The Behavior Analyst Today

A Context for Science with a Commitment to Behavior Change

VOLUME 9, ISSUE 1

ISSN: 1539-4352

TABLE OF CONTENTS

THE BEHAVIOR ANALYST TODAY

PUBLISHER'S STATEMENT

VOLUME NUMBER 9, ISSUE NUMBER 1

ISSN: 1539-4352

Published: April 5, 2007

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 visiting http://www.behavior-analyst-today.com/subscribe.html.

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

Meet Our Staff

Lead Editor:

Michael Weinberg, Ph.D., BCBA - Vice President, Professional Education Resources, Milford, MA, USA
mweinberg@percs.info

Michael Lamport Commons, Ph.D. - ***Harvard University***
camerom3@simmons.edu

Senior Associate Editors:

Cynthia Andersen, Ph.D.

Anne Bramsetter, Ph.D.

Associate Editors:

John Borrero, Ph.D.

Darlene Crone-Todd, Ph.D. - Assistant Professor in Psychology, ***Salem State College***

Paul Malanga, Ph.D. – ***University of South Dakota***
pmalanga@usd.edu

Pat Progar, Ph.D. - ***Cadwell College***

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 "rich text format" (.rtf) to BOTH Michael Weinberg at **mweinberg@percs.info** and Joe Cautilli at **jcautill@astro.temple.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 RTF 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

In Memorium

The Behavior Analyst Online staff and in particular the staff of the Journal of Early and Intensive Behavioral Intervention is deeply saddened over the loss of Jan Handleman who passed away on April 4, 2008. He will be missed!

Licensure as a Postmodern Hero

Joseph Cautilli, Ph.D. and Michael Weinberg, Ph.D.

The idea of the licensed behavior analyst as a post-modern hero is proposed and considered in this editorial commentary. The notion of prediction of events as the idea of licensing behavior analysts is discussed in terms of signal detection theory (Greene & Swets, 1966), regarding backlash of such an endeavor. The authors propose that there will be a reaction. The idea that the licensure movement will be accepted by practitioners and passed by state legistlatures with stand-alone behavior analyst licensing boards is considered.

Keywords: Licensure, Behavior Analysts, Professional Practice, Post-Modern, Signal Detection Theory

---

Initially, when we thought of doing this article, the first author (JC) felt distain. I (JC) must admit that I am not much of a prognosticator. Skinner (1990) argued that we do not "know the future" but we know the past. We attempt to discriminate relevant variables in the present from the past and respond to them. For example, if the reader has an interest in robotics, neural networks (e.g., Thrun, & Mitchell, 1993), and operant conditioning (e.g., Thrun & Schwartz, 1995) models of behavioral development. The world has aging populations. Often this population experiences problems with mobility. Today, this problem is handled by giving them scooters. Scooters have a problem in that once you start using them, the mobility rarely returns and they are not very flexible as to the places that you can go. In the movie Forrest Gump, Forrest had bracers for his legs made from metal to help him walk. Given this set of learning experiences, one can easily suggest envision solving the problem of older people walking as the creation of a device like Forrest's braces only with an operant neural network based chip. When the person is younger (say early 50s), they are made a set of braces. They are instructed to wear the braces for a week or so. Through this wearing, the person's muscles train the chip as to the person's movement range and muscle reactions. The chip is stored for year and when the person is experiencing problems in mobility (maybe 80 or 90 years old), the bracers are taken from the closet, the chip placed back in, and put on the person in effect creating an exoskeleton. This will instantly help the person to walk but it is farm more helpful then that- it can help the person to regain the strength to be independent again. The device can be designed so that each day of consecutive wearing, it gradually transfer .02% of the workload back to the existing muscle structures of the person who is the wearer. Thus, in a year or so, it gradually rebuilds the muscle to walk without the device.

The above sounds like a plan, well maybe or maybe not. Lots of environmental variables might render the device worthless. For example, biological research on stem cells might develop to the point of recreating muscle tissue rapidly regenerating the lost muscle. Another possible break to the plan would be that since scooter technology already exists, it has many people working on its improvement as a technology to increase mobility, as opposed to our suggested exoskeleton. Finally, some unforeseen advance in some other field could change the landscape even further relegating our work to worthless.

Like with the above an ever shifting environment and stimuli emerge, rarely can we make predictions with 100% certainty. The same can be said to be true for licensing. Is it a breath of new air for behavior analysts or a threat and bringer of doom? What we can predict is the community's reaction. When a vague stimuli emerges, signal detection theorists (Greene & Swets, 1966; Swets, 1992) tell us, response can be organized into the following classification system (Table 1)

Table I. Signal Detection Grid

	Signal Present	Signal not present
The subject says present	Hit	False alarm
The subject says not present	Miss	Correct response

So we can ask if the stimulus of licensing is good for the profession. Since it is a vague stimulus our responses will fall into one of the four boxes (Table 1). Yet, our responses are also behavior under operant control and prone to local contingencies that effect the statement. For example, we believe licensing to be good for several reasons (1) it will establish behavior analysis as a profession in a given state. Once it is a profession (a) parents or advocate groups can ask local Universities "Why don't you have such a program?" and thus create pressure on Universities to develop and staff behavior analytic programs. (b) More university programs will mean more faculty positions and granter number of research grants are researchers argue things like "well thousands of clinicians are using stimulus control techniques in work with their clients but the process of stimulus control is not well researched. We see more faculty and research as a good thing. Both authors know plenty of doctoral level behavior analysts who could not secure University positions. We have watched them struggle and seen the agony this create not just for them and their family but also for their mentors. (Most students don't realize that each doctoral student represents a large investment for the primary advisor- years of student and helping develop a dissertation.)[1]. So are we biased? Yes.

In addition, to the above, we believe that many traditional psychology programs will install a behavior analytic track to help their master students hedge who (1) are looking at dwindling options for master level psychology students (2) doctoral students who might have difficulty in finding an academic position. Thus, for us a second set of reinforcers emerges- the reintegration of behavior analytic thinking into psychology.

So who will be biased to say "no." People respond to many contingencies in their lives but several contingencies are relevant here: (1) Those who believe that they will be forced to get a license. (2) Those who currently are licensed in another profession. (3) Those who use behavior analysts as a source of cheap labor because without a license behavior analysts cannot practice. (4) Those who view another master level license as a threat to their finances. (5) Those who generally oppose regulation and government intrusion.

For the first group, the proposed licensing regulations develop by the Behavior Analysis and Public Policy Subcommittee of the Behavior Analyst Online organization are written so as not to be exclusionary. Indeed, you can still practice in states that adopt them as a non-licensed professional. You can still use your Board Certification in Behavior Analysis to treat clients for such groups as TriCare and part of other practices under another licensed professional or clinical practice. Thus your practice should not change unless you choose it too.

For the second group and third, we believe that their will still be many factors that can serve as an incentive to behavior analysts continuing to practice cheaply. The first is if they believe that the environment that you create is a learning environment, so that they can get their two years post graduation experience for licensing. A second would be if your state chooses to adopt a license for Associate Level Licensing, the associates will need to practice under a license behavior analyst.

[1] This is not to argue that the processes that bias choice are necessarily even conscious (Reber, 1992; Bem, 1976). Indeed, as the contingencies shift, people learn, generalize and discriminate without awareness (Reber, 1992).

The fourth group is more difficult to reassure. It is our belief however, that a new license will not create more overall licensed master level people. We strongly believe that people interested in master level practice at this point are becoming counselors, marriage and family therapists, or clinical social workers. Once behavior analysis becomes a profession, it will be able to compete on a level playing field.

The last group will not be swayed no matter what we say. Indeed, their philosophy is a worldview, very resistant to data. It is a matter of training from parents, mentors and those who they deeply trust. No amount of words can change the situation. They believe that the Ph.D. is the terminal degree for any psychologically related field. They have consistently argued that master trained people do not have enough training (see Albee, 1977b). As a group, they have failed to accept a basic fact of life that 48 states license counselors and 26 license marriage and family therapists. When psychology dropped its master level licensing, master level licensing did not disappear it just grew into three new professions (counseling, clinical social work, and marriage and family therapy). All three of which now have considerably more political clout in many states than psychology.

Another argument of the last group is that licensing master level practitioners will hurt the number of students that attend Ph.D. level programs. While this is a possibility, it certainly contradicts what has happened with counseling and school psychology over the last 20 years. Both have practice levels at the master's degree and both divisions of APA have more growth then div 25 the Behavior Analysis Division. In parallel field in the medical profession, nurse practitioners and physician assistants see many of their students go on to achieve Ph.D. in nursing and in public health. It seems that robust master level practice leads to robust Ph.D. level candidates.

What I believe that we can all agree with is the waning influence of behavior analysis on mainstream psychology. While APA grows about 4-5% a year, Div 25 is stagnant. If you look at introductory texts today and compare them to introduction to psychology texts 15 to 20 years ago, you will find far less references to behavior analysis, behavioral psychology or behaviorism. Indeed, most texts still state that Chomsky vanquished behavior analysis. If you look at online journals in psychology or journals with some online psychological content, you will find that currently over 200 such journals exist. Of which 10 are behavior analytic in flavor. Thus assuming all things are equal, psychology out produces behavior analysis at a rate of 20 to one. This means for the 300 articles that the Behavior Analyst Online organization intends to publish next year, psychology will publish over 6,000 articles. We are lead to believe that the situation for print journals is far worse. The stagnation of behavior analysis was even central to Janet Twyman's 2007 presidential address at the Association for Behavior Analysis International.

The postmodernists have redefined the hero in literary work. The POMO hero is a person is fraught with human frailties, unlike the archetypes of the cowboy wearing the white hat and the noble warrior, who saves the day. S/he is gritty and worn by life. We would submit that licensing of behavior analysis is a postmodern hero to the profession and the people who we serve. It is not perfect. No language-based enterprise ever is. It will cause problems- of that we can be assured. It is gritty and unnerving to many. However, it will force us to confront our own values of what our profession is and what we want it to be. And more important then the debate, it is here and offers our students an opportunity that we as aging professors never had- a profession to call our own.

References

Albee, G.W.(1977a). Does including psychotherapy in health insurance represent a subsidy to the rich from the poor? *American Psychologist, 32*, 719-721.

Albee, G.W.(1977b). The uncertain future of the MA clinical psychologist. *Professional Psychology, 8,* 120-122.

Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Reber, A.S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition, 1,* 93-113.

Swets, J.A. (1992). The science of choosing the right decision. *American Psychologist, 47,* 522-532.

Skinner, B.F. (1990) To know the future. *Behavior Analyst, 13*, 103-106.

Thrun, S. & Mitchell, T.M. (1993). *Life long robot learning.* Technical report iai-tr-93-7, Institute for Informatics III. University of Bonn, Bonn Germany

Thrun, S. & Schwartz, A. (1995). Finding structure in reinforcement learning. In advances in *Neural Information Processing Systems,* 7.

Author Contact Information:

Joseph Cautilli, Ph.D.
Children Crisis Treatment Center
1823 Callowhill St.
Philadelphia, PA 19130
Email: jcautilli@cctckids.com

Michael Weinberg, Ph.D., BCBA, LP
Licensed Psychologist - CT, PA, FL
185 Fabyan Rd.
N. Grosvenordale, CT 06255
Tel.: 860-315-7115
Mobile: 860-315-0565
Email: mweinberg@klx.net

(Effect) Size Matters: And So Does the Calculation

Melissa L. Olive & Jessica H. Franco

ABSTRACT

The purpose of this paper is to present the types of measures that may be used to describe intervention effects from single subject designs. A regression approach and several non-regression approaches are described. Non-regression approaches include Standard Mean Difference, Percentage of Non-Overlapping Data, Percent Reduction, and Percentage Exceeding the Median. Researchers are encouraged to combine a non-regression measure along with considerations of methodological rigor and visual analysis to fully appreciate the contributions of single subject intervention data.
Keywords: single subject designs, intervention effects, regression, non-regression

---

In 2001, the American Psychological Association (APA) noted in its publication manual that effect size calculations should be included in manuscripts submitted for publication. However, researchers utilizing single subject designs have not typically embraced the approach of any analyses beyond that of the traditional visual analysis (Marascuilo & Busk, 1988; Parsonson & Baer, 1977).

In visual analysis of single subject data, researchers have examined data for three changes in the data: trend, variability, and level. Using trend analysis, researchers have examined the direction of the data for an increasing (i.e., upward) or decreasing (i.e., downward) trend. Researchers have also inspected for change in data variability or bounce. Finally, researchers have noted changes in level or mean performance.

Recent trends in the field of education have resulted in an increased need to synthesize data sets from single subject studies. For example, the No Child Left Behind Act (NCLBA; 2001) brought considerable attention to the term evidence-based practice. As Odom and colleagues described (2005), some have claimed that only randomized experimental group designs are appropriate for demonstrating scientific evidence. This precluded single subject studies from being included in contributions of scientific evidence on effective intervention methods. However, others have noted that rigorous single subject research has much to contribute when determining scientific knowledge within the field (Horner, et al., 2005). In order to support the use of single subject research as evidence-based, a process of synthesizing single subject data is needed. Additionally, the Individuals with Disabilities Education Act (2004) mandated that teachers use strategies based on evidence based research. It would be tragic for teachers to utilize only teaching strategies proven with group design research; hence a second need to summarize data from single subject studies. Finally, researchers conducting meta-analyses or research syntheses have needed a method for interpreting and comparing intervention effectiveness of single subject studies. Researchers and practitioners in the field have tried to synthesize intervention research and effect sizes have been calculated on single subject data (e.g., Ma, 2006; Parker, Hagan-Burke, & Vannest, 2007; Wanzek, et al., 2006).Therefore, the purpose of this paper is to present the types of measures that may be used to describe intervention effects of single subject research designs. Strengths and limitations of each method will be described. Finally, a recommendation will be made to assist in determining which method should be used with which types of single subject data.

Regression Approaches

Allison and Gorman described the use of regression models to calculate effect sizes with single subject data (Allison & Gorman, 1993; Faith, Allison, & Gorman, 1996). In doing so, the dependent measure in the study (e.g., reading fluency or out of seat behavior) served as the dependent measure in the analysis while the intervention sessions serve as the independent variable. A separate regression equation was then obtained for the baseline and intervention data resulting in two regression equations. Finally, the intervention was subtracted from the baseline and divided by the standard deviation of baseline (Hershberger, Wallace, Green, & Marquis, 1999).

It should be noted that data portrayed in single subject graphs are not independent of one another. Often in single subject research, experimenters visually analyze intervention data following each intervention session. This visual analysis might result in modifications to intervention procedures during the subsequent session resulting in data that are dependent on preceding data. For example, if a child was being taught to exchange a graphic symbol for a preferred item, the dependent variable might be rate of independent exchanges. If, on intervention session four, the child exchanged the symbol at a rate of .2 the experimenter might modify the following session by using a different reinforcer with hopes of increasing the exchange rate. This practice would result in data that are serially dependent. Therefore, regression analyses should be avoided with single subject data.

Non-Regression Approaches

Non-regression analyses, however, may be more appropriate for use with single subject data. A variety of non-regression approaches have been described in the literature. These approaches have manipulated the single subject data resulting in values that quantify the degree of intervention effectiveness above and beyond the traditional approach of visual analysis. Each of these approaches has produced a quantifiable value that must be interpreted.

Percentage of Non-Overlapping Data

A widely used non-regression approach has been Percentage of Non-Overlapping Data (PND; Scruggs & Mastropieri, 2001). This calculation has been described as a "meaningful index of treatment effectives" (p 241). To calculate PND, the percentage of data points during intervention that surpassed the extreme values in pretreatment or baseline was calculated. Specifically, in an intervention to increase the dependent variable, the proportion of treatment data points that exceeded the highest baseline value was calculated. During behavior reduction interventions, the proportion of intervention data points that fell below the lowest baseline was calculated. In either case, the number of non-overlapping intervention points was divided by the total number of intervention data points to determine the PND.

Scruggs and Mastropieri (1998) made special recommendations for using PND with specific types of single subject studies. For example, if a return to baseline design was utilized, the first baseline data set should be used. If multiple treatments were tested, the final phase of intervention data should be used.

Scruggs and Mastropieri also provided suggestions for interpreting PND results (1998). They suggested that PND scores above 90 represented very effective treatments, scores from 70 to 90 represented effective treatments, scores from 50 to 70 were questionable, and scores below 50 were ineffective.

One advantage of the PND score has been that behavioral researchers may be able to readily interpret the data. With extensive practice using visual analysis, behavioral researchers have understood

the meaning of 90% of intervention data not overlapping with baseline data. However, this advantage to behavioral researchers might also serve as a disadvantage to non-behavioral researchers who do not understand single subject research designs. Specifically, a reader without extensive experience with visual analysis would most likely lack an understanding of what 90% of non-overlapping data means.

A second disadvantage of PND is that some studies were not appropriate for the calculation. Specifically, Scruggs and Mastropieri (1998) advised that a PND should not be calculated when a data point in baseline is at the ceiling or floor. Specifically, in a behavior reduction study, if one data point in baseline was zero, then PND would automatically be 0% regardless of the number of data points at zero during intervention.

Percent Reduction

Campbell (2000) termed mean baseline reduction (MBR) using procedures originally described by O'Brien and Repp (1990). In this calculation, the mean baseline and mean intervention measurements were determined for the last three sessions of each. The mean of intervention was subtracted from the mean of baseline and divided by the mean of baseline and multiplying by 100. This produced a mean percent reduction from baseline.

This approach has been helpful in determining how much a behavior has decreased during intervention; however, it has lacked usefulness for determining an effect for interventions that increase behavior, particularly when baseline rates of the behavior are zero.

Percentage Exceeding the Median

A relatively new approach has been the percentage of data points exceeding the median of the baseline phase (PEM; Ma, 2006). For intervention studies focusing on increasing behaviors, Ma suggested that reviewers draw a median line for the baseline data and calculate the percentage of data points in intervention that fall above the median line for behavior reduction studies, the percentage of data points below the median line should be calculated.

Several strengths could be found in the PEM approach. First, there have been no reports of situations where PEM could not be used. Second, PEM has been shown to be correlated with author judgments of intervention effectiveness (Ma, 2006). However, as with PND, the meaning of the percentage calculated may be misconstrued by researchers unfamiliar with single subject design. Finally, as Ma reported, this measure failed to show sensitivity to the magnitude of intervention data points above the median line.

Standard Mean Difference

The standard mean difference is one gauge of intervention effectiveness. Busk and Serlin (1992) presented the standard mean difference (SMD) equation. First, the mean difference from baseline to intervention is calculated. Next a standard is calculated. Many times, the standard deviation of baseline serves as that standard. Finally, the difference is divided by the standard. What results is in an actual effect size value (*d*) that may be more easily understood by readers. Effect sizes should be interpreted as follows: $d = 0.2$ small, $d = 0.5$ medium, and $d = 0.8$ large (Cohen, 1988).

SMD may be calculated in two ways; SMD_{all} and SMD_3. In SMD_{all}, all baseline and all intervention data points are utilized whereas in SMD_3, only the last three data points of baseline and intervention are used. Using only the last 3 data points of baseline and intervention may increase the effect size because, in single subject studies, the last few sessions are usually the best. However, if all the

data points are utilized in the calculation, the variability of the data would be captured in the analysis (although not reflected in the actual results). Therefore readers should recognize that SMD_3 results may be inflated and that SMD_{all} results are most likely more accurate.

Olive and Smith (2005) noted that some rules should be established to create standards for calculating SMD. For example, with a reversal design, the original baseline data and the last intervention data should be used. With an alternating treatments design, the superior treatment data should be used. If a multiple baseline design was employed, an effect should be calculated for each person, setting, or behavior in the study. Finally, in a changing criterion design, the original baseline data and the last intervention data should be used.

The SMD approach offers several strengths. First, average data are used resulting in a formula that may be used in all studies whether the data are increasing in nature (e.g, skill acquisition) or decreasing (e.g., challenging behavior). In this approach, no data need to be discarded due to factors such as overlapping data. The SMD calculation results in an actual *d* score making it more interpretable by readers. Results from other approaches must be interpreted (e.g., is 80% a good effect or an acceptable effect?). Finally, the SMD calculation is simple. Data stored in any spread sheet typically used for graphing may be used without the need for recalculations or re-entry.

Recommendations and Conclusions

Of the non-regression measures, it appears that SMD_{all} may be the most appropriate to use to compare intervention effects during literature reviews or syntheses. Additionally, all data points should be used in the calculation to more accurately describe the true intervention effect and to reduce the likelihood of an inflated effect size. Moreover, the SMD method results in an actual effect size value (*d*) that may be more easily understood than the numbers obtained from calculations of PND or MBR.

It should be noted that all of the non-regressive approaches merely describe changes in the levels of the single subject data. None of the approaches capture the trend or the variability of the data. On the other hand, visual analyses capture all three of these effects. Therefore, non-regression approaches should never be used in lieu of visual analysis, but rather they should be paired with a visual analysis to ensure a comprehensive understanding of the intervention effect.

Moreover, these approaches do not consider the methodological rigor of the study. Horner and colleagues described the quality indicators of methodological rigorous single subject studies (2004). First, they noted that participants and settings should be clearly and operationally described. They noted that it was insufficient to generally describe participant characteristics. Horner and colleagues stressed the importance of operationally defining the dependent variable. The dependent variable should be measured repeatedly and frequently and authors should report a measure of inter-observer agreement on the dependent variable. Horner and colleagues also described the importance of carefully describing the independent variable and presenting data on procedural fidelity. Finally, Horner and colleagues stressed the importance of demonstrating a functional relationship between the independent variable and change to the dependent variable. They noted that a baseline condition was required and that a minimum of three demonstrations of experimental control were necessary.

In summary, researchers are encouraged to combine considerations of methodological rigor and visual analysis with a non-regression measure in order to fully appreciate the contributions of the single subject intervention data. The most appropriate measure depends on the type of research design, the nature of the data collected, and the purpose for the calculation.

References

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behavior, Research, and Therapy*, *31*, 621–641.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington,DC: American Psychological Association.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research designs and analysis: New directions for psychology and education* (pp. 187–212). Hillsdale,NJ: Lawrence Erlbaum Associates.

Campbell, J. M. (2000). *Efficacy of behavior interventions to reduce problematic behaviors in persons with autism: A quantitative synthesis of single-subject research*. Unpublished doctoral dissertation, University of Memphis, TN.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 245–277). Mahwah,NJ: Lawrence Erlbaum Associates.

