The Behavior Analyst Today

A Context for Science with a Commitment to Behavior Change

ISSN: 1539-4352

VOLUME 11 **Number 2, 2010**

TABLE OF CONTENTS

THE BEHAVIOR ANALYST TODAY

PUBLISHER'S STATEMENT

VOLUME NUMBER 11, NUMBER 2

ISSN: 1539-4352

Published: September 30, 2010

The Behavior Analyst Today (BAT) is published quarterly by Joseph Cautilli. BAT is an online, electronic publication of general circulation to the scientific community. BAT's mission is to provide a concentrated behavior analytic voice among voices that are more cognitive and structural. BAT emphasizes functionalism and behavioral approaches to verbal behavior. Additionally, BAT hopes to highlight the importance of conducting research from a strong theoretical base. BAT areas of interest include, but are not limited to Clinical Behavior Analysis, Behavior Models of Child Development, and Community based behavioral analytic interventions, and Behavioral Philosophy. BAT is an independent publication and is in no way affiliated with any other publications.

The materials, articles, and information provided in this journal have been prepared by the staff of the Behavior Analyst Today for informational purposes only. The information contained in this journal is not intended to create any kind of patient-therapist relationship or representation whatsoever.

For a free subscription to The Behavior Analyst Today, send the webmaster an e-mail containing your name, e-mail address, and the word "subscribe" in the subject box, and you will be added to the subscription list. You will receive notice of publication of each new issue via e-mail that will contain a hyperlink to the latest edition. You may also subscribe to the BAT journal by e-mailing baojournals@aol.com.

Our Mission

The Behavior Analyst Today is committed to increasing the communication between the sub disciplines within behavior analysis, such as behavioral assessment, work with various populations, basic and applied research. Through achieving this goal, we hope to see less fractionation and greater cohesion within the field. The Behavior Analyst Today strives to be a high quality journal, which also brings up to the minute information on current developments wihin the field to those who can benefit from those developments. Founded as a newsletter for master level practitioners in Pennsylvania and those represented in the clinical behavior analysis SIG at ABA and those who comprised the BA SIG at the Association for the Advancement of Behavior Therapy, BAT has evolved to being a primary form of communication between researchers and practitioners, as well as a primary form of communication for those outside behavior analysis. Thus the Behavior Analyst Today will continue to publish original research, reviews of sub disciplines, theoretical and conceptual work, applied research, program descriptions, research in organizations and the community, clinical work, and curriculum developments. In short, we strive to publish all which is behavior analytic. Our vision is to become the voice of the behavioral community.

THE BEHAVIOR ANALYST TODAY

Editorial Staff and Board members of The Behavior Analyst Today are volunteers who work in conjunction with and direct interns for The Behavior Analyst Today. To learn more about the Editorial Board or to petition to become a member of the editorial board please contact Joseph Cautilli, M.Ed., M.Ed., BCBA.

Meet Our Staff

Co Lead Editor

Mary Jane Weiss, Rutgers University, Douglas Developmental Center

Co Lead Editor

Thomas Zane, Center for Applied Behavior Analysis at The Sage Colleges

Senior Associate Editor

Cynthia Andersen, Ph.D.

Senior Associate Editor

Anne Bramsetter, Ph.D.

Associate Editors

John Borrero, Ph.D.

Darlene Crone-Todd, Ph.D.

Paul Malanga, Ph.D.

Pat Progar, Ph.D.

Author Submission Information

Most contributions are by invitation and all are then peer-reviewed and edited. The editors, however, welcome unsolicited manuscripts, in which case, we suggest potential authors send an abstract or short summary of contents and we will respond as to our interest in a full manuscript submission. In all cases, manuscripts should be submitted electronically saved in MS Word "DOC" format to Tom Zane at zanet@sage.edu or Mary Jane Weiss at mjweiss@rutgers.edu. Please adhere to APA format and use "Times New Roman" font in 11 pt. throughout. In references, however, please italicize the places where APA format would have you underline. Abstracts and keywords must be included, as well as full contact information for all authors. All submissions are peer-reviewed and must be accompanied by a signed Assignment of Rights (AOR) form. After peer review and follow-up, all articles are copyedited. Authors have an opportunity to review and approve their manuscript prior to publication. Once approved, authors are responsible for all statements made in their work, including changes made by the copy editor prior to approval.

Peer Review Process

All submitted manuscripts are reviewed initially by the Lead Editor. Manuscripts with insufficient priority for publication will be rejected promptly. Other manuscripts will be sent to the Senior Associate Editor, who will distribute them to editorial consultants with relevant expertise. The editorial consultants will read the papers and evaluate (1) the importance of the topic addressed by the paper; (2) the paper's conformity to standards of evidence and scholarship; and (3) the clarity of writing style. Comments provided by the editorial consultants will then be provided to the author(s) for follow up.

Formatting Requirements: To support the electronic copy-editing process, authors must honor all of the following guidelines:

The page set-up for manuscripts must be set for 1-inch margins on all 4 boarders.

All pages must be in portrait orientation. There can be no pages in landscape orientation.

Manuscripts must be typed in single-spacing using size 11 “Times New Roman” Font.

Manuscripts must be submitted as one continuous document rather than in sections or sub-documents.

Each manuscript must include 7 elements in the following order: title, name(s) of author(s), abstract, key words, body, references, author(s)’ contact information.

Do not insert pagination, headers, or footers. (These are inserted in the copy-editing process)

The use of headings is encouraged and should be structured according to the guidelines described in the *Publication Manual of the American Psychological Association* (5th edition).

If graphics, figures and tables are used, they must be created in *.jpg or *.bmp format. No Excel graphs will be accepted.

Graphics, figures, and tables, if used, may be embedded in the body of the manuscript or they may be submitted in a separate MS Word document. If the latter option is chosen, then author(s) must indicate clearly the intended location of each item (graphic, figure, table) within the manuscript so that the copy editor can make the insertions.

Individual graphics, figures, and tables, when used, may not be larger than one page.

The caption for a table must be printed above the table. The caption for a figure must be printed below the figure.

In the references section, please use italics where APA style would allow underlining (e.g., the titles of journals and books).

Author contact information must include the following 4 elements for each author: name, mailing address, phone, and e-mail.

Manuscripts must be saved and submitted in MS Word, “DOC” format.

When there is a conflict between the requirements of *APA* style (see below) and the formatting rules listed here, the **formatting rules will supersede the *APA* requirements**.

Manuscript Style Requirements:

With the exception of the above (formatting) guidelines, authors must write their manuscripts in a style that is consistent with the *Publication Manual of the American Psychological Association (APA Manual)* (5th edition). A copy of this manual may be ordered at http://www.apastyle.org/

Consistent with APA style, authors must use **non-sexist language**. Please refer to Table 2.1 in the *APA Manual* for “Guidelines for Unbiased Language.”

Also consistent with APA style, authors must use **person-first language** for referring to individuals with potentially stigmatizing characteristics. Person-first language requires an author to name the individual first, followed by descriptive information (e.g., "child with autism") rather than to use an adjectival form (i.e., "autistic child") or a nominal form (i.e., "the autistic").

As noted above: When there is a conflict between the requirements of *APA* style and the formatting rules listed in the above section, the **formatting rules will supersede the *APA* requirements.**

General Guidelines for Preparing Abstracts:

The following general guidelines must be honored to insure that JSLP-ABA will be accepted into the major psych databases. (See PsychINFO website: **http://www.apa.org/psycinfo/about/covinfo.html)**

- An abstract may not exceed 960 characters and spaces (approximately 120 words). Characters can be conserved by using digits for numbers (except at the beginning of sentences); by using well-known abbreviations; and by using the active voice.
- Begin the abstract with the most important information, but don't repeat the title.
- Include only the four or five most important concepts, findings, or implications.
- Embed as many key words and phrases in the abstract as possible.
- Include in the abstract only information that appears in the body of the manuscript.
- For the sake of clarity, define all acronyms and abbreviations except for measurements; spell out the names of tests; use generic names for drugs (when possible); and define unique terms.
- Use the present tense to describe results with continuing applicability or conclusions drawn and the past tense to describe variables manipulated or tests applied.
- As much as possible, use the third person rather than the first person.

Abstracts for Empirical Studies: Abstracts for empirical studies are also generally about 100 to 120 words in length. They should include the following information:

- Problem under investigation (in one sentence)
- Pertinent characteristics of participants (e.g., number, type, age, sex, genus and species)
- Experimental method, including apparatus, data-gathering procedures, and complete test
- Names and complete generic names and dosage and routes of administration of any drugs (particularly if the drugs are novel or important to the study)
- Findings, including statistical significance levels
- Conclusions and implications or applications

Abstracts for Literature Reviews and Theoretical Articles

Abstracts for review or theoretical articles are generally about 75 to 100 words in length, and they include the following information:

- The topic (in one sentence)
- The purpose, thesis, or organizing construct and the scope (comprehensive or selective) of the article
- Sources used (e.g., personal observation, published literature)
- Conclusions

Thank you!

The Behavior Analyst Online Journals Department

The Acquisition of Generalized Matching in Children With Developmental Delays

Kristen L. Gaisford & Richard W. Malott

Abstract

The purpose of this study was to assess the extent of a generalized matching repertoire. Three children, ranging from two to four years of age, were selected from an early childhood developmental delay classroom. They were taught identical matching with six objects. After the children mastered those six objects, they were tested for a generalized matching repertoire, and all three children demonstrated a generalized identical-object matching repertoire. One of the children was tested with two-dimensional picture cards and demonstrated a generalized two-dimensional matching repertoire. While the other two children were not formally tested, two-dimensional matching was specifically targeted as part of their individual curricula; and both children mastered two-dimensional matching so quickly as to suggest they had already acquired a generalized two-dimensional matching repertoire as a result of the original training. Additionally, two of the children had a history of emotional responding prior to this study and with procedures conducted concurrently outside of the present experiment when prompted responses were not reinforced, yet they displayed almost no emotional responding in the study itself, suggesting that the reinforcement of prompted responses was the reason for the absence of emotional responding.
Keywords: generalized matching, prompting, emotional responding

---

Identity matching-to-sample consists of matching a sample stimulus to the corresponding identical comparison stimulus from an array of comparison stimuli (Brown, Brown, & Poulson, 1995). For example, if a learner were presented with an array of stimuli such as a block, car, and spoon (the comparison stimuli), then handed a spoon (the sample stimulus), and told to match, the learner should then place that spoon next to the comparison spoon. After the learner has acquired the ability to match all of the three objects, the experimenter could assess whether he or she had also acquired generalized matching-to-sample by using three different objects, such as a plate, sock, and cup to determine if the learner could match the given sample stimulus with the correct comparison stimulus. If the learner is able to correctly match the novel stimuli, then he or she has acquired a generalized identity matching-to-sample repertoire (Brown, et al., 1995).

Identity matching-to-sample has been demonstrated with various animal species such as pigeons (Cummings & Berryman, 1961; Cummings, Berryman, & Cohen, 1965; Wright, Cook, Rivera, Sands, & Delius, 1988), California sea lions (Pack, Herman, & Roitblat, 1991), bottle nosed dolphins (Herman & Gordon, 1974; Herman, Honvancik, Gory, Bradshaw, 1989), infant chimpanzees (Oden, Thompson, & Premack, 1988), and macaque monkeys (Washburn, Hopkins, & Rumbaugh, 1989; as cited by Brown et al., 1995). Not only did all of the aforementioned experimenters intend to determine whether or not the various species of animals could acquire an identical matching-to-sample repertoire, but also, would a generalized repertoire develop as well. Cumming and Berryman (1961) were unable to get generalized matching with pigeons; however, Cumming, Berryman, and Cohen (1965) got low levels of generalized matching with their pigeons, and Wright, Cook, Rivera, Sands, and Delius (1988) got high levels of generalized matching with their pigeons (as cited by Brown et al., 1995). Dolphins demonstrated generalized matching (Herman & Gordon, 1974; Herman et al., 1989), as did infant chimpanzees (Oden et al., 1988; as cited by Brown et al., 1995). Following identity matching-to-sample training, California sea lions demonstrated some generalization (Pack et al. 1991) and it is unclear if

macaque monkeys acquired a generalized matching repertoire (Washburn et al., 1989; as cited by Brown et al., 1995).

Children under five years of age can acquire identity matching-to-sample, but no attempt seems to have been made to assess generalization (Dixon & Dixon, 1978; Lutzer, 1987; Daehler, Lonardo, & Bukatko, 1979; as cited by Brown, et al., 1995). If a skill is going to be targeted for acquisition, it should be taught not only to mastery, but the generalization of that skill needs to be targeted as well. However, Brown, Brown, and Poulson (1995) demonstrated that three typically developing children were able to acquire generalized identity matching-to-sample.

While many manuals recommend training matching-to-sample to young children with developmental disabilities (Lovaas, 1981; Taylor, & McDonough, 1996), we have not found any literature actually showing that traditional matching-to-sample can be achieved by young children with developmental disabilities. And although, Saunders and Sherman (1986) taught matching-to-sample to three developmentally delayed teenagers; and all three children demonstrated generalized matching repertoire (as cited by Brown et al., 1995) there is still a need to demonstrate not only matching-to-sample, but also generalized matching in young children with developmental delays.

And while generalized matching is an interesting theoretical issue, presumably it is also of considerable practical importance, as the children will have little opportunity to use their matching skills with the specific stimulus used in training. The acquisition of an identity matching repertoire will be of little value to the child if that child has not acquired a generalized matching repertoire, so that he or she can use that skill in a variety of educational and practical contexts where the matching task involves novel, untrained stimuli.

When working with clients with developmental delays the main goal is to help them develop skills in order to acquire a functional repertoire. Matching (matching-to-sample) is a common skill taught to children with developmental delays in early intensive behavioral interventions (EIBI). However, neither journal articles nor curriculum guides for EIBI offer recommendations for how many stimuli are necessary to develop a generalized identity matching-to-sample repertoire or what skill should be targeted following skill acquisition of generalized matching-to-sample (Lovaas, 1981; Taylor, & McDonough, 1996; Dixon & Dixon, 1978; Lutzer, 1987; Daehler, Lonardo, & Bukatko, 1979; as cited by Brown, et al., 1995). Perhaps matching-to-sample is so frequently suggested as an early learner skill due to the other skills that are simultaneously being taught such as scanning of an array and also compliance with an instruction. Furthermore, matching-to-sample is relatively easy to prompt.

Another issue is that many people with developmental delays often fail to respond to environmental cues in the same way as typical learners (MacDuff, Krantz, & McClannahan, 2001). Therefore supplemental stimuli must be used to train skills. These stimuli are referred to as prompts, stimuli designed to increase the likelihood that a correct response will occur (Malott, 2008). Therefore added prompts (response prompts) are often used when training individuals with developmental delays.

Most-to-least (MTL) and least-to-most (LTM) prompting strategies are typical response prompt hierarchies and both have drawbacks. The methodology of this study addresses drawbacks for both MTL and LTM that other studies have not considered. Least-to-most prompting is often criticized for producing errors and that high error rates and low levels of reinforcement can lead to emotional responding (Demchak, 1990; Ducharme, 2003; as cited by Mueller, Palkovic, & Maynard, 2007). In an attempt to reduce the probability of emotional responding, both unprompted and prompted correct responses were reinforced in MTL as well as LTM. However, one negative aspect reported about MTL is that a learner can become prompt dependent if prompts are not faded in a timely and systematic manner. Additionally, since prompts are faded out from most to least prompting, learners are sometimes provided with more assistance than necessary to make a correct response (MacDuff, Krantz, & McClannahan, 2001). In order to address both concerns associated with MTL, each session was started with a probe trial to assess the level of assistance necessary for the learner to make a correct response. This increases the probability that the learner is not receiving more assistance than necessary and provides a systematic way to fade prompts. The present study was originally designed to compare least-to-most prompting, most-to-least prompting, and most-to-least with a delay (MTLD) strategies in an attempt to determine the most effective and efficient prompting strategy for skill acquisition. However, there was no demonstrated difference between the prompting strategies, which may be because all of the children acquired a generalized identical three-dimensional matching repertoire. Thus, the main focus of this report is on the acquisition of a generalized matching repertoire, rather than the type of prompting strategies; see Gaisford (2009) for the original manuscript with an extended comparison of the two prompting strategies.

Methods

Participants

Three students, one female and two males, enrolled in an Early Childhood Developmental Delay (ECDD) classroom at Croyden Avenue School, participated in this study. The ages of the children at the start of the study were two years and seven months, three years and nine months, and four years and eight months. To participate, the children had to have the prerequisite skills for matching objects. All three children received discrete trial instruction three hours a day, five days a week and had been enrolled in the classroom for a range of three to nine months (with an average of six months) at the start of the study.

Setting

The author conducted all the sessions in each child's study carrel, which was approximately 3.0 x 2.0 m. The carrels contained a desk and two chairs, with the experimenter sitting perpendicular to the child.

Interobserver Agreement

Two graduate students collected interobserver data on data sheets identical to the ones used by the experimenter. Interobserver agreement was taken for twenty-five percent of the sessions. The percentage of agreement was calculated by dividing the number of agreements by the number of agreements and disagreements and multiplying by 100. The average percentage of agreement was 97.62 percent with a range of 86.11 percent to 100 percent.

Materials

There were two different sets of training materials; one set of materials for each prompting strategy, LTM and MTL or MTLD. The first set of objects included identical pairs of green plastic plates, pink socks, and yellow wooden blocks; the second set included identical pairs of red plastic cups, blue plastic spoons, and white shoes.

Identification of Reinforcers

Twice a week as part of the regular classroom procedures, children selected eight tangible items from two large bins of tangibles. Following this selection, the eight tangibles along with three to four edibles were arranged on the desk in front of the child. Then a brief multiple-stimulus assessment without replacement (MSWO) was conducted to identify and rank the top five of the eleven to twelve preferred items (Higbee, Carr, and Harrison, 2000). At the beginning of each session the experimenter offered the child the two top-ranked items and used the first one the child selected as a reinforcer for that session. However, if during the session that item lost its' reinforcing value (e.g. the child pushed it away, did not engage with the tangible or did not eat the edible) this process was repeated with the two top-ranked items from the remainder of the array of five.

Procedures

General Procedure

Sessions were run five days a week, with both prompting procedures being used once a day and with at least twenty minutes between sessions. The order of the most-to-least (MTL) and the least-to-most (LTM) procedures were alternated across days. Prior to each session of both conditions, LTM probe trials were conducted.

Probe trials

In order to ensure that the child did not receive more assistance than needed, a probe trial was conducted to determine the appropriate prompt level. At the beginning of each session the experimenter began the probe trial by stating, "Match same," and allowed the child two to three seconds to respond independently. If the child did not respond or made an incorrect response, the experimenter provided the next prompt in the least-to-most prompt hierarchy and again allowed the child two to three seconds to respond. The experimenter continued up the prompt hierarchy until the child made a correct response. The prompt hierarchy was a gestural prompt, partial light physical prompt, partial firm physical prompt, and full physical prompt. The experimenter reinforced all correct responses, both prompted and unprompted. This process was repeated for each of the three stimuli in the array before the session began, for both training strategies, LTM and MTL.

Most-to-Least Prompting

Following the probe trial for each of the three items in the array, the prompt levels that were determined for each stimulus were used for the remainder of the session. In order to avoid extinction, a reinforcer was provided for both prompted and unprompted responses, both in the probe trials and training trials. Even as prompts were faded over successive sessions, a reinforcer was still provided if a prompt level that was higher than the probe session was required for a correct response. This was done to address the possibility that the child could have made a correct independent response due to chance

The instruction, "match same" was repeated for each prompt. The child was given two to three seconds to respond before a higher level prompt was used. If the child made an incorrect response, the object was handed back to him or her while the experimenter simultaneously provided a prompt and repeated the instruction until a correct response was made. If the child did not respond to the instruction, after waiting two to three seconds, the experimenter proceeded to provide a higher level prompt and repeated the instruction until a correct response was made. This procedure was the original design; however, some modifications were made along the way; exceptions to this procedure will be discussed.

Most-to-Least Prompting with a Delay

As will be described later, a delay was introduced in some of the MTL procedures. The Most-to-Least with Delay (MTLD) condition also started with a probe trial for each stimulus and the appropriate prompt level was determined as described before. Following the probe trials, the experimenter held up the stimulus and said, "Match same" and gave the child two to three seconds to respond independently. The two to three seconds to respond independently served as the delay in this procedure. If the child did not respond, or responded incorrectly the experimenter provided the predetermined prompt. All responses were reinforced whether they were independent or prompted (Libby, Weiss, Bancroft, & Aheran, 2008).

Least-to-Most Prompting

A probe trial was conducted for the LTM condition so that each condition would have the same number of trials. The experimenter began the trial by stating the instruction, "match same" and allowed the child two to three seconds to respond independently. The experimenter then proceeded through the least-to-most prompt hierarchy until the child made a correct response. The experimenter repeated the instruction, "match same" with every prompt and allowed the child two to three seconds to respond after each prompt. A reinforcer (the most preferred based on a preference assessment that was conducted immediately prior to the session) was delivered contingent upon the correct response regardless of the prompt level.

Matching-to-Sample Training, Data Collection, and Mastery Criterion

After conducting the probe trial for each of the three stimuli in a given set, the training session began with that set of stimuli using the prompting strategy assigned to that set. Three additional trials for each stimulus in the set were conducted in random order, again for a total of twelve trials per session, including the three probe trials. This same method was repeated for the other set of stimuli using the other prompting strategy, again with a total of twelve trials per session. Data were collected on whether or not any emotional responding occurred during training for both prompting strategies.

To be compatible with the standard classroom criterion, the mastery criterion was two consecutive sessions at 89 percent or better or three consecutive sessions at 75 percent or better; however due to the logistics of the experiment, a few sessions were run past the mastery criteria for all three children. Finally, the number of sessions for acquisition was compared between conditions.

Results

Dawson

Using the most-to-least prompting strategy, Dawson met the mastery criterion by session nine and met the mastery criterion by session ten when using least-to-most prompting (Table 1).

Stephen

By session 24 Stephen had not acquired the matching skill. And the classroom's policy required that if a child had not acquired a skill within 20 sessions, the procedure would be modified; therefore the most-to-least prompting strategy was changed to a most-to-least with delay (MTLD). Using the least-to-most prompting strategy, Stephen met a mastery criterion on session 34 (Table 1). He failed to meet a mastery criterion within twenty sessions using MTL, but once the procedure was switched to MTLD, he met a mastery criterion with only 8 more sessions—32 sessions total, suggesting that the addition of the delay to the MTL condition aided in the acquisition of the matching skill (Table 1).

