de Publicaciones APA (2da. Ed., en español, 2002, Editorial El Manual Moderno) y a los lineamientos descritos a continuación.

En la segunda y tercera páginas debe presentarse el título en los dos idiomas, en caso de que el manuscrito este en Español, llevará un resumen con un máximo de 200 palabras, y en Inglés un abstract de 300 a 400 palabras, en caso de que el texto este en Inglés un abstract de 200 palabras y un resumen en Español de 300 a 400 palabras.

Se deberá incluir también 5 palabras clave en español y 5 en inglés. Se recomienda que las palabras claves se refieran a las variables del estudio, la población, la metodología utilizada, al campo de conocimiento, el país donde se llevó a cabo la investigación. Debido a que la revisión editorial se realiza de forma anónima por 2 jueces, es responsabilidad del autor verificar que dentro del cuerpo del artículo no haya elementos que puedan identificar a los autores.

En las páginas siguientes debe aparecer el cuerpo del manuscrito, marco teórico, método, resultados, discusión y referencias. En el mismo archivo, al final del cuerpo del manuscrito, en páginas separadas, deben aparecer las leyendas de figuras y tablas, las figuras, las tablas, los anexos y nota del autor. Dentro del texto del artículo se debe señalar claramente el orden de aparición, y su formato se apegará estrictamente al formato APA.

Dado el corte estrictamente empírico de la publicación, es indispensable que la introducción justifique claramente la importancia del problema de investigación, el cual

debe derivarse directamente de la revisión de la investigación antecedente relevante, incluyendo resultados contradictorios, vacíos en el conocimiento y/o ausencia de conocimiento que el estudio pretenda resolver. En la sección de método deberá incluir la formulación de las hipótesis o las preguntas de investigación en las que se consideraren claramente las variables de estudio y se vinculen directamente con el problema. Las hipótesis o preguntas de investigación deben consideran clara y exclusivamente las variables del estudio, es decir, que se vinculan directa y explícitamente con el problema de investigación, enuncian claramente la dirección de la relación entre las variables y están apoyadas por la revisión de la literatura.

Incluya una descripción amplia y clara de la muestra, procedimientos y mediciones. En el apartado de resultados presente solo datos que se derivan de las hipótesis de estudio y asegure que los análisis estadísticos sean pertinentes. Se ha de proveer información de la magnitud de los efectos, así como de la probabilidad de todos los resultados significativos. Los datos que apoyen los resultados de la investigación deberán conservarse por 5 años después de la publicación, para garantizar que otros profesionales puedan corroborar los argumentos que se sostienen en el trabajo escrito, siempre y cuando al hacerlo no se violen derechos legales o éticos. Por último, la discusión debe derivarse congruente y directamente del marco teórico, la pregunta de investigación y los resultados obtenidos. Finalmente, asegurarse de que cada una de las referencias debe estar citada en el texto y cada cita debe estar en la lista de referencias.

El manuscrito debe enviarse adjunto vía electrónica en un solo archivo nombrado con el primer apellido del primer autor y la (s) inicial (es) del nombre y en formato compatible con PC (.doc, .rtf), a Rolando Díaz Loving al correo electrónico: **actapsicologicaunam@gmail.com**. Los autores deben conservar una copia del manuscrito sometido, en caso de que éste sufra algún daño al enviarlo a la AIP.

Todo manuscrito sometido a AIP se someterá a un filtro inicial, antes de ingresar al proceso editorial. Una vez soslayado este cedazo, se revisarán manuscritos de investigación que cumplan con rigor conceptual y metodológico; esta decisión depende de los miembros del Consejo Editorial, de dictaminadores y en última instancia, del Editor. Los autores de los artículos aceptados deben proveer por escrito las autorizaciones de material con derechos de autor, como pruebas psicológicas, fotografías, figuras, tablas, entre otros, que son utilizados en su artículo.

Proceso editorial

El proceso de recepción, evaluación, dictamen y publicación que se sigue en la revista es el siguiente:

- El Autor principal lee y acepta las políticas de publicación de la revista y será el encargado del seguimiento y comunicación con la misma.
- El Autor principal prepara y envía su artículo y autorizaciones de acuerdo al formato solicitado.
- El Editor recibe el material y revisa que cumpla con los requisitos establecidos (formato, autorizaciones, etc.), de no ser así, se devuelve al Autor para su corrección y posterior postulación. Sí el artículo cumple con todos los requisitos establecidos, el Editor emite confirmación de la recepción y del envío a revisión del artículo. El Editor selecciona a los miembros del Comité Editorial que realizarán la revisión del artículo (entre 2 y 3 miembros).
- Los miembros del Comité Editorial seleccionados, que desconocen la (s) autoría (s) del manuscrito, revisan y emiten un dictamen razonado sobre el artículo basado en la rigurosidad científica, el impacto de la contribución, la congruencia del método de investigación, la sistematicidad y lo adecuado de los resultados, la claridad y contundencia de los argumentos de la presentación (tiempo estimado: 4 semanas máximo).
- El Editor recibe y pondera las evaluaciones de los revisores y emite alguno de los siguientes dictámenes:

    1) Aprobado para publicación.
    2) Aprobado para publicación condicionado a los cambios sugeridos.
    3) Cambios sugeridos mayores que requieren de una nueva evaluación.
    4) La temática, contenido, abordaje o metodología no corresponden a los criterios de evaluación de la revista.