Hershberger, S. L., Wallace, D. D., Green, S. B., & Marquis, J. G. (1999). Meta-analysis of single-case designs. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 109–132). Newbury Park,CA: Sage.

Horner, R. H., Carr, E. G., Halle, J., McGee, G, Odom, S. L., & Wolery, M. (2005). The use of single subject research to identify evidenced-based practice in special education. *Exceptional Children, 37*,165-179.

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 *et seq.* (2004) (reauthorization of the Individuals with Disabilities Education Act of 1990).

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches. *Behavior Modification, 30*, 598-617.

Marascuilo, L. A., & Busk, P. L. (1988). Combining statistics for multiple-baseline AB and replicated ABAB designs across subjects. *Behavioral Assessment, 10*, 1–28.

No Child Left Behind Act of 2001, 20 U.S.C. 70 § 6301 *et seq*. (2002).

O'Brien, S., & Repp, A. C. (1990). Reinforcement-based reductive procedures: A review of 20 years of their use with persons with severe or profound retardation. *Journal of the Association for Persons with Severe Handicaps*, *15*, 148–159.

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 37*, 137-148.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of All Non-Overlapping Data (PAND): An Alternative to PND. *Journal of Special Education, 40*, 194-204.

Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), *Single subject research: Strategies for evaluating change* (pp. 101–166). New York: Academic.

Scruggs, T. E., & Mastropieri, M. A. (1998). Synthesizing single subject studies: Issues and applications. *Behavior Modification*, *22*, 221–242.

Scruggs, T. E., & Mastropieri, M. A. (2001). How to summarize single participant research: Ideas and applications. *Exceptionality*, *9*, 227–244.

Wanzek, J., Vaughn, S., Wexler, J., Swanson, E. A., Edmonds, M., & Kim, A. H. (2006). A synthesis of spelling and reading interventions and their effects on the spelling outcomes of students with LD. *Journal of Learning Disabilities, 6*, 528-543.

Author Contact Information:

Melissa L. Olive, Ph.D., BCBA
The University of Texas at Austin
Department of Special Education
1 University Station D5300
Austin, TX 78712
(512) 475-6585
molive@mail.utexas.edu

Jessica H. Franco, Ph.D. Candidate, CCC-SLP, BCBA
Department of Special Education
1 University Station D5300
The University of Texas at Austin
Austin, Texas 78712
(512) 626-8305
jessicahetlingerfranco@hotmail.com

TREATMENT OF PTSD AND SUBSTANCE ABUSE COMORBIDITY

Theresa Souza & C. Richard Spates

The best evidence suggests that PTSD is often comorbid with other Axis I and Axis II disorders. National estimates hypothesize that over 85% of individuals with PTSD will meet diagnostic criteria for at least one additional disorder. Inpatient substance abuse centers report that up to 50% of their clients also meet criteria for PTSD. The comorbidity of PTSD and substance abuse (PTSD-SA) is unlike that of other comorbid disorders due to the complex relationship between PTSD symptom presentation, the effects of the substance of choice, and the effects of withdrawal symptoms when attempts are made to eliminate the substance from the body. This paper seeks to explore three prominent theories regarding the etiology of PTSD-SA comorbidity as well as the unique relationship between the symptom presentations of these disorders. Also addressed is the controversy between clinical lore and recent empirical evidence regarding efficacious treatments for individuals with this combination of disorders.
Keywords: PTSD, substance use.

Approximately 50% of individuals in inpatient substance abuse treatment centers will also meet criteria for comorbid PTSD (Brown et al., 1999). This combination of disorders has severe consequences for the individual in terms of course, symptom severity, and effectiveness of treatment. When working with a PTSD-SA population, there are several forms of substances which are more likely to be abused when compared to substance users that do not meet criteria for PTSD. Furthermore, these substances appear to be related to the specific symptoms pattern exhibited by the individual (Stewart, Conrod, Pihl, & Dongier, 1999). Research also indicates that some symptoms of PTSD are more likely than others to elicit substance use in general (Sharkansky, Brief, Peirce, Meehan, & Mannix, 1999). Additionally, the negative impact that substance use relapse risk situations have on an individual may further interfere with the individual's ability to cope effectively with the symptoms of PTSD, which would lead to an increase in both PTSD and SA symptoms/behaviors (Sharkansky et al., 1999). The combination of PTSD-SA also poses several barriers to effective treatment. Some of these barriers are based on clinical lore, and have not undergone the rigorous empirical testing pivotal in the field of psychology. Other barriers have been supported in the empirical field, and these must be addressed in order for treatment to be effective.

Post-traumatic Stress Disorder

Post-traumatic Stress Disorder (PTSD) was first recognized by the American Psychiatric Association as a diagnosable condition in 1980 when it was introduced into the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DMS-III) (American Psychiatric Association, 1980). Since that time, PTSD etiology, symptomology, and treatment have been extensively studied.

PTSD is defined as the development of three categories of symptoms following exposure to a traumatic event in which the individual both (1) came into contact with an event that involved actual or threatened death or serious injury to self or others, and (2) responded to this event with intense fear, helplessness, or horror (American Psychiatric Association, 2000). In essence, exposure to a traumatic event is not sufficient to warrant a diagnosis of PTSD. The subjective, emotional experience of the individual in the aftermath of the trauma must also be taken into account (APA, 2000). The three clusters of symptoms that classify PTSD are reexperiencing, avoidance and numbing, and hyperarousal. Each of these symptom clusters is distinct and affects different areas of psychological functioning. Additionally, disturbances in each category can give rise to comorbid diagnoses associated with that cluster of symptoms that will further disrupt the individual's level of functioning (Taylor, 2006). Lastly, the DSM-IV-TR (2000) states that the symptoms must occur for a minimum of one month and cause clinically significant distress and impairment in several areas of functioning.

The first cluster of symptoms, reexperiencing, refers to the persistent emergence of thoughts and feelings associated with the traumatic event. This can occur in several modalities. These include intrusive images, distressing nightmares, acting and feeling as if the event were occurring again, and psychological distress and/or physiological reactivity when confronted with reminders of the traumatic event (American Psychiatric Association, 2000).

The second cluster of symptoms, avoidance and numbing, involves both the persistent avoidance of stimuli associated with the trauma and the numbing of general responsiveness that was not characteristic of the individual prior to the trauma (American Psychiatric Association, 2000). Examples of avoidance include all efforts to keep oneself from coming into contact with thoughts, feelings, conversations, activities, places, or people that remind the individual of the trauma. Symptoms of numbing include the inability to remember aspects of the event, decreased interest in pleasurable activities, feelings of detachment from others, restricted range of affect, and a sense of shortened future. Recent research indicates that avoidance and numbing may be separate clusters of symptoms as they differ in both their clinical correlates and in responsiveness to treatment (Taylor, 2006).

The final cluster of symptoms is hyperarousal. This category refers to persistent symptoms of increased physiological arousal that were not present prior to exposure to the traumatic event. Examples include sleep difficulties, irritability, anger, difficulties concentrating, hypervigilance, and an exaggerated startle response (American Psychiatric Association, 2000).

PTSD Comorbidity

The best evidence suggests that PTSD is often comorbid with other Axis I and Axis II disorders (Barlow, 2002). However, most clinical trials targeting symptoms of the disorder utilize individuals with PTSD as a sole diagnosis, or at least they are silent with respect to other comorbid conditions. Unfortunately, cases of uncomplicated PTSD are likely the exception, not the rule, for individuals exposed to a traumatic event. While PTSD itself is estimated to occur in 8% of the population, the prevalence rates increase dramatically in persons suffering from other mental disorders (Riggs, Volpicelli, Kalmanson, & Foa, 2003). The National Comorbidity Survey (1995) estimates that over 88% of men and 79% of women who meet the criteria for chronic PTSD also meet the criteria for one or more additional psychiatric diagnoses (Kessler, Sonnega, Hughes, & Nelson, 1995). The most frequent diagnoses comorbid with PTSD include Major Depressive Disorder, substance abuse disorders, Borderline Personality Disorder, Paranoid Personality Disorder, and Bipolar Disorder.

The prevalence rates of comorbid disorders vary considerably. Studies indicate that PTSD and depression co-occur in rates between 15% and 53% (Holtzheimer, Russo, Zatzick, Bundy, & Roy-Byrne, 2005). Higher rates are reported in refugee and war veteran samples than in other forms of trauma. Inpatient substance abuse treatment centers report that up to 50% of their clients also meet criteria for PTSD (Brown, Stout, & Mueller, 1999). An average of 17.8% of individuals with PTSD will also suffer from a personality disorders. Of these, 29% will meet criteria for paranoid personality disorder, 26% for borderline personality disorder, and the remaining personality disorder range from 10-14% respectively (Golier, Yehuda, Bierer, & Mitropoulou, 2003). Lastly, approximately 16% of individuals suffering from PTSD will also meet criteria for Bipolar Disorder (Otto, Perlman, Wernicke, Reese, Bauer, & Pollack, 2004).

The effects of comorbid diagnoses negatively affect both treatment course and outcome (Holtzheimer et al., 2005). The previously described symptoms of PTSD become intensified with the existence of further disruptions in psychological functioning introduced by the additional disorders. The specific presentation of symptoms will vary depending on the nature of the comorbid disorder. However, some commonalities to all the comorbid diagnoses are increased distress for the individual, worsened

PTSD symptoms, increased rates of premature therapy termination, and decreased rates of successful treatment (Najavits, Runkel, Neuner, Frank, Thase, Crits-Christoph, & Blaine, 2003).

Substance Abuse

Problems with substance abuse and dependence are among the most prevalent disorders diagnosed in the United States. Prior to determining the comorbid effects of substance use and PTSD, it is important to understand the distinction between abuse and dependence. The DSM-IV-TR defines substance abuse as a maladaptive pattern of substance use leading to impairment and/or distress within a 12 month period in one or more of four different areas of functioning. These areas include (1) failure to fulfill major obligations in work, school, or home; (2) recurrent substance use in hazardous situations such as driving while impaired; (3) recurrent legal problems associated with substances including use, possession, and dealing, and; (4) continued use despite social or interpersonal problems caused or worsened by the substance use (American Psychiatric Association, 2000). In substance abuse, the individual can continue to function in daily life. However, a continuous risk of developing substance dependence is present.

Substance dependence is characterized as a more pervasive pattern of substance use. The time spent engaged in activities associated with obtaining, using, and recovering from the substance creates severe impairment in the individual's daily functioning. The degree of impairment far exceeds that which is seen in substance abuse. The DSM-IV-TR defines substance dependence as a maladaptive pattern of substance use leading to impairment and/or distress within a 12 month period in three or more different areas of functioning. These areas include (1) tolerance (either an increase in the amount of a substance to achieve the desired effect or a decrease in effects associated with use of the same amount of the substance); (2) withdrawal, characteristic symptoms associated with cessation of a substance that vary based on the substance of use; (3) use of the substance in greater amounts over a longer period of time than was originally intended; (4) persistent desire or unsuccessful attempts to control the use of the substance; (5) increasing amounts of time spent obtaining, using, and recovering from the substance; (6) reduction or elimination of social, occupational, and recreational activities not associated with substance use, and ; (7) continued substance use despite recurrent physical or psychological problems caused or worsened by use of the substance (American Psychiatric Association, 2000). For the purposes of this paper, substance use will refer to substance dependence as described by the DSM-IV-TR.

Substance Types

The DSM-IV-TR (2000) refers to a substance as a drug of abuse, a medication, or a toxin. These substances are further grouped into 11 different classes based on their physiological and psychological effects. The classes utilized are alcohol, amphetamines, caffeine, cannabis (marijuana), cocaine, hallucinogens, inhalants, nicotine, opiates, phencyclidine, sedatives, hypnotics, and anxiolytics. Each class was formed based upon the substance's effects on nervous system functioning. Several classes in the DSM-IV-TR can be combined into larger groups based upon these effects, and were separately identified for reasons to be presented shortly.

Scientifically, there are five general categories of substances. The first is depressants. This category of substance results in behavioral sedation. Their purpose is to induce relaxation in the individual. The two DSM-IV-TR classes of substances that can be described as depressants are alcohol, and the sedative, hypnotic, and anxiolytic group (Durand & Barlow, 2006).

The second category of substances is stimulants. These substances result in mood elevations. They cause the individual to become more alert and active. The DSM-IV-TR classes included in this category are amphetamines, cocaine, nicotine, and caffeine (American Psychiatric Association, 2000).

The third category of substances is opioids. These substances produce analgesic effects, as well as euphoria. This class of substance is not combined with any other class in the DSM-IV-TR.

The fourth category is the hallucinogens. These substances alter the sensory perceptions of the user. They can result in delusions, paranoia, and hallucinations. Substances that are included in this class from the DSM-IV-TR are cannabis and other hallucinogens (Durand & Barlow, 2006).

The final category of substances includes all other drugs of abuse. These are substances that do not fit neatly into one of the previous categories. They include inhalants, steroids, and some over-the-counter medications. They produce a wide variety of effects on the brain and body that may span more than one of the previously discussed classes (American Psychiatric Association, 2000).

The prevalence rates of substance dependence are difficult to determine due to the secrecy and illegality of the behaviors associated with the disorder. Alcohol has historically been considered to be the most prevalent drug of choice. It is estimated that over 15 million Americans are dependent on alcohol alone. Twenty-three percent of adults report engaging in binge drinking. This can be defined as consuming five or more alcoholic drinks in one sitting. In regards to college aged adults, the numbers rise to over 40% (Durand & Barlow, 2006). The prevalence rates of other substances vary widely.

PTSD-SA Comorbidity

As previously stated, approximately 50% of individuals in inpatient substance abuse treatment centers will also meet criteria for comorbid PTSD (Brown et al., 1999). This combination of disorders has severe consequences for the individual in terms of course, symptom severity, and effectiveness of treatment.

When working with a PTSD-SA population, there are several forms of substances which are more likely to be abused when compared to substance users that do not meet criteria for PTSD. Furthermore, these substances appear to be related to the specific symptoms pattern exhibited by the individual (Stewart, Conrod, Pihl, & Dongier, 1999). This can occur in two different forms. The first refers to the PTSD symptom clusters of intrusion, arousal, numbing, and avoidance. Self-medication through substances in order to relieve symptoms of PTSD can lead to abuse and dependence for the substance class whose function is associated with this type of symptom. Second, classes of substances that serve to exacerbate the symptoms of PTSD can be negatively reinforced by substance-intoxication-induced or withdrawal-induced intensification (Stewart et al., 1999).

Many classes of abused substances are correlated with PTSD. These include alcohol, opiates, anxiolytics, analgesic, and cocaine. (Stewart et al., 1999). Although it is difficult to obtain accurate data regarding alcohol use, this class of substance is believed to be the most commonly abused substance among individuals with PTSD. Studies indicate that alcohol consumption is correlated with an increase in PTSD symptoms associated with the arousal symptoms.

Through urinalysis, it is possible to determine with increased accuracy the prevalence of illicit substance use in PTSD individuals. Studies indicate that opiate abuse occurs in approximately 23% of PTSD cases. This is followed by marijuana (20%), Benzodiazepines (11%), and cocaine (8%). Other categories of substances, when present, were found in under 5% of the cases (Calhoun, Bosworth, Hertzberg, Sampson, Feldman, Kibry et al., 2000).

Substance dependence on anxiolytics is correlated with the symptoms clusters of arousal and numbing, while analgesic dependence is correlated with the symptom clusters of arousal, intrusion, and numbing (Stewart et al., 1999).

Estimates indicate that the prevalence rates of PTSD and cocaine dependence are high. Studies suggest that approximately 45% of cocaine-dependent individuals will meet criteria of PTSD (8% PTSD individuals (Calhoun et al., 2000)) at some point during their lifetime. Furthermore, 24% of cocaine-dependent individuals will meet criteria for current PTSD (Back, Dansky, Carroll, Foa, & Brady, 2001). In general, cocaine and opiate users report higher rates of exposure to traumatic events when compared to abusers of other groups of substances. Additionally, cocaine users appear more vulnerable to developing PTSD after exposure to a trauma, as well as experience more severe symptoms of PTSD and increased levels of social impairment (Back et al., 2001). PTSD and cocaine dependence appear to be related through the second pathway (Stewart et al., 1999). The effects of cocaine serve to intensify PTSD symptoms, especially while one is withdrawing from cocaine. As a result, any attempts to reduce cocaine use serve to induce both symptoms of cocaine withdrawal and a concurrent increase in distress associated with trauma symptoms (Back et al., 2001).

Research also indicates that some symptoms of PTSD are more likely than others to elicit substance use in general (Sharkansky, Brief, Peirce, Meehan, & Mannix, 1999). Within the substance dependence literature, a taxonomy of situations that lead to relapse has been proposed. This taxonomy includes the following intrapersonal risk situations: negative emotional states, negative physical states, positive emotional states, testing personal control, and giving in to temptations or urges. Additionally, several interpersonal risk situations are also proposed. These include social pressure to use substances and positive interpersonal interactions (Marlatt & Gordon, 1985). Several of these risk situations are associated with the course and symptomolgy of PTSD-SA.

First, consistent with the self-medication hypothesis, PTSD-SA individuals experience an increased likelihood for exposure to high-risk, trauma inducing environmental conditions than the general population (Kellogg & Triffleman, 1998). The relationship between substance use and violent personal histories has been well-documented. Threats to one's self may be presented through gang activity, accidents while under the influence of substances, illegal activity to obtain substances, and violent acts committed while under the influence of substances (Kellogg & Triffleman, 1998). Exposure to these types of situations can result in the risk factors of negative emotional states; and in many instances, negative physical states. As previously stated, the primary requirement for PTSD diagnosis is the existence of an event that involved actual or threatened death or serious injury (American Psychiatric Association, 2000). A cycle of trauma and substance use can be created through these combined diagnoses.

Second, both the intrapersonal risk factors of negative emotional states and negative physical states are characteristic symptoms of PTSD (Sharkansky et al., 1999). It is not essential for the individual to experience repeated traumas in order to be affected by the combined symptoms of PTSD, or to be at increased risk for relapse as a result of them. Several of Marlott's (1985) relapse risk factors exist within the diagnosis of PTSD. Negative emotional states are elicited by the intrusive and hyperarousal symptoms characteristic of the diagnosis (American Psychiatric Association, 2000). Negative physical states are induced by the physiological reactions to the fear and panic associated with stimuli that remind the individual of the traumatic event.

Third, some evidence exists to suggest that PTSD symptomology itself is not solely a result of interpersonal conflict, negative emotional states, and physical illness, but that once PTSD has been established, it can be elicited by these states as well (Sharkansky et al., 1999).

Last, the negative impact that substance use relapse risk situations have on an individual may further interfere with the individual's ability to cope effectively with the symptoms of PTSD, which would lead to an increase in both PTSD and SA symptoms/behaviors (Sharkansky et al., 1999). The intrusive symptoms of PTSD have demonstrated increases in both levels of distress and negative

emotional states. This led to subsequent increases in drug and alcohol cravings for the individuals. Additionally, analog studies indicate increased craving for substances following the presentation of trauma-related stimuli as compared to neutral stimuli (Coffey, Schumacher, Brimo, & Brady, 2005).

To summarize, of the various relapse risk factors proposed in the Marlatt taxonomy of substance use; unpleasant emotions, physical discomfort, and interpersonal conflict appear to elicit higher rates of relapse in PTSD-SA populations than in substance dependent only populations.

PTSD-SA Treatment Considerations

The combination of PTSD and SA not only intensifies the symptoms of both disorders, but also poses several barriers to effective treatment. First, PTSD-SA individuals rarely present for treatment with anxiety concerns. While these individuals present for treatment at a higher rate than SA only individuals, their presenting problem to both hospitals and outpatient treatment centers is substance abuse (Brown, Stout, & Mueller, 1999). As a result, the comorbid PTSD frequently remains undiagnosed. Second, individuals with PTSD-SA demonstrate markedly less improvement during treatment than SA only individuals in several areas. These include the acquisition of more adaptive coping skills, belief in positive cognitions regarding their abilities, and general ability to effective manage psychological distress. The PTSD-SA groups demonstrated greater levels of avoidance coping mechanisms than the SA only groups (Ouimette, Finney & Moos, 1999). Third, PTSD-SA individuals demonstrate increased levels of interpersonal conflict. This leads to a greater level of difficulty developing rapport with the therapist and compliance with the treatment protocol (Sharkensky et al., 1999). The combination of misdiagnosis, decreased therapeutic rapport, increased psychological distress, and decreased ability to learn more adaptive coping skills leads to an increased likelihood of relapse for substance abuse and increased PTSD symptom presentation. This also leads to decreased confidence in one's ability to successfully complete treatment and increase one's quality of life. It is hypothesized that these are all factors in the elevated rates of early treatment termination in this population.

Pathways to PTSD-SA Diagnoses

While it has become apparent that various forms of comorbidity exist when dealing with PTSD and SA separately, simply assessing for and identifying these conditions are not the only problems faced by researchers and clinicians. The pathway through which the disorders emerged jointly in the individual should be determined in order to identify the most appropriate treatment modality for the individual (Jocobsen & Kosten, 2001). Although this area of study is recent, early stages of research indicate that knowledge of the etiological pathway to PTSD and its comorbid disorders may influence the method of therapeutic intervention likely to elicit the greatest degree of changes in psychological functioning. This relationship is present in both successful treatment, and in the exasperation and intensification of existing symptoms (Hien et al., 2004).

At present, there are three proposed pathways through which comorbidity can occur. The first theory views PTSD as a secondary disorder (Otto et al., 2004). This pathway occurs in individuals who suffer from a history of psychiatric illness. The presence of a psychiatric condition with these individuals places them at increased risk of being exposed to a traumatic event, and they also have an increased risk of developing PTSD following exposure to the traumatic event. One such example involves individuals with a diagnosis of schizophrenia. Epidemiological studies conducted with this population found rates of comorbid PTSD ranging as high as 42% (Muenzenmaier, Castille, Shelly, Jamison, Battaglia, Opler, & Alexander, 2005). The theory posits that the symptoms and lifestyles associated with schizophrenia predispose an individual to victimization. Such factors include psychosis, paranoia, homelessness, and physical and psychological abuse.

This pathway can relate directly to substance use. Individuals with substance abuse disorders frequently engage in high-risk behaviors in order to obtain their substance of choice. These behaviors include drug dealing, theft, assault, prostitution, gang activity, and more (Kellogg & Triffleman, 1998). These behaviors have the potential of placing the individual in a variety of traumatic situations that could subsequently lead to the development of PTSD. Once established, the symptoms of PTSD can increase the individual's use of substances in order to cope with the additional stressors that tax their psychological functioning.