Kiarra

Because of Stephen's success when the delay was added to the most-to-lest procedure, only MTLD and no simple MTL, was used with Kiarra. However, by the 11th session the experimenter noticed that Kiarra always waited for a second statement of the instruction "Match same," before responding in both the MTLD and LTM conditions. Therefore from the twelfth session on, the experimenter presented the second instruction "Match same," without an accompanying prompt. And for the next two sessions she responded one hundred percent correct to that second instruction with no prompting, suggesting that she had already mastered the matching skill under both MTLD and LTM conditions (Table 1).

Table 1. Number of sessions to mastery

Child	Least-to-Most Prompting	Most-to-Least Prompting/Most-to-Least Prompting with Delay
Dawson	10	9
Stephen	34	32
Kiarra	13	13

After the original matching training, the children's generalized matching repertoires were assessed with different novel objects and two-dimensional stimuli.

Generalized Identical-Object Matching.

When a child met a mastery criterion for object matching, the child was tested for a generalized identical-object matching repertoire. To assess this generalized repertoire, the child was presented with a set of three novel comparison objects, then handed a sample object that matched one of the comparison objects and told, "Match same." This was done three times with each of the three comparison objects, for a total of nine trials. Then the procedure was repeated with three different, novel objects. This testing was done in extinction. All three children performed above the generalization-mastery criterion (80 percent correct) when presented with the six novel identical objects, indicating that all three had acquired a generalized identical-object-matching repertoire (Fig 1.).

Generalized Similar-Object-Matching.

The same testing procedure was used with non-identical, but similar, objects. Both Stephen and Kiarra performed above the generalization-mastery criterion when presented with similar non-identical objects indicating a generalized similar-object-matching repertoire. While Dawson did not meet the generalization-mastery criterion, he still demonstrated better than chance performance, suggesting some generalization (Fig. 1).

Generalized Simple Picture Matching.

The same testing procedure was used with identical, simple, two-dimensional pictures (pictures consisting of a single large shape). The cards were four inches wide by six inches high. For logistic reasons, only Kiarra was tested, and she met mastery criterion (Fig. 1).

Generalized Complex Picture Matching.

Finally, the same testing procedure was used with Kiarra, with identical, complex, two-dimensional pictures (pictures consisting of four or five smaller shapes). The cards were four inches wide by six inches high. Again, Kiarra met mastery criterion (Fig. 1).

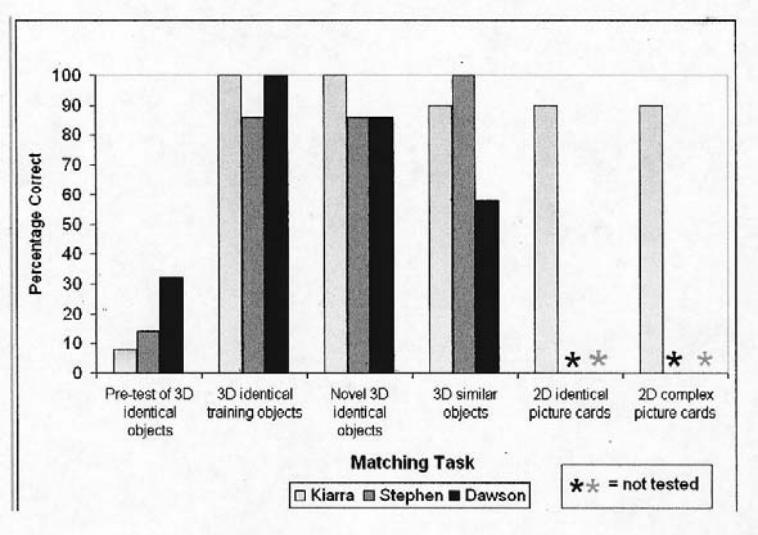

Figure 1. Generalized matching performance

While Stephen and Dawson were not tested for a generalized two-dimensional picture repertoire, Stephen completed the simple and complex matching procedures prior to leaving the classroom and did so with no difficulty during either procedure. Dawson completed the simple and complex matching procedures and three additional two-dimensional matching procedures as part of his curriculum and did so without any difficulty. For all participants, the acquisition of six identical objects led to a generalized identical-matching repertoire and a generalized two-dimensional matching repertoire for one of the participants.

The Reinforcement of Prompted Responses and the Elimination of Emotional Responding

Prompted responses were reinforced in both prompting conditions. This component was done because of the observation of emotional responding by children in the classroom when those children made several errors in a row and thus experienced low rates of reinforcers. Furthermore, one child, Dawson, had a history of emotional responding when the reinforcement density was low, prior to the start of this study. His emotional responding included, crying, whining, swiping instructional materials and slapping his face. After this study, he was in the classroom for one year and nine months and had had over ten procedures that had either been removed from his curriculum or had to be modified because of emotional responding that occurred within those procedures. However, in the present experiment, the reinforcement of prompted responses meant that he always received a high rate of reinforcers; and in this experiment, he emitted almost no emotional responding. In addition, reinforcing prompted responses during other procedures outside of this experiment was also associated with the elimination of emotional responding with those procedures. Stephen also demonstrated emotional responding prior to this study when experiencing low rates of reinforcement. Stephen had near zero levels of emotional responding during this study. Kiarra did not demonstrate high rates of emotional responding prior to this study, and she did not exhibit any emotional responding during the study. This suggests that tutors should reinforce prompted responses if a child has a history of emotional responding during instruction.

Conclusions

This study demonstrated that generalized, identical, three-dimensional matching can reliably result from matching training with as few as six three-dimensional objects; and, at least in some cases, such training can produce generalized matching between similar, non-identical objects and both simple and complex identical pictures. However, there were a few limitations of this study that warrant future research. One such limitation was that only Kiarra was formally tested for a generalized simple and complex identical picture matching repertoire. While both Stephen and Dawson performed well on these tasks when they were a programmed part of their curriculum, we can only speculate as to whether or not they had developed a generalized two-dimensional matching repertoire as a result of the original identical object matching training. It is also possible that all three students could have developed an even more extensive matching repertoire than was tested such as three-dimensional to two-dimensional matching. The design of the original experiment included a total of six identical objects for training matching to sample; however, it is not known if the training of six objects is needed to obtain the level of generalized matching that was achieved by the three children in this study or if similar results can be achieved by training with fewer objects. Future research should systematically determine

how many objects are needed to acquire a generalized matching repertoire and if the training of just identical object matching leads to a generalized two-dimensional matching repertoire. Such information could assist practitioners in selecting curriculum and should also prompt assessments to determine if students have acquired a generalized matching repertoire. Furthermore, this study strongly suggested that reinforcing prompted responses can essentially eliminate emotional responding for children who tend to respond emotionally when prompted responses are not reinforced; however, replication is needed to determine if the results found in this experiment can be generalized to other young learners with developmental delays.

References

Brown, A. K., Brown, J. L., & Poulson, C. L. (1995). Generalization of children's identity matching-to-sample performances to novel stimuli. The Psychological Record, 45, 29-43.

Demchak, M. (1990). Response prompting and fading methods: A review. American *Journal on Mental Retardation, 94*(6), 603-615.

Gaisford, K. L. (2009). *A comparison of prompting strategies in an early childhood developmental delay classroom.* Retrieved December 16, 2009, from Web site: http://dickmalott.com/autism/articles.html

Higbee, T.S., Carr, J.E., Harrison, C.D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities, 21*, 61-73.

Libby, M.E., Weiss, J.S., Bancroft, S., & Ahearn, W. H. (2008). A comparison of most-to-least and least-to-most prompting on the acquisition of solitary play skills. *Behavior Analysis in Practice, 1*, 37-43.

Lovaas, O. I., Ackerman, A., Alexander, D. D., Carr, E. G., Firestone, P., Newsom, C., Perkins,

M., & Young, D. B. (1981). *Teaching developing disabled children: The me book*. Austin, TX: PRO-ED, Inc.

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* (37- 50). Austin, TX: PRO-ED.

Malott, R.W. (2008). *Principles of Behavior* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

McConville, M. L., Hantula, D. A., & Axelrod, S. (1998). Matching training procedures to outcomes: A behavioral and quantitative analysis. *Behavior Modification, 22*(3), 391-414.

Mueller, M. M., Palkovic, C. M., & Maynard, C. S. (2007). Errorless learning: Review and practical applications for teaching children with pervasive developmental disorders. *Psychology in the Schools*, 44(7), 691-700. doi:10.1002/pits.20258

Taylor, B. A., & McDonough, K. A. (1996). Selecting teaching programs. In Maurice, C., Green, G., & Luce, S. C. (Eds.), *Behavioral interventions for young children with autism* (pp. 63-177). Austin, TX: PRO-ED, Inc

Author's Note

This study is based on a thesis conducted by Kristen Gaisford in partial fulfillment of the requirements for the PhD degree in Behavior Analysis at Western Michigan University. She is now at Evidence-Based Behavioral Consultation in Elko, Nevada.

Author Contact Information:

Dr. Kristen Gaisford, BCBA-D
1525 Opal Drive
Unit B203
Elko, Nevada 89801
(860) 671-1367 or (775)753-8469
kgaisford@hotmail.com

Dr. Richard Malott
Department of Psychology
3700 Wood Hall
Western Michigan University
Kalamazoo MI 49008
(269) 372-1268
dickmalott@dickmalott.com

ADVERTISEMENT

A Review of Timeout Ribbons

Douglas E. Kostewicz

Abstract

Researchers and practitioners often employ timeout procedures to manage inappropriate classroom behavior. When implemented inappropriately, however, timeout can result in dangerous situations and have received increased scrutiny (i.e., seclusion). The timeout ribbon procedure can prevent some of the dangerous situations associated with other forms of the punishment procedure. This review examines how researchers have used the timeout ribbon to affect change in the behaviors of school-aged children. A summary of the findings indicates that the timeout ribbon procedure effectively reduced inappropriate behaviors but did not increase compliance. The review provides both implications for practitioners and additional research directions.
Keywords**:** Timeout ribbon, classroom management, punishment

---

Timeout, when used effectively, is a powerful behavior management tool (Turner & Watson, 1999). Timeout is defined as "the withdrawal of the opportunity to earn positive reinforcement or the loss of positive reinforcers for a specified time, contingent upon the occurrence of a behavior; the effect to reduce the future probability of that behavior" (Cooper, Heron, Heward, 2007, p. 357). Thus, timeout has two necessary conditions. First, the current environment must have reinforcing qualities. Second, a removal of those qualities must be less reinforcing than a removal from that environment. In other words, there must be a discrepancy between time-in (i.e. the environment with reinforcement) and timeout (i.e., the environment without reinforcement; Friman & Finney, 2003; Harris, 1985; Marlow, Tingstrom, Olmi, & Edwards, 1997). In early studies, researchers demonstrated timeout by placing an animal on extinction following some behavior, which subsequently decreased that behavior's probability (Anderson & King, 1974). However, as timeout was applied in more and more settings, variability rather than conformity appeared (Friman & Finney, 2003).

Even with response variability, timeout is now one of the most common disciplinary tactics used with children in the United States (Friman & Finney, 2003). There are three types of timeout: isolation or total removal from a reinforcing environment, exclusion from reinforcement within an environment, and non-exclusionary or reinforcement is stopped (Harris, 1985). Additionally, three types of non-exclusionary timeout include a removal of the reinforcing stimulus (i.e., withholding food or the cessation of music), ignoring the subject (i.e., turning away from the subject), and contingent observation (i.e., the subject must sit out and watch the appropriate behaviors of peers; Harris, 1985). With different variations available, considerations must be made when choosing a timeout procedure.

For a timeout to be effective it must be applied immediately following each occurrence of the target behavior, which is not always possible with isolation and exclusion (Hugenin & Mulick, 1981). Additionally, moving an individual during isolation, exclusion, or contingent observation timeout procedures usually involves physical guidance, which has been shown to reinforce misbehavior (Kern, Delany, Hilt, Bailin, & Elliot, 2002). Recently, the Council of Children with Behavioral Disorders (2009) has released a position statement concerning the use of seclusion and isolation. The considerations include secluding the individual too long (i.e., the loss of a considerable amount of educational time), the potential for abuse, and additional paradoxical effects (i.e., timeout as a positive or negative reinforcer for inappropriate behavior). In summary, timeouts have heightened detrimental effects when used ineffectively by inexperienced people (Harris, 1985).

On the other hand, non-exclusionary timeout procedures do not have the same negative concerns as other timeout techniques. A modified non-exclusionary timeout, the timeout ribbon procedure,

combines contingent observation and the removal of a reinforcing stimulus. The availability of reinforcement is contingent upon the presence of some discriminative stimulus of which reinforcement has been paired. That stimulus is removed contingent upon the appearance of inappropriate or target behaviors and returned after a short period of time.

The timeout ribbon procedure controls for some of the negative side effects of timeout, but also raises additional concerns. The potential risk of abuse and paradoxical effects are reduced. During timeout, the child is not touched or removed from the educational environment. Also, earning a timeout does not allow the child to escape from educational demands; they are expected to continue working. However because the child remains in the educational setting, the environment must be able to maintain higher intensity inappropriate behaviors. After weighing the benefits and costs, the timeout ribbon procedure may be helpful for common occurring lower intensity, higher frequency behaviors (e.g., noncompliance) in educational settings (Ford, Olmi, Edwards, & Tingstrom, 2001).

With the many varieties, implications, and concerns regarding timeout, a critical literature review will lead to a greater understanding of one specific type: the timeout ribbon. Thus, the purpose of this review is to address the question: How has the timeout ribbon or modified ribbon procedure been used to affect change in behaviors of school-aged children? Specific questions include:

1. What timeout ribbon procedures have researchers used?

2. What target behavior outcome measures were used to evaluate the effectiveness of the timeout ribbon procedure?

Methods

Studies included in this review were located through two steps. First, a computerized search of PsychINFO and ERIC databases was conducted. Descriptors used were timeout, time-out, and classroom. Second, an ancestral search was conducted of articles identified in specified databases.

The computerized search generated 397 articles, five of which met all of the article inclusion criteria located below. An ancestral search of five articles resulted in identification of one additional article meeting criteria. Overall, the literature search process identified six articles (Alberto, Heflin, & Andrews, 2002; Fee, Matson, & Manikam, 1990; Foxx & Shapiro, 1978; Salend & Gordon, 1987; Yeager & McLaughlin, 1994; Yeager & McLaughlin, 1995) published in six different psychology and education journals (Table 1).

For an article to be included, the following had to be met:

1. The article was published in a peer-reviewed journal.

2. The article was an empirical study using group or single subject design.

3. The article included as participants school-aged children (3-18 years old).

4. The article used as a dependent measure behaviors that either interfered with or improved instruction in an educational setting.

5. The article examined a timeout ribbon procedure or a modified timeout ribbon procedure as an intervention.

***Table 1.*Timeout Ribbon Studies**

Study	*Students*	*Independent Variables*	*Dependent Variable*	*Results*
Alberto et al, 2002	Two males aged 10-11	TI: Token delivery FI 5 min for appropriate behavior as long as wristband is present; TO Wristband procedure: 5 min nonexclusionary TO (loss of wristband) for appearance of target behavior; IN behavior during TO: TO extended until 30 seconds of appropriate behavior; Conclusion of completed TO: Wristband returned	Number of IN vocalizations when initiating a task, uninvited approaches of strangers within 3 feet, IN self touches, and/or yells	Behaviors reduced to 0 occurrences for both participants across two settings and maintained at 0 occurrences during maintenance
Fee et al., 1990	59 children ages 4-5 years old	TI: Verbal praise 2 times normal (only for children with their wristbands) TO Wristband: Appearance of target behavior, warning first, if compliance, no TO, if not, 3 min nonexclusionary TO (loss of wristband); IN behavior during TO: Timer is reset and a minute is added; Conclusion of completed TO: Wristband returned	Matson Evaluation of Social Skills with Youngsters (MESSY) Preschool Behavior Questionnaire (PBQ) Six item semantic differential Out-of-seat and talking out of turn (Rated pairs of children (20 minute pre- and post-test) for six consecutive 10 sec intervals)	Treatment group showed significant decreases in both out-of-seat and talking out of turn behaviors and perceived as better group members.
Foxx & Shaprio, 1978	Five males (aged 9-18) with MR	TI: Social and edible reinforcement VI 2.5 min (only for children with their ribbons) TO Ribbon: Appearance of target behavior, 3 min nonexclusionary TO (loss of ribbon); IN behavior during TO: TO extended slightly until misbehavior ceases; Conclusion of completed TO: Ribbon returned	% of 30sec intervals of disruptive behaviors (e.g., Out-of seat, banging objects on table, throwing objects, hitting others, crying, yelling)	IN behaviors intervals reduced from a range of 70%-7% to a range of 10%-1%

Table 1. **(continued)**

Study	*Students*	*Independent Variables*	*Dependent Variable*	*Results*
Salend & Gordon, 1987	5 males aged 6-9 years; 4 with LD and one with ED and 3 males and 1 female aged 9-10 with LD	TI: Token delivery FI 2.5 min (only for groups with their ribbons) TO Ribbon: Appearance of target behavior within group, 1 min nonexclusionary TO (loss of ribbon); IN behavior during TO: TO extended until 1 min of appropriate group behavior displayed Conclusion of completed TO: Ribbon returned	IN vocalizations per minute	Group reduced IN vocalizations from 1.65 and 3.1 per minute to 0.1 and 0.4 per minute, respectively and low levels maintained during follow-up.
Yeager & McLaughlin, 1994	4 year-old child	TI: Praise (when ribbon was present) TO ribbon procedure: TO for target behavior Additional TI: Chose story to be read	% of compliance	Compliance increased from 7% to 22% of the time
Yeager & McLaughlin, 1995	4 year-old male with Tuberous Sclerosis	TI: edible for each instance of compliance TO happy face: Instance of noncompliance, loss of happy face until next chance for compliance End of TO: Compliance with next request; TO happy face and precision requests: noncompliance, warning 5 seconds, warning, 5 seconds, 4 min TO in TO chair (Any time before TO compliance earned edible) Conclusion of TO: Happy face returned when compliance demonstrated	% of compliance	Compliance increased from 2.2% to 62% of the time.

Note. MR = mental retardation, LD = learning disabilities, ED = emotional disturbance, TI = time-in, TO = timeout, VI = variable interval, FI = fixed interval, IN = inappropriate

Results

Timeout Ribbon Procedures

Each of the studies reported using a timeout procedure to reduce inappropriate behavior. All of the studies, except Yeager and McLaughlin (1995), did not move the participant while in timeout. Length of timeouts varied across the studies. Timeouts ranged from one minute (Salend & Gordon, 1987), three minutes (Fee et al., 1990; Foxx & Shapiro, 1978), four minutes (Yeager & McLaughlin, 1995), to five minutes (Alberto et al., 2002). The remaining studies either had an experimental phase in which the

timeout ended with the next chance for compliance (Yeager & McLaughlin, 1995) or only specified that the participant was in timeout for a period of time (Yeager & McLaughlin, 1996). Five studies adjusted the length of timeout when inappropriate behaviors occurred or continued during a timeout. Criteria included resetting the time (Salend & Gordon, 1987; Yeager & McLaughlin, 1995), resetting the time and adding one minute (Fee et al., 1990), extending timeout 30 seconds (Alberto et al., 2002) or extending the time an unspecified period (Foxx & Shapiro, 1978). Prior to starting a timeout, two studies (Fee et al., 1990; Yeager & McLaughlin, 1995) allowed for a warning before the loss of the timeout ribbon. In both cases, compliance with the request after the warning continued availability of reinforcement.

A characteristic similar across studies was the creation of a reinforcement rich time-in (i.e. an increased amount of supposed reinforcers delivered more frequently than normal). Two studies delivered reinforcement on a variable interval (VI) of either 2.5 minutes (Foxx & Shapiro, 1978) or 5 minutes (Alberto et al., 2002). Salend and Gordon (1987) delivered reinforcement on a fixed interval (FI) of 2 minutes, while Yeager & McLaughlin (1995) provided continuous reinforcement (CRF) the target behavior. Two studies did not specify their reinforcement procedures, but one (Fee et al., 1990) did state that reinforcement occurred twice as often. The reinforcers used included social praise (Fee et al., 1990), tokens (Alberto et al., 2002; Salend & Gordon, 1987) and edibles (Foxx & Shapiro, 1978; Yeager & McLaughlin, 1995). Within each of the studies some type of discriminative stimulus was used to signal the availability of reinforcement. Three of the studies (Foxx & Shapiro, 1978; Salend & Gordon, 1987; Yeager & McLaughlin, 1994) administered a ribbon, two (Alberto et al., 2002; Fee et al., 1990) used a wristband and one (Yeager & McLaughlin, 1995) used a ribbon and moved to a happy card. In all studies, reinforcement was available only when the participant's or group's discriminative stimulus (i.e., ribbon, wristband, or happy face) was present.

Target Behavior Outcome Measures and Short\Long Term Effects

All six studies reported dependent measures that either interfered with or improved educational situations. Four of the studies (Alberto et al., 2002; Fee et al., 1990; Foxx & Shapiro, 1978; Salend & Gordon, 1987) targeted decreasing inappropriate behaviors. Behaviors included yelling or inappropriate vocalizations (Alberto et al., 2002; Fee et al., 1990; Foxx & Shapiro, 1978; Salend & Gordon, 1987), out-of-seat behaviors (Fee et al., 1990; Foxx & Shapiro, 1978), inappropriate touching or hitting of others, self, or property (Alberto et al., 2002; Foxx & Shapiro, 1978), and uninvited approach of strangers (Alberto et al., 2002). Two studies (Yeager & McLaughlin, 1994; Yeager & McLaughlin, 1995) targeted increasing an appropriate behavior; namely compliance.

Additional dependent measures included teacher's perceptions of children's behavior recorded on two standardized tests: Matson Evaluation of Social Skills with Youngsters (MESSY) and Preschool Behavior Questionnaire (PBQ; Fee et al., 1990), a six item semantic differential (Fee et al., 1990), and a teacher questionnaire regarding acceptability of the timeout ribbon procedure (Fee et al., 1990; Foxx & Shapiro, 1978).

Findings from the six studies showed differing results. Researchers who targeted reducing inappropriate behaviors reported marked decreases. Fee et al. (1990) found significant decreases in talking out turn and out-of-seat behaviors as compared to the control group. Alberto et al. (2002) disclosed decreases in all inappropriate behaviors to zero across two participants and four settings. Salend and Gordon (1987) saw a reduction from 1.5-3.5 inappropriate vocalizations per group to under 0.5 inappropriate vocalizations per group. Foxx and Shapiro (1978) demonstrated a decrease to roughly zero intervals of disruptive behaviors across five participants.