- En el caso 2, el Editor hace del conocimiento del Autor los cambios sugeridos al artículo para su publicación.
- El Autor recibe y realiza los cambios sugeridos al artículo, y en un plazo máximo de 4 semanas a partir de conocer los cambios sugeridos remite el artículo corregido al Editor.
- El Editor revisa los cambios y en caso de requerirse sugiere tantas modificaciones como sean necesarias. El Autor las realiza y lo reenvía al Editor
- En el caso 3, el Autor realiza los cambios sugeridos y lo reenvía al Editor quien a su vez lo envía a evaluación por el Comité Editorial.
- Una vez aceptado un manuscrito sin cambios adicionales, el Editor informará a todos los autores el número de la revista donde será publicado su artículo, conciliando la composición y tamaño de cada uno.
- Cuando el número es publicado, se proporcionarán dos revistas a cada autor.

Guidelines for Authors

The purpose of Psychological Research Records (PRR) is to publish original empirical scientific articles in all fields of psychology, simultaneously in hard copy and electronically. Contents of submitted manuscripts should be approved by all authors and have not appeared in other publications. In addition, manuscripts should not be sent to consideration in other journals while in the process of evaluation.

Articles describing original empirical research may be submitted either in English or in Spanish. In any case, the cover page should include title in both languages, no longer than 85 characters with spaces included. The title should be clear, precise and include variables under study, complete names of authors and institutional affiliation. As a footnote to this first page, interested parties should include the full name of author to whom correspondence should be directed, phone number, e-mail and full address.

Manuscripts must be sent in one single document (**actapsicologicaunam@gmail.com**), double spaced, Arial type 12, and should not exceed 25 pages including tables and figures. Text format should strictly adhere to APA Publication Manual stipulations and to the norms described below.

Second and third pages should include titles in both languages. When the paper is in Spanish, an abstract in this language of maximum 200 words and an abstract in English of minimum 300 and maximum 400 words should be presented. When the submission is in English, then the abstract should be no longer than 200 words and a Spanish abstract of minimum 300 and maximum 400 words should be presented. 5 key words in each language should also be provided. It is recommended that key words include study variables, population characteristics, methodology and field of knowledge referred to. Since the editorial revision is conducted by two judges blind to authors identity, it is the authors responsibility to insure that no identification clues are in the body of the paper.

The following pages must include the main body of the manuscript, theoretical framework, methodology, results, discussion and references. At the end of the same file, in separate pages, authors should insert tables, figures, attachments and author's notes.

Given the strict empirical orientation of the journal, it is essential that the introduction clearly justifies the weight of the study, which should be directly derived from relevant previous research, including contradictory results, omissions, or lack of knowledge which the study intends to rectify. The methods section must include clear research questions, hypothesis and include all conceptual and operational definitions of variables under scrutiny. In addition, an ample description of the sample, procedures, and research design and measurement instruments should be included.

In the results section, only present data that respond to hypothesis and make sure that statistical analysis are appropriate and justified. Give information on significance and effect sizes. Data for the study should be kept for 5 years after the publication, to insure

that other researchers can revise them if needed, unless ethical or legal rights preclude this action. For the discussion section, it is imperative that it strictly address only content that is derived from the introduction, the research question and the results. Finally, insure that all cited references from the body of the text are included in the reference list.

All manuscripts submitted to PRR will go through an initial screening before entering the formal editorial process. Once APA format and minimum research specifications have been met, research manuscripts will be sent to 2 to 3 members of the Editorial Board for who will assess the conceptual and methodological rigor of the proposal. The decision will be informed to the authors by the Editor, and in cases of acceptance, the authors should provide written consent of any materials under publishers rights used in the article.

Editorial Process

The reception, evaluation, verdict and publication for the journal are as following:

- Principal Author should read and accept the journals publication norms and will be assigned to follow up and communicate with the editor.
- Prepares and submits manuscripts and required authorizations in adherence to specified formats and norms.
- Editor confirms receiving the manuscript and revises text for adequate form; if the paper does not meet the standards the Editor sends the manuscript back to the Authors for corrections before it can enter the editorial revision.
- If Authors consider it adequate, they resubmit with proper format.
- Editor confirms receiving manuscript and sends it to 2 to 3 members of the Editorial Board who are blind to Author's identity. Editorial board members revise and give a reasoned judgment on article based on scientific rigor, importance of contribution, congruence of research method, adequacy of results and clarity and impact of arguments and discussion (estimated time, one month).
- Editor receives evaluation, considers strengths and weaknesses and gives one of the following verdicts:

    1) Approved for publication.
    2) Approved if suggested changes are made.
    3) Major changes require resubmission and a new evaluation.
    4) Theme, content or methodologies do not match the journals evaluations standards.

    - For case 2, Authors makes changes and sends manuscript to the Editor (time limit one month). Editor reviews changes and suggests as many additional changes as necessary. Once all issues are resolved, the article is approved for publication.

    - For case 3, Authors make required changes and resend manuscript to the Editor who assigns new judges from the Editorial Board.

    - Once an article is fully approved, the Editor informs the Authors in what date and number the text will be published. When the journal appears, each author receives 2 copies of the journal where the articles came out.

Contenido:

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE PSICOLOGÍA
www.psicologia.unam.mx