The second pathway involves PTSD as the primary disorder. These individuals displayed no signs of severe psychological distress prior to their exposure to the traumatic event and the subsequent emergence of PTSD. It is hypothesized that the intrusive thoughts, nightmares, avoidance behaviors, hyperarousal, dissociation, and depersonalization symptoms characteristic of PTSD serve as catalysts for further deterioration in psychological functioning (Brown, Stout, & Mueller, 1999). For example, the increased physiological arousal in public settings associated with PTSD frequently gives rise to the avoidance of an increasing number of situations. The decreased anxiety associated with this avoidance negatively reinforces such behavior. As a result, the individual becomes more and more secluded from previously enjoyable activities and the social support that has demonstrated positive effects on psychological functioning. Given time, these actions may develop into Major Depressive Disorder, which is the most frequent comorbid diagnosis to PTSD (Otto et al., 2004).

Similarly, the individual may turn to alcohol or other substances in order to reduce the negative impact of the intrusive symptoms of PTSD, or to calm themselves from the hyperarousal symptoms (Riggs, Rukstalis, Volpicelli, Kalmanson, & Foa, 2003). Unfortunately for the individual, the use of substances will not permanently eliminate the symptoms. As the drug leaves the system, the symptoms return and or intensify. A negative reinforcement cycle develops. The individual continues to utilize substances in order to reduce or avoid the symptoms of PTSD. As with most substances, physical tolerance to the drug develops. The individual's body will require increasing amounts of the substance in order to achieve the desired effects. In time, physical dependence on the substance develops as well (Coffey et al., 2005). This is most likely in instances where the primary disorder has gone untreated or not treated successfully. Research on alcoholism has demonstrated that individuals, who suffered from intense anxiety and panic symptoms prior to their use of alcohol, were more likely to consume alcohol in 8 out of 12 situations compared to alcoholic without a history of anxiety (Sharkansky et al., 1999).

The third and final pathway is referred to as reciprocal increased vulnerability. This theory recognizes that individuals with histories of trauma, particularly in childhood, are at heightened risk for several psychological disorders as adults. Each of these disorders increases the likelihood of further trauma and decreases levels of healthy psychological functioning. As a result, proponents of this theory posit that the primary diagnosis is inconsequential for the individual. Only the presentation of symptoms should be considered when treating the individual, and this treatment should be tailored towards that person's specific needs (Meenzenmaier et al., 2005).

Treatment Target Rules: Clinical Lore

In regards to efficacious treatment for these comorbid disorders, current practices may be inconsistent with empirical studies due to commonly held beliefs and clinical lore on how to best treat substance dependence. Two main theories exist in the therapeutic community which may be limiting the amount of research available as well as the dissemination of alternate forms of treatment. These are the Pandora's Box hypotheses and the self-medication hypothesis.

The Pandora's Box hypothesis states that any attempts to address trauma related material in the incipient stages of substance dependence treatment would severely interfere with treatment effectiveness.

Encouraging the individual to recall trauma material is believed to result in an influx of negative thoughts and emotions with which the individual is not equipped to cope at this stage in therapy. In essence, "Pandora's box" would be opened and neither the client nor the therapist would be capable of dealing with the resultant effects. As a result, proponents of this theory believe that the individual will be incapable of maintaining abstinence from substances and would either resume or intensify their drug behaviors (Hien, Cohen, Miele, Litt, & Capstick, 2004).

The Pandora's Box hypothesis is sustained by a number of beliefs. The first is that most PTSD-SA individuals initially seek treatment for substance dependence. Advocates state that it would be inappropriate to then treat the clients for a problem for which they did not seek treatment. Secondly, substance abuse is associated with a decreased tolerance for negative affect. It is believed that addressing trauma-associated material early in treatment will increase their level of negative affect (Triffleman, Carroll, Kellogg, 1999). Third, many proponents of exposure therapy for PTSD state that this type of treatment requires the ability to cope with powerful affective and cognitive material. They believe that substance users do not meet these requisite criteria (Brady, Dansky, Back, Foa, & Carroll, 2001). Fourth, anecdotal evidence is presented that suggests PTSD or combined PTSD-SA treatment will be ineffective unless abstinence is first achieved. Last, clinical lore suggests that PTSD symptoms will reduce or remit with the cessation of substance use (Triffleman et al., 1999).

The second theory, self-medication, posits a differing view of PTSD-SA comorbidity and occurs in two parts. Proponents of this view suggest the substance dependence as a secondary diagnosis. In the first part, it is believed that the individual began to abuse substances in order to cope with the symptoms associated with PTSD. In essence, the substance abused served, in part, as a self-induced medical intervention. The effects of the substance on the brain and body reduced the anxiety and/or depression present when the symptoms were at full intensity (Hien et al., 2004).

The second aspect of the self-medication hypothesis concerns withdrawal symptoms. By this point in the substance use, the individual has developed physical dependence on the substance. Any attempt on the part of the addict to reduce this substance use results in a vicious cycle of withdrawal symptoms, increased PTSD symptoms, and increased cravings for the substance associated with each. The individual has previously been conditioned to escape the PTSD symptoms by use of a substance, so this pattern is repeated. Any further attempt to halt substance use only serves to reinforce the existing pattern (Jacobsen et al., 2001).

One common theme in both of the theories found in clinical lore is the idea that the substance dependence diagnosis should be the primary focus of treatment. The main reason for this stems from a multitude of ideas, few of which were formulated based on empirical data. Prior to addressing the validity of therapeutic focus in this population, it is essential to understand the method by which PTSD is treated. This is especially important because much of the clinical lore cited above is advanced by professionals who may have no background in empirically supported treatments for PTSD, even if they are trained substance abuse counselors. Their perspective is therefore based on their experience but not informed by evidence based practice for treating this complex disorder.

PTSD Treatment

Although various forms of treatment exist for PTSD, Exposure Therapy has shown the highest levels of effectiveness (Taylor, Thordarson, Federoff, Maxfield, Lovell, & Ogrodniczuk, 2003). Exposure therapy can be administered in three forms; imaginal, interoceptive, and in vivo. Each of these involves the client coming into contact with some form of the pivotal traumatic memory or event. In imaginal exposure, the client may be asked to recount the traumatic memory to the therapist multiple times in order to process the events. In time, the intense emotional and physiological arousal states

associated with the memory loose their potency. Interoceptive exposure involves exposing the client to the physiological reactions, sensations, and subjective perceptions of distress associated with the traumatic reaction. The client learns how to monitor and down-regulate their bodily reactions to trauma-relevant imaginal or interoceptive sources of stimulation. In vivo exposure involves the client physically confronting some physical or environmental context of the trauma and remaining in that situation until their arousal level declines (Taylor, 2006).

Other treatments utilized for PTSD include cognitive therapy, relaxation training, and EMDR. However, an element of exposure is present in each of these therapeutic modalities, and research suggests that the exposure component alone is responsible for the positive effects of treatment (Spates & Koch, 2004; Davidson, 2001, Van Etten & Taylor, 1998). Over the past 30 years, several meta-analytic reviews have been conducted on the various therapies utilized in the treatment of PTSD (Van Etten & Taylor, 1998; Waller, Mulick, & Spates, 2000). All of these have concluded that Exposure Therapy and Eye Movement Desensitization and Reprocessing (EMD/R) demonstrate the most consistent empirical evidence regarding effectiveness in the treatment of PTSD. While EMD/R contains strong connections to Exposure Therapy, there is evidence to suggest that clients find this type of therapy to be less aversive than other forms of exposure (Waller et al., 2000). This information is essential when considering treatment options, as the drop-out rates for PTSD treatment are high. When combining the PTSD drop-out rate to the substance abuse relapse rates, the need to identify intervention strategies that are both effective and palatable for the clients is increased exponentially.

Exposure Therapy, like any therapy, does not come without side-effects to the client. While the procedure is quite safe, and the client is never placed in a dangerous situation, Exposure Therapy may increase levels of arousal and the frequency of nightmares associated with the traumatic event during the course of treatment. These effects dissipate quickly during treatment, but many clients report discomfort associated with them (Taylor, 2006). Clinical lore suggests that deliberately subjecting substance dependent individuals with PTSD to a treatment that is known to increase arousal symptoms is analogous to encouraging them to abuse substances in increasing amounts (Hien et al., 2004). However, as previously stated, this theory appears to stem from "educated guesses" about the responses of substance abusing individuals and does not stem from empirical data. One guess is that it arises in the context of substance treatment where therapy is delivered by individuals who are trained only to treat the SA disorder and not trained in treating the trauma reaction and thus only witness increased arousal with no available effective intervention at their disposal.

Substance Use Treatment

Several types of therapies have been proposed for the treatment of substance abuse. The most frequently utilized treatment strategy involves a 12-Step program, which typically falls under the caveat of Alcoholics Anonymous (AA) or Narcotics Anonymous (NA). There exists no clear definition or single explanation as to the treatment modality and techniques of 12-Step programs. While there are certain core concepts, assumptions, hypotheses, and ideas that can be used to identify the approach in a broad sense, there are no clear guidelines as to the process for conducting the meetings (Rotgers, Morgenstern, & Walters, 2006). The wide varieties of substance abuse programs that follow a 12-Step model are heterogeneous. The program itself can best be described as a social movement or fellowship. Within the program itself, there is no formal assessment process, diagnosis, case conceptualization, dispensation of medication, treatment plan, case management, individual therapy, or group therapy (Rotgers et al., 2006).

The practice of 12-Step meetings began in 1953 and views substance abuse from the disease model. The belief inherent in the program is that total abstinence from the substance is the only method through which control can be taken over one's life. The meetings are generally conducted by community members who have achieved abstinence. Stories of the trials and tribulations of other abusers are related

to show that no individual is alone in their battle against substance dependence. Individuals that have maintained abstinence for some length of time serve as sponsors for newer members, and offer support during the beginning stages of abstinence. Each member is taught a series of 12 steps designed to admit powerlessness over the substance, make amends for past wrongs, and grow closer to a spiritual entity (Rotgers et al.; 2006).

Another treatment strategy for substance abuse involves the use of pharmacological agents to decrease cravings for substances, decrease the impact of withdrawal symptoms on the client (Vocci, Acrid, Elkashef (2005), decrease the physiological effects that substances have on the body and/or create unpleasant side effects for the individual who utilizes substances while taking an alternate form of medication (Kleber, 2004).

The first pharmacological intervention is known as agonist substitution. This involves providing the substance abuser with a drug that has a similar chemical make-up as their drug of abuse. The new drug is less harmful on the body of the user. The goal of this form of therapy is to eliminate the body's dependence on the drug of abuse without causing the body to simultaneously cope with the symptoms of withdrawal. The individual can then be titrated off of the "safe" drug until complete abstinence is reached (Miller & Carroll, 2006). Unfortunately, this method has limitations. While the first stage of the treatment has shown effectiveness in substance abuse treatment, the second has demonstrated multiple problems. The clients frequently become dependent on the "safe" drug. There is also a high relapse rate for the original drug once the "safe" drug is eliminated (Miller & Carroll, 2006). One example of this type of pharmacological intervention is methadone maintenance therapy for opiate use. Methadone is a long-lasting synthetic opiate that has shown effectiveness in achieving abstinence, as well as decreasing the craving for, drugs such as heroin and hydromorphone. This has subsequently resulted in reductions in relapse, decreases in criminal activity, and reduced mortality rates as compared to those actively taking the original substance (Vocci, Acri, Elkashef (2005).

Antagonist drug therapies are designed to counteract, or block, the euphoric effects of substances on the user. In essence, if an individual were to use the addictive substance while taking the medication, they would not achieve the "high" that they are accustomed to experiencing when taking the drug (Miller & Carroll, 2006). The goal is to break the stimulus-response pairing between ingestion of the substance and the resultant feelings of euphoria and thus extinguish the drug taking behavior. This method does not result in dependence on the new drug. However, the disadvantage of this type of pharmacological therapy is in the side effects. Ingestion of a drug antagonist will elicit immediate withdrawal symptoms from the drug of abuse. This experience is highly aversive for the client, which decreases the probability that they will continue to take the antagonist when the doctor or therapist is not present (Kleber, 2004). One example of this form of pharmacological therapy is Naltrexone. Naltrexone is a long-acting opioid antagonist that will block the effectiveness of opioids for up to 72 hours (Kleber, 2004).

The final form of pharmacological therapy for substance abuse is aversive treatment. These drugs both block the euphoric effects of the substance of abuse (antagonistic properties) and make the ingestion of the abused substance highly unpleasant for the individual (Miller & Carroll, 2006). One example of this type of therapy is disulfram, or Antabuse, for alcohol addiction. If an individual were to drink alcohol while taking disulfram, the result would be nausea, vomiting, elevated heart rate, and respiration. Other subjects have reported feelings of anxiety and paranoia as well. As a result, treatment compliance to this form of pharmacotherapy is poor (Vocci et al., 2005). O'Brian (1996) conducted an analysis of pharmacological treatments for substance abuse. The analysis concluded that abstinence from drugs is generally not achieved with initial courses of treatment, regardless of treatment modality. However, the outcomes for substance abuse treatment utilizing pharmacological medication was comparable to those of other chronic conditions such as diabetes and hypothyroidism (O'Brian, 1996).

PTSD-SA Treatment Implications

To summarize thus far, relapse is common in substance abuse treatment, and the rates of relapse only increase when PTSD occurs in conjunction with substance use. Current treatments for each of these disorders separately are vastly different. The main component of PTSD treatment is exposure therapy, while the treatment of choice for substance use involves a 12-Step program targeting abstinence from the substance. Clinical lore suggests that prior to treatment for PTSD, the substance use must be addressed or exposure therapy would at best fail, and at worse incite increased substance use (Jocobsen & Kosten, 2001). However, this belief is based on "hunches" and not on empirical data. Few studies exist to date testing the validity of these strongly-held beliefs, and those that do exist suggest an alternate view of PTSD-SA comorbidity.

An emerging view states that PTSD treatment should be introduced earlier in the therapy. These individuals note that purposefully ignoring the symptoms of PTSD while targeting substance use leaves the clients vulnerable to increasing levels of distress with no appropriate coping mechanisms. By addressing PTSD symptoms early in treatment, much suffering could be alleviated. This could serve as a motivating factor for continuing treatment and increase confidence that abstinence can be both achieved and maintained (Triffleman et al., 1999).

Empirical Evidence Regarding PTSD-SA Treatment

Early studies have shown that joint PTSD-SA treatment, regardless of treatment modality, is essential in achieving and maintaining abstinence from substances. Those individuals who receive PTSD treatment in addition to SA treatment in the first three months of intervention are 3.7 times more likely to be free of substances at a 5-year follow-up than individuals who receive only SA treatment (Ouimette, Moos, Finney, 2003). The same study tracked the clients yearly for 5 years. Those clients who continued to experience problems with substances in the 5th year of the study received PTSD treatment at that time. Those clients were 4.6 times more likely to be free of substances at follow-up than those who did not receive the treatment (Ouimette et al., 2003). This indicates that SA treatment alone was not sufficient to maintain abstinence from substances. However, it is important to note that of the total sample, only 27% were in remission from substances at follow-up. There remains much work to be done in the area of PTSD-SA treatment.

One recent area of research addressed the question of treatment tolerance. Can individuals with substance dependence tolerate exposure therapy prior to the complete cessation of substance use, or will exposure therapy lead to increased drop-out rates and/or increased substance use?

Preliminary studies indicate that exposure therapy is tolerated well in the substance abusing population. Brady et al., (2000) studied 39 cocaine dependent individuals. Of those who prematurely terminated treatment, 75% did so prior to the onset of the exposure component, which was in session 6 of the treatment protocol. Of those available for a follow-up survey, none of the participants reported exposure therapy as the reason for termination. Reasons included admittance to an inpatient facility, transportation and scheduling difficulties, and relocation (Brady et al., 2000).

Hien et al., (2004) studied 96 women with PTSD-SA comparing cognitive behavioral therapy, including an exposure component, relapse prevention, and community care. The researchers found no differences between the active treatment groups in client drop-out rate. Furthermore, both active treatment groups demonstrated significant decreases in symptom severity. The results continued after treatment was completed and follow-up data indicated sustained improvement in functioning (Hien et al., 2004). Furthermore, meta-analyses report treatment completion rates for exposure based PTSD-SA treatments as 39%, which is higher than the 28% completion rate reported be the National Institute on Drug Abuse for substance abuse treatment alone (Coffey et al., 2005).

The aforementioned meta-analysis by O'Brian (1996) included the effectiveness of pharmacological interventions in relation to comorbid disorders. While the analysis concluded that abstinence from drugs is generally not achieved for SA only treatment, when comorbid conditions were considered, research was very poor. Pharmacological treatment interventions have demonstrated efficacy with comorbid depression and opiate addiction, but effectiveness studies with other substances were lacking. No treatment outcome studies on pharmacological interventions for comorbid substance dependence and anxiety disorders were found at that time (O'Brian, 1996).

Subsequent to the O'Brian (1996) review, there have been several studies conducted on pharmacological interventions, few of which included substance abusers with comorbid PTSD. Several studies support the use of pharmacological interventions for opiate dependence (Sigmon, Wong, Chausmer, Liebson, & Bigelow, 2004; Sigmon, 2007; Trafton, Minkel, & Humphreys, 2006).

When including comorbid PTSD participants, opioid substitution therapy demonstrated positive results. Subjects with and without PTSD treated for SA with Methadone showed comparable, significant reductions in use of heroin (Trafton, Minkel, & Humphreys, 2006). The number of days of heroin use decreased by 91% for individuals with comorbid PTSD, and these results were maintained at a 12 month follow-up. Use of cocaine, marijuana, and prescription opioids also decreased in both groups (Trafton et al., 2006). However, in order to achieve these SA symptom reductions in the comorbid PTSD group, higher doses of opiate medication and additional psychological treatment sessions were required. Furthermore, the PTSD-SA patients demonstrated increased levels of psychiatric symptoms both at intake as well as throughout SA treatment. These symptoms did not diminish regardless of the patients' responsiveness to SA treatment (Trafton et al., 2006).

The second question that has received much needed attention in the literature is the effectiveness of PTSD only or concurrent PTSD-SA treatment on comorbid populations. Two exposure-based treatments have been developed for PTSD-SA comorbidity. The first is concurrent treatment of PTSD and cocaine dependence (CTPCD), and the other is substance dependence PTSD therapy (SDPT) (Coffey et al., 2005).

CTPCD is an 8 week PTSD-SA protocol consisting of twice weekly sessions, each of 90 minute duration. The program was initially developed for cocaine-dependent individuals, but can be adapted for use with other substances (Back et al., 2001). The treatment package contains components of the cognitive skills training, relapse prevention, cognitive restructuring, in vivo, and imaginal exposure therapies. The primary goals of the program are to (1) educate clients about the relationship between substance abuse and PTSD, (2) achieve and maintain abstinence from substances, (3) decrease severity of PTSD symptoms and (4) reduce high-risk behaviors that could lead to future traumatization and substance abuse (Back et al., 2001). The first 5 sessions focus on Psychoeducation regarding exposure therapy, PTSD, and SA. Additionally, core coping skills for the cessation of substance use are presented. These include sessions dealing with cravings for substances, motivation and commitment to cease SA, assertiveness skills, and decision making (Coffey et al., 2005). Exposure therapy is begun in session 6 and continues throughout the remainder of treatment.

SDPT is a manual-based individual therapy for PTSD-SA. It is a compilation of a variety of treatments for both PTSD and SA. The goals of SDPT are threefold. They include abstinence from substance use, maintenance of abstinence during the course of trauma work, and overall reduction of PTSD severity (Triffleman, Carroll, Kellogg, 1999). SDPT consists of twice weekly individual therapy sessions over the course of approximately 20 weeks. Treatment is presented in two phases. The first phase, lasting 12 weeks, focuses on psychoeducation for PTSD, SA, and the effects of comorbidity as well as SA treatment. Specific skills taught during this phase of treatment include examination of craving and dysphoria associated cognitions, generalization of alternate cognitions, and increasing management of

emotional and physical states (Triffleman et al., 1999). The final 8 weeks, phase two, focuses on PTSD treatment. This phase is a combination of Stress Inoculation Training and Exposure Therapy. Both phases of treatment include in-session tasks and out-of-session homework assignments.

Future Directions

On the basis of this review, several changes are needed in the arena of PTSD-SA treatment. First, accurate assessment of psychological functioning prior to the onset of treatment is a necessity. Studies indicate that the majority of PTSD-SA individuals that seek treatment, will do so for their SA problem only. These individuals may then receive SA treatment on multiple occasions before the PTSD is diagnosed. Research supports the notion that PTSD-SA individuals utilize hospital and outpatient services at a much higher rate than SA only individuals, yet they do not receive differing forms of treatment, primarily because the PTSD remains undiagnosed (Brown, Stout, & Meuller, 1999). This is disturbing as approximately 50% of individuals with PTSD will also meet criteria for a substance use disorder (Brown et al., 1999).

Once a PTSD-SA individual is accurately diagnosed, additional barriers exist to effective treatment. The long-held opinion that psychopathology of any type cannot be addressed until the substance use problem has been addressed is currently under debate in the psychological community, especially in relation to comorbid PTSD. Recent empirical studies surrounding PTSD and Depression (O'Donnell et al., 2004), PTSD and OCD (Gershuny et al., 2002), as well as PTSD and SA (Coffey et al., 2005) have indicated that treating the disorder that emerged first (primary pathway diagnosis) can result in remission, or improvement, of the secondary diagnosis without specifically targeting this diagnosis in treatment. The psychological field is in need of additional research to replicate these results and to determine if altering current SA treatments to address the primary pathway diagnosis first will increase treatment success. Currently, the relapse rate for substance abuse is high, and this rate increases dramatically as co-occurring psychological stressors and diagnoses are identified. PTSD appears to be one of the most difficult comorbid disorders to treat with SA due to the intensification of PTSD symptoms while withdrawing from substances and learning alternate coping skills to address the PTSD symptoms (Sharkansky et al., 1999).

Within the past decade, several studies have addressed the untested belief that PTSD-SA individuals cannot tolerate exposure therapy, the preferred treatment for PTSD, until they have achieved abstinence from their substance of abuse (Brady et al., 2001; Triffleman et al., 1998; Coffey et al., 2005). The results of these studies have indicated that exposure therapy is well-tolerated. The drop-out rates are comparable to PTSD only individuals. Furthermore, the treatment outcomes for those individuals that complete PTSD-SA treatment are better than for those who receive SA only treatment (Brady et al., 2001).