When the aim was to improve appropriate behaviors, the outcomes suggested less effective results. Yeager and McLaughlin (1994) showed an increase from a mean of 4.5% intervals of compliance

in baseline to a mean of 27% intervals of compliance during treatments. Yeager and McLaughlin (1995) also demonstrated a greater increase of compliance intervals (2.2% in baseline to 54.2% during timeout ribbon phase to 74.6% during timeout ribbon/precision request phase).

Additional dependent measures showed changes in teachers' perceptions and acceptability of approaches. Fee et al. (1990) found that children within the timeout ribbon group were perceived as better group members, took turns more frequently, and stayed in their seat. Additionally, the timeout ribbon procedure was considered highly acceptable (Fee et al., 1990; Foxx & Shapiro, 1978), less restrictive than other methods for decreasing behaviors (Foxx & Shapiro, 1978), and a preferable choice if equally effective with exclusionary timeout (Foxx & Shapiro, 1978).

Three studies (Alberto et al., 2002; Foxx & Shapiro, 1978; Salend & Gordon, 1987) also had maintenance measures. Alberto et al. (2002) made three measurements 14 days after completion to find zero occurrences of targeted inappropriate behavior. They also thinned the token schedule from FI5 to FI10 minutes. Foxx and Shapiro (1978) took one measurement 21 days after completion to find the same low levels of disruptive behavior. Salend and Gordon (1986) made three measurements after 14, 21, and 28 days following the study's completion and found the same low levels of inappropriate vocalizations (<0.5 per minute).

Discussion

Timeout Ribbon Procedures

Extending from Foxx & Shapiro (1978), the reviewed studies reported a basic structure for the timeout ribbon procedure. Researchers established a clear discriminative stimulus with the students, removed that stimulus (and all access to reinforcement) when students displayed the targeted inappropriate behavior and returned that stimulus after a set amount of time. While some specifics (e.g., length of discriminative stimulus removal) varied across the six studies one aspect held consistent: the creation of a reinforcement rich time-in. Necessary for the effective use of all timeout, a reinforcement-rich time-in increases the reinforcing properties of the reinforcers denied through timeout (Friman & Finney, 2003). None of the reviewed studies reported varying the type of reinforcement used; just that reinforcement was delivered frequently. Thus, students' behaviors effectively differentially reinforced teacher's delivery of reinforcement, both in type and schedule. Should inappropriate behavior fail to show a decrease, teachers can examine how time-in was created (i.e., provided consequences, wording of praise statements, rewards associated with a token system, etc.) and the density of reinforcement adjusting accordingly. If students do show a marked decrease in the targeted inappropriate behavior, the teacher can consider thinning the schedule to promote generalization (Stokes & Baer, 1977).

While satiation might raise concerns, teachers often present only minimal amounts of reinforcing contingences (Sutherland, Wehby, & Yoder, 2002). Educators often provide little positive attention for appropriate behavior, whether academic or social, that continues to decrease as students advance through their academic career (White, 1975). Additionally, low intensity but high frequency inappropriate student behaviors such as non-compliance, the types of behavior suited for the timeout ribbon procedure, often frustrate teachers into over-reaction increasing the likelihood of managing student behavior with coercion (Sidman, 1989). The time-out ribbon procedure forces teachers to provide potential reinforcers at a rate higher than normally observed, while also providing an intervention that may replace ineffective and damaging coercive contingencies.

Target Behavior Outcome Measures

Four of the studies (Alberto et al., 2002; Fee et al., 1990; Foxx & Shaprio, 1978; Salend & Gordon, 1987) reported a decrease in inappropriate behaviors with two (Yeager & McLaughlin, 1994;

Yeager & McLaughlin, 1995) targeting increases in compliance. As timeout is a behavior reduction technique (Cooper et al., 2007), researchers reported effective results when used as such. However, fewer gains were noted when researchers examined subsequent increases in appropriate behavior. These differences suggest the nature of timeout and punishment techniques in general. The individual learns what not to do, rather than what to do. However, the slight gains to appropriate behavior noted by Yeager and McLaughlin in 1994 and 1995 might be the result of providing reinforcement for appropriate, alternative behaviors demonstrating the potential versatility of the timeout ribbon procedure. It might be surmised that when a teacher focuses more of the necessary reinforcing contingencies on certain appropriate behaviors, an associated increase in appropriate alternative behaviors occurs between ribbon removals.

Results from the current, yet limited, literature base do display generality across the domains suggested by Stokes and Baer (1977). Researchers measured and noted effective results for a wide range of behaviors (i.e., inappropriate approaches and vocalizations, out-of-seat behavior, etc.) across students with different exceptionalities (e.g., learning disabilities, mental retardation). Additionally, three of the studies (Alberto et al., 2002; Foxx & Shapiro, 1978; Salend & Gordon, 1987) reported follow-up measures that demonstrated positive outcomes. Considering individuals maintain behaviors in their repertoire post-intervention, successfully demonstrating the continued effects speaks to the time-in/timeout nature of the ribbon procedure. To maintain a generalized effect, behavior reduction techniques such as the timeout ribbon must be used consistently on each instance of targeted inappropriate behavior and involve a systematic coordinated fading during both time-in and timeout.

Implications for Practitioners

Posing clear advantages and disadvantages, the timeout ribbon procedure has applications for teachers and clinicians working with school-aged children. Those who choose to implement a timeout procedure should use a defined, consistent, effective approach to target a well defined set of behaviors. Advantageously, the timeout ribbon procedure has certain set characteristics (e.g., a reinforcement rich time-in, a conspicuous discriminative stimulus, targeting low-intensity inappropriate behaviors) and empirical support. These aspects guide the practitioner during implementation, tethering them to the procedure, but also provide the ability to modify some of the “looks” of the procedure. For example, a teacher can employ different discriminative stimuli, how and when to deliver preferred consequences as potential reinforcers, and what low-intensity inappropriate behaviors to target. Teachers can also rely on another clear advantage; the students remains in the educational or clinical setting during the timeout. Students have the opportunity to increase their exposure to instructional situations as compared to both exclusion or isolation forms of timeout. Finally, the procedure itself demands a positive educational or clinical environment (i.e. reinforcement rich time-in). This not only sets the stage for an effective non-exclusionary timeout, but also for effective instruction, learning, and student experiences.

A clear disadvantage of the timeout ribbon procedure involves the ability of the procedure to deal with the escalation of student inappropriate behaviors. Some students, such as those with emotional or behavioral disorders, may have a history of behaviors that, when displayed, no longer allow them to safely remain in the current educational setting (Kauffman & Wong, 1991). Behaviors such as physical damage to self, property, or others, place an undue stress on many educational environments. Planning ahead, teachers can use the timeout ribbon procedure as an initial intervention for low intensity/ high frequency inappropriate behaviors. Then, as necessary and following appropriate guidelines (CCBD, 2010), implement additional timeouts (i.e., exclusion or seclusion) as a back-up intervention targeting the more dangerous behaviors only (Foxx & Shapiro, 1978).

Future Research Directions

One specific area for researchers to explore involves the amount of time students remain in timeout or in the current case without a ribbon. The current body of timeout ribbon literature suggests a range from one to five minutes, however many students remain in timeout for longer periods (CCBD, 2010). Future researchers can experimentally manipulate timeout duration with the goal of determining the minimal amount necessary to garner effective results.

Without providing a consensus, each of the studies reported a different approach to resetting or adding time to a timeout. Continued prompting or adding additional time can create adverse situations in which some students may increase the intensity and frequency of inappropriate behavior past the point of remaining in the educational setting during a non-exclusionary timeout (Gunter, Denny, Jack & Shores, 1993; Gunter, Denny, Shores, & Reed, 1994). Future research can compare the different methods to determine effectiveness.

Other avenues for research include additional examinations of the timeout ribbon procedures in a variety of settings and populations. Researchers can also vary the schedule of reinforcement (both contingent and non-contingent) and/or the discriminative stimuli used to signal time-in to examine the effect of timeout. Each of the replications/modifications builds a greater case of the power, utility, and generality of the timeout ribbon procedure.

Conclusions

Often misused, timeout still plays a role in the management of student inappropriate behaviors. Versions of timeout, even when applied appropriately, can create situations that increase the likelihood of adverse effects. Timeout ribbons present a balance. The proper use of the procedure forces a teacher to identify and provide reinforcing contingencies for all students; which may decrease escape motivated behaviors while also increasing the effect of ribbon removal (i.e., non-exclusionary timeout) on common inappropriate classroom behaviors such as non-compliance. Teachers understanding that no matter the effect students do not specifically learn what to do rather only what not to do can implement the timeout ribbon in its intended situations accompanied by other teaching strategies.

References

*Alberto, P., Heflin, L. J., & Andrews, D. (2002). Use of the timeout ribbon procedure during community-based instruction. *Behavior Modification, 26*, 297-311.

Anderson, K. A. & King, H. E. (1974). Time-out reconsidered. *Journal of Instructional Psychology, 1*, 11-17

CCBD Position Summary on the Use of Seclusion in School Settings (2009). *Behavioral Disorders, 34*, 235-243.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis (2nd ed.)*. Upper Saddle River, NJ: Pearson.

*Fee, V. E., Matson, J. L., & Manikam, R. (1990). A control group outcome study of a nonexclusionary time-out package to improve social skills with preschoolers. *Exceptionality, 1*, 107-121.

Ford, A. D., Olmi, D. J., Edwards, R. P., & Tingstrom, D. H. (2001). The sequential introduction of compliance training components with elementary-aged children in general education classroom settings. *School Psychology Quarterly, 16*, 142-157.

*Foxx, R M. & Shaprio, S. T. (1978). The timeout ribbon: A nonexclusionary timeout procedure. *Journal of Applied Behavior Analysis, 11*, 125-136.

Friman, P. C. & Finney, J. W. (2003) Time-out (and time-in). In W. O'Donohue, J. E. Fisher, & S. C. Hayes (Eds.), *Cognitive behavior therapy: Applying empirically supported techniques in your practice* (pp. 429-435). Hoboken, NJ: John Wiley & Sons.

Gunter, P. L., Denny, R. K., Jack, S. L., & Shores, R. E. (1993). Aversive stimuli in academic interactions between students with serious emotional disturbance and their teachers. *Behavioral Disorders, 18*, 265-274.

Gunter, P. L., Denny, R. K., Shores, R. E., & Reed, T. M. (1994). Teacher escape, avoidance, and counter control behaviors: Potential responses to disruptive and aggressive behaviors of students with severe behavior disorders. *Journal of Child and Family Studies, 3*, 211-223.

Harris, K. R. (1985). Definitional, parametric, and procedural considerations in timeout intervention and research. *Exceptional Children, 51*, 279-288.

Huguenin, N. H. & Mulick, J. A. (1981). Nonexclusionary timeout: Maintenance of appropriate behavior across settings. *Applied Research in Mental Retardation, 2*, 55-67.

Kauffman, J. M., & Wong, K. L. H. (1991).Effective teachers of students with behavioral disorders: Are generic teaching skills enough? *Behavioral Disorders, 16*, 225–237.

Kern, L., Delaney, B. A., Hilt, A., Bailin, D. E., & Elliot, C. (2002). An analysis of physical guidance as reinforcement of noncompliance. *Behavior Modification*, 26, 516-536.

Marlow, A. G., Tingstrom, D. H., Olmi, D. J., & Edwards, R. P. (1997). The effects of classroom based time-in/time-out of compliance rates in children with speech/language difficulties. *Child & Family Behavior Therapy, 19(2),* 1-15.

*Salend, S. J. & Gordon, B. D. (1987). A group-oriented timeout ribbon procedure. *Behavioral Disorders, 12*, 131-137.

Sidman, M. (1989). *Coercion and its fallout.* Boston: Authors Cooperative, Inc.

Stokes, T. F. & Baer, D. M. (1977) An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Sutherland, K. S., Wehby, J. H., & Yoder, P. J. (2002). Examination of the relationship between teacher praise and opportunities for students with EBD to respond to academic requests. *Journal of Emotional and Behavioral Disorders, 10*, 5-13.

Turner, H. S. & Watson, T. S. (1999). Consultant's guide for the use of time-out in the preschool and elementary classroom. *Psychology in the Schools, 36*, 135-148.

White, M. A. (1975). Natural rates of teacher approval and disapproval in the classroom. *Journal of Applied Behavior Analysis, 8*, 367-372.

*Yeager, C. & McLaughlin, T. F. (1994). Use of a timeout ribbon with and without consequences as procedures to improve a child's compliance. *Perceptual and Motor Skills, 79*, 945-946.

*Yeager, C. & McLaughlin, T. F. (1995). The use of a time-out ribbon and precision requests to improve child compliance in the classroom: A case study. *Child & Family Behavior Therapy, 17*, 1-9.

** Denotes studies included in the review.*

Author Contact Information:

Douglas E. Kostewicz, PhD., BCBA
University of Pittsburgh
5146 Wesley W. Posvar Hall
230 South Bouquet Street
Pittsburgh, PA 15260
Office: (412) 648-7113
E-mail: dekost@pitt.edu

Morphosyntactic Learning: A Neurobehavioral Perspective

Jean A. Rondal

Abstract

Structural linguistic and psycholinguistic approaches to morphosyntax are plagued with a major logical-empirical caveat that prevents the definition of a plausible explanatory theory. An alternative approach is presented viewing morphosyntactic regulations as operating on line according to sequential and associative principles learned implicitly.
Keywords: morphosyntactic regulations, semantics, pragmatics, generative linguistics, implicit learning.

Introduction

In his last book, Ernst Moerk laments:

> How could a field that is between 100 and 200 years old, whose data are so abundantly and so readily at hand, and which has produced impressive evidence for the wealth of input and its effects, be at present still in a state where almost everything is controversial and where misleading conclusions are so predominant? While year in and year out about two billion young people acquire the various levels of widely differing, and therefore learned, mother tongues, learnability of language has been seriously questioned and rejected in some quarters. (2000, p.179)

It can be argued that the major reason for this situation is that the linguistic grammatical classes are still viewed as psychologically real and necessary for language acquisition. Whereas the formal concepts forged by linguists may be appropriate for describing sentence relationships, it is dubious that they are used by native speakers. A sequential-associative theory of morphosyntactic functioning, rooted in pragmatics and semantics, may be proposed as a plausible alternative.

Tongue and Language

Theoretical writing in linguistics witnesses a confusion between tongue (the language of a nation, country, etc.) and language (a neuropsychological function). The major caveat arises in keeping with the tradition of generative grammar. Over a period of 50 years, the aim of Chomsky and followers has been to account for a human faculty of language, defined as the ability to produce and understand an infinite number of grammatical sentences.

Few people have realized the unrealistic character of such a research agenda. Linguistics is a hermeneutic of the tongues. It lacks the methodological tools to go beyond description. Linguists have no experimental control over the situations in which language behaviors occur and have no objective methods for validating empirically alternative theoretical models. There exists a belief in that field (assumed uncritically in psycholinguistics) that what is descriptively relevant must be ipso facto appropriate for explaining how real people proceed when producing

sentences. However, to the extent that language functioning is concerned, one is addressing a neuropsychological question calling for a behavioral methodology.

Asking people is enough to convince oneself that native speakers (non-language specialists) ignore grammatical notions. They rely on semantic categories. For example, grammatical subjects are agents or topics of state, verbs specify states, actions or events, clauses express "complete" ideas, etc. Compare with the geometrical definitions in structural linguistics (for example, the reverse-tree scheme for sentence representation): the grammatical subject is the noun head of the noun phrase, located immediately below the symbol for the sentence and there is only one noun in this position.

Such a state of affairs is not alien to the generative linguist. Chomsky (1965) warns:

> Thus a generative grammar attempts to specify what the speaker actually knows, not what he may report about his knowledge. (p. 8)

Assuming for the sake of discussion that the typical native speaker tacitly has at her/his disposal the formal machinery described by generative grammar (disregarding differences between successive versions of the theory), where could such a knowledge come from? Generative linguists and psycholinguists (e.g., Pinker, 1994) insist that syntactic categories cannot be induced from the input given that they are not overtly marked and have no one to single correspondence with the semantic categories. Syntactic categories, it is assumed, must be supplied innately or elaborated under the guidance of innate representations. The trouble is that representational inneism has no empirical foundation. Genes coding for universal grammatical representations have yet to be discovered. It is even doubtful that the genome has sufficient capacity for encoding the huge number of binary decisions that would be necessary to account for a linguistic grammar (Kurzwiel, 2006). One language gene has been identified. FOXP2 is a single, autosomal and dominant gene located on chromosome 7 (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). The FOX genes (forkhead box) are a big family of genes coding for proteins binding to a specific area of DNA and regulating the expression of a number of target genes. The null mutation of a gene from the FOX family can affect a potentially large number of other genes. Many forkheads are critical regulators of embryonic development. In the KE family studied by Lai et al. (2001), members over several generations had a variety of problems with spoken and written language. Corresponding indications emerge from Stromswold's (2001) review of a number of genetic studies of language (concordance analyses of disorders in twins, adoption studies, and linkage studies). Genetic factors account for much of the variance in language abilities among people with disorders and some of the variance in normal people. Nowhere, however, is there a demonstration that genetic factors are involved in coding for the abstract linguistic notions implied by representational inneism. It is more likely that they play a role in the development of the neural structures supporting language functioning. If ordinary speakers have no clear awareness of the descriptive linguistic categories and these do not develop in the brain as a consequence of some particular genetic blueprint, there is no reason to consider that they have a role in the neurobehavioral organization called language.

Morphosyntactic Functioning and Learning

Language pragmatics and semantics interfacing respectively with social and conceptual cognition are assumed to supply the early stages in contemporary models of language production (for example, Levelt, 1989,1999). They are very much part of the behavioral process with the proviso that the message needs to be further patterned according to the requirements of the tongue.

Any communicative intention presupposes a "theory of mind" (Gazzaniga, 2008). In a handful of milliseconds, the communicative intention activates a pragmatic framework specifying the objective of the message and the contrast between old information (what the speaker may reasonably hold to be known or immediately accessible to the addressee) as opposed to new information. This contrast controls the use of the ellipsis and the emphasis (prosodic or syntactic, e.g., the use of the passive voice instead of the active), and new information. Other aspects of the future utterance are also programmed at this stage, such as the illocutionary type (declaration or request), person or participant deixis (distinction between first, second, or third person, or speaker, recipient, and "bystander", respectively), optionally social deixis (e.g., polite forms), time and place deixis (e.g., proximal v. distal reference), polarity (positive v. negative), aspect (e.g., completion, duration, frequency of a given action, state or process), and mood (expressing probability, usuality, obligation, presumption, plausibility, degree, intensity, conditionality, obligation, permission, prohibition, exemption).

Lexical and semantic relational concepts are called upon as soon as the pragmatic framework is activated. Various semantic theories have been proposed (e.g., Fillmore, 1968; Chafe, 1970; Van Valin, 1999). They concord in viewing semantic relations as organized around a small number of verb types (e.g., state, process, action, action-process) and a series of subtypes (see, for example, Chafe, 1970, for more details).

A morphosyntactic component operates on the output of the pragmatic-semantic component for organizing the expression according to the sequential requirements of the tongue. Sahin, Pinker, Cash, Schomer, and Halgren (2009), using intracranial electrophysiology with epileptic patients for clinical evaluation, show that a sequence is implemented in the brain Broca's area revealing distinct neuronal firing for lexical (more or less 200 milliseconds following initiation of the language task), morphosyntactical (320 milliseconds), and phonological activity (450 milliseconds).

Specifying the psychological nature of the morphosyntactic component is one of the major challenges of today language psychology. As semantic relations are not ordered sequentially, a mechanism is needed for translating the semantic fabric into utterances. This is where the so-called functional categories of the linguists were supposed to operate (for example, the notions of "grammatical subject or object of the verb). The problem, as mentioned, is that common language users are not in possession of these notions. Consequently, they must have other means available to the same purpose. One may want to go back to suggestions by Skinner (1957) regarding morphosyntactic patterning as a purely sequential process. Most relevant is the suggestion that:

Verbal responses cannot be grouped or ordered until they have occurred or at least are about to occur (p.332). Osgood (1971) expresses the same opinion that utterances are patterned on line, “Grammars are not time-bound in generating sentences; speakers and hearers operate within time constraints and sentences must be created and understood on a ‘left-to-right’ basis. “ (p. 521). Skinner (1957) also suggested that inflectional morphology correspond to both particular semantic characteristics of the entities referred to and proximal and/or distal associations between words.

Moreover, language production, according to Skinner, is largely formulaic. Some sentences are standard comments or responses to common situations and events. Others are nearly complete “skeletal frames” upon which a specific indication or two may be hung.

Skinner’s proposals were flatly rejected by Chomsky (1959). Formulaism is indeed anathema to the generative school which has always put to the fore the so-called creative aspect of language. Actually, the fact that a language makes an infinite use of finite means in no way rules out a formulaic account of morphosyntactic functioning. Open formulas exploited by language users along with paradigmatic substitutions (see below) supply the computational power necessary to account for creativity in language use. According to Chomsky (1957), the idea that morphosyntactic patterning operates sequentially cannot be accepted because it implies a “finite state Markov process.” (i.e., in producing a sentence, the speaker begins in the initial state, produces the first word, thereby switching into a second state which limits the choice of the second word, etc.; each state encompasses the syntactic restrictions that limit the choice of the next word at this point in the utterance). Chomsky claims that natural languages are not finite state languages, not the least because they exhibit recursive mechanisms of diverse types. It follows that a Markov-process conception of language is irrelevant at least for the purpose of explaining grammar. Actually, there is no obligation for a “left-to-right” syntax to operate strictly linearly, i.e., to pertain to the “simple state grammar” to which Chomsky was alluding in his 1957 essay. Nor is it necessary for a finite state grammar to conform to a simple Markov chain. A finite state syntax may include an unlimited number of recurrent loops generating an infinite number of utterances of infinite length. Moreover, assimilating Markov processes to a simple device only able to generate concatenations of first order, i.e. ruled by transitional probabilities holding between adjacent items, is misleading. Sophisticated Markov processes can deal with the production of sequences of higher order. The system can include probability networks dealing with sequences of words or groups of words. The selective probability of an item in a sequence may depend on the presence or the absence of one or several preceding items according to the notion of “limited horizon”. The states the “machine” passes through may be known (so-called visible Markov models) or not (only some probabilistic function of the state sequence is known; hidden Markov models as currently exploited in speech recognition devices; Kurzwiel, 2006). The reader is referred to Manning and Schütze (1999) for mathematical details.