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition- Text Revision.* Washington, DC, Americal Psychiatric Association.

Back, S. E., Dansky, B. S., Carroll, K. M., Foa, E. B., Brady, K. T. (2001). Exposure therapy in the treatment on PTSD among cocaine-dependent individuals: Description of procedures. *Journal of Substance Abuse Treatment*, 21, 35-45.

Brady, K. T., Dansky, B. S., Back, S. E., Foa, E. B., Carroll, K. M. (2001). Exposure therapy in the treatment of PTSD among cocaine-dependent individuals: Preliminary findings. *Journal of Substance Abuse Treatment*, 21, 47-54.

Brown, P. J., Ouimette, P. C. (1999). Introduction to the special section on substance use disorder and posttraumatic stress disorder comorbidity. *Psychology of Addictive Behaviors,* 13 (2), 75-77.

Brown, P. J., Stout, R. L., Mueller, T. (1999). Substance use disorder and posttraumatic stress disorder comorbidity: Addiction and psychiatric treatment rates. *Psychology of Addictive Behaviors,* 13, 115-122.

Brown, T. A., Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and DSM-IV. *Journal of Consulting and Clinical Psychology,*60 (6), 835-844.

Calhoun, P. S., Bosworth, H. B., Hertzberg, M. A., Sampson, W. S., Felfman, M. E., Kirby, A. C., Wampler, T. P., Tate-Williams, F., Moore, S. D., Beckman, J. C. (2000). Drug use and validity of substance use reports in veterans seeking help for posttraumatic stress disorder. *Journal of Counseling and Clinical Psychology,* 68 (5), 923-927.

Coffey, S. F., Stasiewicz, P. R., Hughes, P. M., Brimo, M. L. (2006). Trauma-focused imaginal exposure for individuals with comorbid posttraumatic stress disorder and alcohol dependence: Revealing mechanisms of alcohol craving in a cue reactivity paradigm. *Psychology of Addictive Behaviors,* 20 (4), 425-435.

Cohen, J. A., Mannarino, A. P., Knudsen, K. (2004). Treating childhood traumatic grief: A pilot study. *Journal of American Academy of Child and Adolescent Psychiatry*, 43: 1225-1234.

Coffey, S. F., Schumacher, J. A., Brimo, M. L., Brady, K. (2005). Exposure therapy for substance abusers with PTSD: Translating research to practice. *Behavior Modification,* 29, 10-38.

Deters, P. B., Novins, D. K., Fickenscher, A., Beals, J. (2006). Trauma and posttraumatic stress disorder symptomatology: Patterns among American Indian adolescents in substance abuse treatment. *American Journal of Orthopsychiatry,*76 (3), 335-345.

Gershuny, B., S., Baer, L., Jenike, M. A., Minichiello, W. E., Wilhelm, S. (2002). Comorbid posttraumatic stress disorder: Impact on treatment outcome for obsessive-compulsive disorder. *American Journal of Psychiatry,* 159 (5), 852-854.

George, W. H. (1989). Marlatt and gordon's relapse prevention model: A cognitive-behavioral approach to understanding and preventing relapse. In *Relapse: Conceptual, Research, and Clinical Perspectives,* New York, Hawthorne Press Inc.

Golier, J. A., Yehuda, R., Beirer, L. M, Mitropoulou, V. (2003). *American Journal of Psychiatry,* 160 (11), 2018-2024.

Grella. C. E., Stein, J. A., Greenwell, L. (2005). Associations among childhood trauma, adolescent problem behaviors, and adverse adult outcomes in substance abusing women offenders. *Psychology of Addictive Behaviors,* 19, 43-53.

Hien, D. A., Cohen, L. R., Miele, G. M., Litt, L. C., Capstick, C. (2004). Promising treatments for women with comorbid PTSD and substance use disorders. *American Journal of Psychiatry,* 161, 1426-1432.

Holtzheimer, P. E., Russo, J., Zatrick, D., Bundy, C., Roy-Burne, P. P. (2005). The inpact of comorbid posttraumatic stress disorder on short-term clinical outcome in hospitalized patients with depression. *American Journal of Psychiatry,* 162 (5), 970-976.

Jocobsen, L. K., Steven, M., Kosten, T. R. (2001). Substance use disorders in patients with posttraumatic stress disorder: A review of the literature. *American Journal of Psychiatry,* 158 (8), 1184-1190.

Kilpatrick, D. G., Ruggiero, K. J., Acierno, R., Saunders, B. E., Resnick, H. S>, Best, C. L. (2003). Violence and risk of PTSD, major depression, substance abuse/dependence, and comorbidity: Results from the national survey of adolescents. *Journal of Consulting and Clincial Psychology,* 71 (4), 692-700.

Kleber, H. D. (2004). Treatment of drug dependence enters a new high technologic al era, *Addiction,* 99, 1475-1478.

Lellogg, S., Triffleman, E. (1998). Treating substance-abuse patients with histories of violence: Reactions, perspectives, and interventions, *Psychotherapt,* 35 (3), 405-414.

Muenzenmaier, K., Castille, D. M., Shelly, A., Jamison, A. (2005). Comorbid posttraumatic stress disorder and schizophrenia. *Psychiatric Annals,* 35 (1), 50-56.

Najavits, L. M., Runkel, R., Neuner, C., Frank, A. F., Thase, M. E., Crits-Christoph, P., Blaine, J. (2003). Rates and symptoms of PTSD among cocaine-dependent patients, *Journal of Studies on Alcohol,* 64, 601-606.

O'Brian, C. P. (1996). Recent developments in the pharmacotherapy of substance abuse. *Journal of Consulting and Clinical Psychology,* 64 (4), 677-686.

O'Donnell, M. L., Creamer, M., Pattison, P. (2004). Posttraumatic stress disorder and depression following trauma: Understanding comorbidity. *American Journal of Psychiatry,* 161 (8), 1390-1396.

Otto, M. W., Perlman, C. A., Wernicke, R., Reese, H. E., Bauer, M. S., Pollack, M. H. (2004). Posttraumatic stress disorder in patients with bipolar disorder: A review of prevalence, correlates, and treatment strategies. *Bipolar Disorders,*6, 470-479.

Ouimette, P. C., Finney, J. W., Moos, R. H. (1999). Two-Year posttreatment functioning and coping of substance abuse patients with posttraumatic stress disorder. *Psychology of Addictive Behaviors,* 13, 105-114.

Ouimette, P., Moos, R. H., Finney, J. W. (2003) PTSD treatment and 5-year remission among patients with substance use and posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology,* 71 (2), 410-414.

Riggs, D. S., Rukstalis, M., Volpicelli, J. R., Kalmanson, D., Foa, E. B. (2003). Demographic and social adjustment characteristics of patients with comorbid posttraumatic stress disorder and alcohol dependence: Potential pitfalls to PTSD treatment. *Addictive Behaviors,* 28, 1717-1730.

Sharkansky, E. J., Brief, D. J., Peirce, J. M., Meehan, J. C., Mannix, L. M. (1999). Substance abuse patients with posttraumatic stress disorder: Identifying specific triggers of substance use and their associations with PTSD symptoms. *Psycholog of Addictive Behaviors,* 13 (2), 89-97.

Sigmon, S. C., Wong, C. J., Chausmer, A. L., Liebson, I. A., Bigelow, G. E. (2004). Evaluation of an injection depot formulation of buprenorphine: Placebo comparison. *Society for the Study of Addiction*, 99, 1439-1449.

Sigmon, S. C. (2007). Investigating the pharmacological and nonpharmacological factors that modulate drug reinforcement. *Experiential and Clinical Psychopharmacology,* 15 (1), 1-20.

Stewart, S. H., Conrod, P. J., Pihl, R. O., Dongier, M. (1999). Relations between posttraumatic stress symptom dimensions and substance dependence in a community recruited sample of substance abusing women. *Psychology of Addictive Behaviors,* 13 (2), 78-88.

Taylor, S., Thordson, D. S., Federoff, I. C., Maxfield, L., Lovell, K., Ogrondniczuk, J. (2003). Comparative efficacy, speed, and adverse effects of three PTSD treatments: Exposure therapy, EMDR, and relaxation training. *Journal of Counseling and Clinical Psychology*, 71 (2), 330-338.

Taylor, S. (2006). *Clinician's guide to PTSD: A cognitive behavioral approach.* New York, Guilford Press.

Trafton, J. A., Minkel, J., Humphreys, K. (2006). Opioid substitution treatment reduces substance use equivalently in patirents with and without posttraumatic stress disorder. *Journal of Studies on Alcohol,* 67 (2), 228-235.

Triffleman, E., Carroll, K., Kellogg, S. (1999). Substance dependence posttraumatic stress disorder therapy: An integrated cognitive-behavioral approach. *Journal of Substance Abuse Treatment,* 17, 3-14.

Van Etten, M., L., Taylor, S. (1998). Comparative efficacy of treatments for post-traumatic stress disorder: A meta-analysis. *Clinical Psychology and Psychotherapy*, 5 (3), 126-144.

Vocci, F. J., Acri, J., Elkashef, A. (2005). Medication development for addictive disorders: The state of the science. *American Journal of Psychiatry,* 162 (8), 1432-1440.

Waller, S. A., Mulick, P. S., Spates, C. R. (2000). Interventions for PTSD: A meta-analysis and quantitative review. *Proceedings of the Third World Conference of the International Society for Traumatic Stress Studies,* Melbourne, Australia.

Author Contact Information:
Theresa M. Souza, MS
Western Michigan University
230 Morhouse St
Galesburg, MI 49053
269-665-9915

Responsible for major organization of the manuscript, major writing and final editing of the manuscript

C. Richard Spates, Ph.D.
Professor & Director Anxiety Disorders Laboratory
Department of Psychology
Western Michigan University
Kalamazoo, MI 49008
Tel: 269 387-4329
Fax: 269 387-4550

Assisted with the organization, direction to selected resources, and assisting with editing of the manuscript

Extinction-induced response resurgence: A selective review

Adam H. Doughty and Gabriella Oken

Resurgence refers to the recovery of previously extinguished responding when a recently reinforced response is extinguished. Although the topic of resurgence has received limited experimental attention, there recently have been an increased number of investigations involving the topic. This increased experimental attention also has been accompanied by conceptual analysis. This increased interest in resurgence by both basic and applied behavior analysts is noteworthy because the topic relates to several different areas in and outside behavior analysis. This paper is a brief and selective review of resurgence, and its aim is to illustrate the importance of continuing to investigate the topic. More specifically, the role of resurgence in understanding topics such as behavioral history, drug relapse, severe problem behavior, communication disorders, and cognition, is described.
Keywords: Behavioral history, extinction, reinforcement, resurgence, review

---

Resurgence refers to the recovery of previously extinguished responding when a recently reinforced response is extinguished (e.g., Epstein, 1983, 1985). Figure 1 illustrates resurgence using hypothetical data. In the first condition, only Response A is reinforced. Response A is extinguished in the second condition concurrent with the reinforcement of Response B. Response B is extinguished in the third condition, and the subsequent recovery of Response A is resurgence. The degree of recovery of Response A can be compared to a different (control) response that never was reinforced (Response C). Measuring Response C in this third condition serves as a measure of extinction-induced response variability (e.g., Antonitis, 1951). Thus, resurgence provides a more refined characterization of extinction-induced behavior. That is, responses predicted to occur in extinction are those responses that previously were reinforced under comparable stimulus conditions.

Although resurgence and related extinction-induced phenomena were investigated in the 1970s (e.g., Mulick, Leitenberg, & Rawson, 1976; Pacitti, & Smith, 1977; Rawson, Leitenberg, Mulick, & Lefebvre, 1977), as well as earlier (see Epstein, 1985), Epstein (1983, 1985) usually is viewed as being the most forceful commentator regarding the utility of understanding the determinants and consequences of resurgence (see also Epstein & Skinner, 1980). Epstein argued that resurgence provides behavior analysts with a more descriptive means of understanding findings sometimes attributed to Freud's concept of *regression*. For Freud, regression involved an individual *regressing* to previously demonstrated behavior considered acceptable earlier in the individual's development. In addition to linking resurgence and regression, Epstein also summarized much of the experimental findings related to resurgence. The aim of the present paper, therefore, is to describe more recent experimental findings related to resurgence. By showing the relevance of the topic to a variety of topics in and outside behavior analysis (behavioral history, drug relapse, severe problem behavior, communication disorders, and cognition), this paper urges both basic and applied behavior analysts to continue the analysis of resurgence.

Behavioral history

Appealing to the behavioral history of an organism often provides behavior analysts with a means of explaining complex behavior (e.g., Barrett, 1986; Doughty, Cirino, Mayfield, da Silva, Okouchi, & Lattal, 2005; Lattal, & Neef, 1996; Tatham, & Wanchisen, 1998; Wanchisen, & Tatham, 1991). Broadly speaking, resurgence is a behavioral-history effect in that current behavior only can be understood by appealing to contingencies of reinforcement exposed to the organism previously. A primary issue,

therefore, in the study of resurgence is: what are the necessary and sufficient conditions to produce resurgence, in terms of both past experiences and current environmental conditions?

Lieving and Lattal (2003) addressed the question posed above across several experiments. In one experiment, key pecking of pigeons was reinforced via grain delivery in the first condition. In the second condition, key pecking was extinguished concurrent with the reinforcement of treadle pressing (also via grain delivery). In the third condition, treadle pressing was extinguished, and the resurgence of key pecking was measured. By repeating these three conditions, and consistently observing key-peck resurgence in the third condition, it was confirmed that resurgence is a general and replicable finding. In two subsequent experiments, the necessity of extinguishing treadle pressing in producing key-peck resurgence was investigated. In one experiment, instead of arranging extinction in the third condition, food delivery was made response independent. Not only was key-peck resurgence absent in this condition, but in a subsequent condition key pecking resurged when extinction replaced the response-independent, food-delivery schedule. These latter results suggest that resurgence necessitates a prior reinforcement history and the *extinction* of a more recently reinforced response. In a final experiment, the issue of whether *local periods of nonreinforcement* (i.e., conditions *approximating* extinction) resulted in resurgence of key pecking was addressed. In the third condition of this experiment, the reinforcement schedule maintaining treadle pressing was made *leaner*. Resurgence of key pecking occurred, though it was limited in amount. These results have compelled investigators to investigate the necessity of *complete* extinction in producing resurgence (Doughty & Reed, 2007; see below under *Summary and conclusions*).

Drug relapse

The clinical significance of analyzing resurgence, if not clear already, is illustrated well by *naming* Response A and Response B. That is, by suggesting that Response A can represent drug taking and Response B a therapeutically established alternative response, it is obvious that resurgence sometimes is undesirable from a therapeutic standpoint. Podlesnik, Jimenez-Gomez, and Shahan (2006) used resurgence as a model of clinical relapse in the context of *drug seeking* (i.e., an operant response maintained by alcohol reinforcement).

Podlesnik et al. (2006) first reinforced right-lever pressing of rats with alcohol administration. In a second condition, not only was right-lever pressing extinguished, but chain pulling was reinforced with food pellets. In the final condition, when food-pellet delivery was discontinued, and chain pulling decreased, there was a resurgence of right-lever pressing despite the absence of alcohol administration. The recovery of right-lever pressing is described appropriately as resurgence because there was minimal responding to the left (control) lever in this final (extinction) condition. Resurgence was observed by Podlesnik et al. despite a difference in procedure compared to Lieving and Lattal (2003). In Lieving and Lattal, grain delivery was the reinforcer across all conditions, whereas in Podlesnik et al., different reinforcers (i.e., alcohol and food) maintained the different responses (i.e., lever pressing and chain pulling).

In proposing their animal model of relapse, Podlesnik et al. (2006) noted that food-maintained chain pulling may be analogous to therapeutically established alternative environment-behavior relations. These latter relations may be drug treatment, the establishment of healthier family relationships, friendships, and/or employment opportunities. The cessation of food-pellet delivery may be discontinued treatment, family strife, divorce, and/or unemployment.

Severe problem behavior

A considerable number of applied behavior analysts treat problem behavior in individuals with intellectual disabilities. Because applied behavior analysts often are not requested to treat problem

behavior until it becomes severe, they necessarily must acknowledge and combat the behavioral history of their clients (e.g., Progar et al., 2001). Lieving, Hagopian, Long, and O'Conner (2004) reported one way resurgence can be considered when assessing severe problem behavior.

Lieving et al. (2004) studied two participants, a 7-year-old girl (Christine) diagnosed with moderate to severe intellectual disabilities who engaged in disruption, aggression, and self injury, and one 9-year-old boy (Sam) diagnosed with mild intellectual disabilities (as well as other disorders) who engaged in disruption, dangerous acts, inappropriate language, and aggression. For both participants, it was hypothesized at the start of the study that the specific problem-behavior topographies listed above made up a response-class hierarchy of aggression maintained by positive reinforcement in the form of tangibles. Lieving et al. began the study by reinforcing any form of aggression. Extinction then was implemented across conditions for one specific response topography after another, while the occurrence of all these topographies was measured. For example, Sam first was exposed to extinction of disruption, then extinction of disruption and dangerous acts, and finally extinction of disruption, dangerous acts and cursing. Both participants exhibited resurgence at various times across conditions by emitting a different form of aggression after being exposed to extinction for some other form of aggression.

Lieving et al. (2004) did not report the eventual treatment they implemented for these two participants. However, their results force practitioners to consider carefully the role resurgence may play in the *assessment* of severe problem behavior. Furthermore, in linking Lieving et al. and Podlesnik et al. (2006), practitioners also must consider resurgence when establishing *treatment* effects that will persist best outside the clinic setting. That is, it is desirable to prevent problem behavior from resurging upon the client's return to the target setting (i.e., after the problem behavior has been eliminated in the clinic).

Communication disorders

In keeping with the notion that effective treatment must be resilient to conditions that promote the resurgence of undesirable behavior, discussed next is the role resurgence may have in understanding communication breakdowns and problem behavior in individuals with communication disorders (e.g., Halle, Brady, & Drasgow, 2004).

Communication breakdowns occur when speakers (for example, young children with limited communicative repertoires) initiate communication but do not receive their requested consequences. Halle et al. (2004) describe *repairs* as the subsequent responses of speakers following these breakdowns, and the function of these repairs is to procure the requested consequence. Halle et al. classify three types of communication breakdowns and note their varying effects on repair frequency. *Requests for clarification* reliably occasion repair attempts in the form of a slightly different response topography. *Nonacknowledgements* do not occasion repairs reliably in that the speaker does not receive a response from the listener. That being said, requesting may persist if maintained by valuable consequences. Finally, *topic shifts* are when listeners attempt to manipulate the speakers' attention to a different subject, and children have been shown to repair least frequently to these breakdowns. That being said, the reliability of this latter effect depends on a variety of variables (e.g., reinforcement factors correlated with the initial request as well as with the changed-to topic).

Halle et al. (2004) also classify two different types of repairs, repetitions and modifications. In a repetition, the response of the speaker is an exact replica of the original communication attempt, and in a modification, the response is different from the original attempt. Three modifications were described: additions, in which the original attempt is augmented in some way; reductions, in which something present in the original attempt is removed; and substitutions, in which the original attempt is changed considerably. Observing the occurrence of a desirable repair on the behalf of the child is important because in the absence of such a repair problem behavior may occur. It seems reasonable to suggest that

the emission of problem behavior following a communication breakdown is the result of that behavior being reinforced in the history of the organism (i.e., the resurgence of problem behavior following the extinction of an appropriate communication attempt). It, therefore, is imperative that practitioners arrange conditions to maximize the probability of observing *effective* repairs. Halle et al. describe well the reinforcement factors responsible for establishing persistent repairs in the face of breakdowns.

Cognition

An essential aspect of cognitive therapy is the instruction from the therapist to the client to *refrain* from thinking about the targeted problem (e.g., the obsession in obsessive-compulsive disorder). These instructions sometimes include the establishment of an alternative response (i.e., the client is told to think of X instead of Y, where Y is the targeted problem). If the recurrence of the targeted thought is problematic, then effective treatment must reduce its probability of occurrence (i.e., its resurgence under *unwanted* circumstances). Several cognition studies with non-clinical populations, usually college students, have sought to examine the conditions under which people are more or less successful at effective thought suppression.

Wegner, Schneider, Carter, and White (1987) studied the so-called *paradoxical effects of thought suppression*. The focus was on why, when attempting to suppress a particular thought, the content of this thought tends to recur at relatively high levels. In Phase 1, one group of students was asked to relay a stream of consciousness into a tape recording in several 5-min periods. After a number of sessions, the students were asked to continue reporting these streams but to *abstain from thinking of a white bear*. They were told to ring a bell if they thought of a white bear. The students supposedly found it extremely difficult to suppress white-bear thoughts. In Phase 2, the same students were asked to report their streams of consciousness as they had done *before* the request was given to suppress white-bear thoughts. Students supposedly still showed a white-bear preoccupation. With a different group of students, Wegner et al. repeated the aforementioned procedures; however, these students were told in Phase 2 to think of a *red Volkswagen* if white-bear thoughts arose. Although white-bear thoughts supposedly were present in this group of students as well, they were significantly less likely than the first group of students to think about the white bear. Evidently, the establishment of a specific alternative response reduced the recurrence of the unwanted thought.

Although the differences between Wegner et al. (1987) and the studies discussed above are too numerous to list, it nevertheless may be noteworthy to consider the results of Wegner et al. in terms of behavioral resurgence. That is, when placed in a stimulus context previously correlated with a target response, there are times when that discriminative context occasions the target response, as opposed to some alternative, more recently learned response. Wegner et al. describe their results in the following way, "some reminder occurs, [which] in a moment of weakness, the person gives rise to the rumination [of the formerly suppressed thought]." In the context of behavioral work involving resurgence, some investigators (e.g., Doughty, da Silva, & Lattal, 2007) have analyzed the conditions that promote *better* this response recurrence.

Summary and conclusions

Resurgence is a topic that has not received extensive empirical attention in behavior analysis. That being said, behavior analysts recently have shown an increased interest in the topic. This interest is evidenced in the basic non-human animal laboratory (e.g., Doughty et al., 2007; Doughty & Reed, 2007; Lieving & Lattal, 2003), in the translational non-human animal laboratory (Podlesnik et al., 2006), in the basic human laboratory (e.g., Wilson & Hayes, 1996), and by applied behavior analysts (e.g., Lieving et al., 2004). The topic also has been included in novel conceptual analyses (e.g., Shahan & Chase, 2002). Although the present paper did not include detailed discussion of all these issues (e.g., Shahan & Chase,

2002; Wilson & Hayes, 1996), it did provide a glimpse into some of the areas, both in and outside behavior analysis, related to resurgence.

Regarding resurgence in the basic non-human animal laboratory, the procedures and results of Lieving and Lattal (2003) were presented in detail. This set of experiments was described partly because it analyzed systematically the necessary and sufficient conditions that promote resurgence. An intriguing aspect of these data was that resurgence may not be confined to *complete* extinction. That is, by reducing the rate of reinforcement rate correlated with Response B, these investigators found some, albeit limited, evidence of resurgence. Doughty and Reed (2007) extended Lieving and Lattal by assessing resurgence following a reduction in *reinforcer magnitude* for Response B. Doughty and Reed also found some evidence of resurgence. The conditions necessary to uncover resurgence clearly need to be studied further.

Regarding resurgence in the translational non-human animal laboratory, Podlesnik et al. (2007) introduced the topic to experimental psychologists interested in drug relapse. Because animal models of relapse have focused on extinction-related phenomena other than resurgence (i.e., reinstatement; Doughty, Reed, & Lattal, 2004; Shaham, Shalev, Lu, de Wit, & Stewart, 2003), much work still is needed to assess the utility of resurgence in this area. Similarly, because applied behavior analysts still have not examined resurgence systematically (cf. Lieving et al., 2003), there remains considerable work in this area as well. For example, Doughty et al. (2007) observed that the onset and amount of resurgent key pecking in pigeons depended on how the key pecking was eliminated. The response-elimination techniques studied mirror procedures used by applied behavior analysts to eliminate problem behavior (i.e., extinction, differential reinforcement of *other* behavior, and differential reinforcement of an *alternative* behavior).

Arguably the most intriguing questions still unanswered involving resurgence relate to its ability to provide connections between behavior analysis and other fields of study. Two such fields discussed above were communication disorders (Halle et al., 2004) and cognitive therapy (Wegner et al., 1987). During a time in which behavior analysts sometimes find themselves isolated from other fields of study, resurgence may provide a means of collaboration. More importantly, behavior analysts have much to offer to these fields by showing how behavioral concepts and principles discovered in the non-human animal laboratory extend so well to socially significant human behavior outside the laboratory.

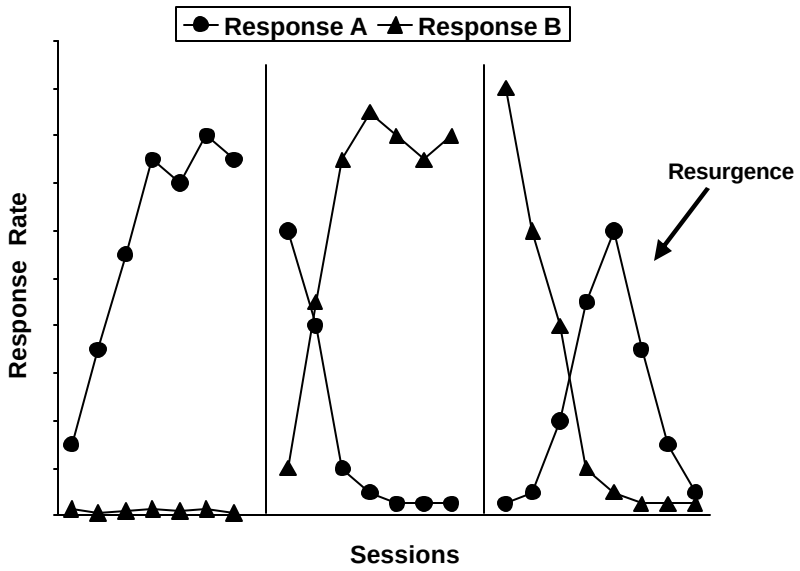

Figure 1. Hypothetical data illustrating resurgence.

References

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42,* 273-281.

Barrett, J. E. (1986). Behavioral history: Residual influences on subsequent behavior and drug effects. In N. Krasnegor & D. B. Gray (Eds.), *Advances in behavioral pharmacology,* Vol. 5: Developmental behavioral pharmacology (pp. 99-114). Hillsdale, NJ: Erlbaum.

Doughty, A. H., Cirino, S., Mayfield, K. H., da Silva, S. P., Okouchi, H., & Lattal, K. A. (2005). Effects of behavioral history on resistance to change. *The Psychological Record, 55,* 315-330.

Doughty, A. H., da Silva, S. P., & Lattal, K. A. (2007). Differential resurgence and response elimination. *Behavioural Processes, 75,* 115-128.

Doughty, A. H., & Reed, P. (2007). Resurgence under decreased reinforcer magnitude. *Unpublished data.*

Doughty, A. H., Reed, P., & Lattal, K. A. (2004). Differential reinstatement predicted by preextinction response rate. *Psychonomic Bulletin and Review, 11,* 1118-1123.

Epstein, R. (1983). Resurgence of previously reinforced behavior during extinction. *Behaviour Analysis Letters, 3,* 391-397.

Epstein, R. (1985). Extinction-induced resurgence: Preliminary investigations and possible applications. *The Psychological Record, 35,* 143-153.

Epstein, R., & Skinner, B. F. (1980). Resurgence of responding after the cessation of response-independent reinforcement. *Proceedings of the National Academy of Sciences, 77,* 6251-6253.

Halle, J., Brady, N. C., & Drasgow, E. (2004). Enhancing socially adaptive communicative repairs of beginning communicators with disabilities. *American Journal of Speech Language Pathology, 13,* 43-54.

Lattal, K. A., & Neef, N. A. (1996). Recent reinforcement-schedule research and applied behavior analysis. *Journal of Applied Behavior Analysis, 29,* 213-230.

Lieving, G. A., Hagopian, L. P., Long, E. S., & O'Conner, J. (2004). Response-class hierarchies and resurgence of severe problem behavior. *The Psychological Record, 54,* 621-634.

Lieving, G. A., & Lattal, K. A. (2003). Recency, repeatability, and reinforcer retrenchment: An experimental analysis of resurgence. *Journal of the Experimental Analysis of Behavior, 80,* 217-233.

Mulick, J. A., Leitenberg, H., & Rawson, R. A. (1976). Alternative response training, differential reinforcement of other behavior, and extinction in squirrel monkeys (Saimiri sciureus). *Journal of the Experimental Analysis of Behavior, 25,* 311-320.

Pacitti, W. A., & Smith, N. F. (1977). A direct comparison of four methods for eliminating a

response. *Learning and Motivation, 8,* 229-237.

Podlesnik, C. A., Jimenez-Gomez, C., & Shahan, T. A. (2006). Resurgence of alcohol seeking produced by discontinuing non-drug reinforcement as an animal model of drug relapse. *Behavioural Pharmacology, 17,* 369-374.

Progar, P. R., North, S. T., Bruce, S. S., DiNovi, B. J., Nau, P. A., Eberman, E. M., Bailey, J. R., Jr., & Nussbaum, C. N. (2001). Putative behavioral history effects and aggression maintained by escape from therapists. *Journal of Applied Behavior Analysis, 34,* 69-72.

Rawson, R. A., Leitenberg, H., Mulick, J. A., & Lefebvre, M. F. (1977). Recovery of extinction responding in rats following discontinuation of reinforcement of alternative behavior: A test of two explanations. *Animal Learning and Behavior, 5,* 415-420.

Shaham, Y., Shalev, U., Lu, L., de Wit, H., & Stewart, J. (2003). The reinstatement model of drug relapse: History, methodology and major findings. *Psychopharmacology, 168,* 3-20.

Shahan, T. A., & Chase, P. N. (2002). Novelty, stimulus control, and operant variability. *The Behavior Analyst, 25,* 175-190.

Tatham, T. A., & Wanchisen, B. A. (1998). Behavioral history: A definition and some common findings from two areas of research. *The Behavior Analyst, 21,* 241-251.

Wanchisen, B. A., & Tatham, T. A. (1991). Behavioral history: A promising challenge in explaining and controlling human operant behavior. *The Behavior Analyst, 14,* 139-144.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53,* 5-13.

Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior, 66,* 267-281.

Address correspondence to:

Adam H. Doughty, Ph.D., Assistant Professor
Department of Psychology, College of Charleston
66 George Street
Charleston, SC 29424
Telephone: 843.953.7300
Fax: 843.953.7151
Email: doughtya@cofc.edu

Gabriella Oken, Assistant Professor
Department of Psychology, College of Charleston
66 George Street
Charleston, SC 29424
Telephone: 843.953.7300
Fax: 843.953.7151
goken@edisto.cofc.edu).

Can it be it was all so simple then? A Reply to Von Hilsheimer and Quirk (2006)

Joseph Cautilli

The 1970s was a wonderful time for those studying operant based procedures. The future of the field looked very promising in many areas to name but a few vision (Cooper & Feldman, 1980), epilepsy (Lubar, & Bahler, 1976) and hyperactivity. In particular, operant based neurofeedback seemed to have promise then and still has promise now for treating offenders (Martin & Johnson, 2005). While I enjoyed Hilsheimer and the late Quirk's (2006) personal trip down memory lane, I found some aspects of the article to not be consistent with my knowledge of this area.
Keywords: operant procedures, offenders, treatment

Indeed, the neurofeedback conceptualization remains an appealing and compelling one: simple shaping of EEG patterns by operant conditioning that correlate with problem behavior and looking for adjunctive (Falk, 1971) decreases in the problem behavior. Indeed, there was much to enjoy about the article; however, one claim from the article appears to need some citation as support or proof. The authors claim:

> "In a series of reports to Corrections officials, Quirk indicated that the three year recidivism rate from the biofeedback unit (using these techniques) had never been more then 45% and in some groups had fallen to 15%. Overall, among nearly 3000 felons trained by Quirk's volunteers, the rate of recidivism was consistently close to 15%." (p. 501)

This claim is unreferenced. Indeed, in the only other publication that I was able to find (e.g., Quirk, 1995) the best results indicate that subjects had 20% recidivism. In this study, the number of people trained was approximately 300 not 3000. It is unclear if the three thousand represents the overall training number of O.C.I or if it is the number of people in published studies. From review of studies, I conclude it must be the former and not the latter. While 20% could be considered a good recidivism rate, the study offers neither control group nor a reference to the recidivism rate for the Ontario area for the type of offenses that Von Hilsheimer and Quirk (2006) worked to correct. Without that information, the number remains an unsubstantiated claim. In addition, it is important to remember that as Quirk (1995) states "The O.C.I. serves adult males sentence to less than 2 years of incarceration" (p. 10). Thus, Quirk was not working with highly recidivistic or violent offenders.

In addition, O.C.I. was a treatment facility for offenders and it is unclear if offenders were receiving other treatments while there. It is not unreasonable to think that offenders who engaged in high rates of neurotherapy as a treatment also engaged in high rates of other treatments.

The methodology in the Quirk (1995) used a titration design to show that those who received more therapy improved more; however, with no objective standards to compare the recidivism rate- the number must remain questionable. Indeed, the paper would have been greatly enhanced by citing a report on recidivism for that area of Ontario. In addition, since these studies occurred in the 1970s, the fact that they have not been replicated in over 35 years increases my skepticism that they will ever be replicated.

Interestingly the 15% number has become part of the clinical lore, which seems to have encompassed most of the biofeedback and neurofeedback community. It has been repeated so many times, I am sure that the author and many will believe it true. As Bem (1966, 1967) pointed out when looking at memory, the context can play a powerful role in what is remembered. Bem further stated,

"Saying becomes believing only when we feel the presence of truth, and certainly only when a minimum level of inducement of and the mildest and most subtle forms of coercion are used. "(pp. 23-24). At conferences and with other clinicians, belief in interventions may be especially strong, coupled with legitimate success and the belief that the field has gone unnoticed, can help explain the emergence of clinical lure and the adherence to such lore.

In closing, I would like to repeat that I found this report of work in the 1960s and 1970s to be fascinating and very interesting. I hope that it becomes replicated many times over and ways to strengthen and enhance the package are found. However, it seems to me that the paper lacked critical references that would have made the claims more understandable and consistent with the literature.

References

Bem, D. J.(1966). Inducing belief in false confessions. *Journal of Personality and Social Psychology, 3,* 707-710.

Bem, D. J.(1967). When saying is believing. *Psychology Today, 1(2),* 21-25.

Cooper, J., & Feldman, J.(1980). Operant Conditioning of fusional convergence ranges using random dot stereograms. *American Journal of Optometry and Physiological Optics,* 205-213.

Falk, J.L. (1961). The nature and determinants of adjunctive behavior. *Physiology and Behavior, 6,* 577-588.

Lubar, J. F., & Deering, W. M. (1981). *Behavioral approaches to neurology*. Academic Press

Lubar, J. F., & Bahler, W. W. (1976). Behavioral management of epileptic seizures following EEG biofeedback training of the sensorimotor rhythm. *Biofeedback and Self-Regulation, 1,* 77-104.

Martin, G. & Johnson, C.L. (2005). The Boys Totem Neurofeedback Project: A pilot study of EEG biofeedback with incarcerated juvenile felons. Journal of Neurotherapy, 9(3), 71-86.

Quirk, D.A. (1995). Composite Biofeedback Conditioning and Dangerous Offenders: III Journal of Neurotherapy, 1 (2),

Von Hilsheimer & Quirk, D.A. (2006). Origin of EEG biofeedback for remediating misbehavior. *The Behavior Analyst Today, 7(4),* 492-507

Author Contact Information
Joseph Cautilli, Ph.D.
St. Joseph's University
Department of criminal justice
Philadelphia, PA 19131
(215) 496-0707 x1196
Jcautilli2003@yahoo.com

Evaluating Curriculum-Based Measurement from a Behavioral Assessment Perspective

Scott P. Ardoin, Claire M. Roof,
Cynthia Klubnick & Jessica Carfolite

Curriculum-based measurement Reading (CBM-R) is an assessment procedure used to evaluate students' relative performance compared to peers and to evaluate their growth in reading. Within the response to intervention (RtI) model, CBM-R data are plotted in time series fashion as a means modeling individual students' response to varying levels of instruction and the interpretation of these data is used as a source of information for making special education eligibility decisions. While substantial evidence exist demonstrating the reliability and validity of CBM-R procedures from a classical test theory (CTT) perspective, little evidence exist demonstrating the quality of CBM-R from a behavioral assessment perspective. This paper discusses (a) the necessity of evaluating CBM-R from a behavioral assessment perspective and (b) those studies which have evaluated CBM-R from a perspective other CTT, and (c) recommendations for future research.
Keywords: Curriculum-based measurement, behavioral assessment, absolute versus relative performance, accuracy, sensitivity, treatment utility.

Curriculum-based measurement for reading (CBM-R) procedures were developed in the late 1970's and early 1980's as a set of standardized assessment tools for gauging students' academic performance in reading. It was developed in order to provide teachers with an efficient, easily understood measurement system yielding relevant data about students' level of performance, as well as their reading growth over time (Deno, 1985). Educators and researchers have noted that an important characteristic of CBM-R is its ability to measure both inter-individual differences in groups of students as well as intra-individual change within specific students (Fuchs & Fuchs, 1998; Fuchs, Fuchs, & Speece, 2002). While CBM-R data were initially used exclusively to guide low-stakes educational decisions (Deno, 1985; Deno, 1986; Deno, Marston, & Tindal, 1985; Deno & Shinn, 1989), CBM-R data are now being used for making high-stakes decisions (i.e., special education eligibility) within Response to Intervention (RtI) models.

Several features distinguish CBM-R procedures from other standardized measures used to assess students' reading. First, the assessment materials are relatively cheap and it requires little time to administer probes to students. Second, CBM-R is meant to be a measurement of students' global reading performance, which allows for practitioners to evaluate how students' are progressing toward towards long-term goals (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993). Finally, as described by Deno et al.

> "CBM-R departs from conventional psychometric applications by integrating the concepts of standardized measurement and traditional reliability and validity with features from behavioral and observational assessment methodology: repeated performance sampling, fixed time recording, graphic display of times-series data, and qualitative descriptions of performance" (Deno et al., 2001, p. 508)

These characteristics makes CBM-R ideal for use within an RtI model, as an instrument used to both identify students at-risk for academic problems and to evaluate individual students' response to instruction.

Within the Fuchs and Fuchs (1998) dual discrepancy RtI model, CBM-R data are first used to identify students who are at-risk for academic problems based upon a comparison of their performance to

that of their normative group (i.e., nomothetic context). Such decisions are relatively low-stakes decisions, given that the identification of a student at-risk simply results in the student being provided with supplemental instruction. CBM-R data modeling individual student's response to the supplemental instruction is then used as a primary source of data for making high-stakes decisions. Students' response to supplemental instruction is generally evaluated using CBM-R progress monitoring procedures, which entails the frequent administration of CBM-R probes and plotting of collected data in time-series fashion. Progress monitoring data are evaluated by comparing the plotted data to a pre-established goal line or an estimate of weekly growth calculated using ordinary least square regression techniques. Regardless of the method of comparison, these evaluations of data are within an idiographic context in which an individual's data are compared to his/her previous performance(s). Based upon the evaluation of data, one of the following high-stakes decisions is made: (a) the intervention was successful, therefore the student is not eligible for special education and the intervention should be terminated, (b) the intervention was not adequate and a more intense intervention is needed, or (c) supplemental interventions of varying levels of intensity have not been successful and the student is eligible for special education.

The importance of the psychometric adequacy of CBM-R is especially salient now that CBM-R data are being used as a primary source for making high-stakes educational decisions (Ardoin & Christ, in press; Christ & Ardoin, 2007). As previously noted, CBM-R combines features from CTT as well as behavioral assessment. Ample evidence is available demonstrating the reliability and validity of CBM-R from the CTT perspective. Fuchs, Tindal, and Deno (1984) reported that CBM-R probes demonstrated adequate criterion validity when compared to traditional measures of reading, such as the Stanford Diagnostic Reading Test and the Woodcock Reading Mastery Test. CBM-R has also demonstrated a high level of correspondence with other widely used reading tests (Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988; Shinn, 1989), tests of reading comprehension, and teachers' perceptions of reading competence (Fuchs & Fuchs, 1992). Furthermore, CBM-R data have been shown to effectively discriminate populations such as general and special education students (Deno, 1985; Deno, Shinn, Marston, & Tindal, 1983) and students from different grade levels (Deno, 1985; Fuchs et al., 1993).

Despite proponents of CBM-R frequently citing studies employing CTT as proof that CBM-R is psychometrically sound and thus appropriate for use within an RtI framework, CTT is not necessarily the most appropriate model for evaluating its adequacy within an RtI framework. Fuchs and Fuchs (1998) note that traditional psychometric standards provide insufficient evidence for using a measure to model individual growth. CTT evaluates the stability of group level data in which it is ideal for each student's performance to remain stable relative to the sampled group. Any changes in performance relative to the sampled group are treated as error and either explained away or ignored. Observed scores are considered to represent an individual's true score plus error. In contrast to these principals of CTT, when evaluating CBM-R data plotted in time series fashion (a) data are being compared to the previous performances of the same student, rather than to a sample group; and (b) change in student performance is desired, rather than treated as error. It is the departure of CBM-R from traditional psychometrics into behavioral assessment that permits modeling of an individual student's growth and thus it is also necessary that CBM-R be evaluated from a behavioral assessment perspective in order to demonstrate its appropriateness within an RtI framework.

Evaluation of CBM-R from a Behavioral Assessment Perspective

Nelson and Hayes (1979) state that the goal of behavioral assessment is to understand and modify behavior through the identification of meaningful behavioral responses and the environmental and internal factors that control them. Practitioners dealing with behavioral issues are motivated by the utility of measurement in generating change in all relevant environments, and are therefore interested in what makes an individual's behavior fluctuate from situation to situation. For these reasons, behavioral assessment is conducted from an idiographic approach in which the individual is perceived as an entity

with qualities that distinguish him or her from others, rather than as part of a cohort or category. Due to the fundamental differences in the nature of the traditional and behavioral approaches to assessment, Hayes, Nelson, and Jarrett (1986) maintain that the two areas should have separate standards for evaluation. Hayes et al. specified three distinct areas in which behavioral assessment tools should be evaluated: accuracy, sensitivity, and treatment utility.

Accuracy. Cone (Cone, 1981) defined accuracy as "how faithfully a measure represents objective topographic features of a behavior of interest" (p. 59). The first step towards establishing the accuracy of a measure is to (a) develop explicit instructions to accompany the measure in order to establish consistency in measurement and (b) select a behavior that is verbally definable (Cone, 1977, 1992; Nelson & Hayes, 1979). Accuracy of a measure can then be established by comparing measurement results to those of some standard (similar to criterion-related validity). Since the targets of behavioral assessment are directly observed, accuracy is generally construed as how a measure compares with "reality" (Hayes et al., 1986). This definition contrasts with the interchangeability of the concepts of accuracy and reliability (i.e., consistency among scores across time, settings, items, etc.) that is often encountered in the literature on traditional assessment, where the focus is on inferred traits (Hayes et al., 1986). Accuracy of a behavioral measure can be established by weighing responses against those obtained with a mechanical record, or by emitting a scripted set of responses to ensure that the measure accurately detects them (Hayes et al., 1986).

Accuracy of CBM-R. CBM-R procedures clearly address the first steps towards establishing accuracy (i.e., explicit instructions, verbally definable target behavior). Explicit administration and scoring procedures have been available since the early development of CBM-R. These procedures involve having students read a passage for one minute while an examiner records misread words and words which the examiner provides to students when they hesitate on a word for three seconds. A word is marked as misread if a student mispronounces a word given the context of the story, skips a word, or transposes two words. Examiners provide words on which students hesitate for three seconds, but examiners do not correct students if they misread a word. After one minute elapses, student are asked to stop reading and the examiner calculates words read correctly by subtracting misread words from total words read. Evidence of the importance of adhering to the explicit administration directions has been provided by studies which have found variation in student performance as a function of directions given to students (Eckert, Dunn, & Ardoin, 2006; Fuchs, Tindal et al., 1984). Fortunately, researchers have consistently used the same explicit administration and scoring procedures across studies evaluating the reliability and validity of CBM-R.