Utterance production is necessarily sequential but its processing need not be strictly linear. Chomsky’s formulations rule out linear syntax. A sequential syntax is different to the extent that even if the words follow each other, semantic dependencies between them (captured by transitional probabilities) may hold in proximity as well as at a distance. In this respect, it may be relevant (and amusing) to spot another confusion in Chomsky’s 1957 contribution. Grammatical but nonsensical sentences, such as “Colorless green ideas sleep furiously” were

taken therein to suggest that, “Probabilistic models give no particular insight into some of the basic problems of syntactic structure” (p.17).

The alleged reason was that probabilities holding between words are nil or extremely low. In utterances of the kind, however, highly probable associations do exist not between individual words but between the semantic categories to which they belong. In the example above, *Colorless* and *green* are qualities, *ideas* are entities, *sleep* a state, and *furiously* the characteristic of a state (or an “action”). This is these categorical semantic (and, of course, corresponding syntactic) dependencies that render the utterance grammatical.

A reasonable hypothesis is that syntactic patterning proceeds on line following the distributional regularities of the tongue. Pragmatic-semantic preferences may also influence word ordering. For example, in many languages there is a progression in nominal phrases from the kind of element that has the greatest specifying potential (the deictic) to that which has relatively less (for example, the quantifier or the qualifier). In French, the qualifier (in function of epithet as opposed to the attribute that is introduced by a copula) may be placed either before or after the entity (*une belle maison*; *une maison belle*) ; with sometimes noticeable differences in meaning (for example, *un home grand* means a man of an elevated stature, whereas *un grand homme* means a man of exceptional moral qualities). Often the anteposited epithet expresses a meaning that is not literal. There is a formal parallelism between the nominal and the verbal phrase. The verbal phrase begins with the finite verb (i.e., the verbal operator expressing person, number, tense, polarity, and modality, e.g., *can*, *must*), which is the verbal equivalent of the deictic, relating the meaning of the verb to the speaker at the time of enunciation. Corresponding suggestions may be made for other types of phrases. Phrases are ordered in various ways to form clauses depending on the tongue. There is no need to posit complicated linguistic mechanisms operating at hierarchical levels in order to account for the patterning of complex sentences. The editing mechanism controlling concord across clauses in complex sentences proceeds by proximal or distal association. Longer and/or more complex or composite sentences may force the speaker to go back to the departure semantic matrix. Behavioral correlates of this monitoring have been identified (Goldman-Eisler, 1968).

Inflectional morphology reinforces cohesion within and between phrases, clauses, and sentences. The forms vary across formal categories and there are many exceptions. This constrains the language user to a learning case by case or close to as converging analogies can be verified for the regular and some irregular forms. As suggested by Skinner (1957), inflectional morphology proceeds on line. The word inflected at the beginning of the utterance (for pragmatic/semantic reasons) serves as a discriminative flag reminding the speaker to operate cohesively with the following words if the norms of the tongue require that they must be inflected too. Given the arbitrary complexity of that part of the grammar and the high speed of language production (approximately 3.3 words - 12 to 15 phonemes - per second; around 200 words per minute), it is not conceivable that inflectional marking could operate through anything else but a sequential-associative process rendered automatic by repeated practice.

Paradigmatic Substitution

Form - and meaning-based analogies play an important role in phrase, clause, and sentence production. They pertain to the paradigmatic or substitutive axis of language storing the expressive potential of the tongue. The speaker may choose among a large number of alternatives for translating her/his pragmatic-semantic plans and contents into patterned sequences of lexemes. Substitutions are dealt with on line with the proviso that there must be a distributional correspondence between substitutable lexical elements or groups.

Language production is enhanced by its formulaic character. At least a part of the utterances produced are idiomatic to some extent (Wray, 2002), meaning that the lexicon and the seqential pattern are frozen. Frequent lexical combinations are stored in memory as chunks or fomulas which allows to activate them faster. Communicators do not just have words and isolated grammatical devices; rather they have prepackaged linguistic constructions already available. Such constructions are not only idiomatic in the sense of being frozen. They may also correspond to what Skinner (1957) labels “skeletal frames”, involving predetermined slots where to insert newly selected lexical material. Such frames are flexible prepatterned strings. Zipf’s mathematical formula (together with Mandelbrot’s; Manning & Schütze, 1999) show that a limited set of lexical elements with a high frequency of occurrence in a corpus of language accounts for the major part of the token distribution (corresponding to a hyperbolic function). Lexemes are related to each other in groups of limited size (horizontal or hierarchical relations) – for example, the associative norms compiled by Cramer (1968), for American English. Networks of elements are constituted in this way. They supply the material for the skeletal frames.

Learning Morphosyntactic Regulations

In order to explain the existence of implicit morphosyntactic procedural knowledge in humans, one must, almost logically, appeal to implicit learning. Other components interact with this type of learning including:

(a) Innate brain structures devoted to language treatment language from the time of late fetal life.

(b) A proper theory of mind serving as a basis for Interpersonal development and language pragmatics

(c) Relevant cognitive development supporting lexical and relational semantic development.

(d) Parents supplying children with an input adapted to their abilities and adequate corrective feedbacks contingent upon the children’s verbal productions.

Innate Devoted Brain Structures

Neonates recognize their mother’s voice and tongue based on prosodic characteristics (Nazzi, Bertoncini, & Mehler,1998). This ability is a direct consequence of the fetus’ exposure to maternal speech during the last three months of pregnancy. By that time the peripheral and

central auditory systems are fully functional (there is a loss of 40 decibels on the incoming speech frequencies due to the aquatic milieu). Newborns can discriminate categorically between all pairs of sounds of natural languages (Jusczyk, 1997). They react differentially to short sequences of variegated syllables (e.g., *bagaba* v. *babaga*; Marcus, Vijayan, Bandi Rao, & Vishton, 1999). Dehaene-Lambertz, Dehaene, and Hertz-Pannier (2002) have adapted the functional magnetic resonance technology for use with babies. Their data suggest that the left hemisphere is already dominant for the perception of speech sounds in neonates. The neuronal circuits involved in speech perception are functional right from the start. Cerebral electrophysiological data (e.g., mismatch negativity measures, brain event-related potentials) point to the same conclusion (Dehaene-Lambertz, 1997, 2000).

Numerous data document further analytical capacities in slightly older infants (for example, 8-month-olds differentiating newly presented words from older ones based on the serial order of syllables, Saffran, Aslin, & Newport,1996; six-month-old babies differentiating prosodically well- versus ill-formed English clauses even when embedded in sentences; Nazzi, Kemler, Nelson, P. Jusczyk, & A. Jusczyk, 2000; infants discriminating between lexical items randomly selected from maternal speech belonging either to closed linguistic classes, e.g., articles, prepositions, auxiliaries, conjunctions, or to open classes, e.g., nouns, verbs, qualifiers; based on relative length and prosodic characteristics; Shi, Werker, & Morgan,1999).

Infants develop sensitivity to nonadjacent input regularities in the course of the first year (Gomez & Gerken, 1999). Eighteen-month-olds are able to identify sequential dependencies at short left-to-right distances corresponding to groups of words (Gomez, 2002). They are reactive to the relationship between the auxiliary *is* and the morpheme *ing* in the following verb providing that the two elements are not separated by more than three syllables.

The statistical structure of the natural tongues supplies an enormous source of distributional information for syntactic learning that is only beginning to be properly evaluated. (Manning & Schütze, 1999). Sequential regularities can be use for identifying words, groups of words, sequences of groups forming clauses, and for inducing clause organization in complex sentences.

Parental Assistance

The task of breaking the code's regulations is facilitated by the ways parents address their language-learning children. An important literature exists on this topic (see Moerk, 2000). Maximal formal simplification is observed around 12 months of age. Virtually at any moment in development the child is confronted with a language input finely tailored to her/his receptive and productive ability (Cross, 1977; Rondal, 1978). For example, over the course of language acquisition, maternal mean length of utterances (MLU) directed to the child exceeds that of the child by roughly 2.5 points for the younger and one point for the older child (Rondal, 1985). Mother-child evolution in MLU is linearly correlated (product-moment + .69 in Moerk's data - 1975 - between 2 and 5 years; + .55 in Rondal's data – 1978 – between 1 and 3 years; both correlation coefficients being statistically significant). Parental input to language-learning children is found to be almost perfectly grammatical (Marcus, 1993). This observation annihilates one of the favorite arguments of the generative tradition against the mere possibility of syntactic learning, i.e., the poverty of stimulus argument (Pinker, 1994).

A wealth of convergent data leaves little doubt that parental dynamic adaptation facilitates morphosyntactic learning. Children can concentrate their distributional analyses on series of utterances properly calibrated for length and intrinsic complexity; utterances that are proposed within adequate pragmatic and semantic interactive frameworks.

Some authors have suggested that there is no clear indication in the literature regarding the existence of negative feedbacks contingent upon the morphosyntactic dimension of children's utterances. Baker and McCarthy (1981) defined the "logical problem" of language development as the acquisition of a grammatical system in the absence of negative evidence. There is no question, however, that such evidence does indeed exist. As any other language user, parents are guided primarily by truth value and contextual adequacy, less by grammatical preoccupations. Recent data indicate, however, that they reformulate children's grammatically ill-formed utterances much more often than repeating correct ones (for example, Chouinard & Clark, 2003).

Implicit Morphosyntactic Learning

The paradigm of implicit learning and tacit knowledge was first introduced by Reber (1967). In implicit (or procedural) as opposed to explicit (or declarative) learning, the acquisition phase is incidental (i.e., subjects are unaware of the regularities governing the material that they learn) and learning leads to knowledge that is difficult or impossible to access consciously and/or to report verbally. Implicit morphosyntactic learning is computational in the sense that it exploits the stochastic characteristics of data fed into to the system and the relative frequencies of various sequential and associative patterns appearing in the input. Numerous observations confirm that the calendars of grammatical acquisitions correspond to the cumulative effects of the following parameters: (a) the relative frequencies of relevant structures in the input; (b) the intrinsic complexity of the forms to be acquired; and (c) the availability of the corresponding cognitive notions. Numerous observations support point 1 (the other two being self-evident). For example, Moerk (1980) has shown that the inflectional morphemes most often used by the parents of the Harvard children are the first to reach 90% of correct production in obligatory contexts. Lahey, Liebergott, Chesnick, Menyuk, and Adams (1992), and Wilson (2003) concur that differences in parental frequency of use of the grammatical structures play a decisive role in the acquisition calendar. Rowland and Pine (2000) and Rowland, Pine, Lieven, and Theakston (2003) report that the frequency of interrogative and relative WH-terms as well as the inversion or noninversion of the verbal element and the pronoun in parental interrogative clauses predict the corresponding evolution in children's productions.

Implicit learning is concerned with surface features of the material. Some sort of abstraction may be involved with repeated exposure. There is no need to refer to tacit knowledge remote from regularities existing in the learning material. The principles governing associative learning are sufficient to account for the acquisition of complex cognitive structures (Perruchet & Gallego, 1997). Knowledge derived from the regularities present in the input through a general process of implicit learning regards proximal and distal associations, sequential patterns, and distributional and other on line statistics. The present theory implies a direct mapping of semantic relationships (given a series of preliminary pragmatic choices) upon sequences of words and inflectional morphemes. Research on the implicit learning of graphotactic and

morphological regularities in written French and English (Pacton, Perruchet, Cleeremans, & Fayol, 2001; Pacton, Perruchet, & Fayol, 2005; Deacon, Pacton, & Conrad, 2008) shows that rules are not abstracted even after massive amounts of exposure to a rule-based material. Instead subjects (children as well as adults) keep relying on statistical regularities in the material even when they have the opportunity to develop symbolic rules. They deal with associative regularities between orthographic features relating to co-occurrences of letters and morphemes within words, words within clauses, as well as with graphotactic analogies holding between words. I believe that the same conclusion holds for the morphosyntactic aspects of spoken languages.

Conclusion

There is no need for descriptive linguistic categories and mysterious genetic hypotheses in order to account for the human morphosyntactic ability. This core instrument of language is the object of a sequential-associative type of implicit learning enhanced in the way the input is delivered to the children. There exists an innate support to language acquisition but it is not of the representational type as claimed in the generative tradition. Instead, and more consistent with present-day tenets in neurosciences, it corresponds to an evolved sensitivity of devoted brain structures to patterns of word concatenations.

References

Baker, C., & McCarthy, J. (1981). *The logical problem of language acquisition.* Cambridge, MA: MIT Press.

Chafe, W. (1970). *Meaning and the structure of language*. Chicago, IL: Chicago University Press.

Chomsky, N. (1957). *Syntactic structures*. The Hague, The Netherlands: Mouton.

Chomsky, N. (1959). Review of *Verbal behavior* by B.F. Skinner. *Language*, 35, 26-58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chouinard, M., & Clark, E. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30, 637-669.

Cramer, P. (1968). *Word association.* New York, NY: Academic.

Cross, T. (1977). Mother's speech adjustments: The contribution of selected child's listener variables. In C. Snow & C. Ferguson (Eds.), *Talking to children* (pp. 151-188). New York, NY: Cambridge University Press.

Deacon, H., Pacton, S., & Conrad, M. (2008). A statistical learning perspective on children's learning about graphotactic and morphological regularities in spelling. *Canadian Psychology*, 49, 118-124.

Dehaene-Lambertz, G. (1997). Assessment of perinatal pathologies in premature neonates using a syllable discrimination task. *Biology of the Neonate*, 71, 299-305.

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience*, 12, 449-460.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298, 2013-2015.

Fillmore, C. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory* (pp. 1-87). New York, NY: Holt, Rinehart & Winston.

Gazzaniga, M. (2008). *Human*. New York, NY : HarperCollins.

Goldman-Eisler, F. (1968). *Psycholinguistics. Experiments in spontaneous speech*. New York, NY : Academic.

Gomez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431-436.

Gomez, R., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109-135.

Hauser, M., Chomsky, N., & Fitch, W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1559-1569.

Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge, Massachusetts, MA: MIT Press.

Kurzwiel, R. (2006). *The singularity is near. When human transcend biology*. London: Duckworth.

Lahey, M., Liebergott, J., Chesnick, M., Menyuk, P., & Adams, J. (1992). Variability in children's use of grammatical morphemes. *Applied Psycholinguistics*, 13, 373-398.

Lai, C., Fisher, S., Hurst, J., Vargha-Khadem, F., & Monaco, A. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413, 519-523.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. (1999). Producing spoken language: A blueprint of the speaker. In A. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). New York, NY: Oxford University Press.

Manning, C., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Marcus, G. (1993). Negative evidence in language acquisition. *Cognition*, 46, 53-85.

Marcus, G., Vijayan, S., Bandi Rao, S., & Vishton, P. (1999). Rule learning by seven month-old infants. *Science*, 283, 77-80.

Moerk, E. (1975). Verbal interaction between children and their mothers during the preschool years. *Developmental Psychology*, 11, 788-794.

Moerk, E. (1980). Relationships between parental input frequencies and children's language acquisition: A reanalysis of Brown's data. *Journal of Psycholinguistic Research*, 7, 105-118.

Moerk, E. (2000). *The guided acquisition of first language skills*. Stamford, Connecticut: Ablex.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1-11.

Nazzi, T., Kemler, P., Nelson, D., Jusczyk, P., & Jusczyk, A. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, 1, 123-147.

Osgood, C. (1971). Where do sentences come from? In D. Steinberg & L. Jakobovits (Eds.), *Semantics* (pp. 497-529). New York: Cambridge University Press.

Pacton, S., Perruchet, P., Cleeremans, A., & Fayol, M. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, 130, 401-426.

Pacton, S., Perruchet, P., & Fayol, M. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, 76, 324-339.

Perruchet, P., & Gallego, J. (1997). A subjective unit formation account of implicit learning. In D. Berry (Ed.), *How implicit is implicit learning*? (pp. 124-161). Oxford, United Kingdom: Oxford University Press.

Pinker, S. (1994). *The language instinct*. New York : Morrow.

Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855-863.

Rondal, J.A. (1978). Maternal speech to normal and Down's children matched for mean length of utterance. In E. Meyers (Ed.), *Quality of life in severely and profoundly mentally retarded people: Research foundations for improvement* (pp. 193-265). Washington, D.C.: American Association on Mental Deficiency, Monograph Series N°3.

Rondal, J.A. (1985). *Adult-child interaction and the process of language acquisition*. New York: Praeger Press.

Rowland, C., & Pine, J. (2000). Subject-auxiliary inversion errors and WH-question acquisition: What children do know. *Journal of Child Language*, 27, 157-181.

Rowland, C., Pine, J., Lieven, E., & Theakston, A. (2003). Determinants of acquisition order in WH-questions: Re-evaluating the role of caregiver speech. *Journal of Child Language*, 30, 609-635.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.

Sahin, N., Pinker, S., Cash, S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science*, 326, 445-449.

Santelman L., & Jusczyk, P. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, 69, 105-134.

Shi, R., Werker, J., & Morgan, J. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72, B11-B21.

Skinner, B.F. (1957). *Verbal behavior*. Englewood Cliffs, New Jersey: Prentice-Hall.

Stromswold, K. (2001). The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language*, 77, 647-723.

Van Valin, R. (1999). Generalized semanticx roles and the syntax-semantic interface. In F. Corblin, C. Dobrovie-Sorin, & J. Marandin (Eds.), *Empirical issues in formal syntax and semantics* (Vol. 2; pp. 373-389). The Hague, The Netherlands: Thesus.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, United Kingdom: Cambridge University Press.

Acknowledgments

This essay is dedicated to the memory of Ernst Moerk. I am indebted to James Welch for his careful checking of the English language.

Author contact information:

Jean A. Rondal
Department of Cognitive Sciences
University of Liège, Bat. 32; Sart Tilman
4000 Liege, Belgium
E-mail: jeanarondal@skynet.be

Motivational Traps

Ryan Olson

Abstract

The current paper describes an approach for talking with safety professionals or other practitioners about the behavioral foundations of common health and safety problems. Four kinds of Motivational Traps for bad habits are reviewed (Pay-Off, Rare Events, Effort, and Sneaky Traps). This practical taxonomy lives in the introductory slides of the author's presentations for lay audiences, and over the years, most technical language has been "selected out," while more catchy language has survived. The topic is helpful for getting audience members on the hook early during presentations, and given the common origins of many behavioral problems, the approach may be useful for other behavior analysts who are engaged in the process of marketing behavioral management techniques.
Keywords: motivation, health, safety, prevention, practical advice, technical language

---

Introduction

As a behavior analyst working in the area of occupational health and safety part of my job is to persuade people to implement behavioral interventions. When speaking to safety professionals I try to follow the heuristic that my language should only be as technical as the audience requires. In general, the audience does not need to know about establishing operations, discriminative stimuli for punishment, un-signaled avoidance behavior, or analogs to reinforcement. These are things that I should know about in order to design effective interventions. The audience needs something much different: they need to learn in a memorable way that behavior is powerfully affected by its consequences, and to be persuaded that sustained behavioral management processes are necessary to prevent disease and injury. The purpose of the current paper is to share a practical way of talking with people about the common behavioral underpinnings of health and safety problems. Instead of teaching technical operant relations and principles, I describe four Motivational Traps for bad habits (Sneaky, Rare Events, Pay-Off, and Effort Traps). My goal is to use memorable language and examples to convince audience members that their organizational safety and health problems can be solved with behavioral management techniques.

The Case for Using Plain English with Clients

When speaking with potential clients our goal should be to communicate in ways that positively affect their consumption of behavioral technology. There are several good reasons why speaking to people in plain English is better than using technical jargon. The following paragraphs review logical, traditional, and empirical reasons to use language like the Motivational Traps taxonomy in the current paper.

Logical Rationale. Imagine a simplified bridge building scenario that includes a bridge engineer/designer, construction workers, and end users. The engineer/designer must understand physics and mathematics in order to design a safe structure; the construction workers need the skills to follow plans and build a structure that meets technical specifications, but do not need to know how to describe the physics behind the design; the end users don't need to know the science or engineering behind the bridge design or any technical specifications for components, but instead, just need access to an appealing and functioning structure that gets them from point A to point B. The end user is most likely to be concerned with what the bridge will look like and how it will affect traffic, not with the stress-strain

curves for specific construction materials or the size and thickness of strip footings. The citizen wants a "socially acceptable" bridge.

Behavior analysts are like the engineers/designers in the analogy. We need to understand why and how a certain procedure will work, and we use a technical language to communicate with each other about the behavioral principles and functional relations that make our interventions effective. However, unlike the bridge engineer/designer, behavior analysts sometimes require that practitioners who implement our interventions (construction workers in the analogy) speak and write like behavioral engineers, when all a technician needs to know is how to skillfully implement a prescribed intervention process. And, even further removed from the logic of a construction project, behavior analysts sometimes expect their clients (end users) to learn and understand technical terms for operant functional relations, when all the user wants is to get from point A to point B in the most comfortable and convenient way possible. Our clients want a socially acceptable technique for changing behavior, not a behavioral science education. Admittedly, behavioral technology is much harder for an end user to "consume" than a bridge, but end users don't need to speak like a behavior analyst to train their dog, shape their child's behavior, or get their workers to wear personal protective equipment. They just need to know what they must do to generate the behavior they want. The point of the bridge analogy above is that we need to speak in ways that make our techniques easy and appealing to use. In Jon Bailey's (1991) words, "…the problem with behavior analysis is not that we are too technological but rather that we have not realized that we are ultimately in the business of developing a 'consumable' product that must be 'user friendly'…" (p. 445)

Traditional Rationale. The impact of technical language on the lay public has been a concern since the formative years of applied behavior analysis, and notable behavior analysts have developed useful ways of communicating with the public. When Lindsley and colleagues established the first human operant research laboratory at the Metropolitan State Hospital in Waltham, Massachussetts, they named their operation "Studies in Behavior Therapy" (Lindsley, 1991). Lindsley explained that they were concerned about negative reactions to more technical names for the laboratory (e.g., Experimental Analysis of Behavior Laboratory), so they used a simple empirical method to select an appropriate "plain English" alternative. Their method involved having patients, parents, and physicians rank order potential names according to their descriptive value and potential offensiveness, and then the researchers selected the lab name on the basis of those social validity data. The use of plain English with clients became an important subsequent theme in Lindsley's work, and between the years of 1953 and 1991, he and his colleagues used a 10-step empirical method to generate plain English translations for about 40 jargon terms (Lindsley, 1991). Examples of translations included using "before" instead of "baseline," "pinpoint" instead of "target behavior," and "relief" instead of "negative reinforcement." Lindsley gathered data on the effectiveness of using plain English translations with clients during some of his workshops. For example, when comparing two jargon words to their translations, only about 20% of workshop groups correctly matched the jargon words to their Skinnerian meanings, while 90% correctly matched the translations to their appropriate meanings (Lindsley, 1991).