Despite having clearly addressed the first steps of establishing the accuracy of a measure, researchers have not established the accuracy of CBM-R by comparing it to a known standard. Stoner (1992) suggested that accuracy was a "questionable metric" for evaluating CBM-R because no "gold standard" has been developed against which these data can be compared. Cone (1992) proposed using a tape of words read into a recorder at a speed paced by a metronome or using a computer with an audio digitizer that could count the number of words read at different speeds as a means of independently evaluating the accuracy with which examiners record words read correctly per minute (WRCM). Fifteen years later, researchers have not adhered to the suggestions made by Cone. Instead, the accuracy of CBM-R observations is based upon assessments of inter-rater agreement. Using audio recorders, numerous studies have evaluated the accuracy of observers on a word by word basis, comparing each word the observer reported as correct and incorrect to each word a blind rater scored as correct and incorrect. Inter-rater agreement is calculated as the percentage of agreement over disagreements and generally exceeds 90% (e.g., Ardoin, Suldo, Witt, Aldrich, & McDonald, 2005; Ardoin et al., 2004; Graney & Shinn, 2005). High rates of inter-rater reliability resulting from an easily quantified target behavior and standardized instructions, administration, and scoring procedures, support the measure's capacity to accurately detect and report the behavior of interest.

Instead of evaluating accuracy using procedures recommended by Cone, accuracy within the CBM-R literature typically is referred to in terms of standard error of measurement (*SEM*). *SEM* is the amount of error associated with the measurement method, and is influenced by the quality of instrumentation (e.g., equivalence in level of difficulty for parallel probes), as well as administration conditions (Christ & Ardoin, 2007). A method for evaluating this form of accuracy is to conduct similar but independent measurements (Hayes et al., 1986) in a short period of time during which meaningful growth is unlikely to have occurred (Ardoin & Christ, in press; Christ & Ardoin, 2007; Poncy, Skinner, & Axtell, 2005). An examiner may accurately record the WRCM for a student in several administrations of parallel probes during this time; yet we would still expect scores to be somewhat variable if, for example, the probes used are inconsistent in difficulty. Although consumers of CBM-R data rarely report the standard error associated with scores, it has been found that the *SEM* associated with CBM-R performance level estimates in typical testing conditions is actually quite substantial, approximating 10 WRCM (Christ & Silberglitt, 2007). This number is significant when considering that the expected performance level of a student in the spring of the first grade is only around 40 WRCM (Good, Simmons, & Kame'enui, 2001). Due to this performance variability, it has been recommended that the median score from a minimum of three probes administered at once should be used for screening purposes (Ardoin & Christ, in press; Poncy et al., 2005).

Because the majority of research evaluating CBM-R has been conducted from a CTT perspective, in which error is typically either explained away or ignored, only a few studies have evaluated the sources of error associated with student performance. In one of the few studies specifically evaluating CBM-R from a behavioral perspective, Derr-Minneci and Shapiro (1992) investigated the potential sources of error related to various environmental factors and found main effects for setting, tester, and timed versus untimed administrations. More recently efforts have been made to quantify the sources of error associated with CBM-R. Using generalizability theory researchers have separated the variance in CBM-R scores into that which is attributed to the person (i.e., student learning), the item (i.e., CBM-R probe passages), and residual error (Christ & Ardoin, 2007; Poncy et al., 2005). Results of Poncy, et al. indicated that up to 19% of the total variance in students' performance level could be explained by either the item (up to 10%) or residual error. Studies by Ardoin and colleagues (Ardoin & Christ, in press; Ardoin et al., 2005; Christ & Ardoin, 2007) suggest that the inability of readability formulas to adequately control passage difficulty is likely the primary source of variance (i.e., error) associated with items.

Depending on the context in which data are being evaluated, variability in student performance as a function of variation in passage difficulty can lead to misinterpretations of observed performance. For instance when evaluating relative performance, a student's WRCM is compared to other students administered the same probe(s) and thus the difficulty of the probe will be constant across students resulting in little to no change in relative performance. However, if performance is being evaluated at the individual child level, a student's performance on a probe is likely to be compared to a pre-established criterion or to the child's previous/future performances. Neither comparison considers the possibility of variation in passage difficulty, which is likely to result in variation in student performance that is not a function of change in the student's global reading skills. Given variability in student performance as a function of tester, setting, and passage difficulty, it is important that schools recognize that each observed score accounts for a student's true score plus error. In recognizing the lack of precision of an individual's performance confidence intervals should be placed around observed student CBM-R performance (Christ & Coolong-Chaffin, 2007).

Sensitivity. Psychometric theory cannot be unequivocally applied to behavioral assessment because behavior is more specific to situations than it is consistent across situations (Mischel, 1968). Inconsistencies in results could be due to actual changes in behavior, not weaknesses of the measure (Nelson & Hayes, 1979). This issue marks a key difference between the behavioral approach to assessment and its traditional counterpart (i.e., CTT) in which consistency is highly prized and

performance variability is labeled "error" (Cone, 1977; Hayes et al., 1986). A measure designed to detect a particular behavior should be sensitive to factors expected to have an impact on the occurrence of that behavior; without such consideration, concerns about the accuracy, reliability, or validity of the behavioral assessment tool are rendered "meaningless" (Hayes et al., 1986, p. 493).

In effect, the analysis of environmental influences on behavior constitutes the premise of behavioral assessment (Derr-Minneci & Shapiro, 1992; Nelson & Hayes, 1979). Nevertheless, sensitivity can be viewed as a double-edged sword precisely because the assessment process plays an integral role in the portion of variance that is due to the "situation" (Nelson & Hayes, 1979). For example, while an educator would want a test of academic performance to be sensitive to changes in learning, it would generally not be desirable if performance was affected by variables such as characteristics of the administrator. It is therefore imperative to identify and eliminate sources of unnecessary and uncontrollable systematic or random error and to foster an awareness of the potential causes of variation in assessment results (Christ, 2006; Christ & Ardoin, 2007; Hayes et al., 1986).

There are several means by which the sensitivity of a behavioral measure can be evaluated (Hayes et al., 1986). One method is to conduct idiographic correlations between different measures of the same construct over time. This technique allows for the recognition of changes in data at the *individual* level, taking into account the error associated with testing over time for each participant. Another method to evaluate the sensitivity of a measure is to implement an intervention that theoretically should result in changes in behavior and then to evaluate whether the measure detects differences in behavior (e.g., Ardoin & Martens, 2004). Additionally, a criterion measure known to demonstrate adequate sensitivity (i.e., a gold standard) can be used for comparison purposes (Hayes et al., 1986). When assessing sensitivity using these procedures, small-n-design procedures must be employed because the issue of sensitivity is often temporal in nature (i.e., involves an idiographic comparison of an individual's performance to his or her past performances). Understanding the nature of the data at the individual level is critical when the data are to be used for making decisions within an idiographic context (i.e., comparing a person's to his/her past performances). Nomothetic or group-level analyses, in contrast, do not take into account the fact that situational factors affect individuals differentially.

Sensitivity of CBM-R. Several intervention based studies provide evidence of the sensitivity of CBM-R by demonstrating that effective interventions can result in changes in a student's level of performance on a specific probe. Using multi-element designs, researchers have found that students performance on intervention passages and generalization passages increases as a function of varying degrees of intervention intensity when evaluated using CBM-R procedures (Ardoin, McCall, & Klubnik, 2007; Eckert, Ardoin, Daisey, & Scarola, 2000; McCurdy, Daly, Gortmaker, Bonfiglio, & Persampieri, 2007). While these studies demonstrate that CBM-R procedures are sensitive to direct intervention effects, they fail to provide evidence that CBM-R data used to model student growth are sensitive to variations in instruction.

Evidence used to support CBM-R as a measure of global gains in student achievement across time has largely centered on the idea that a steeper slope for WRCM signifies a more sensitive measure. From this perspective, when compared to traditional achievement measures (e.g., norm referenced individually/group administered achievement tests), CBM-R progress-monitoring techniques are clearly more sensitive to fluctuations in student learning (Deno, 1985; Fuchs, Fuchs, & Hamlett, 1989a; Marston, Fuchs, & Deno, 1986). While traditional reading achievement tests have been unable to differentiate between low-achieving and students with learning disabilities, CBM-R does in fact show a meaningful difference in the academic gains of these two groups of students (Shinn, Ysseldyke, Deno, & Tindal, 1986).

Studies have also been conducted demonstrating the sensitivity of CBM-R to model the growth of students by demonstrating differences in rates of growth as a function of variables that theoretically should result in different rates of gain. Some of the first studies in the area of the sensitivity of CBM-R involved longitudinal research on the relative effectiveness of general education and special education services for students at risk for reading problems (Marston, 1987). Weekly progress-monitoring data were used to show that academic growth, on average, was higher in the special education settings. Researchers have also used CBM-R to compare the effects of instruction in different classrooms, providing evidence that the quality of instruction provided to a child is a critical consideration in the evaluation of student performance (Fuchs & Fuchs, 1998; Speece & Case, 2001). Studies have also been conducted evaluating the sensitivity of CBM-R to different methods of reading instruction (literature-based vs. traditional based). Hintze, Shapiro, and Lutz (1994) found that on average CBM-R slopes were greater for the group of students instructed in the traditional curricula. It was concluded that steeper slopes may have signified the greater sensitivity of CBM-R to learning in traditional curricula, possibly as a result of greater overlap between the traditional curricula and assessment materials. Further evidence of the sensitivity of CBM-R progress monitoring data to variables that theoretically should alter rates of growth is provided by Allinder and Eicher (1994). In this study a decline in rates of growth was detected during the summer months due to lack of active instruction.

While results of the above cited studies (e.g., Allinder & Eicher, 1994; Hintze et al., 1994; Shinn et al., 1986) provides some evidence of the sensitivity of CBM-R progress monitoring procedures used to model students global gains in reading, it is important to attend to the context in which these studies have been conducted. These studies have evaluated the sensitivity of CBM-R to global gains in reading using CTT procedures to examine the growth of groups of students. Analyses of students' CBM-R progress monitoring data in order to evaluate their response to instruction within an RtI framework are not made based upon group level data. Rather, analyses within an RtI framework are conducted within an idiographic context to determine whether across time the observed performances of an individual student increases, decreases, or stays relatively the same.

Results of recent studies bring into question whether CBM-R might in fact be too sensitive to environmental factors for high-stakes decisions to be made with confidence outcomes (Christ, 2006; Christ & Ardoin, 2007; Christ & Silberglitt, 2007; Poncy et al., 2005). Variability in student performance as a function of inconsistencies in probe difficulty negatively impacts the reliability, accuracy, stability, and thus sensitivity of CBM-R progress monitoring (Christ, 2003; Hintze, Owen, Shapiro, & Daly, 2000; Hintze & Shapiro, 1997; Hintze, Shapiro, & Daly III, 1998; Hintze et al., 1994). Using estimates of error associated with rates of growth calculated for individual students from previously published studies, Christ (2006) found that the magnitude of error often exceeds expected rates of growth. These results indicate that CBM-R procedures are extremely sensitive to variables other than changes in rates of growth, suggesting that high-stakes decisions based upon estimates of CBM-R progress monitoring data should be made with extreme caution.

The behavioral assessment literature does offer several directions for evaluating data used in an idiographic context that should be applied to CBM-R. For example researchers could conduct idiographic correlations (Hayes et al., 1986) using two sets of CBM-R probe sets. Although significant correlations between individual students' performances on the two probe sets would not necessarily indicate that daily change was a function of growth, the data would provide evidence that changes were systematic and not a function of random error. Additional research must be conducted to identify and eliminate sources of error associated with CBM-R progress monitoring data to ensure that change in student performance is a function of change in global reading achievement.

Treatment Utility. Treatment utility was selected as an area of appraisal for behavioral assessment because the objective of behavioral assessment is to inform intervention (Hayes et al., 1986). Quality

assessment measures help practitioners develop goals that allow them to predict and control behavior in the most efficient manner possible, ultimately enhancing treatment outcomes (Nelson & Hayes, 1979). The first step towards establishing the treatment utility of a measure is selecting a target behavior that has meaning and can efficiently discriminate between individuals who need intervention and those who do not. Once a target behavior and a procedure for detecting it have been selected, the best way to evaluate treatment utility is to consider the results of decisions based on assessment data (Messick, 1980). This can be done after assessing the effectiveness of a treatment strategy by using a single-case experimental design with many subjects and correlating individual assessment results with individual treatment effects. Hayes et al (1986) and Nelson and Hayes (1979) provide several examples of idiographic methods for evaluating treatment utility when the relationship between assessment and treatment outcomes are predicted a priori. For example the "manipulated assessment" is described as systematically varying an aspect of assessment (e.g., the target behavior or the assessment method), and implemented treatment according to the assessment data. Treatment outcomes are then evaluated to assess the effectiveness of each. In a second method ("manipulated use"), all participants receive the same assessment, but the correspondence between assessment and treatment is manipulated. Finally, the "observed differences" method involves all participants receiving the same treatment which is evaluated for effectiveness, but differences in assessment are noted (Hayes et al., 1986; Nelson & Hayes, 1979).

Treatment Utility of CBM-R. CBM-R, consistent with a behavioral approach to assessment, was developed precisely for the purpose of providing information that teachers could use to better understand how a child functions in the curriculum (Deno, 1985). CBM-R procedures inform teachers by discriminating between typical and atypical performance (Deno, 2002; Shinn, Thomas, & Grimes, 2002), thus illuminating a student's need for intervention. Once a student is determined to be in need of academic assistance, CBM-R is designed to show the student's progress while receiving intervention (Fuchs, 2003), this information is helpful in making instructional decisions. Numerous studies have been conducted to investigate the ability of CBM-R to enhance teacher planning and most importantly improve student achievement (Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, & Hamlett, 1989b; Fuchs, Fuchs, Hamlett, & Stecker, 1991).

Evidence of the treatment utility of CBM-R is often provided by studies in which the students of teachers who view their students CBM-R progress monitoring data make greater academic gains than those students whose teachers do not view CBM-R progress monitor data. Fuchs and Fuchs (1986) conducted a meta-analysis of this research and found an effect size of .7 for progress-monitoring. In one study included in this review, Fuchs, Deno, & Mirkin (1984) investigated the effect of using CBM-R during the course of 18 weeks. A single CBM-R probe was administered twice weekly to each student in the experimental group. Teachers graphed data and used graphed data as a means of determining whether students were making progress toward IEP goals and thus whether instructional modifications were needed. In contrast, control group teachers used traditional informal procedures to track student growth (i.e., a "manipulated assessment" design). Results revealed increased student goal revisions initiated by the teacher, instructional structure, student awareness of their own learning, and student reading achievement for those students and teachers in the experimental condition. This study provided evidence supporting teachers' use of CBM-R when accompanied with explicit guidelines for modifying instruction. Moreover, Wesson (1991) found that special education students, whose teachers referred to CBM-R progress monitoring data and participated in group consultation regarding student performance, achieved higher levels of reading achievement than those students for whom progress-monitoring data were not collected.

Similar studies, conducted in general education settings where teachers are typically allotted fewer resources for attending to individual students, have found conflicting results. Graney and Shinn (2005) examined whether CBM-R progress monitoring information provided to a teacher about individuals and groups of students had an effect on reading performance. At five weeks of progress-

monitoring, teachers in the experimental groups viewed trend lines for their students and were provided the opportunity to discuss possible ways to adapt instruction with a consultant. Students in the control condition were tested with CBM-R, but teachers did not view these data (i.e., "manipulated use" design). Results indicated that feedback (concerning individuals or entire classrooms) had no impact on the reading achievement of the students at the group level, and that individual feedback may have actually had a negative effect. It appears that providing information to general education teachers on students' slopes, and even altering instruction based on these slopes, may not be enough to effect positive student change. This evidence is supported by other findings showing that when teachers are provided with CBM-R data without specific guidelines on how to use this information to guide instruction (i.e., "change the program decision rules"), the effects on student learning are negligible (Fuchs et al., 1989a).

Consistent with those studies evaluating the accuracy and sensitivity of CBM-R, studies evaluating the treatment utility of CBM-R at the individual student level are sparse. Although studies by Daly and colleagues (Daly, Persampieri, McCurdy, & Gortmaker, 2005; McCurdy et al., 2007) provide evidence of the use of CBM-R for measuring the effects of brief instructional trials on individual students, researchers have not yet evaluated the impact of progress monitoring on individual students. Caution should therefore be taken when using CBM-R progress monitoring data to inform treatment and decisions about individual students' responsiveness to instruction within an RtI framework. Within-subject design studies must be conducted correlating individual assessment results with individual treatment outcomes in order to account for error at this level of analysis. An additional issue that should be considered in future research examining the treatment utility of CBM-R progress monitoring data is whether improvements in performance over control students are in fact a function of making decisions based upon progress monitoring data. While most previous studies have employed a control condition for comparison purposes, these studies did not use a yoked control group whose instruction was changed on the basis of a student in the progress monitoring condition. It is possible that merely being exposed to dynamic instruction characterized by frequent changes benefitted students academically. Adding a yoked condition with these characteristics would greatly enhance the strength of the "manipulated assessment" and "manipulated use" designs (Hayes et al., 1986).

Conclusions

Numerous studies have been conducted for the purpose of exhibiting the strengths of CBM-R as an innovative academic assessment tool, and the results of these studies are often used to applaud CBM-R for the improvements that it represents over the traditional IQ/achievement test discrepancy model. Influenced by the proliferation of the RtI for determining special education eligibility, the educational context in which CBM-R is used has evolved to the present state, where both low-stakes and high-stakes decisions are made about individual students based on time-series data. Abundant evidence has been provided for the psychometric adequacy of CBM-R from a CTT perspective supporting its many uses for making low-stake educational decisions for students. Substantially fewer studies have however been conducted demonstrating the quality of CBM-R from a behavioral assessment perspectives in which the quality of a measure is based upon its accuracy, sensitivity, and treatment utility for making decisions within an idiographic context. It is within this context that high stake decisions are being made regarding student's special educational eligibility.

In order for CBM-R data to be used as a primary source of information for high-stakes decision making several steps must be taken. First, equivalent sets of CBM-R reading probes must be developed. We know that readability formulas are inadequate for selecting passages equivalent in level of difficulty and that variation in student performance is a primary source of error. Ignoring this fact will result in students being misidentified for special education. Second, we must begin to identify sources of error other than variability in passage difficulty and determine procedures for minimizing this error. Considering studies have illustrated that variation in CBM-R directions, performance feedback, goal

setting, and reinforcement influences student performances these may be areas to consider (Eckert et al., 2006; Fuchs, Tindal et al., 1984). For example, standard CBM-R directions simply inform students to do their "best reading." It is possible that the meaning of "best reading" changes as students observe themselves being timed repeatedly (Colon & Kranzler, 2006). Defining best reading for students and providing them with a constant source of reinforcement of doing their best reading might results in more consistent performance. Finally, researchers need to address the question of how many data points across, how many weeks are needed to make a reliable and accurate decision regarding the effectiveness of instruction for students. Current sources used to provide evidence that 20 data points are sufficient for making decisions based upon CBM-R data were designed for making low-stake decisions (Good & Shinn, 1990; Shinn, Good, & Stein, 1989). It is also questionable whether these studies actually support the use of 10-20 data points as the magnitude of error in predicting future student performance within these studies is greater than expected rates of student gain (Good & Shinn, ; Shinn et al.), which is consistent with recent research (Christ, 2006).

The purpose of this manuscript was to evaluate CBM-R from a behavioral perspective and thus to highlight the fact that there is a considerable amount of work that needs to be conducted before decisions using CBM-R data can be made with great confidence. The intention of this paper was not, however, to suggest that CBM-R should not be used within an RtI framework. CBM-R data are invaluable for schools to evaluate the quality of instruction for their student body and to identify students in need of supplemental instruction. CBM-R data are also useful as one component of multiple sources of assessment data that can be used for evaluating a student's response to instruction. Schools should, however, consider alternative means of assessing intervention effects that directly evaluate the impact of instruction; this might allow decisions regarding intervention effects to be made within relatively short periods of time (e.g., 5 weeks). It is essential to remember that CBM-R is a global measure of reading achievement. Another way to explain this is that CBM-R evaluates generalization effects. In the same way that the effects of an intervention targeting classroom behavior should not be evaluated by only evaluating generalization effects, we should not only evaluate the effects of a reading intervention by evaluating generalization effects. Generalization does not occur naturally and it is likely that generalization effects will require longer to appear than schools allot for evaluating intervention effects. Measures must be used that evaluate whether a student is mastering the individual skills being taught. Only with mastery of component skills is generalization likely to occur, resulting in improvements in global/composite skills (Ardoin & Daly, 2007; Binder, 1996).

References

Allinder, R. M., & Eicher, D. D. (1994). Bouncing back: Regression and recoupment among students with mild disabilities following summer break. *Special Services in the Schools, 8*(2), 129-142.

Ardoin, S. P., & Christ, T. J. (in press). Evaluating curriculum-based measurement slope estimates using data from tri-annual universal screenings. *School Psychology Review*.

Ardoin, S. P., & Daly, E. J., III. (2007). Introduction to the Special Series: Close Encounters of the Instructional Kind--How the Instructional Hierarchy is Shaping Instructional Research 30 Years Later. *Journal of Behavioral Education, 16*(1), 1-6.

Ardoin, S. P., & Martens, B. K. (2004). Training children to make accurate evaluations: Effect on behavior and the quality of self-ratings. *Journal of Behavioral Education, 13*, 1-23.

Ardoin, S. P., McCall, M., & Klubnik, C. (2007). Promoting Generalization of Oral Reading Fluency: Providing Drill versus Practice Opportunities. *Journal of Behavioral Education, 16*(1), 55-70.

Ardoin, S. P., Suldo, S. M., Witt, J. C., Aldrich, S., & McDonald, E. (2005). Accuracy of readability estimates' predictions of CBM performance. *School Psychology Quarterly, 20*(1), 1-22.

Ardoin, S. P., Witt, J. C., Suldo, S. M., Connel, J. E., Koenig, J. L., Resetar, J. L., et al. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probe when conducting universal screening. *School Psychology Review, 33*, 218-233.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*, 163-197.

Christ, T., J. (2003). The effects of passage-difficulty on CBM progress monitoring outcomes: Stability and accuracy. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 62*(7), 3414.

Christ, T., J. (2006). Short term estimates of growth using curriculum-based measurement of oral reading fluency: Estimates of standard error of the slope to construct confidence intervals. *School Psychology Review, 35*(1), 128-133.

Christ, T. J., & Ardoin, S. P. (2007). Curriculum-based measurement: Passages difficulty. *Manuscript submitted*.

Christ, T. J., & Coolong-Chaffin, M. (2007). Interpretations of curriculum-based measurement outcomes: Standard error and confidence intervals. *School Psychology Forum: Research in Practice, 1*(2), 75-86.

Christ, T. J., & Silberglitt, B. (2007). Curriculum-based measurement of oral reading fluency: The standard error of measurement. *School Psychology Review*, (36), 130-146.

Colon, E. P., & Kranzler, J. H. (2006). Effect of Instructions on Curriculum-Based Measurement of Reading. *Journal of Psychoeducational Assessment, 24*(4), 318-328.

Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. *Behavior Therapy, 8*, 411-426.

Cone, J. D. (1981). Psychometric considerations. In M.Herson & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (2nd ed., pp. 38-68). Elmsford,NY: Pergamon Press.

Cone, J. D. (1992). Accuracy and curriculum-based measurement. *School Psychology Quarterly, 7*(1), 22-26.

Daly, E. J., III, Persampieri, M., McCurdy, M., & Gortmaker, V. (2005). Generating reading interventions through experimental analysis of academic skills: Demonstration and empirical evaluation. *School Psychology Review, 34*(3), 395-414.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15*(3), 358-374.

Deno, S. L. (2002). Problem-solving as "best practices." In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (Vol. 1, pp. 37-56). Washington, DC: National Association of School Psychologists.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum - based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*, 507-524.

Deno, S. L., Marston, D., & Tindal, G. (1985). Direct and frequent curriculum-based measurement: An alternative for educational decision making. *Special Services in the Schools, 2*(2), 5-27.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-45.

Deno, S. L., Shinn, J., Marston, D., & Tindal, G. (1983). Oral reading fluency: A simple datum for scaling reading disability. *Topics in Reading and Reading Disabilities, 2*, 53-59.

Deno, S. L., & Shinn, M. R. (1989). *Curriculum-based measurement and special education services: A fundamental and direct relationship*. New York, NY, US: Guilford Press.

Derr-Minneci, T. F., & Shapiro, E. S. (1992). Validating Curriculum-based measurement in reading from a behavioral perspective. *School Psychology Quarterly, 7*, 2-16.

Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. *Psychology in the Schools, 37*, 463-473.

Eckert, T. L., Dunn, E. K., & Ardoin, S. P. (2006). The effects of alternate forms of performance feedback on elementary-aged students' oral reading fluency. *Journal of Behavioral Education, 15*(3), 148-161.

Fuchs, D., & Fuchs, L. S. (1992). Limitations of a feel-good approach to consultation. *Journal of Educational and Psychological Consultation, 3*(2), 93-97.

Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice, 18*, 172-186.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on student achievement, pedagogy, and student awareness of learning. *American Educational Research Journal, 21*, 449-460.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*(3), 199-208.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13*(4), 204-219.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989a). Effects of alternative goal structures within curriculum-based measurement. *Exceptional Children, 55*(5), 429-438.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989b). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *Journal of Educational Research, 83*(2), 103-110.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teaching planning and student achievement in mathematics operations. *American Educational Research Journal, 28*(3), 617-641.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*, 27-48.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20-28.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*(1), 33-45.

Fuchs, L. S., Tindal, G., & Deno, S. L. (1984). Methodological issues in curriculum based reading assessment. *Diagnostique, 9*, 191-207.

Good, R. H., III, & Shinn, M. R. (1990). Forecasting accuracy of slope estimates for reading curriculum-based measurement: Empirical evidence. *Behavioral Assessment, 12*, 179-193.

Good, R. H., III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision making utility of continuum of fluency-based indicators of foundational reading skills for third grade high-stakes outcomes. *Scientific Studies of Reading, 5*(2), 257-288.

Graney, S. B., & Shinn, M. R. (2005). Effects of Reading Curriculum-Based Measurement (R-CBM) Teacher Feedback in General Education Classrooms. *School Psychology Review, 34*(2), 184-201.

Hayes, S. C., Nelson, R. O., & Jarret, R. B. (1986). Evaluating the quality of behavioral assessment. In R. C. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 463-503). New York: Guilford.

Hintze, J. M., Owen, S. V., Shapiro, E. S., & Daly, E. J., III. (2000). Generalizability of oral reading fluency measures: Application of G theory to curriculum-based measurement. *School Psychology Quarterly, 15*(1), 52-68.

Hintze, J. M., & Shapiro, E. S. (1997). Curriculum-based measurement and literature-based readings: Is curriculum-based measurement meeting the needs of changing reading curricula? *Journal of School Psychology, 35*, 351-375.

Hintze, J. M., Shapiro, E. S., & Daly III, E. J. (1998). An investigation of the effects of passage difficulty level on outcomes of oral reading fluency progress monitoring. *School Psychology Review, 27*(3), 433-445.

Hintze, J. M., Shapiro, E. S., & Lutz, J. G. (1994). The effects of curriculum on the sensitivity of curriculum-based measurement in reading. *The Journal of Special Education, 28*, 188-202.

Marston, D. (1987). The effectiveness of special education: A time series analysis of reading performance in regular and special education settings. *The Journal of Special Education, 21*(4), 13-26.

Marston, D., Fuchs, L. S., & Deno, S. I. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique, 11*, 71-90.

McCurdy, M., Daly, E., Gortmaker, V., Bonfiglio, C., & Persampieri, M. (2007). Use of Brief Instructional Trials to Identify Small Group Reading Strategies: A Two Experiment Study. *Journal of Behavioral Education, 16*(1), 7-26.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35*(11), 1012-1027.

Mischel, W. (1968). *Personality and Assessment.* Hoboken, NJ: John Wiley & Sons Inc.

Nelson, R. O., & Hayes, S. C. (1979). Some current dimensions of behavioral assessment. *Behavioral Assessment, 1*, 1-16.

Poncy, B. C., Skinner, C. H., & Axtell, P. K. (2005). An investigation of the reliability and standard error of measurement of words read correctly per minute using curriculum-based measurement. *Journal of Psychoeducational Assessment, 23*(4), 326-338.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guilford Press.

Shinn, M. R., Good, R. H., III, & Stein, S. (1989). Summarizing trends in student achievement: A comparison of methods. *School Psychology Review, 18*, 356-370.

Shinn, M. R., Thomas, A., & Grimes, J. (2002). *Best Practices in Using Curriculum-Based Measurement in a Problem-Solving Model.* Washington, DC, US: National Association of School Psychologists.

Shinn, M. R., Ysseldyke, J. E., Deno, S. L., & Tindal, G. A. (1986). A comparison of differences between students labeled learning disabled and low achieving on measures of classroom performance. *Journal of Learning Disabilities, 19*(9), 545-552.

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93*(4), 735-749.

Stoner, G. (1992). Validating curriculum-based measurement: Essential concerns from a behavioral perspective. *School Psychology Quarterly, 7*(1), 17-21.

Wesson, C. L. (1991). Curriculum-based measurement and two models of follow-up consultation. *Exceptional Children, 57*(3), 246-256.

Author Contact Information

Scott P. Ardoin, Ph.D.
Assistant Professor, School Psychology
University of South Carolina
Department of Psychology
Columbia, SC 29208
spardoin@sc.edu

803 777-7616

Claire M. Roof
USC School Psychology Graduate Student
University of South Carolina
Department of Psychology
Columbia, SC 29208
RoofCM@gwm.sc.edu
803 730-1210

Cynthia Klubnik
M.S. Psychology from Univ. of South Carolina
University of South Carolina
Department of Psychology
Columbia, SC 29208
klubnik@mailbox.sc.edu
860 798-9560

Jessica C. Williams
USC School Psychology Graduate Student
University of South Carolina
Department of Psychology
Columbia, SC 29208
jcarfolite@yahoo.com
803 422-9837

Modelling Social Learning in Monkeys

Jeremy R. Kendal

The application of modelling to social learning in monkey populations has been a neglected topic. Recently, however, a number of statistical, simulation and analytical approaches have been developed to help examine social learning processes, putative traditions, the use of social learning strategies and the diffusion dynamics of socially transmitted information. Here, I review some of the recent advances and show how they influence and combine with empirical studies of social learning.
Keywords: cultural transmission, modelling, social learning, traditions.

Quantifying Social Learning Processes

In the laboratory, paired demonstrator-observer designs are used primarily to investigate whether animals are *capable* of specific forms of social learning (Galef 1988). Laboratory experimental have also been used to investigate the diffusion dynamics of learned behaviour through populations in controlled conditions (Lefebvre & Palameta 1988; Whiten et al. 2005). The results of such experiments do not tell us much about social learning in the wild and often lack ecological validity in terms of the behaviour being learned. They have the potential, however, to provide quantitative data that can be used in the parameters for mathematical models of social learning.

For instance, Kendal et al. (2007) presented captive groups of callitrichid monkeys with novel extractive foraging tasks. By measuring the proximity of each monkey to the task and noting any food extractions, they quantified the effects of two social learning processes, 'stimulus enhancement' and 'observational learning', and two asocial processes, 'intrinsic movement to the task' and 'asocial learning of the task' on the adoption of a novel foraging behavior. The values of these processes from the observed data were fed into a set of models for the spread of a novel behaviour. Model selection was used to discern that the model best-fit to the monkey diffusion data only required asocial processes. Nonetheless, quantification of the processes provided statistical evidence for a small positive effect of stimulus enhancement, where a demonstrator manipulating the task attracts an observer to move to the task, but not for observational learning at the task.

Derived parameter values can also be used in competing models to predict the shapes of diffusion curves. Theoretical models predict that the diffusion of cultural traits will typically exhibit a sigmoidal ('S' shaped) pattern over time (Boyd & Richerson 1985; Cavalli-Sforza & Feldman 1981), while asocial learning has been expected to result in a linear, non-accelerator, or at least non-sigmoidal increase in frequency (Lefebvre 1995). In the absence of parameter values to feed into diffusion models, the observed shape of a diffusion curve is unlikely to be reliable, as more recent models predict that under certain conditions asocial learning can generate acceleratory curves while social learning can generate decelerator curves (Laland & Kendal 2003; Reader 2004). Many competing sets of assumptions can generate very similar diffusion curves using different parameter values. Thus, if parameter values can be estimated, it may be possible to select between competing hypotheses.

Population-Level Homogeneity of Behaviour

While social learning experiments in the laboratory may provide estimates of the limitations of social learning behaviour, an understanding of social learning will be incomplete without analysis of natural populations. Field translocation experiments have been performed to identify traditions in the wild. For instance, Helfman and Shultz (1984) translocated French grunts *Haemulon flavolineatum* between

populations and found that fish placed into established populations adopted the same schooling sites and migration routes as the residents. Control fish however, that were introduced into regions from which all residents had been removed, did not adopt the behaviour of former residents.

In monkeys, however, this type of approach would be impractical and often unethical. The identification of traditions in natural monkey populations will often rely on observational data without recourse to manipulating groups of animals. The ethnographic method is one such approach, identifying behavioural variation between populations that cannot be accounted for by local ecology or genetic variation. Whiten et al. (1999) identified 39 different behaviour patterns across populations of chimpanzees (*Pan troglodytes*) in Africa, including tool usage, grooming and courtship behaviours, that were common in some communities but absent in others. Critics of the method have pointed out that ecological explanations may have been discounted erroneously and that some populations have been separated for sufficient time for genetic variation to contribute to the observed behavioural variation (Laland and Janik 2006). Unfortunately, the method excludes cases where selection has resulted in covariance between cultural and (i) ecological factors, because culture is a source of adaptive behaviour, and and/or (ii) genetic factors enabling animals to learn about and exploit environmental resources because animal learning is influenced by evolved predispositions (Laland & Janik 2006; Laland et al. In press).

An alternative, the 'option-bias' method, is being developed to identify traditions in the wild, and like the ethnographic approach, assumes that social learning results in within-population homogeneity of behaviour (R. Kendal et al. In prep.). The option-bias method, however, does not require comparison between populations, as instead, the observed level of homogeneity is compared against a distribution of expected homogeneity-levels generated by stochastic simulation that assumes asocial but not social learning. Thus, the simulated distribution represents the probability that populations of monkeys behaving independently would exhibit the observed level of homogeneity in their option choices. Social learning can be inferred if the observed level of homogeneity lies above the critical level (alpha = 0.05) in the tail of the simulated distribution. Kendal et al. applied the method to data collected from groups of callitrichid monkeys provided with novel extractive-foraging tasks, with each task exhibiting two equivalent means (or options) of obtaining a food reward (using different coloured doors to access food). The analysis distinguished social learning in cases where the monkeys were presented with a particularly difficult task; here the acquisition of asocial information is likely to have been costly in terms of time and effort. The method does not assess any genetic influence on behaviour and requires an a priori estimate of the maximum probability of acquiring the behavioural variant through asocial learning (something that could be estimated in the field or taken from laboratory experiments). Unlike the ethnographic method, however, a putative cultural variant in each population can be assessed independently from other populations and can be used to examine variants that covary with the local ecology.

Social Learning and Patterns of Association

Another approach to identify social learning is to look for cases of 'directed social learning', that is, non-random paths of information transmission through a population (Coussi Korbel & Fragaszy, 1995). Often, social structure may influence the spread of information, influencing whether innovations spread at all, and if so, the rate of spread relative to rate of change in the environment. In a changing environment, information becomes more likely to become obsolete with time, so the efficacy of socially transmitted information depends on an animal's location in the chain of acquisition, which may consequently affect the propensity for the animal to use information that is learned socially.

Most approaches start by compiling a matrix of association metrics between all pairs of individuals as a measure of the opportunity for information transmission; if social learning is directed it may often be as a function of the level of association and thus the opportunity to observe the demonstration of a novel behaviour. If social learning is vertically transmitted (parent to offspring),

information may be directed according to genealogy. Laland and Kendal (2003) show that a correlation between the coefficient of association and the order of information acquisition can provide evidence to distinguish directed social learning from asocial learning, while cases of acquisition by offspring before parents may often render a genetic account implausible (unless there are developmental constraints on the adoption of the new behaviour). Furthermore, the concordance of the behaviour between parents and offspring can help dissociate cases of vertical transmission (from parents to offspring) from oblique (from non-parents, between parent and offspring generations) and horizontal transmission (within-generation). More sophisticated approaches are currently under development by Laland et al. (In prep.), using a path analysis approach, to predict the relative influence of genetic, environmental and cultural effects on the behavioural phenotype.

Whitehead (In press) suggests using partial regression to test how much inter-individual variation in behaviour is caused by 'social similarity', over and above any effects of ecological and/or genetic similarity. Here, social similarity is a measure of the opportunity for social transmission of information, typically measured by some form of association index. The inter-individual variation in behaviour, scored as behavioural similarity between each pair of individuals, is non-independent, so Whitehead addresses the significance of the partial regression coefficient by bootstrapping a distribution of partial regression coefficients, where the similarity scores of the independent variables are randomized across individuals. Social learning, or more precisely, the effect of the social similarity, is invoked if the original partial regression coefficient lies beyond the critical point in the upper tail of the bootstrapped distribution (e.g. $P < 0.05$).

Boogert et al (In press) present a randomization approach to examine whether the order to solve extractive foraging tasks in captive groups of starlings is a function of the level of association between individuals. They compared the association scores between individuals, summed along a linear chain in the order of task acquisition, and compared this 'path strength metric' against a distribution of random path strength metrics, generated from randomly ordered chains of acquisition. Directed social learning would be in evidence if the observed path strength metric lay in the upper tail of the distribution; that is, if the sum of associations in the chain of acquisition were significantly larger than expected in a population where each observer's choice of demonstrator was independent of their strength of association. As naïve individuals could potentially learn from any informed individual and not just the adjacent informed individual in the chain of acquisition, they also repeated the analysis assuming a nonlinear chain of acquisition, where the association strength for each individual is the average across all individuals that had already learned the task. Although they did not find evidence for directed social learning, this method should be appropriate for future investigations of directed social learning.

Voelkl and Noë (submitted manuscript) use the association metrics taken from a population of rhesus monkeys (*Macaca mulatta*) to determine the strength of connections, or edges, between individuals, or nodes, in a social network. They simulated the spread of information across this network and compared the rate of spread and the average path length with that of stylised artificial network structures. Although they did not analyse the structure of the rhesus network directly, they explored how diffusion dynamics can be a function of network structure. Further investigation using this approach will surely yield insight on constraints affecting the spread of information as it can exploit the substantial body of theory on small-world networks (Watts 1999). The approach should also help us understand the relationship between directed social learning and social hierarchy.

Couzin and Krause (2003) used agent based models to show that the non-random assortment of individuals within a population can occur without any capacity for individual recognition, as individuals naturally 'self-sort' to form close associations with those of similar behavioural or motivational states. This process is likely to provide a structured pathway for the non-random spread of information across a group. It has been shown in fish that the identity of leaders and followers can be predicted by state

dependent factors such as hunger (Reader and Laland 2000). Rands et al. (2003) use a dynamic programming game model to show that if it pays not to forage alone, the foraging behaviour of two individuals can become highly synchronised, from which a difference in energetic state between the two foragers can emerge. The identity of a leader and follower in the pair of individuals is determined by each adopting the ESS (evolutionarily stable strategy) rule of thumb that is 'to forage if either their own reserves have fallen below a certain threshold value, or if their partner chooses to forage'.

Social Learning Biases

Social learning may also be directed as a function of biased attention on the part of the observer, including direct bias, indirect bias and conformity. Much of the theory concerns the evolutionary consequences of social learning biases in humans (Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985). Direct bias occurs when an observer chooses to copy a behaviour that is demonstrated to be particularly successful. If there is noisy feedback about the success of other individuals, however, game-theoretical modelling has found that the most effective strategy is to copy in proportion to the observed payoff that results from the demonstrator's performance (Schlag 1998). In contrast, occurs when an individual copies a behavioural trait exhibited by a demonstrator who is deemed successful (e.g. foraging success) regardless of whether the demonstrator's 'success' was due to that behaviour. Indirect bias may, however, result in erroneous copying of ineffective behaviour. The most intuitive examples are in humans, where advertising companies invoke indirect bias to enhance sales of for example brands of razor, sunglasses or perfume by showing, often falsely, that they are used by particularly successful sportsmen, singers or actors; the adopted behaviour itself, can become a marker for prestige (Henrich and Gil-White 2001).

Studies of direct and indirect bias in animals are rare and evidence for these psychological dispositions is sparse. Wilkinson (1992), however, found evidence of copying successful individuals in nursing colonies of evening bats (*Nycticeius humeralis*); adult females that had been unsuccessful in foraging tended to follow successful individuals, and females that followed successful individuals from the nursing colony were subsequently more successful than females that departed alone (Galef and Laland 2005). Also, in domestic fowl (*Gallus gallus*), high-status individuals are more effective demonstrators than are low-status individuals as flock members are more likely to pay attention to the former than the latter (Nicol and Pope 1999). Henrich and Gil-White (1999) however, would argue that this is not prestige bias, as dominant individuals may attract high levels of attention for reasons other than prestige (i.e. held in high esteem), such as the threat of aggression. Furthermore, they argue that prestige bias might be rare in non-humans, in part, because it requires the capacity for direct transmission of information, for instance, through imitation. Direct transmission ensures that the benefit derived by the copier is a function of the utility of the demonstrator's information, and may be a prerequisite to the evolution of prestige bias. In monkey populations, indirect forms of social learning may be prevalent, such as stimulus enhancement, where a demonstrator attracts the attention of an observer to a task. Once at proximity to the task, the observer may then acquire the solution to the task through asocial learning.

Conformist bias can maintain within-group homogeneity and between group differences and thus may lead to the persistence of between-population variation in behaviour (Boyd & Richerson 1985; Henrich and Boyd 1998). Henrich and Boyd (1998) used a numerical approach to show that the propensities for social learning and for conformity bias may coevolve. Recent theoretical studies however, have not found this to be the case (Eriksson et al. 2007; Wakano and Aoki 2007). There is scant evidence for conformity in animals besides experimental evidence in fish and chimpanzees (Day et al. 2001; Whiten et al. 2005). Nonetheless, one might anticipate conformity in fission-fusion monkey populations, where in a variable environment, it might pay for immigrants to copy locally common patterns of behaviours.

Laland et al. (In prep.) are currently using game theory and individual-based models to examine the evolution of rules-of-thumb (or 'social learning strategies') that animals may use to assess when and from whom to learn socially (Laland 2004; Kendal et al. 2005). The game theory develops the classic producer-scrounger paradigm to consider the producing and scrounging of information rather than food. The individual-based models allow the examination of spatially explicit effects such as local clustering of types of learning strategists. The models are complemented with experimental approaches applied to fish and monkeys. The results should give a comparative insight into the evolutionary processes affecting the use of social learning strategies.

Local Interactions and Collective Behaviour

Individual-based models have also been used to examine the effects of individual decision rules on group spatial dynamics. Couzin (2007) describes how the formation of spatial patterns may facilitate the spread of information through a group; in particular, an individual's 'effective range' of perception can be larger than their actual sensory range as a result of combined local interactions within the group. For instance, a fish might detect a predator sooner due to the behaviour of shoal members rather than seeing the predator themselves. The scaling from actual to effective sensory range is non-linear: too small a range limits group cohesion and limits the spread of information through the group, while too large a range risks misdirecting individuals as they obtain information from distant models that may not be relevant. These results follow the same principles that underpin the evolution of social learning in general, where an intermediate scale of environmental variation is likely to favour social learning. For example, birds at the head of a migrating flock may experience a new thermal that is absent at the back of the group. Here, too long a sensory range may result in the tail-end birds adopting an inappropriate change in course by copying the lead birds.

To date, this theory has been developed with large animal groups in mind, such as insects, birds, bats and humans. In principle, the theory could also be applied to monkeys. For example, a predator alarm call may result in monkeys escaping by copying the route of their neighbours. If there is variation in the optimal escape route depending on variation in the location of the monkeys, too small or too large a sensory range may result in potentially fatal errors in the escape route taken. The optimal level of group cohesion is likely to be context dependent and allows individuals to acquire information that otherwise would elude them.

Conclusions

It is striking how little integration there has been between theoreticians and primatologists. Given the depth of knowledge about life histories, social systems and the local ecology of monkey populations, it would be entirely reasonable to develop models to help identify traditions, examine psychological dispositions underlying social learning, and make predictions concerning the use of social learning and the dynamics of socially transmitted information. In the phylogenetic tree, the origin of human culture lies far deeper than the common ancestor of apes and developing mathematical models of social learning specifically for monkey species will help to write an important chapter in this story.

References

Boogert, N. J. Reader, S. M. Hoppitt, W. & Laland, K.N. (In Press). The origin and spread of innovations in starlings. *Animal Behaviour*.