While technical jargon may have the same precise effects on behavior analysts, consumers of behavioral technology can easily misunderstand or react badly to our technical language. In Lindsley's (1991) words, "…*you should select words for their effects on the listener, not for their effects on the speaker* (Skinner, 1957)" (p. 449). For example, to a behavior analyst the word "rate" means the frequency of behavior over time. However, for a member of the lay public the word "rate" has many other popular meanings, including a fixed price, a subjective evaluation of value, or interest percentages. Other terms we use such as manipulate, control, and intervene may cause people to withdraw from rather than approach our technology (Bailey, 1991). While Lindsley's systematic and empirical approach to selecting words to use with clients is unparalleled, other behavior analysts have argued for the importance of tailoring our language for public audiences (e.g., Bailey, 1991) and have developed acronyms and terms that make behavioral technology easier for users to "consume" (Daniels, 1989; Gilbert, 1978;

Haughton, 1981). A good example is Daniel's (1989) PICNIC system for evaluating the power of behavioral consequences by rating them as Positive or Negative, Immediate or Future, and Certain or Uncertain, with the most powerful consequences being PICs or NICs.

Empirical Rationale. The way we talk with clients is more of a professional tradition than it is an empirical field of study. However, communication and information processing research suggests that we should use simple language and pictures when marketing behavioral technology to the public. In the area of occupational safety, Larkin (2007) found that the average reading level of safety communications in several oil and chemical companies was the 16th grade. Based on literacy levels in the US, he estimated that only 4% of supervisors and workers would read, understand, and have their behavior affected by the average safety bulletin. In other words, the bulletins were being written in the language of the speakers (highly educated safety professionals) and not in the language of the listeners (supervisors and line workers). Larkin reviewed data showing that if the text was reduced to the 5th grade level then about 70% of readers would understand the material (Aldridge, 2004; Brownson, 1990), and that including pictures could improve understanding above written text alone by over 100% (Mayer, Bove, Bryman, Mars, & Tapango, 1996).

Research on the cognitive concept of "information overload" supports the general strategy of using pictures and small amounts of plain English text when communicating about technical topics. In this area of study, information overload is said to occur when the quantity and/or difficulty of information presented to a user exceeds the information processing capacity of that user (O'Reilly, 1980). It has been hypothesized that a user experiencing overload must shift thinking resources from *understanding* the information to *selecting* which parts of the information to attend to, which results in less total information "processed" and poorer performance (Iastrebova, 2006). Although the effects of information overload on hypothesized private events are difficult to study, the behavioral effects include decreased information usage and recall, and poorer performance on subsequent tasks related to the information. For example, Mayer et al. (1996) conducted a series of experiments where they measured the effects of different types of information on recall and generalized problem solving. The information focused on teaching the steps of lightning formation to learners with limited knowledge of meteorology. In the first two experiments the authors showed that an illustrated summary with about 50 words produced 50% better recall than a 600 word textbook style passage, and that the illustrations in the short summary were essential for producing good generalized problem solving scores. In a final experiment the authors demonstrated how overloading learners produced stepwise reductions in performance. Conditions evaluated included the original short illustrated summary, the illustrated summary plus 50 extra words, and the illustrated summary plus an 550 extra words. Adding 50 words reduced recall and problem solving by 15% and 24% respectively, and adding 550 words reduced recall and problem solving by 66% and 44% respectively.

Preface to the Taxonomy of Motivational Traps

The way I talk about Motivational Traps has grown from an early interest in behaviors that prevent delayed aversive consequences (Olson, 1997) and exposure to the work of Platt and colleagues in the area of Social and Individual Traps (Platt, 1973). The material lives in the introductory slides of my presentations, and helps me get audience members "on the hook" early in my talks. Over the years, most technical language has been "selected out," while more catchy and memorable language has survived. Many behavioral problems share a fundamental mechanistic origin, which is that at-risk behavior is supported by immediate pay-offs, while desirable prevention behaviors have only small, cumulative, or delayed benefits. Therefore, a memorable way of talking about this dilemma might be useful to behavior analysts working on a range of special problems. On this note, the categories of Motivational Traps were organized to serve a practical function, and not for the purposes of technically defining types of behavioral contingencies or principles that could be tested with experiments. However, there may be

some experimental heuristic value in thinking about the high-level themes associated with behaviors that prevent disease and injury.

My general goal when speaking with safety professionals (or teachers, parents, clients) is to convince them that environmental consequences have powerful effects on behavior, and to do so in the most compelling and memorable way possible. I want them to remember my points, talk about behavior amongst themselves, and seek out more information after I am gone. In Malcom Gladwell's (2000) terms, I am trying to design a speech that is "sticky" (i.e., memorable, useful, repeatable). Within my sphere of practice, the way I have come to talk about Motivational Traps seems to have some stickiness. Audience members have responded well and remember parts of my speech after we meet. In presentations I generally cover the topic in less than five minutes, so the descriptions of Traps in the paragraphs below are not meant to be a verbatim summary of my presentation. In addition, I have written each section as if I were describing the Trap to a safety professional, but with "asides" that are meant for behavior analyst colleagues. This is a somewhat awkward writing style, but it seemed to be the most appropriate given the current topic. In addition, safety professionals have a cultural value that their top priority is to physically engineer hazards completely out of the work environment when it is possible. Therefore, when talking about Motivational Traps, I support this value by emphasizing that engineering hazards out of the environment is the best practice because it prevents Motivational Traps from taking control over behavior. And finally, most at-risk behaviors are encouraged by multiple complex contingencies, so in most cases more than one Trap applies. The point of my pitch, however, is not to describe the technicalities of complex behavior-consequence relationships, but rather, to help the audience understand that *behavioral consequences matter*.

Motivational Traps

Nature is in Control

My first goal is to set the stage that nature is in control. I introduce the topic of Motivational Traps by arguing that nature has stacked the odds in favor of production. The word production could be replaced with "chronic disease" or "risk-taking," or one could write that the odds are stacked *against* "prevention." The point is that *if* an organization has no active process or management system to prevent injury and illness, *then* nature will win, and workers will be injured at predictable rates. In my slides I have pictures of two "tug-of-war" teams, with one side having several team members and the other with just one person who is clearly outmatched and in agony (See Figure 1; See acknowledgment at end of article for source of photos). I argue that we are all built to respond best to the demands of our immediate environment. In general, people work to maximize short-term gain, while discounting the importance or significance of the delayed consequences of their actions (short-term maximizing, delay discounting) (Herrnstein, Loewnstein, Prelec, & Vaughan, 1993). While this orientation is adaptive in many situations, it does not naturally encourage behaviors that prevent disease and injury over the long-term. After emphasizing that nature is in charge I am ready to talk about four types of Motivational Traps for bad habits. The use of both words is purposeful. First, "Motivational" Traps is better than "Contingency" Traps. Contingency has multiple meanings that can be confusing, including planned or organized responses to emergency events (Poling, 2010). And, although behavior analysts might like to argue about how Motivation is a cognitive epiphenomenon resulting from current environmental contingencies, genetics, and learning history, a lay audience member will take the term at face value. A person who hears the word "Motivational" is likely to think, "OK, we're talking about how badly people want to do things." Second, the Trap concept is compelling because people can think about natural forces luring us into harmful situations. Traps are also easily illustrated with analogies for both acute events (traumatic injuries) and chronic conditions (heart disease). For example, it is easy for someone to think of a mouse trap snapping violently shut (acute), or of a slow hike down a deep canyon that goes horribly wrong and results in no easy way out (chronic).

Figure 1. Nature in Control Slide.

Pay-Off Traps

Most behaviors that increase the risk of injury or chronic disease have immediate pay-offs. In other words, it is often more comfortable, convenient, or pleasurable to take risks than it is to engage in health promoting or prevention behavior. For example, Ben and Jerry's ice cream tastes great, a drag on a cigarette delivers a bolus of nicotine to the brain within seconds, and driving fast in a car can be fun and exhilarating. In the workplace the Pay-Off Trap causes trouble because people generally seek out the quickest path to reinforcement, which is evident when people skip steps or short cut safety procedures. While observing in a manufacturing plant, I once watched a worker walk across a steel ground-level conveyor belt to get to his destination instead of walking 30 feet to bypass the hazard. Getting to the destination just a few seconds sooner was enough to cause the worker to take a short cut over a significant slip hazard. The point is that people are sensitive to maximizing even minor pay-offs in the immediate environment, and that these small behavioral shortcuts can substantially magnify hazard exposure. I rarely, if ever, use the word reinforcer when talking about pay-off traps. The word pay-off is good enough and has more memorable associations, such as winning a lottery or experiencing a natural or chemical high. Moreover, I think it is an easy way for professionals to start thinking diagnostically about workplace behavioral problems. They can ask themselves, "What's the pay-off when our workers do (insert your own problem behavior)?"

Rare Events Traps

Bad things just don't happen that often, even in the most dangerous occupations. After sharing this general truism, I will usually illustrate the Rare Events Trap with statistics from an occupation with elevated injury or fatality rates. For example, almost 20% of *all* annual workplace fatalities are occupational drivers, however, the chances of any individual driver being killed on the job is about 28 in 100,000 each year (Bureau of Labor Statistics, 2009). Drivers who wear safety belts are 5 times less likely to be killed in a serious crash than unbelted drivers (Sivak, Schoettle, & Rupp, 2010), but most individuals never have a serious crash where they personally experience a belt's protective benefits. Rare Events Traps are also relevant to slowly developed chronic conditions. For example, smoking is responsible for 85% of all lung cancer cases, but only 15 out of 100 smokers eventually develop lung

cancer (The Society of Thoracic Surgeons, 2010). On this theme, my undergraduate mentor Carl Cheney used to ask us "what is the consequence for avoidance behavior?" Most of the students in class were, of course, afraid to answer one of Carl's loaded rhetorical questions, so he would answer himself by saying "Nothing bad happening! Is nothing a consequence? No. Can nothing shape behavior? Not very well." Later on, Carl taught us about the complexities of avoidance behavior, including unsignaled acquisition of avoidance responses (Sidman, 1953), but the lesson he pushed in his introductory classes was that prevention behaviors often have no perceptible effect on the immediate environment and no perceptible relation to the delayed aversive consequences they prevent. In other words, avoidance behaviors are generally on extinction or exceedingly lean reinforcement schedules. So, the take home point for lay audience members is that we just aren't built to respond very well to non-events. I don't have to tell them about extinction, reinforcement schedules, or anything of the sort. People understand that "nothing bad happening" or "business proceeding as usual" is not going to have powerful effects on behavior.

Effort Traps

People are built to find the path of least effort for reaching a goal. In most situations, it is good for people to conserve energy by finding the most efficient way to get things done. In pre-industrial ages this was adaptive and had survival value, but in the modern petroleum age where we have easy access to calorie dense foods and motorized transport, it is a recipe for obesity, diabetes, heart disease, sleep apnea, and so on. Exercise is the best example of an Effort Trap because by definition it burns more calories than more sedentary behavioral options such as sitting or driving. In occupational settings our susceptibility to this *Principle of Least Effort* is dangerous because it can discourage the use of personal protective equipment. For example, 47% of fatal falls in construction are due to failure to use or improper use of fall protection equipment *that was available at the worksite at the time of the event* (National Institute for Occupational Safety and Health, 2000). Why does this happen? Well, from a behavioral perspective, it is more effortful to work while wearing a fall protection harness. The worker must clip, unclip, and then re-clip the tie off as he or she moves about the work area. In this regard it is useful to think of the Effort Trap producing its effects in tandem with the Pay-Off Trap, where people engage in low effort at-risk behaviors that produce small but meaningful pay-offs.

Another example of an Effort Trap is using recommended body mechanics when lifting, which is usually described as bending at the knees, holding the object close to your trunk, and lifting with the legs. In graduate school I took a "Physiology of Work" course from Tycho Fredericks who demonstrated the Effort Trap by having us measure each other's oxygen consumption as we lifted boxes using different body mechanics techniques. The prescribed "safe" technique for repeatedly lifting a box from the floor to another level required significantly more oxygen than bending at the waist. It should therefore be no surprise that poor body mechanics are common during manual material handling; it feels and literally is easier to bend at the waist. *The Effort Trap is a compelling reason to place a priority on engineering the physical environment so that the risky behavior is not possible.* In other words, rather than relying on people to overcome the Effort Trap and use proper body mechanics, it is better to design an environment where people are not required to lift boxes from the floor to another level, or where they can use tools to move material without bending over. Smart engineering of the physical environment can often generate better behavior than reinforcement-oriented management processes. In fact, environmental alteration is the best kind of behavioral technology because it is almost as easy to use as a bridge. Our end users can get from point A to point B without having to manage or self-manage behavior, which as we all know, is "effortful" work!

Sneaky Traps

Many of the most devastating diseases and injuries arise gradually in tiny, imperceptible steps. If the audience seems receptive (I detect no SD^{P+}), I sometimes refer to the Sneaky Trap as the "boiling frog

problem." This is a politically incorrect but extremely memorable analogy I once heard that goes as follows. If you want to boil a frog, you have to put the frog in tepid water and then gradually ramp up the heat. If you throw him in boiling water, he'll jump right out. Nobody in their right mind would throw themselves into "boiling water," but many of us are sitting in that tepid water and ramping up the heat with our bad habits. For example, 1 lb of fat represents 3500 stored kilocalories. Exceeding energy requirements by 500 calories (about one large soda) on about two days each week could cause a person to gain about 1 pound of fat per month. Over a year, a person could gain about 12 lbs. Over 10 years, a person could end up 120 lbs heavier and meet the criterion for class III obesity (See Figure 2 for an example of a Sneaky Traps slide). In the occupational domain cumulative trauma injuries have a similar slow onset. For example, a professional driver with a poor quality seat could be exposed to elevated levels of harmful whole body vibrations (Bovenzi & Betta, 1994; Bovenzi & Zadini, 1992; Pope, Wilder, & Magnusson, 1998). However, back pain and detectable damage to the body might not be evident for many years down the road. Sneaky Traps are perhaps the most deadly and difficult to overcome because once you have been trapped the way out is long and hard. Sometimes it is simply impossible to reverse the aversive condition. These traps epitomize the analogy I shared in the Introduction where a person slowly hikes down a beautiful and cool canyon only to find out later that there is no easy way (or no way) to climb out.

Figure 2. Sneaky Traps Slide.

When Leaders get Trapped

When a leader falls victim to a Motivational Trap, the effects are magnified throughout an organization. Because workers are injured or made sick more often than managers, I spend a lot of time talking about the plight of workers on the front line of production. However, the upstream decisions made by leaders have significant downstream consequences for workers. In this sense line worker behavior is the "final common pathway" (p. 19) rather than the fundamental cause, of injuries and illness (Krause, 1997). What this means is that complex factors at every level of the organization ultimately

converge upon and create the environmental and social conditions experienced by workers on the ground, and that line worker behaviors are therefore only the final step in organizational causal chains. Do leaders make investments in engineering controls for physical hazards? Do leaders set a pace of production that does not harm workers' bodies? Do leaders purchase convenient and comfortable personal protective equipment and reinforce its use? Do leaders make healthy food choices available, and incorporate opportunities for exercise into the daily work routine? The answer to these questions is often "no" because leader behavior is also subject to Motivational Traps. Leaders can fail to make investments in safety equipment and machinery maintenance in order to maximize short-term profits (Pay-Off Trap). A leader may cancel an expensive training program on handling hazardous chemicals because there has never been a chemical injury at a plant (Rare Events Trap). A leader might sit in his or her office instead of spending time interacting with workers, monitoring performance, and reinforcing healthful and safe practices (Effort Trap). A leader might tolerate a hazardous condition, such as elevated exposure to whole body vibration among workers, because the negative consequences are only experienced gradually over decades (Sneaky Trap). The point is that leaders, who often attend our talks as behavior analysts, should consider how their own behavior is affected by Traps.

Discussion and Conclusion

Future Directions

As noted above, the way we talk with potential clients about behavioral technology is a professional tradition more than it is an empirical area of study. As described above in the Introduction, I developed the taxonomy of Motivational Traps unsystematically through trial and error, and have only anecdotal evidence that it is effective for helping clients understand and get interested in behavioral management techniques. However, we would probably become much better at disseminating behavioral technology if we used the scientific method to understand how our professional speech affects our listeners' behavior. In this regard, we should encourage more use of Lindsley's (1991) 10-step empirical approach to selecting plain English translations of behavioral terms and tactics. In addition, technologies are available that make it possible to collect in vivo evaluation data during presentations. For example, radio frequency (RF) cards can be used with Turning Point® and PowerPoint® software to ask audience members multiple choice questions as a presentation is taking place. With RF cards the effects of a Motivational Traps talk could be easily compared to the effects of an alternative more technically oriented presentation. In addition to gathering reaction criteria, behavior analysts should be encouraged to evaluate their professional presentations and workshops with all four levels of Kirkpatrick's (1959) criteria for evaluating training effectiveness: (1) reaction, (2) learning, (3) behavior change, and (4) results. With regard to learning, how many concepts can participants list and define at the end of a talk? How many questions can they answer correctly before and after the training? What can participants recall several weeks after a presentation? With regard to behavior, direct measurement of change is difficult. However, attendees could be surveyed several weeks later about whether they applied any recommended techniques or sought out additional behavioral reading material. And finally, do our presentations produce results of value to the presenter or the attendees' organizations? Did it result in any actual implementations of behavioral interventions? Future research in this area seems essential if we want our technology to be widely recognized and applied by people around the globe (Poling, 2010).

Conclusion

It is useful and important for behavior analysts to package their expertise for the consumption of practitioners and end users. In the area of occupational health and safety, I've found it useful to package information about the behavioral underpinnings of diseases and injuries as Motivational Traps. The goal is to deliver a "sticky" speech that convinces audience members of the power of behavioral consequences,

and hopefully causes them to seek out more information. I hope some of my colleagues will find this "practical taxonomy" useful and relevant to their own behavioral engineering practice. The take home points of my approach are:

- *Nature is in Control:* Nature has stacked the odds in favor of risky behavior
- *Getting Trapped*: Motivational Traps lure people into behaviors that generate sudden or violent aversive consequences like a mousetrap (acute injury), or that generate cumulative aversive consequences like a hike down a canyon that results in no easy way out (chronic disease).
- *Pay-Off Trap*: Most behaviors that increase the risk of injury or chronic disease have immediate pay-offs. Even small pay-offs can trap people in bad habits that substantially increase hazard exposure.
- *Rare Events Trap:* Bad things just don't happen that often, even in the most dangerous occupations. Our behavior is not affected powerfully by "non-events."
- *Effort Trap:* People are built to find the path of least effort for reaching their goals. It is best to engineer the environment in ways that eliminate Effort Traps altogether.
- *Sneaky Trap:* Many of the most devastating diseases and injuries arise gradually in tiny, imperceptible steps. Sneaky Traps are very hard, and sometimes impossible, to escape.
- *Leaders are Human Too*: When a leader falls victim to a Motivational Trap the effects are magnified throughout an organization. Although worker behavior is the final common pathway to illness and injury, leaders ultimately have the most power to shape the work environment and the behavior it generates.

References

Aldridge, M. D. (2004). Writing and designing readable patient education materials. *Nephrology Nursing Journal, 31*(4), 373-377.

Bailey, J. S. (1991). Marketing behavior analysis requires different talk. *Journal of Applied Behavior Analysis, 24*(3), 445-448.

Bovenzi, M., & Betta, A. (1994). Low-back disorders in agricultural tractor drivers exposed to whole-body vibration and postural stress. *Applied Ergonomics, 25*(231-240).

Bovenzi, M., & Zadini, A. (1992). Self-reported low back symptoms in urban bus drivers exposed to whole-body vibration. *Spine, 17*, 1048-1058.

Brownson, K. (1990). Literacy: A problem managers must handle. *Hospital Material Management Quarterly, 20*(1), 37-47.

Bureau of Labor Statistics. (2009). Worker safety and health. *Spotlight on Statistics, June 2009*. Retrieved from http://www.bls.gov/spotlight/2009/safety_and_health/

Daniels, A. C. (1989). *Performance Management*. Tucker, GA: Performance Management Publications.

Gilbert, T. F. (1978). *Human Competence*. New York: McGraw-Hill.

Gladwell. (2000). *The tipping point*. New York: Little, Brown and Company.

Haughton, E. C. (1981). REAP. *Data Sharing Newsletter*.

Herrnstein, R. J., Loewnstein, G. F., Prelec, D., & Vaughan, W. (1993). Utility maximization and melioration: Internalities in individual choice. *Journal of Behavior and Decision Making, 6*, 149-185.

Iastrebova, K. (2006). *Information overload: The impact of coping strategies on decision-making performance.* Unpublished Dissertation, Erasmus University, Rotterdam.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the Society for Training and Development, 13*(9), 3-9.

Krause, T. R. (1997). *The behavior-based safety process: Managing involvement for an injury-free culture*. New York: Van Nostrant Reinhold.

Larkin, T. J., & Larkin, S. (2007). You know safety but admit it...you don't know communication. New York City: Larkin Communication Consulting.

Lindsley, O. R. (1991). From technical jargon to plain English for application. *Journal of Applied Behavior Analysis, 24*(3), 449-458.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapango, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*(1), 64-73.

National Institute for Occupational Safety and Health. (2000). *Worker deaths by falls: A summary of surveillance findings and investigative case reports*. Cincinatti, OH: NIOSH.

O'Reilly, C. A. I. (1980). Individuals and information overload in organizations: Is more necessarily better? *Academy of Management Journal, 23*(4), 684-696.

Olson, R. (1997). *The contingency trap: An analysis of pervasive problems and the application of measures to avoid delayed aversive consequences.* Unpublished Undergraduate Honors Thesis, Logan, Utah.

Platt, J. (1973). Social traps. *American Psychologist, 28*, 641-651.

Poling, A. (2010). Looking to the future: Will behavior analysis survive and prosper? *The Behavior Analyst, 33*(1), 7-17.

Pope, M. H., Wilder, D. G., & Magnusson, M. (1998). Possible mechanism of lower back pain due to whole body vibration. *Journal of Sound and Vibration, 215*(687-698).

Sidman, M. (1953). Two temporal parameters of the maintenance of avoidance behavior by the white rat. *Journal of Comparative and Physiological Psychology, 46*, 253-261.

Sivak, M., Schoettle, B., & Rupp, J. (2010). Survival in fatal road crashes: Body mass index, gender, and safety belt use. *Traffic Injury Prevention, 11*(1), 66-68.

Skinner, B. F. (1957). *Verbal Behavior*. Englewoood Cliffs, NJ: Prentice-Hall.

The Society of Thoracic Surgeons. (2010). What percentage of smokers get lung cancer? *Ask Dr. Oz* Retrieved July 30, 2010, from http://ask.doctoroz.com/question/smokers-with-lung-cancer

---

Acknowledgment

Photos retrieved from www.eefjedehaan.nl and reproduced with the photographer's permission.

Author Contact Information

Ryan Olson, PhD
Assistant Scientist
Center for Research on Occupational & Environmental Toxicology
Oregon Health & Science University
3181 SW Sam Jackson Park Rd, L606
Portland, OR 97239-3098
Phone: (503) 494-2501
Fax: (503) 494-4278
E-mail: olsonry@ohsu.edu

A Quantitative Analysis of Language Interventions for Children with Autism

Meghan Kane, James E. Connell and Melanie Pellecchia

Abstract

Research and services continue to expand to community-based programs serving individuals diagnosed with autism. A focus of great interest in those efforts is that of language acquisition and functional usage. For the purpose of this evaluation, language acquisition interventions are grouped into two broad categories, contrived and naturalistic. Contrived interventions include those in tightly controlled settings with specific instructional sequences (e.g. discrete trials), whereas naturalistic interventions are defined as less structured and with instructional sequences that change with the child's interest. This meta-analysis focuses on interventions of spoken language developed for children diagnosed with autism. Contrived teaching approaches were compared to naturalistic approaches during intervention, generalization, and maintenance conditions by calculating the percentage of non-overlapping data (PND). The results illustrate that naturalistic interventions are more effective than contrived interventions for teaching language skills and for maintaining the skills learned. The data also indicate that contrived approaches result in greater generalization than naturalistic approaches. Implications and future directions for the acquisition and generalization of acquired language skills and the importance of treatment integrity are discussed.
Keywords: autism, contrived teaching, discrete trial teaching, DTT, language interventions, meta-analysis, natural environment teaching, naturalistic, NET

Autism is not only of the most prevalent developmental disabilities, but it is also the fastest growing according to the Autism Society of America, (2008). The Center for Disease Control and Prevention statistics indicate that 1 in 150 8-year-old children in the United States have an autism spectrum disorder (CDC; 2007). The rising incidence may be due to increased awareness, early identification markers and screenings, and more sensitive and specific assessment diagnostic instruments. As the number of children with the disorder rises, so too does the need for qualified therapists and effective interventions to maximize each child's full potential. As such, the CDC recommends that individuals diagnosed with autism receive evidence-based, early intervention services as soon as possible (CDC; 2007).

According to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; 1994), autism spectrum disorder (ASD) is described by significant deficits in three behavioral domains: 1) qualitative impairment in social interaction, 2) qualitative impairment in communication, and 3) restricted, repetitive and stereotyped patterns of behavior, interests, and activities. Communication deficits include a delay in the development of spoken language. When language does develop, impairments in conversational language occur with high frequency (DSM-IV; 1994). Many instructional models (e.g. discrete trial teaching, incidental teaching, pivotal response training) specifically teach spoken language as part of the instructional sequence to remediate these debilitating communication deficits.

As stated above, there are multiple approaches used to teach individuals with developmental disabilities spoken language just within the field of applied behavior analysis (ABA). Historically, the approach most associated with ABA is discrete trial teaching (DTT). Discrete trial teaching is a systematic and structured teaching methodology, consisting of "discrete" trials. A discrete trial consists of one concise instruction, a learned response, and a consequence highly controlled by an instructor. Discrete trial teaching sessions generally occur

at an isolated table in a designated area of a home or school and thus the model has received significant criticism over the years (Steege, Mace, Perry, & Longnecker, 2007). Therefore, for purpose of this investigation, all research using DTT is referred to as *contrived approaches* because the instructional strategy is not "typical" of a naturalistic setting.

In addition to the setting and approach being contrived, discrete trial critics have argued that there is a lack of skill generalization, that the instructional approach only produces rote responding, and that there is an inability to teach sequential chains since instruction only occurs as discrete trials (Steege et al., 2007; Sundberg & Partington, 1998). For these reasons, critics have referred to DTT as an analog training condition and not likely to generalize to natural contingencies of reinforcement.

Autism interventions have also evolved to address some of the criticisms of DTT and to better address the needs of the population. Some researchers worked toward altering teaching strategies in hopes of achieving more promising outcomes. For instance, Koegel, O'Dell, and Koegel (1987) conducted a study in which they manipulated teaching variables to include more functional teaching stimuli, naturalistic reinforcers, and teaching within the natural environment. The results of this study suggest that these teaching methods resulted in greater generalization of language skills. In addition, other ABA approaches have emerged that specifically focus on training in the natural environment and are thus called, *naturalistic approaches*. Naturalistic approaches include incidental teaching (Hart & Risley, 1975, 1982), natural environment teaching (Sundberg & Partington, 1998), pivotal response training (Koegel, Koegel, Harrower, & Carter, 1999), and enhanced milieu teaching (Hancock & Kaiser, 2002). The term naturalistic is used because the stimuli, prompts, and reinforcer availability occur naturally in the student's environment. Naturalistic programming is generally considered more child-centered than contrived approaches. For example, instructors will often follow the child's lead in a play situation and look for occasions to provide instruction in language skills not yet secure in the student's repertoire. Additionally, there is an assumption with this approach that the acquired skills will generalize more readily than contrived approaches because the reinforcement for correct responding is accessible in the natural environment (Cowan & Allen, 2007; Sundberg & Partington, 1998). Finally, proponents of naturalistic approaches argue that acquired language is less robotic and rote than language acquired through a contrived approach.

There is little extant research evaluating the relative effectiveness of the approaches mentioned above. Additionally, a comprehensive search reveals no quantitative analyses comparing these approaches. Comparative reviews typically fall into two broad categories, literature reviews and meta-analyses. Literature reviews synthesize and summarize previous evaluations but offer no quantitative analyses. Whereas, a meta-analysis is a method of synthesizing, quantifying and summarizing the results from independently conducted evaluations. Glass (1976) argued that a meta-analysis is the statistical approach to review extant data that is superior to narrative reviews. He explained that meta-analysis is the "analysis of analyses."

Scruggs, Mastropieri, and Castro (1987) outlined a procedure for quantitatively evaluating the effectiveness of treatments for single-subject research. The percentage of non-overlapping data points (PND) is a non-parametric measure that examines the difference between

the baseline and treatment phases of individual studies. Kazdin (1978) suggested that evaluating non-overlapping data between baseline and intervention phases is a reliable method of evaluating treatment effectiveness.

Scruggs, Mastropieri, Forness, and Kavale (1988) conducted a meta-analysis using PND to assess early language interventions for children identified as handicapped. The investigation focused on the effects of language interventions on frequency of verbalizations. The researchers concluded that interventions which included specific generalization procedures produced more favorable outcomes than interventions involving the "train and hope" philosophy for increasing frequency of verbalizations for that population.

Other researchers have used PND to evaluate other topics related to skill acquisition and programming for individuals diagnosed with autism. Bellini, Peters, Benner, and Hopf (2007) analyzed social skills interventions for children with autism in school settings. The authors concluded that social skills interventions in schools were minimally effective. Bellini, and Akullian (2007) compared video modeling and video self-modeling for the ASD population by computing PND. They concluded that video modeling and video self-modeling were both effective for social skills acquisition and maintenance. An evaluation of the Picture Exchange Communication System (PECS) literature reported effectiveness for mand initiations (Devis & Tincani, 2008). Using the percentage of non-overlapping data (PND) method, the authors reported that of the studies that were included, PECS proved to be effective in establishing mands (i.e., functional communication).

The present meta-analysis evaluates the relative effectiveness of interventions for children with autism that lead to the acquisition of spoken language. The extant literature included in the investigation was categorized into the two broad approaches described above, contrived and naturalistic. Intervention, maintenance, and generalization effects were measured by calculating the percentage of non-overlapping data in order to answer three questions: a) do contrived approaches demonstrate a greater change from baseline to intervention as expected? b) do naturalistic approaches show comparatively greater generalization?, and c) do the two approaches differ comparatively over time?

Method

Search Criteria

Relevant articles were searched using PsycINFO and Educational Resources Information Center (ERIC) databases by using various combinations of the following search terms: *analog, Asperger's, autism, autoclitic acquisition, communication, contrived, discrete trial, discrete trial instruction, discrete trial teaching, DTI, DTT, echoic acquisition, incidental learning, incidental teaching, intraverbal acquisition, language, language acquisition, mand acquisition, milieu, natural, natural environment teaching, naturalistic teaching, PDD, pervasive developmental disorder, pivotal response, pivotal response training, structured teaching, tact acquisition, and verbal behavior.* In addition, manual searches of the *Journal of Applied Behavior Analysis* (JABA), and the *Journal of Autism and Developmental Disorders* were conducted in order to locate additional studies that were not captured by the original database searches. Finally, the

reference sections of the studies captured by the original searches were reviewed, and all research relevant to this meta-analysis were retrieved.

Next, the first author ensured that the evaluations found from the searches described above met the focus of the study (i.e., inclusion/exclusion criteria) which was established a-priori. The inclusion/exclusion criteria was established to include only those studies that fell within the scope of the questions raised above, and to include: only research demonstrating experimental manipulations of language acquisition; only research using single-subject designs; multiple baseline, multiple probe, or ABA designs; research participants between the ages of 2 years and 13 years in order to specifically focus on children; research participants that were diagnosed with ASD, Pervasive Developmental Disorder (PDD), or Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS); and finally each study had to present a graphic display of the data in order for PND values to be calculated (see Figure 1).

Authors/Year	Research Design	Participants Number/Age/Diagnosis	Dependent Measures	Type of Intervention	Phases Included in Analysis
Krantz, Zalenski, Hall, Fenske, & McClannahan (1981)	Multiple baseline	4 children with autism (age 7-13)	use of label, color, size, shape, & verb	contrived	Baseline Intervention
Woods (1984)	Multiple baseline	2 children with autism (age 6)	verbal tacting	compare contrived & naturalistic	Baseline Intervention
McGee, Krantz, & McClannahan (1985)	Multiple baseline	3 children with autism (age 8-11)	preposition use	compare contrived & naturalistic	Baseline Intervention Generalization
Yamamoto, & Mochizuki (1988)	Multiple baseline	3 children with autism (age 10-11)	"give me" and "that's not it" responses	contrived	Baseline Intervention Generalization Maintenance
Miranda-Linne, & Melin (1992)	Multiple baseline	2 children with autism (age 10-12)	use of color adjectives	compare contrived & naturalistic	Baseline Intervention Maintenance
McGee, Almeida, Sulzer-Azaroff, & Feldman (1992)	Multiple baseline	3 children with autism (age 3-5)	peer manding	naturalistic	Baseline Intervention
Pierce, & Schreibman (1995)	Multiple baseline	2 children with autism (age 10)	initiating and maintaining conversation with peers	naturalistic	Baseline Intervention Generalization Maintenance
Taylor, & Harris (1995)	Multiple baseline	3 children with autism (age 5-9)	mands for information	compare contrived & naturalistic	Baseline Intervention

Pierce, & Schreibman (1997)	Multiple baseline	2 children with autism (age 7-8)	initiating and maintaining conversation with peers	naturalistic	Baseline Intervention Generalization Maintenance
Koegel, Camarata, Koegel, Ben-Tall, & Smith (1998)	ABA design*	5 children with autism (age 3-7)	intelligibility of various speech sounds	compare contrived & naturalistic	Baseline Intervention
Buffington, Krantz, McClannahan, & Poulson (1998)	Multiple baseline	4 children with autism (age 4-6)	pairing verbal responses with gestures	contrived	Baseline Intervention Generalization
Finkel, & Williams (2001)	multiple baseline	1 child with autism (age 6)	intraverbal responses	contrived	Baseline Intervention Maintenance
Sarokoff, Taylor, & Poulson (2001)	multiple baseline	2 children with autism (age 8-9)	conversational exchanges	naturalistic	Baseline Intervention
Hancock, & Kaiser (2002)	multiple baseline	4 children with autism or PDD (age 2-4)	conversational utterances	naturalistic	Baseline Intervention Maintenance
Jennett (2005)	multiple probe	6 children with autism or PDD-NOS (age 3-5)	frequency of mands	compare contrived & naturalistic	Baseline Intervention Generalization
Koegel, Carter, & Koegel (2003)	multiple baseline	2 children with autism (age 4-6)	correct use of targeted morphemes	naturalistic	Baseline Intervention Generalization
Esch, Carr, & Michael (2005)	multiple baseline	3 children with autism (age 6-8)	echoic responses	contrived	Baseline Intervention
Endicott, & Higbee (2007)	multiple baseline	4 children with autism (age 3-5)	mands for information	compare contrived & naturalistic	Baseline Intervention Generalization
Perez-Gonzalez, Garcia-Asenjo, Williams, & Carnerero (2007)	multiple baseline	2 children with autism (age 6-8)	use of intraverbal antonyms	contrived	Baseline Intervention
McGee, & Daly (2007)	multiple baseline	3 children with autism (age 4-5)	use of social phrases	naturalistic	Baseline Intervention Generalization
Naoi, Yokoyama, & Yamamoto (2007)	multiple baseline	3 children with autism (age 4-7)	verbal tacting	contrived	Baseline Intervention Generalization

Jones, Feeley, & Takacs (2007)	multiple probe	2 children with autism or PDD-NOS (age 3)	intraverbal responses	contrived	Baseline Intervention Generalization

Figure 1 – Details pertaining to studies included in the meta-analysis. PDD = pervasive developmental disorder; PDD-NOS = pervasive developmental disorder-not otherwise specified.

Some of the most well known discrete trial evaluations (Lovaas, 1987; McEachin, Smith, & Lovaas, 1993) did not meet the inclusion criteria for the meta-analysis. These studies were excluded largely because they used a group design, rather than a single-subject design. Additional studies were excluded if they didn't specify the diagnosis of the research participants, didn't provide graphic displays of data, didn't involve measures of spoken language, or focused on research participants that were dually diagnosed.

Categorization of Included Evaluations (Independent Measures)

Each study was then classified as either "contrived", "naturalistic" or both "contrived and naturalistic". Contrived approaches (as described above) were operationally defined as highly structured teaching that consists of 1:1 instruction that occurred in an isolated location in a tightly controlled instructional environment. Studies using this approach involved the key components of a discrete trial as outlined by Koegel, Russo, and Rincover (1977). Those components include the presentation of a discriminative stimulus (Sd), a target response (or approximation), the programmatic and systematic delivery of a reinforcing consequence, and a specified inter-trial interval. In this approach, reinforcement for correct responding is not necessarily found in the student's natural environment and the instructor controls access to reinforcers (praise, tangibles, etc.). In contrived approaches, the emphasis is usually placed on the topography of the response (i.e. what the child says), and not necessarily on the function of the response.

"Naturalistic" teaching was operationally defined as a student-centered approach in which the instructor follows the student's lead, and bases instructions on the student's interests. Compared to contrived approaches, naturalistic approaches occur in a more natural setting (e.g., school, home, playground) with an array of play items available for the child to sample. The instructional arrangements are then "naturally occurring" in that the instructor waits for the student to engage in one of the play items, and then uses that item to shape or maintain verbal repertoires. For example, a naturalistic approach to teaching mands for information might include hiding a piece from a student's favorite puzzle. Once the child asks, "where is the puzzle piece?" the instructor provides information leading to the missing item. The function, rather than topography, of language is heavily emphasized in naturalistic approaches. Pivotal Response Training (PRT), Natural Environment Teaching (NET), mand training, milieu teaching, and incidental teaching are all classified as "naturalistic" for the purposes of this study.

Dependent Measures

Examples of dependent variables for contrived approaches included labeling items such as color, size, shape, and verbs (Krantz, Zalenski, Hall, Fenske, & McClannahan, 1981), and

acquisition of intraverbal antonyms (Perez-Gonzalez, Garcia-Ansenjo, Williams, & Carnerero, 2007). For naturalistic approaches, dependent variables ranged from conversational utterances (Hancock & Kaiser, 2002) to peer manding (McGee, Almeida, Sulzer-Azaroff, & Feldman, 1992). Some researchers compared both contrived and naturalistic approaches with the same dependent measures. Examples include the study by Woods (1984) which compared both approaches to evaluate the effects on verbal tacting, and the Endicott and Higbee (2007) study which compared both approaches to evaluate the effects on mands for information (see Figure 1).

Data Analysis: Percentage of Non-overlapping Data Points

The PND values were calculated using the procedure outlined by Scruggs et al. (1987). That is, a ruler was horizontally placed at the highest baseline data point. Then, the number of data points in the treatment phase that exceeded that point were counted. The value of the number of data points in the treatment phase exceeding the highest baseline data point was then divided by the number of total data points in the treatment phase. That value was then multiplied by 100 in order to obtain a percentage (PND = number of treatment data points exceeding the highest baseline data point / total number of treatment data points * 100). For example, if three points in the treatment phase exceed the highest data point in the baseline, and there are four data points total in the treatment phase, the PND value is 75% (3/4 * 100 = 75%). Scruggs and Mastropieri (1998) establish cutoff criteria for evaluating quantitative data using PND: very effective treatments (over 90 PND), effective treatments (between 70 and 90 PND), questionable effectiveness (between 50 and 70 PND), and ineffective treatments (below 50 PND).

Percentage of non-overlapping data scores were calculated from baseline to intervention for each research participant, across all dependent variables, in each of the 22 studies. Additionally, the PND scores were calculated for the generalization and maintenance (i.e., follow-up) conditions when applicable by comparing the appropriate condition (i.e. generalization or maintenance) to baseline conditions (see Figure 1). Research studies that compared both types of approaches were included in the mean PND by first categorizing the interventions in each study, then calculating the values from that intervention using the procedure described above.

Inter-observer Agreement

Two researchers independently evaluated each of the 22 studies to ensure that the inclusion/exclusion criteria were met. The researchers were in 100% agreement that all studies met criteria. Next, both researchers independently classified each experiment into one of three categories (based on the operational definitions above): 1) contrived programming, 2) naturalistic programming, or 3) both contrived and naturalistic programming. The researchers agreed 100% on the categorical grouping.

Finally, PND inter-observer agreement (IOA) was calculated by dividing the number of agreements by the sum of agreements and disagreements, and multiplying that value by 100 (agreements/agreements + disagreements * 100). IOA was calculated for all intervention, generalization, and maintenance effects. The IOA for PND calculations was 90.2%.

Results

A total of 22 studies were found using the search keywords and then met the inclusion/exclusion criteria. Of the 22 studies, 21 were published in peer-reviewed journals between the years of 1981 and 2007. One doctoral dissertation met the inclusion criteria (Jennett, 2005). The 22 studies included an aggregate total of 65 research participants. The mean age of the research participants for the included studies was 6.2 years, with a range of 2 – 13 years (see Figure 1).

The majority of the studies included were multiple-baseline designs (n=19). Two studies used a variation of the multiple baseline design, described as multiple probe designs. One study used an ABA design. Ten studies were classified as contrived, six were classified as naturalistic, and six of the studies compared both types of interventions within the same experiment (Figure 1).

Quantitative Analysis

A total of 103 PND scores were calculated, across participants and dependent variables, for contrived interventions from the 22 included studies (Figure 1). The PND scores for contrived interventions ranged from 0 - 100%, with a median of 90%. The mode for this distribution was 100% (see Figure 2). The mean PND score for contrived interventions was 65% (see Figure 3).

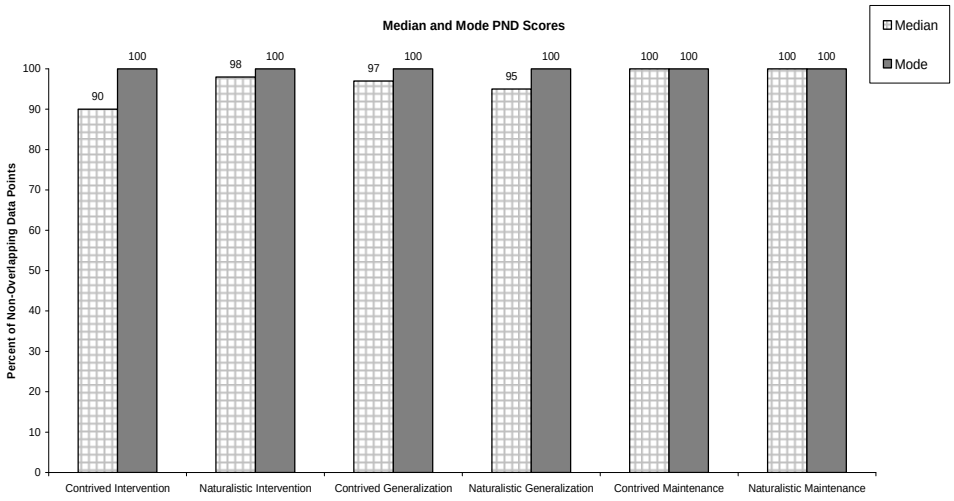

Figure 2 – Median and Mode PND scores for contrived and naturalistic approaches to teaching language skills. Data were collected from the distribution of intervention, generalization, and maintenance effects. PND = Percent of Non-overlapping Data points.

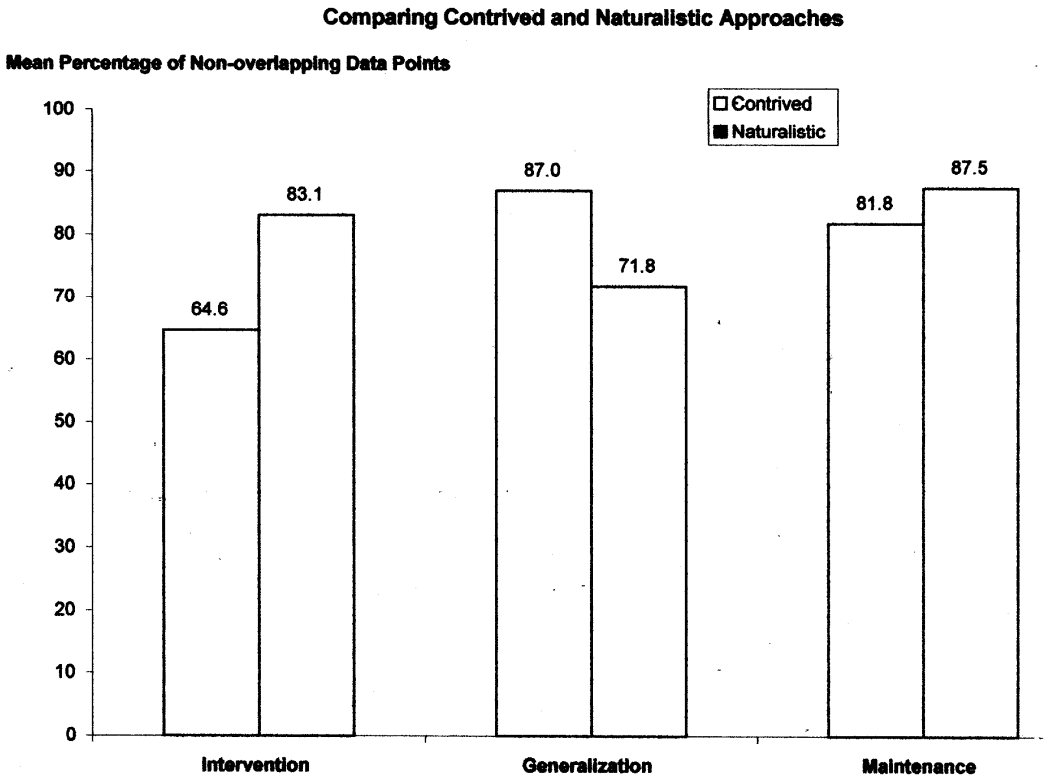

Figure 3 – Mean PND scores for both contrived and naturalistic approaches to teaching language skills across intervention, generalization, and maintenance effects. PND = Percent of Non-overlapping Data.

For naturalistic interventions, a total of 49 PND scores were calculated, across participants and dependent variables, from the 22 included studies (see Figure 1). The PND scores for naturalistic interventions ranged from 0 – 100%. The median for the distribution of PND scores for naturalistic interventions was 98%, and the mode was 100% (see Figure 2). The mean PND value for naturalistic interventions was 83% (see Figure 3).

PND scores were also calculated to determine generalization effects for both contrived and naturalistic approaches to teaching spoken language skills. Seven of the 22 included studies involved graphic displays of generalization effects for contrived programming (see Figure 1). Within those seven studies, 35 PND scores were calculated across participants and dependent variables. The range of PND values for generalization for contrived approaches was 0 – 100%, with a median of 97%. The mode for this distribution was 100% (see Figure 2). The mean PND score for generalization for contrived teaching was 87% (see Figure 3).

Six of the 22 included studies provided graphic displays of generalization results for naturalistic programming (see Figure 1). Of those six research studies, 30 PND scores were calculated across subjects and dependent variables. The range of PND scores for generalization of naturalistic approaches was 0 – 100%, with a median of 95%. The mode for this distribution was 100% (see Figure 2). The mean PND for generalization for naturalistic approaches was 72% (see Figure 3).

Finally, PND scores were calculated for maintenance effects for both contrived and naturalistic approaches to teaching spoken language skills. Only 3 of the 22 included studies provided measures of maintenance for contrived approaches (see Figure 1). Of those three studies, 11 PND scores were calculated across subjects and dependent variables. The range of PND scores for maintenance for contrived approaches was 0 – 100%, and the mode was 100%. The median for this distribution was 100% (see Figure 2). The mean PND for maintenance for contrived programming was 82% (see Figure 3).

Four of the 22 included studies provided measures of maintenance for naturalistic approaches to teaching spoken language skills to children with autism (see Figure 1). Twenty PND scores were calculated from those four studies. The range of PND scores for maintenance for naturalistic approaches was 0 -100%. The median PND score was 100%, and the mode was 100% (see Figure 2). The mean PND score for maintenance for naturalistic programming was 88% (see Figure 3).

Figure 3 compares the mean PND for intervention, generalization, and maintenance effects for contrived and naturalistic approaches to teaching spoken language skills. For intervention effects, contrived approaches scored a mean PND of 65%, while the mean PND for naturalistic language interventions was 83%. For generalization effects, the mean PND for contrived programming was 87%, and the mean PND for naturalistic programming was 72%. The mean PND score for maintenance for contrived approaches was 82%, compared to a mean PND of 88% for maintenance with naturalistic approaches.

Discussion

The results for this meta-analysis reveal that when comparing mean PND scores, naturalistic interventions were more effective than contrived interventions from baseline to intervention and baseline to follow-up conditions, but less effective when comparing the baseline to generalization conditions. These results were somewhat surprising. It was anticipated that contrived approaches would show the greatest change from baseline to intervention, but less change from baseline to generalization. That is, it was expected that the tightly controlled environment of the contrived settings would demonstrate a superior instructional arrangement in that the new skills would be acquired faster and with less variability than the naturalistic approaches. And, that the variability of the naturalistic approaches would promote the generalization of the newly acquired skills, after skill mastery was established, and therefore be apparent in the PND between baseline and generalization conditions comparatively. In other words, naturalistic approaches are designed to mimic conditions in which language is learned naturally, and therefore generalize to natural settings more readily. For example, the first time an infant babbles a sound that resembles “mama” in the presence of his/her mom, mom reacts favorably, and that sound is selected over others in the future. This natural contingency is deliberately mimicked in the naturalistic approaches.

The generalization effects reported for the meta-analysis suggest that the ASD population generalizes spoken language skills more readily with contrived teaching approaches, rather than naturalistic approaches. The mean PND scores for both naturalistic and contrived generalization measures were considered effective by the scale provided by Scruggs and

Mastropieri (1998), although the mean for contrived generalization was 15% higher than the mean for naturalistic generalization. This finding is contrary to what the authors predicted at the onset of this analysis. It was anticipated that language skills learned through naturalistic approaches would readily generalize across settings, people, and/or stimuli, and that contrived approaches would not. The reason for this conjecture is that the instructions, stimuli, and reinforcers used in contrived interventions are highly specialized; thus, the skills learned through contrived programming would not likely generalize to natural contexts.

One possibility for the unexpected results for generalization effects found in the meta-analysis is that a bias exists in the studies that included measures of generalization for contrived teaching. The fact that researchers were testing for generalization effects could imply that programming included specific features conducive to generalization (e.g. fading contrived reinforcers for naturalistic ones; including responding in other settings in mastery criteria). At the very least, the researchers recognized the importance of generalization; hence, generalization effects were being measured. For example, Jones and Feeley (2007) focused on teaching spontaneous responses to verbal and non-verbal cues. The authors clearly describe the intervention approaches as discrete trial teaching. However, the generalization procedures consisted of fading edibles (i.e., contrived reinforcers) for more naturalistic ones, such as the teacher speaking louder in response to the student saying “what?”. Generalization probes were done in a novel setting with novel teachers, and correction procedures were not used. The measures taken for generalization purposes were conducive to generalization of spoken language skills, and all contrived teaching should include such measures. In fact, the generalization procedures described in Jones et al. (2007) closely resemble naturalistic teaching, and the outcome for the research participants was favorable. Another possibility for these surprising results is that the repetition that is characteristic of contrived approaches is more conducive to generalization. In naturalistic teaching methods, the instructor often follows the student’s lead and capitalizes on any teaching opportunities. It is unlikely that there are many repetitive opportunities to practice target skills in naturalistic approaches. However, repetition is intrinsic to contrived teaching approaches.

Although contrived approaches scored a higher mean PND for generalization, naturalistic approaches scored a higher mean PND for maintenance effects. Maintenance measures involve the occurrence of a response over time (Alberto & Troutman, 2006), which is a critical factor when teaching spoken language skills. Children with autism have a tendency to lose language skills over time (Goldberg, Osann, Filipek, Laulhere, Jarvis, Modhal, Flodman, & Spence, 2003). Therefore, research involving acquisition of spoken language should include maintenance conditions and measures of skill maintenance.

A possibility for stronger maintenance effects for naturalistic programming is that the naturally occurring contingencies used during intervention phases continue to be available within the environment over time. The reinforcers in a contrived teaching approach would be difficult to sustain over time, which may lead to lower rates of the target behavior. However, for this analysis, the mean PND for maintenance effects for both types of programming were considered effective according to the scale provided by Scruggs and Mastropieri (1998). The difference between the two means was 5.7%.

Future Directions

Children with autism do not readily learn spoken language without a specific intervention. Today, many interventions for children with autism begin with contrived teaching approaches during acquisition. Clinicians are more likely to move to a more naturalistic approach to train for generalization of language skills. However, the results of the current meta-analysis suggest this method is not optimal for teaching spoken language skills. These findings suggest that perhaps it would be more effective to devise language interventions utilizing naturalistic teaching approaches for acquisition with a progression to more contrived approaches to train for generalization. At the least, further investigation is warranted in order to examine the differences between naturalistic and contrived interventions for learners with autism. It may also be useful to examine different variables (i.e. attention span, stereotypical behaviors, number of established reinforcers, etc.) for this population in order to determine the most appropriate language interventions for individuals. Ultimately, clinicians and practitioners should use research in this area to develop an assessment tool to determine which teaching strategies are most beneficial for individual learners.

Another direction for future research is to investigate to what extent researchers are actually measuring generalization in studies involving spoken language acquisition. For instance, can the child respond to various instructions? Or, are the discriminative stimuli for the target response a set of highly specialized instructions? Another area of interest is for researchers to address the functional nature of spoken language learned through contrived teaching. When a child learns to say “cookie” after being asked, “what do you want?” in a contrived approach, is the word “cookie” functioning as a mand with a specific reinforcer (i.e., access to the cookie)? Or is the word “cookie” functioning as a learned intraverbal response? Another aspect of examining the function of spoken language learned through contrived programming is to address the criticisms that rote, robotic language emerges.

An additional direction for future research is to consider the intervention, generalization, and maintenance effects of various skills taught through contrived and naturalistic approaches. For instance, it would be interesting to examine if skills such as eye contact or imitation are more readily learned through naturalistic programming. It is possible that the best approach varies depending on the skill being taught.

Although the findings of this meta-analysis are important to consider when devising a language intervention for a child with autism, one must consider that outcomes can be idiosyncratic. That is, each child benefits the most from an individualized approach, based on the principles of applied behavior analysis. It is possible that a comprehensive approach is best, including aspects of both naturalistic and contrived teaching approaches (see Sundberg & Partington, 1999). However, each intervention should consist of continuous assessment, empirically validated teaching methods, and data collection. These interventions should also heavily emphasize generalization and maintenance of skills learned. As with any teaching strategy, the skill level of the practitioners is of upmost importance, as well as treatment integrity.

Limitations: Analysis Methods

A standard outcome metric for analyzing single-subject research has not been identified. PND has been criticized in the literature for not representing treatment outcomes accurately (Allison & Gorman, 1994; Levin, 1992). However, as Scruggs and Mastropieri (1987) explain, PND can be easily computed through graphic displays of data, and it is an evaluative method that provides a measure of treatment effectiveness in most cases. PND also allows for a greater number of studies to be included in the analysis than parametric approaches because there aren't constraints regarding normality, homogeneity, and independence of single-case data. Scruggs et al. (1987) remind the reader that there is not one single approach for evaluating data of single-subject designs, just as there is not one approach for evaluating research involving group designs. PND measures are valuable because they can be easily interpreted by the reader. For the purposes of this meta-analysis, PND was the most appropriate metric for evaluating treatment effectiveness.

References

Allison, D.B., Gorman, B.S. (1994). "Make things as simple as possible, but no simpler": A rejoinder to Scruggs and Mastropieri. *Behaviour Research and Therapy, 32,* 885-890.

Alberto, P.A., & Troutman, A.C. (2006). *Applied Behavior Analysis for Teachers, Seventh Edition*. Upper Saddle River, NJ: Pearson Prentice Hall.

Autism Society of America (2008). Facts and Statistics. Retrived November 24, 2008, www.autism-society.org

Bellini, S., Peters, J.K., Benner, L., & Hopf, A. (2007). A Meta-Analysis of School-Based Social Skills Interventions for Children With Autism Spectrum Disorders. *Remedial and Special Education, 28*, 3, 153-162.

Bellini, S., & Akullian, J. (2007). A Meta-Analysis of Video Modeling and Video Self-Modeling Interventions for Children and Adolescents With Autism Spectrum Disorders. *Council for Exceptional Children*, 73, 3, 264-287.

Buffington, D.M., Krantz, P.J., McClannahan, L.E., & Poulson, C.L. (1998). Procedures for Teaching Appropriate Gestural Communication Skills to Children with Autism. *Journal of Autism and Developmental Disorders*, 28, 6, 535-545.

Center for Disease Control and Prevention (2007). Autism information center. Retrieved October 26, 2008, www.cdc.gov/ncbddd/autism/

Cowan, R.J., & Allen, K.D. (2007). Using Naturalistic Procedures to Enhance Learning in Individuals with Autism: A Focus on Generalized Teaching Within the School Setting. *Psychology in the Schools, 44*, 7, 710-715.

Devis, K.Z., & Tincani, M. (2008). *Quantitative Synthesis of Single-Subject Studies on the Picture Exchange Communication System*. Manuscript submitted for publication.

Diagnostic and Statistical Manual of Metal Disorders, Fourth Edition (DSM-IV; 1994). American Psychiatric Association.

Endicott, K., & Higbee, T.S. (2007). Contriving Motivating Operations to Evoke Mands for Information in Preschoolers with Autism. *Research in Autism Spectrum Disorders*, 1, 210-217.

Esch, B.E., Carr, J.E., and Michael, J. (2005). Evaluating Stimulus-Stimulus Pairing and Direct Reinforcement in the Establishment of an Echoic Repertoire of Children Diagnosed with Autism. *Analysis of Verbal Behavior*, 21, 43-58.

Finkel, A.S., & Williams, R.L. (2001). A Comparison of Textual and Echoic Prompts on the Acquisition of Intraverbal Behavior in a Six-Year-Old Boy with Autism. *Analysis of Verbal Behavior*, 18, 61-70.

Glass, G.V. (1976). Primary, Secondary, and Meta-analysis of Research. *Educational Researcher*, 5, 3-8.

Goldberg, W.A., Osann, K., Filipek, P.A., Laulhere, T., Jarvis, K., Modahl, C., Flodman, P., & Spence, M.A. (2003). Language and Other Regression: Assessment and Timing. *Journal of Autism and Developmental Disorders*, 33, 6, 607-616.

Hancock, T.B., & Kaiser, A.P. (2002). The Effects of Trainer-Implemented Enhanced Milieu Teaching on the Social Communication of Children with Autism. *Topics in Early Childhood Special Education*, 22, 1, 39-54.

Hart, B., & Risley, T.R. (1975). Incidental Teaching of Language in the Preschool. *Journal of Applied Behavior Analysis*, 8, 411-420.

Hart, B., & Risley, T.R. (1982). How to use incidental teaching for elaborating language instruction. Austin, TX: PRO-ED.

Jennett, H.K. (2005). The effects of discrete trial instruction and mand training for teaching children with autism to make requests for items. Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.

Jones, E.A., Feeley, K.M., & Takacs, J. (2007). Teaching Spontaneous Responses to Young Children with Autism. *Journal of Applied Behavior Analysis*, 40, 3, 565-570.

Kazdin, A.E. (1978). Methodological and Interpretive Problems of Single-Case Experimental Designs. *Journal of Consulting and Clinical Psychology*, *46*, 629-642.

Koegel, R.L., Camarata, S., Koegel, L.K., Ben-Tall, A., & Smith, A.E. (1998). Increasing Speech Intelligibility in Children with Autism. *Journal of Autism and Developmental Disorders*, 28, 3, 241-251.

Koegel, L.K., Carter, C.M., Koegel, R.L. (2003). Teaching Children with Autism Self-Initiations as a Pivotal Response. *Top Language Disorders*, 23, 2, 134-145.

Koegel, L.K., Koegel, R.L., Harrower, J.K., & Carter, C.M. (1999). Pivotal Response Intervention I: Overview of Approach. *The Journal of the Association for Persons with Severe Handicaps*, 24, 3, 174-185.

Koegel, R.L., O'Dell, M.C., & Koegel, L.K. (1987). A Natural Language Teaching Paradigm for Nonverbal Autistic Children. *Journal of Autism and Developmental Disorders*, 17, 2, 187-200.

Koegel, R.L., Russo, D.C., & Rincover, A. (1977). Assessing and Training Teachers in the Generalized Use of Behavior Modification with Autistic Children. *Journal of Applied Behavior Analysis*, 10, 2, 197-205.

Krantz, P. J., Zalenski, S., Hall, L.J., Fenske, E.C., & McClannahan, L.E. (1981). Teaching Complex Language to Autistic Children. *Analysis and Intervention in Developmental Disabilities*, 1, 259-297.

Kroeger, K.A., & Nelson, W.M. (2006). A Language Programme to Increase the Verbal Production of a Child Dually Diagnosed with Down Syndrome and Autism. *Journal of Intellectual Disability Research, 50*, 2, 101-108.

Levin, J.R. (1992). Single-case research design and analysis: Comments and concerns. In T.R. Kratchowll & J.R. Levin (Eds.), Single-case research design and analysis: New directions for psychology and education (pp. 213-242). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lovaas, O.I., (1987). Behavioral Treatment and Normal Educational and Intellectual Functioning in Young Autistic Children. *Journal of Consulting and Clinical Psychology*, 55, 1, 3-9.

McEachin, J.J., Smith, T., & Lovaas, O.I. (1993). Long-term Outcome for Children with Autism Who Received Early Intensive Behavioral Treatment. *American Journal on Mental Retardation*, 97, 4, 359-372.

McGee, G.G., Almeida, C., Sulzer-Azaroff, B., & Feldman, R.S. (1992). Promoting Reciprocal Interactions Via peer Incidental Teaching. *Journal of Applied Behavior Analysis*, 25, 1, 117-126.

McGee, G.G., & Daly, T. (2007). Incidental Teaching of Age-Appropriate Social

Phrases to Children with Autism. *Research & Practice for Persons with Severe Disabilities*, 32, 2, 112-123.

McGee, G.G., Krantz, P.J., & McClannahan, L.E. (1985). The Facilitative Effects of Incidental Teaching on Preposition Use by Autistic Children. *Journal of Applied Behavior Analysis*, 18, 1, 17-31.

Miranda-Linne, F., & Melin, L. (1992). Acquisition, Generalization, and Spontaneous Use of Color Adjectives: A Comparison of Incidental Teaching and Traditional Discrete-Trial Procedures. *Research in Developmental Disabilities*, 13, 191-210.

Naoi, N., Yokoyama, K., Yamamoto, J. (2007). Intervention for Tact as Reporting in Children with Autism. *Research in Autism Spectrum Disorders*, 1, 171-184.

Perez-Gonzalez, L.A., Garcia-Asenjo, L., Williams, & Carnerero (2007). Emergence of Intraverbal Antonyms in Children with Pervasive Developmental Disorder. *Journal of Applied Behavior Analysis*, 40, 4, 697-701.

Pierce, K., & Schreibman, L. (1995). Increasing Complex Social Behaviors in Children with Autism: Effects of Peer-Implemented Pivotal Response Training. *Journal of Applied Behavior Analysis*, 28, 3, 285-295.

Pierce, K., & Schreibman, L. (1997). Multiple Peer Use of Pivotal Response Training to Increase Social Behaviors of Classmates with Autism: Results from Trained and Untrained Peers. *Journal of Applied Behavior Analysis*, 30, 1, 157-160.

Sarokoff, R.A., Taylor, B.A., Poulson, C.L. (2001). Teaching Children with Autism to Engage in Conversational Exchanges: Script Fading with Embedded Textual Stimuli. *Journal of Applied Behavior Analysis*, 34, 1, 81-84.

Scruggs, T.E., & Mastropieri, M.A. (1998). Synthesizing Single Subject Research: Issues and Applications. *Behavior Modification*, 22, 221-242.

Scruggs, T.E., Mastropieri, M.A., & Castro, G. (1987). The Quantitative Synthesis of Single-Subject Research: Methodology and Validation. *Remedial and Special Education*, 8, 24-33.

Scruggs, T.E., Mastropieri, M.A., Forness, S.R., & Kavale, K.A. (1988). Early Language Intervention: A Quantitative Synthesis of Single-Subject Research. *The Journal of Special Education*, 22, 3, 259-283.

Steege, Mace, Perry, & Longnecker (2007). Applied Behavior Analysis: Beyond Discrete Trial Teaching. *Psychology in the Schools*, 44, 91-99.

Sundberg, M., & Partington, J. (1998). *Teaching language to children with autism and other developmental disabilities*. Pleasant Hill, CA: Behavior Analysts Inc.

Sundberg, M. L., & Partington, J.W. (1999). The need for both discrete trial and natural environment language training for children with autism. In P.M. Ghezzi, W.L. Willaims, & J.E. Carr (Eds.), *Autism: Behavior analystic perspectives* (pp. 139-156). Reno, NV: Context Press.

Taylor, B., & Harris, S.L. (1995). Teaching Children with Autism To Seek Information: Acquisition of Novel Information and Generalization of Responding. *Journal of Applied Behavior Analysis*, 28, 1, 3-14.

Woods, T.S. (1984). Generality In the Verbal Tacting Of Autistic Children As A Function Of "Naturalness" In Antecedent Control. *Journal of Behavior Therapy & Experimental Psychiatry*, 15, 1, 27-32.

Yamamoto, J, & Mochizuki, A. (1988). Acquisition and Functional Analysis of Manding with Autistic Students. *Journal of Applied Behavior Analysis*, 21, 1, 57-6

Author Contact Information

Meghan Kane, M.S.Ed., BCBA
meghan_kane@elwyn.org

James E. Connell, PhD., BCBA-D
jconnell@temple.edu

Melanie Pellecchia, M.S.Ed., BCBA
melaniep@temple.edu

Delay Discounting of Self-Determined and Experimenter-Determined Commodities

Jeffrey N. Weatherly, Jennifer Gudding, and Adam Derenne

Abstract

Research suggests that individuals prefer self-determined reinforcers over experimenter-determined ones. The present study had 518 college students complete a delay-discounting task in which the commodity was cigarettes, a grocery store gift card, casino tokens, cash, or the choice of the four. The least amount of delay discounting was observed for the group that was given a choice, although the difference was not significantly different from two of the individual commodities (casino tokens & cash). The results support the conclusion that participants place a high value on having choice. These results have practical implications for teachers, researchers, and/or businesses who want to provide incentives for their students, participants, or customers, respectively. The results also highlight the usefulness of measuring delay discounting.
Keywords: Delay discounting, Self determination, Value, College sample

---

Self determination has long been a concept of great interest to psychologists. For instance, self determination was a critical concept in the person-centered therapy and theory of personality that was forwarded by Carl Rogers (Patterson & Joseph, 2007). More recently, self-determination theory has emerged (e.g., Deci & Ryan, 2008) as a general theory of human behavior and motivation that has been applied in a wide variety of situations (e.g., education, Kaufman & Dodge, 2009; worker satisfaction & productivity, Kuvaas, 2009). The common theme related to self determination is that it is a good thing; people are happier, more motivated, and more productive when they can exercise self determination than when they cannot.

Although behavioral psychologists have historically eschewed internal or personality variables as explanations for behavior, they too have explored the idea that situations that involve self determination are functionally different than situations that do not involve self determination. For instance, Graff, Libby, and Green (1998) found that participant-selected reinforcers maintained higher rates of free-operant responding, and produced less challenging behaviors, than did experimenter-selected reinforcers in two male participants with severe developmental disabilities. Geckeler, Libby, Graff, and Ahearn (2000) failed to replicate this effect on free-operant responding in three boys with Autism, but did find that when participant- and experimenter-selected reinforcers were available in a concurrent-choice procedure, all three boys showed a response preference for the alternative that allowed them to choose their own reinforcer. These results replicated those of a previous study (i.e., effect of self-determined choice only in the concurrent-schedule situation; Graff & Libby, 1999), which had studied four boys with either developmental disabilities or attention-deficit disorder.

More recently, Tiger, Hanley, and Hernandez (2006) studied the effect of reinforcer choice on the behavior of preschool children. Results indicated that five of the six children showed an initial preference for choosing their own reinforcer, although this preference did not persist throughout the entire condition for several of the children. Tiger et al.'s fourth study demonstrated that the children continued to choose the reinforcer-choice option despite the fact that the response requirement for doing so was higher than the no-choice reinforcer option.

Overall, these studies support the idea that an outcome that allows the individual to determine his/her own reinforcing consequences can be a more effective or preferred reinforcer than the identical outcome that is not chosen by the individual.

Determining whether outcome choice is indeed a more powerful reinforcing consequence than a predetermined outcome has a number of potential implications, especially if that outcome can be demonstrated in an adult sample. For instance, the implication for individuals in the field of marketing would be that offering potential customers a self-chosen prize for visiting a business or website might be a more effective promotion than simply offering a predetermined prize. Instructors might find that students' work improves if their efforts result in a self-chosen outcome rather than an instructor-chosen outcome. Researchers who employ human participants in laboratory studies with the incentive of winning a prize (e.g., a gift card) for good performance may find improved performance if participants are offered a self-determined reward rather than an experimenter-chose one.

The study of delay discount affords one measure of the value of an outcome or commodity. Delay discounting occurs when the value of an outcome is devalued because it is delayed in time. For instance, if someone owed you $100 but was not going to be able to pay you for a month, you might accept $95 immediately rather than waiting a month for the full amount. If so, that outcome would indicate that the delay of one month has discounted the value of the $100 by at least 5%. The typical outcome is that the longer the delay to the full amount of a particular outcome, the more individuals tend to discount its value (e.g., see Chapman, 1996; Beck & Triplett, 2009).

There are multiple ways to measure delay discounting. One is to ask participants to make a series of binary choices (e.g., would you prefer $95 today or $100 in one month; e.g., see Smith & Hantula, 2008) where the amount of the immediately available commodity is adjusted across choices. The subjective value of the commodity at that delay is determined at the point at which the participant switches from preferring the delayed commodity to the immediate one. By making such determinations across different delays, one can calculate a "discounting curve," which determines the "rate" of discounting. Other methods include having the participant identify the indifference point at each delay from a series of experimenter-presented choices (e.g., Beck & Triplett, 2009) or having the participant self generate the indifference point (i.e., the fill-in-the-blank method; e.g., Chapman, 2009).

One factor known to systematically alter rates of delay discounting is the absolute value of the commodity or outcome, a finding that has been labeled the *magnitude effect* (Chapman, 1996; Thaler, 1981). For instance, in the above example, the delay of one month decreased the relative value of $100 by at least 5%. However, if the amount of money you were owed was $10,000 rather than $100, research suggests that you would be less likely to discount that amount by 5% (i.e., $500) over a month. In general, the research literature supports the finding that the higher the value of the commodity or outcome, the less individuals discount it when it is delayed. With that being the case, the measure of delay discounting can be considered a dependent variable for the value of a particular commodity or outcome.

If a self-determined choice of outcome has a higher value than a predetermined outcome, then that difference should be reflected in the rate at which individuals discount those outcomes when they are delayed. The present study was designed to investigate this prediction. A sample of college students was recruited to complete a delay-discounting task. The hypothetical commodity involved in the task varied across groups. For four of the groups, there was one experimenter-determined commodity, which was always valued at $100. The commodity for the fifth group was their choice of the four commodities. Given the theoretical and empirical research that suggests that self-determined choices have a higher value than predetermined ones, we predicted that participants in the choice group would display the least delay discounting relative to participants in the other groups, who were faced with a predetermined commodity.

Method

Participants

The original sample of participants was 571 undergraduates psychology students enrolled at the University of North Dakota. Participants from this original sample were excluded if they failed to complete all of the questions on the delay-discounting task (or provided a value of $0 for every option). When these respondents were excluded, the final sample consisted of 518 participants (330 females, 188 males). The mean age of these participants was 19.61 years (SD = 2.10 years) and they reported a mean grade point average of 3.17 out of 4.00 (SD = 0.55). The vast majority of respondents self identified as Caucasian (472; 91.1%), with 14 (2.7%) individuals self identifying as American Indian, 23 (4.4%) as Asian, Black, or other ethnic minority, and 9 (1.7%) not providing a response. Four hundred ninety five (95.6%) participants reported being single or in a relationship, 10 (1.9%) reported being married, 2 (0.4%) reported being divorced or widowed, and 11 (2.1%) failed to provide a response. In terms of annual income, 453 participants (87.5%) reported an annual income of less the $10,000 per year, 38 (7.3%) reported an annual income between $10,000 and $25,000, 9 (1.7%) reported an annual income above $25,000, and 18 (3.5%) did not provide a response.

Materials and Procedure

Participants completed the study in their psychology class. Each participant received a packet that included three items. The first was an informed consent form, as approved by the Institutional Review Board at the University of North Dakota, that outlined the study and the expectations / risks involved. The second was a demographic survey that asked about the information reported above. The third was a delay-discounting task.

The commodities involved in the delay-discounting task were $100 in cigarettes, a grocery store gift card worth $100, $100 in casino tokens, $100 in cash, or the participant's choice of those four commodities. The exact phrasing of the question(s) used in the discounting task can be found in the Appendix. The present study employed the fill-in-the-blank method for measuring delay discounting (Chapman, 1996; Smith & Hantula, 2008; Weatherly, Derenne, & Terrell, in press). With this particular method, the participant is asked to generate and provide a specific amount of the particular commodity that s/he would accept immediately rather than waiting a certain amount of time for the full amount. The present study utilized five time delays (one week, one month, six months, one year, and five years). Thus, the delay-discounting task

consisted of five questions. For each commodity (group), the order of the questions was randomly determined (independently across commodities / groups) and all participants completing the task for that particular commodity answered the questions in the same random order. Further, distribution of the different questionnaire packets was done randomly within each class. In other words, when the packets were distributed within a particular class, approximately an equal number of respondents were completing the delay-discounting task for each of the five commodities / groups.

Data Preparation

Several legitimate methods exist to analyze data from delay-discounting tasks. One is to fit the indifference points (i.e., in the present study, the values provided by the respondents) with the following hyperbolic equation (e.g., Mazur, 1987):

$$V = A / (1 + kD) \quad \text{(Equation 1)}$$

In Equation 1, V is the subjective value of the delayed outcome, A is the amount of the commodity, D is the delay to the full amount of the commodity, and *k* is a free parameter that describes the rate of delay discounting. High values of *k* indicate steep rates of (i.e., more) discounting; low values indicate little or no discounting.

A second method is to determine the area under the curve (AUC) that is created by the indifference points across the different delays. AUC can be determined with the following equation (Myerson, Green, & Warusawitharana, 2001):

$$(x_2 - x_1)\ [(y_1 + y_2)/2] \quad \text{(Equation 2)}$$

When using Equation 2, the value of AUC can vary between 0.0 and 1.0. Low values of AUC indicate a great amount of discounting of the commodity; high values represent little or no (i.e., 1.0) discounting.

Although Equation 1 has been used in numerous studies on delay discounting, the present study utilized Equation 2 and AUC as the dependent measure. We did so for several different reasons. First, as outlined by Myerson et al. (2001), Equation 1 assumes that the pattern of delay discounting will follow a certain (hyperbolic) form. Equation 1 has successfully described numerous published data sets, but we had no theoretical reason to assume that the present data would be hyperbolic in nature and Equation 2 does not presume the data will follow a certain pattern. Second, *k* in Equation 1 has a lower, but no upper, bound, which commonly results in a skewed distribution that requires data transformation before parametric analyses can be conducted. AUC in Equation 2, on the other hand, does not pose this problem. Third, and perhaps most importantly, Equation 1 did not provide a good fit for many of the respondents' data. Specifically, Equation 1 accounted for less than 70% of the variance in the delay-discounting data for 365 of the 518 participants. It accounted for over 90% of the variance for only 63 participants. Because of this poor fit, we determined that utilizing Equation 2 and AUC would be the most conservative approach.

Results

The initial step in data analysis was to apply the exclusion criteria to the sample (described above). After applying the criteria, the number of participants who completed the delay-discounting task for cigarettes, the grocery story gift card, casino tokens, cash, or their choice of those four commodities was 94, 117, 106, 110, and 89, respectively. Figure 1 presents the mean AUC that was observed for each commodity / group. As predicted, the least amount of discounting (i.e., highest AUC value) was observed for the group who had their choice of the four commodities. However, the absolute difference between the AUC values for this group and for several of the other commodities / groups was not large.

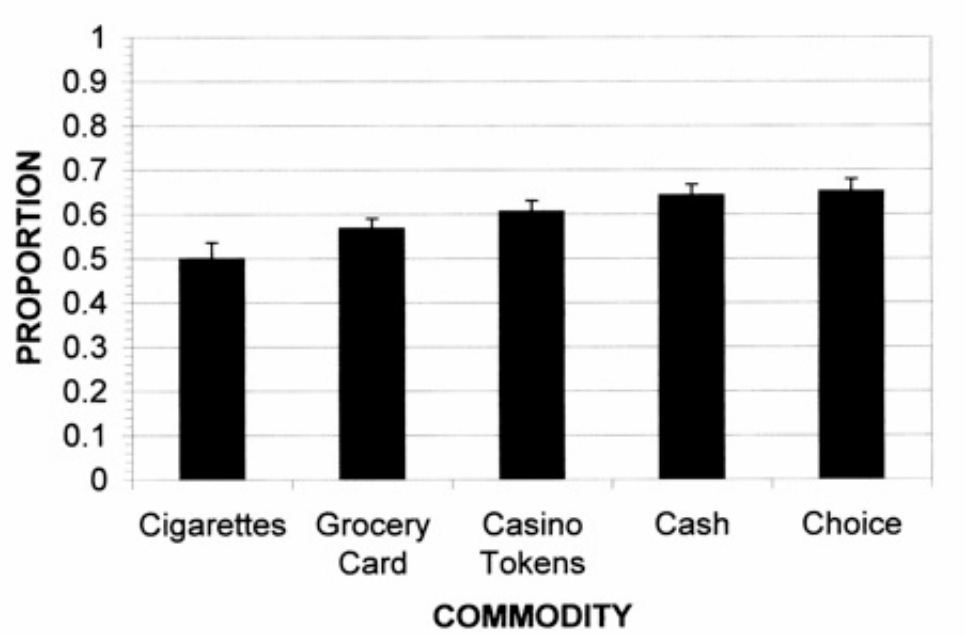

Figure 1. Presented is the mean AUC value for each commodity / group. The error bars represent one standard error of the mean for that particular commodity / group.

Results from statistical analyses were consistent with these visual impressions. A one-way analysis of variance conducted on participants' AUC values in each group resulted in a significant main effect of group, $F(4, 513) = 4.78$, $p = .001$, $\eta^2 = .036$. Planned pairwise comparisons confirmed that participants in the choice group displayed significantly less delay discounting than participants in the cigarette ($p < .001$) and grocery story gift card ($p = .033$) groups, but did not significantly differ in their amount of delay discounting relative to the casino tokens ($p = .255$) or cash ($p = .824$) groups. Results from these analyses were considered significant at $p < .05$.

Discussion

The present study was designed to determine whether less delay discounting would occur when the commodity involved a choice of items than when the commodity was a single item.

Such an outcome was observed, although the difference in discounting between the choice commodity and two of the individual items was not statistically significant. Overall, these results support the conclusion that choice (i.e., self determination) of an outcome can have a greater value to individuals than the same outcome when it is not self determined.

One could potentially argue that the difference between the choice group and the cigarettes group may be linked to the fact that few of the participants smoke and therefore the cigarettes had little or no value to begin with. Other studies from our laboratory have also found that participants discount cigarettes to a greater extent than many other commodities (including money; see Weatherly et al., in press b). However, the counterargument that can be made with the current procedure is that all items were scaled in terms of the same monetary value (i.e., $100). Additionally, the indifference points that the participants provided were in dollars, not in cigarettes. That is, participants were not asked how many cigarettes they would accept, but rather how much money they would accept immediately rather than waiting for the $100 in cigarettes.

Finding that significantly more discounting was observed for the grocery store gift card than for the choice group has several practical implications. Offering gift cards as an incentive (e.g., for research participation; e.g., Dillen & Dixon, 2008) has become relatively common practice. The present results suggest that such an incentive may be less than its face value. Phrased differently, if researchers want to continue to use this practice, they might find that an increased incentive is a choice of different types of gift cards (e.g., to different retailers). In fact, by offering a choice of different cards, researchers may find that the monetary value of the gift cards can be reduced (e.g., $20 cards instead of $25 cards) without decreasing the incentive that they provide. However, given the present results, perhaps a better solution when possible would be to provide participants with cash as an incentive (e.g., McDougall, Terrance, & Weatherly, in press).

As for why the gift card was discounted to a greater extent than the cash or choice options, the explanation may itself be linked to delay discounting. Whereas cash can be spent immediately, gift cards present an additional delay in that one must go to the grocery store to collect the gift card's value. However, two issues should be noted when considering the legitimacy of this possible explanation. One is that participants differed in how they discounted the $100 gift card and the $100 in casino tokens, both of which would seem to invoke a similar additional delay in obtaining their actual value. Next, as with cigarettes, participants were asked how much cash they would accept immediately rather than waiting for the gift card.

Failing to find a significant difference in discounting between the choice group and the casino tokens group may be somewhat surprising for several reasons. For one, the vast majority of the present sample was below the legal age to gamble in North Dakota (i.e., 21 years of age). Furthermore, the nearest casino is over 50 miles from campus. Despite these facts, the participants basically treated the casino tokens as equivalent to cash money. Then again, given the popularity of gambling (see Petry, 2005, for a review), perhaps such a result should not be considered unexpected.

The failure to find a significant difference between the choice and cash groups may be less than surprising. In fact, one could possibly argue that the cash group was also a choice group. That is, with cash, one could obtain any of the other three commodities employed in the present study. Thus, if one is trying to provide an incentive and cannot reasonably provide a self-determined choice of incentive, then offering cash may be the best option.

The failure to find a significant difference between the discounting rates in the self-determined commodity group and the casino-token and cash groups may in fact help address the issue of what is the underlying value of self determination. One argument would be that, by allowing self determination across successive trials / choices, one can negate the decrease in reinforcer effectiveness that would be expected through the process of habituation (e.g., McSweeney & Murphy, 2009). A second argument is that by allowing choice among alternative reinforcers, the respondent always has the option of choosing his/her preferred reinforcer (e.g., participants in the choice group in the present study were mentally choosing cash). Although not definitive, given that the participants in the present experiment A) were choosing among hypothetical outcomes and B) made only five choices overall, the present results would seem to favor the latter, rather than the former, explanation.

For researchers in the field of delay discounting, the failure of Equation 1 to adequately fit the present data is worth noting. The present study utilized the fill-in-the-blank method for measuring discounting. Prior studies have used this technique and found that Equation 1 provided an adequate fit to the data (e.g., Smith & Hantula, 2008). However, several studies from our laboratory (e.g., Weatherly, Derenne, & Terrell, 2010, in press; Weatherly, Terrell, & Derenne, 2010) have utilized this method and, to date, Equation 1 has failed to provide adequate fits to any of those data sets. These failures may represent weaknesses in the fill-in-the-blank method, in Equation 1 to describe delay discounting, or both. Regardless, researchers who are intent on using Equation 1 as their dependent measure may want to choose another technique for collecting delay-discounting data (e.g., the binary-choice method; see Smith & Hantula, 2008). Researchers who are intent on using the fill-in-the-blank procedure should be forewarned that they may need an alternative measure of discounting other than Equation 1.

The present study may have translational value for individuals or businesses who are trying to provide an incentive for people. First of all, studying how people discount certain consequences or commodities may provide an indirect measure of what people value, which may provide a more accurate piece of information than asking the people directly. Secondly, the present results suggest that one may be able to save resources. That is, by offering people a choice of reward, one may find that smaller rewards maintain just as much behavior as larger rewards when only one reward is offered. With that said, the present results also support the conclusion that providing a choice will not always increase the value of the outcome, as was documented by the non-significant differences in discounting between the choice, casino-tokens, and cash groups. It is also the case that the present study employed only college students, so research on other adult populations would be warranted before broad conclusions are made. Further, it should be noted that although significant differences were found in the present study, the effect size that was observed was small (Cohen, 1988). Thus, choice of the outcome was only a small, albeit significant, influence on the observed rates of delay discounting.

With that said, the present procedure was easily conducted, taking participants less than five minutes to complete. People interested in determining what incentives might work in their particular situation may wish to adopt a delay-discounting procedure. Doing so could potentially provide them, both quickly and easily, with the information they desire.

References

Beck, R.C., & Triplett, M.F. (2009). Test-retest reliability of a group-administered paper-pencil measure of delay discounting. *Experimental and Clinical Psychopharmacology, 17*, 345-355.

Chapman, G.B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 771-791.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, *Second Ed.* Lawrence Erlbaum Associates.

Deci, E.L., & Ryan, R.M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology, 49*, Aug. 2008 Special issue: Social psychology and self-determination theory: A Canadian contribution. 182-185.

Dillen, J., & Dixon, M.R. (2008). The impact of jackpot and near-miss magnitude on rate and subjective probability of slot machine gamblers. *Analysis of Gambling Behavior, 2*, 121-134.

Geckeler, A.S., Libby, M.E., Graff, R.B., & Ahearn, W.H. (2000). Effects of reinforcer choice measured in single-operant and concurrent-schedule procedures. *Journal of Applied Behavior Analysis, 33*, 347-351.

Graff, R.B., & Libby, M.E. (1999). A comparison of presession and within-session reinforcement choice. *Journal of Applied Behavior Analysis, 32*, 161-173.

Graff, R.B., Libby, M.E., & Green, G. (1998). The effects of reinforcer choice on rates of challenging behavior and free operant responding in individuals with severe disabilities. *Behavioral Interventions, 13*, 249-268.

Kaufman, A., & Dodge, T. (2009). Student perceptions and motivation in the classroom: Exploring relatedness and value. *Social Psychology of Education, 12*, 101-112.

Kuvaas, B. (2009). A test of hypotheses derived from self-determination theory among public sector employees. *Employee Relations, 31*, 39-56.

Mazur, J.E. (1987). An adjusting procedure for studying delayed reinforcement. In M.L.

Commons, J.E. Mazur, J.A. Nevin, & H. Rachlin (Eds.), *Quantitative Analyses of Behavior: Vol. 5. The Effect of Delay and Intervening Events on Reinforcement Value* (p. 55-73. Hillsdale, NJ: Erlbaum.

McDougall, C.L., Terrance, C., & Weatherly, J.N. (in press). The effect of male confederate presence, betting, and accuracy of play on males' gambling on Blackjack. *The Psychological Record*.

McSweeney, F.K., & Murphy, E.S. (2009). Sensitization and habituation regulate reinforcer effectiveness. *Neurobiology of Learning and Memory, 92*, 189-198.

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior, 76*, 235-243.

Patterson, T.G., & Joseph, S. (2007). Person-centered personality theory: Support from self-determination theory and positive psychology. *Journal of Humanistic Psychology, 47*, 117-139.

Petry, N.M. (2005). *Pathological Gambling: Etiology, Comorbidity, and Treatment*. Washington, D.C.: American Psychological Association.

Smith, C.L., & Hantula, D.A. (2008). Methodological considerations in the study of delay discounting in intertemporal choice: A comparison of tasks and modes. *Behavior Research Methods, 40*, 940-953.

Thaler, R.H. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters, 8*, 201-207.

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.

Weatherly, J.N., Derenne, A., & Terell, H.K. (2010). College students discount money "won" more than money "owed." *The Psychological Record, 60*, 463-472.

Weatherly, J.N., Derenne, A., & Terrell, H.K. (in press). Testing the reliability of delay discounting of ten commodities using the fill-in-the-blank method. *The Psychological Record*.

Weatherly, J.N., Terrell, H.K., & Derenne, A. (2010). Delay discounting of different commodities. *Journal of General Psychology, 137*, 273-286.

Author Contact Information:

Jeffrey N. Weatherly, Ph.D.
Department of Psychology

University of North Dakota
Grand Forks, ND 58202-8380
Phone: (701) 777-3470
Fax: (701) 777-3454
Email: jeffrey.weatherly@und.edu

Jennifer Gudding (jennifer.gudding@und.edu) and Adam Derenne (adam.derenne@und.edu) can also be contacted at Department of Psychology, University of North Dakota, Grand Forks, ND 58202-8380.

Appendix

in X = (A) in cash, (B) in casino tokens, (C) in cigarettes, (D) gift card for a local grocery story, or (E) in cash, $100 in casino tokens, $100 worth of cigarettes, or a $100 gift card for a local grocery store.

Y time = one week, one month, six months, one year, or five years

Question

You have won a raffle in which the prize is $100 *in X*. However, it will be *Y time* before you receive the prize. What is the smallest amount of money you would accept today rather than having to wait *Y time* for your prize?