Boyd, R. & Richerson, P. (1985). *Culture and the Evolutionary Process.* (Chicago University Press.

Cavalli-Sforza, L. L. & Feldman, M.W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.

Coussi-Korbel S., Fragaszy D. M. (1995). On the relation between social dynamics and social learning. *Animal Behav*iour, 50, 1441–1453.

Couzin, I.D. & Krause J. (2003). Self-organization and collective behavior in vertebrates. *Advances in the Study of Behav*ior, 32, 1-75.

Couzin, I.D. (2007). Collective minds. *Nature,* 455, 715.

Day, R. L., MacDonald, T., Brown, C., Laland, K. N. & Reader, S. M. (2001). Interactions between shoal size and conformity in guppy social foraging. *Animal Behav*iour, 62, 917-925.

Galef, B. G., Jr. (1988). Imitation in animals: history, definition, and interpretation of data from the psychological laboratory. In T.R. Zentall & B.G. Galef, Jr. (Eds.) *Social Learning: Psychological and Biological Perspectives* (pp. 3-28). Erlbaum, New Jersey.

Helfman, G. S. & Schultz, E. T. (1984). Social transmission of behavioural traditions in a coral-reef fish. *Animal Behaviour,* 32, 379-384.

Henrich, J. & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution Human Behavior,* 19, 215-241.

Henrich, J. & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22, 165–196.

Kendal J. R. Kendal R. L. & Laland K. N. (2007). Quantifying and modelling social learning processes in Monkey Populations. *J. Psychology and Psychological Therapy*, 7(2), 123-138. Also see Erratum: vol. 8(1).

Kendal R. L., Coolen I, van Bergen Y, & Laland K. N. (2005). Tradeoffs in the adaptive use of social and asocial learning. *Advances in the Study of Behavior,* 35, 333-379.

Laland K. N. (2004). Social learning strategies. *Learning and Behavior,* 32, 4–14.

Laland K. N. Kendal J. R. (2003). What the models say about animal social learning. In: Traditions in nonhuman primates: models and evidence. Eds. Fragaszy D & Perry S. Chicago University Press.

Laland, K. N., & Janik, V. (2006) The animal cultures debate. *Trends in Ecology & Evolution*, 21, 542-547.

Laland K. N., Kendal J.R. & Kendal R. L. (In Press). Animal culture: problems and solutions. In: *The Question of Animal Culture.* Eds. KN Laland BG Galef Jr., Harvard University Press.

Lefebvre, L. & Palameta, B. (1988). Mechanisms, ecology and population diffusion of socially learned, food-finding behaviour in feral pigeons. In T.R. Zentall & B.G. Galef, Jr. (Eds.) *Social Learning: Psychological and Biological Perspectives* (pp. 141-163). Erlbaum, New Jersey.

Lefebvre, L. (1995). Culturally-transmitted feeding behaviour in primates: evidence for accelerating

learning rates. *Primates*, 36, 227-239.

Rands S.A., Cowlishaw G., Pettifor R.A., Rowcliffe J.M. & Johnstone R.A. (2003). The spontaneous emergence of leaders and followers in a foraging pair. *Nature*, 423, 432–434.

Reader S. M. & Laland K. N. (2000). Diffusion of foraging innovations in the guppy. *Animal Behav*iour, 60, 175-180.

Schlag, K. H. (1998). Why imitate, and if so how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78 (1), 130–156.

Voelkl, B. & Noë, R. (Submitted). The influence of social structure on the propagation of social information in artificial primate groups: a graph-based simulation approach. *Journal of Theoretical Biology*.

Watts, D. J. (1999). *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W. & Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399, 682-685.

Whitehead H. (In Press). How might we study culture: a perspective from the ocean. In: *The Question of Animal Culture.* Eds. K.N. Laland B.G. galef Jr., Harvard University Press.

Whiten, A., Horner, V. & de Waal, F.B.M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, 437, 737-740.

Author Contact Information:

Jeremy R. Kendal, Ph.D.
Dept. of Anthropology, University of Durham
43 Old Elvet,
Durham DH1
3HN, U.K.
Tel.: +44 (0) 191 334 6185
jeremykendal@gmail.com

FAPRS MANUAL

Manual for the Functional Analytic Psychotherapy Rating Scale

Glenn M. Callaghan & William C. Follette

ABSTRACT

The Functional Analytic Psychotherapy Rating Scale (FAPRS) is behavioral coding system designed to capture those essential client and therapist behaviors that occur during Functional Analytic Psychotherapy (FAP). The FAPRS manual presents the purpose and rules for documenting essential aspects of FAP. The FAPRS codes are exclusive and exhaustive for FAP essential behaviors but also include codes for generally effective therapy behaviors by both client and therapist. Client behaviors identified include those that are FAP-specific such as Clinically Relevant Behaviors (in-session improvements and problems), specification of controlling variables, and discussion of outside problems and improvements that have been identified as targeted behaviors. Therapist behaviors that have been identified as theoretically essential for conducing FAP are included such as discussions about the therapeutic relationship, responding effectively and ineffectively to in-session client behaviors, and evoking client behavior in-session. For each behavioral code a definition is provided along with examples and counter examples of how the code might be applied to client or therapist behaviors. A decision hierarchy is provided for those cases when a client or therapist behavioral event (called a turn) may receive more than one possible code. The FAPRS can be used as a tool in research (e.g., to provide evidence for the proposed mechanism of change for FAP) or as a method for assisting the training of psychotherapists. The FAPRS has demonstrated acceptable psychometric properties (demonstrated by Callaghan, Follette, Ruckstuhl, & Linnerooth, this issue).

Keywords: Functional Assessment, Behavioral Therapy, Functional Analytic Psychotherapy, Behavioral Rating Scale, Behavioral Coding.

TABLE OF CONTENTS

I. INTRODUCTION

This Ratings Manual includes the Functional Analytic Psychotherapy Rating Scale (FAPRS). It is designed for rating audio tapes or videotapes of psychotherapy conducted according to the principles of functional analytic psychotherapy (FAP) to code the existence of particular therapist and client behaviors and to document the changes in client behavior as a function of the therapist's contingent responding.

It is essential that the rater be familiar with the material in the Ratings Manual and the principles of FAP as outlined by Kohlenberg & Tsai (1991; see also Callaghan, Naugle, & Follette, 1996; Follette, Naugle, & Callaghan, 1996) before making ratings on the scale. Although it is the intention of the author that this manual be used in conjunction with training FAP, it is not to be used as a replacement for the original text by Kohlenberg & Tsai. The manual presumes a familiarity with the Kohlenberg & Tsai text and a basic understanding of a functional analysis and the principles of behaviorism.

The Manual begins with *General Comments* and instructions to raters which are important in rating using this system. The remainder of the Manual is organized according to codes for effective, ineffective, and neutral impact therapist behaviors, as well as codes for client behaviors. Each item contains (when applicable):

(1) The exact wording and format of the item as it appears in the scale.
(2) A restatement of or elaboration on the item's purpose.
(3) Definitions of terms used in the item.
(4) General guidelines for rating a turn using that code.
(5) Important distinctions to be made between codes through the use of marginal examples of the code.
(6) Counter example(s) of the rule.

II. GENERAL COMMENTS

1. *Coding Therapist Behaviors*: This scale is designed to rate the behavior of both client and therapist. It is important to distinguish therapist behavior (as much as possible) from client response. That is, in rating therapist behavior, the rater should consider the function of client and therapist behavior (i.e., the actual effect that behavior has on the other member of the dyad). Raters should not code what it appears the therapist attempted to do but should instead code whether those attempts met with success or failure.

2. *Prerequisite Knowledge for Rating This Scale*: Raters are required to be familiar with the principles of Functional Analytic Psychotherapy (FAP; Kohlenberg & Tsai, 1991) and should have a basic knowledge of behavior analysis. It is also recommended that coders have additional experience with FAP outside of knowledge based on readings through seeing clients while implementing this therapy and being supervised by a knowledgeable supervisor. This Ratings Manual will not provide all of the specific background needed in order to code client and therapist behavior. In the event this manual is used to supplement training, the therapist must continue to obtain supervision as is appropriate according to the ethical principles of psychologists and consistent with the training guidelines set forth by the training facility or institution.

   When using the scale, the rater must be careful and conscientious in listening to and rating the taped therapy sessions. Because rating is a complex task, it requires the rater to be thoughtful and to exercise good judgment. Tapes are not to be listened to or viewed by individuals not bound legally and ethically by the rules of confidentiality and should not be viewed or listened to by individuals not involved in the coding of tapes.

3. *Coding Each Turn when Using Transcripts using the Decision Hierarchy*: While it is likely that each functional unit corresponds to a floor change or turn (i.e., an utterance, sentence, paragraph, or paragraphs by either member of the client-therapist dyad), where distinct codeable units exist during a turn, the rater is required to code the turn using the decision hierarchy included in this manual.

   The use of transcripts is not a requirement of this manual, but likely helps reliability between raters. In the event that coders rate audio or video tapes without transcripts, it is highly recommended that all raters adhere to a rule about what units are to be coded as events (i.e., floor changes, the last statement made by the speaker).

4. *Two Word Utterances.* One and two word utterances by either the client or therapist when they are affirmations (uh-huh, um-hmh, no, yeah, yes, etc.) have been embedded in the turn of the other speaker and are not coded. When a one or two word utterance is determined to be a codeable response it will remain in the body of the text. In the event a transcript is either not used or does not have the utterance embedded, the rater should ignore such one and two word utterances.

5. *Avoid Haloed Ratings*: The scale was designed to describe therapist and client behavior in session. To use the scale correctly, it is essential that raters code what is heard or seen on the transcript, NOT what raters think OUGHT to occur (i.e., code each unit independently).

   The rater must be sure to apply the same standards for coding a unit regardless of:
   (a) what other behaviors the therapist or client emitted during the session;
   (b) what ratings were given to other behaviors that occurred just prior;
   (c) how skilled the rater believes the therapist to be;
   (d) how much the rater likes or dislikes the therapist or the therapist's style;
   (e) how skilled or ineffective the rater believes the client *should* be;
   (f) how much the rater likes or dislikes the client or the client's style;
   (g) whether the rater believes he or she would have done something similarly or differently than either the client or therapist did.

   a) *Rater halo resulting from a consideration of other therapist behaviors during the session*: In deciding what rating to assign a turn, the rater erroneously bases his/her ratings on behaviors similar to the target behavior or on behaviors that are likely to covary with the target behavior. For example, if the therapist engaged in a high number of effective therapist behaviors during the session, the rater must continue to evaluate each therapist behavior as it occurs. The rater should

not allow early behaviors to unduly influence the code assigned to later behaviors based on the positive or negative valence of those previous behaviors.

b) *Rater halo resulting from ratings given to other items*:
In deciding what code to assign a behavior, the rater erroneously bases her/his rating on ratings given to other events. This is likely to occur when the rater believes the code assigned to another item affects the rating given to the item currently being rated. *Each turn must be rated independently*. For example, in this system, when a client emits a problem behavior in-session (a CRB1), the rater should not assume that the therapist will respond to that client behavior (e.g., TRB1) in a particular way.

c) *Rater halo resulting from rater's judgment of the therapist's level of skill*:
The rater may feel that the therapist is highly skilled in a particular strategy and assumes that the turn being rated should be reflected in what the therapist has done while implementing a particular strategy. Based on these assumptions, which may be erroneous, the item could be coded inaccurately. Similarly, if the rater judges the therapist to be less skilled, the item might also be coded inaccurately. *Each turn must be coded according to the turn itself, not according to the skill level of the therapist*. This is especially true when rating behaviors of therapists who are at either end of the skill continuum (e.g., those who are particularly advanced in training and those who are more novice).

d) *Rater halo resulting from how much the rater likes the therapist*:
The rater might assign codes inaccurately simply because he or she has a positive affective reaction to the therapist while the same behavior by a different therapist would not receive the same code. Turns must be coded according to the behavior, as specified in this Ratings Manual, regardless of whether or not the rater likes the therapist.

e) *Rater halo resulting from how skilled or ineffective the rater believes the client should be*:
The rater might assign codes inaccurately if he or she erroneously believes the client should be more or less skilled or effective in an interpersonal interaction than the client actually is. This is especially likely to happen when a therapist codes his or her own tapes without additional reliability data from another rater or when the rater has information about the order of sessions being coded.

f) *Rater halo resulting from how much the rater likes the client*
The rater assigns codes inaccurately simply because he or she has a positive or negative affective reaction to the client while the same behavior by a different client would not receive the same code. As with criterion (d) above, Events must be coded according to the behavior, as specified in this Ratings Manual, regardless of whether the rater likes the client.

(g) *Rater halo resulting from rater believing he or she would have done something similarly or differently*:
The rater might assign a code to a turn because the rater believes he or she would have done a similar thing in therapy or something differently. This coding system is designed to assess the impact a therapist has on a client and vice versa. It is not designed to measure whether coders agree or disagree with a therapist's behavior. *Judgments of similarity* to what a rater would have done are *not* part of this coding process and will corrupt the data. The rater must avoid these mistakes assiduously.

This needs to be conceptually distinct from noticing how a client or therapist behavior is believed to impact the other member of the dyad. If, for example, the rater determines the client behavior

functions as a problem behavior in session and should have impacted the therapist this way, but the therapist did not respond to that behavior, the rater should code this behavior accordingly. If, on the other hand, the rater notices the therapist responds to a client behavior topographically differently from the way the coder would, the rater must code the therapist's behavior as it impacted the client, not as a dissimilar response from one the rater would have chosen to emit.

6. *Use of Guidelines*: The descriptions and definitions of items in this Ratings Manual are intended to be guidelines for use in coding. In all cases, there is information about what code should be given. The guidelines also specify marginal or "borderline" cases and how to determine if a particular code should be assigned. This information is very important.

7. *Use of Examples*: For many of the items in this Manual, we have given examples of therapeutic exchanges as guidelines for rating therapist and client behavior. Nevertheless, the examples are only guidelines. This is because the examples are only brief interchanges that might occur in the larger context of an on-going session and are absent of a case conceptualization. When providing code to behaviors in a taped session, the rater must consider the context of the behavior and, most importantly, the case conceptualization. This is not to confuse the importance of avoiding halo effects. Raters are to code the behavior of the client and therapist based on the specific information they are given about each person and the potential functional classes of behaviors in which clients might engage. The examples will not be based on this idiographic data particular to the client-therapist behavior being coded.

   The examples are provided merely as illustrations of the code. Raters will realize that each of the examples may have other interpretations of which code should be applied if a more thorough description of the client were provided. Still, lacking this information, the rater should let the example illustrate the code given the way the function of the client or therapist behavior is described.

   Most importantly, the examples must serve as roughly functional examples of the client or therapist behavior relevant to the code. If the rater begins to look in tapes for topographical similarities to the examples provided, the rater will likely code the turn in question inaccurately. The rater needs to attend to the impact the behavior of the client or therapist has on the other person, not to how the behavior appears or is similar to the examples provided.

8. *Examples in the Manual can occur in three different forms*:
   (1) list of relevant aspects of the code which should be considered;
   (2) therapy exchange or exchanges which are marginal examples of the code;
   (3) therapy exchange or exchanges which are counter examples of the code.

   When dialogue is given in an example, it is italicized. The letter "T" indicates the therapist as speaker, and the letter "C" indicates the client as speaker.

9. *Response Class Decision Hierarchy*: A case conceptualization will accompany any tapes that you code. These will outline the different functional classes that are believed to result in the client's distress. Sometimes, these classes of clinically relevant behaviors are not mutually exclusive. Rather, there is a functional overlap whereby the success of one class of behaviors is dependent upon the client achieving success in a more basic repertoire class. If success at the more base level has not been achieved, then client improvement in the more advanced class of behaviors cannot be coded. An example would be a client who has difficulty producing clear speech when talking about emotional issues. When their speech is unclear, it is considered to mitigate or reduce the impact of their interpersonal disclosures *even if those disclosures would be considered improvements for another response class!*

10. *Multiple Code Decision Hierarchy*: Because turns can sometimes appear to be assigned more than one code, a decision hierarchy has been developed. These rules are abbreviated by the number for the code and are explained to the right in text. It is essential that this hierarchy be followed if it is unclear which of two (or more) codes should be assigned. This decision hierarchy is included at the end of this Manual. Only one code is given even if a turn appears to have multiple functions.

11. *Overlap Between Current and Prior Sessions*: Often an issue that was discussed in an earlier session is implicitly or explicitly referred to in the current session. For example, a client might refer to a turn in a previous session when the therapist objected to her coming late. Code only therapist and client behavior that takes place in the current session as clinically relevant behaviors (CRBs). For example, if the therapist pursued the matter of lateness further, after the client mentioned it, such behavior has indeed occurred in the current session and should be coded based on the impact it has during the current session. Otherwise, talk about past therapy interactions is simply focusing on the therapy relationship and not actively engaging in treatment.

III. INSTRUCTIONS TO RATERS

1. *RATE EVERY TURN*. This scale is designed so that every turn is rated using one of the codes. DO NOT LEAVE ANY TURNS UNRATED.

2. *CODE THE APPROXIMATE FUNCTION OF EACH TURN*. This rating system is designed to be functional, or at least quasi-functional with respect to the codes that are assigned to each client or therapist behavior. Therefore, when coding a turn, watch the impact that that turn has on the other person to determine the approximate function that the behavior might have. If the client appears to be engaging in a problem behavior, note that on the transcript as a CRB1 and then continue to watch how the behavior functions on the therapist. If the client's behavior appears to have the effect of a CRB1, then leave the code as written, if the client's behavior instead actually functioned as a CRB2 based on how the therapist was impacted by this, adjust the code to accurately reflect the function that behavior had. Similarly, if the therapist appears to engage in behavior coded as responding to a CRB1, note this on the transcript and watch the following turn to determine the impact the therapist had on the client. If the client responds in such a way as to support the previous code of effectively responding to a CRB1, the code remains. If , however, the client's response indicates that the behavior the therapist emitted was ineffective (perhaps too affect laden) and serves to confuse the client or stop all responding, the rater should correct the code to indicate that the therapist's behavior is coded as Ineffective Contingent Responding.

   By attending to the function that the client's and therapist's behaviors actually have on the other person, this system approximates a functional coding system and better identifies the behaviors of interest to functional analytic psychotherapy.

   The number of turns examined to determine a code should be kept to a minimum. A total of six turns, three separate turns for client and therapist, is set as the limit to provide context to the rater. If, for example, a rater determines that a code was given inaccurately to a therapist's behavior two turns earlier the rater could go back and adjust this code. The purpose of limiting the number of turns is to help keep raters on task, coding behaviors as they occur during a session, and so that a limited amount of session information is provided about later events (to prevent biasing the coding of earlier behaviors).

Natural responses by either therapist or client may be subtle. Notice how the impact of a therapists' seemingly subtle or even casual response functions as a specifically codeable FAP response.

3. *READ CODE DEFINITIONS WHEN THERE IS A QUESTION ABOUT CODE ASSIGNMENT*. Raters should periodically review the manual, particularly when the coders has been informed of potential criterion drift on codes, when there is a question about which code to assign. When there is a question about code assignment, the coder should review the relevant code(s), marginal and counterexamples, and review the decision hierarchy.

   Careless errors may result when raters code a turn based on reading only the code name and not the code as defined in the manual. This is especially important when raters are being trained to use this rating system. Because of the complexity of the codes, it is also essential that the rater be completely familiar with the information in the Manual for each code before applying it. It is important that the rater continually refer to the Manual, even after she/he has become familiar with it, in order to prevent rater drift.

4. *ATTEND TO MANUAL NOTES*. Commentary following each scale item in the manual may specify conditions under which an item should be given, examples where it is difficult to know how to apply a code, and counter examples of each code.

5. *LISTEN BEFORE RATING*. Do not apply and code to a behavior until the turn has occurred (i.e., do not anticipate what the client or therapist is about to say and decide how that should be coded).

6. *TAKE NOTES*. We recommend that the rater take notes while listening to the session. This procedure enhances the accuracy of ratings both because it helps remind raters of information relevant to rating, and because it helps keep the rater focused. Because the rater is required to make many fine distinctions, it is essential that the rater not attempt to do any other task when listening to tapes to be rated. This is especially valuable when training.

ABBREVIATION OF CODES, NAMES, NOTES, AND LOCATION IN MANUAL

LIST OF CLIENT CODES

Code Abbr.	Full Name of Code	Brief Description	Special Notes	Page # in Manual
CRB1	Clinically Relevant Behavior 1 (problems in session)	client engages in problematic behavior in-session in the context of the therapeutic relationship	subscript as A,B, C, etc. or Z for class of CRB	13
CRB2	Clinically Relevant Behavior 2 (improvements in session)	client engages in improved behavior in-session in the context of the therapeutic relationship	subscript as A,B, C, etc. or Z for class of CRB	16
CRB3	Clinically Relevant Behavior 3 (description of important controlling variables)	client describes how different controlling variables impact the client's behavior and when the client makes these functional descriptions		18
CTR	Client Focus on the Therapeutic Relationship	client focuses on the therapeutic relationship	Behavior is **not** codeable as a specific FAP response	20
O1	Discussion of Clinical Problems Outside the Therapeutic Relationship ("outside CRB1s")	client discusses or describes problem behaviors that have been the focus of treatment but that occur in other situations outside of session	subscript as A,B, C, etc. or Z for class of "outside" CRB	22
O2	Discussion of Clinical Improvements Outside the Therapeutic Relationship ("outside CRB2s")	client discusses or describes improvements that have been the focus of treatment but that occur in other situations outside of session	subscript as A,B, C, etc. or Z for class of "outside" CRB	24
CPR	Client Positive Session Progression	client discusses or describes problems as they occur in situations other than the therapeutic relationship, or clarifies or provides context about problems		26

The above list of Client codes should be used only as reference and reminders for code abbreviations and special notes for each code. When raters have questions about codes, they must consult the manual.

LIST OF THERAPIST CODES

Code Abbr.	Full Name of Code	Brief Description	Special Notes	Page # in Manual
TTR	Therapist Focus on Therapeutic Relationship	therapist continues focus on therapeutic relationship including sharing the therapist's feelings in response to the client		28
ERB	Therapist Evokes a CRB by Client	therapist evokes a clinically relevant behavior by the client, either CRB1, 2, or 3	subscript as 1,2, or 3 for type of CRB	30
TRB1 (effectively)	Therapist Responds Effectively to CRB1 (problems in session)	therapist's response is to in-session client problem behavior	subscript as A,B, C, etc. or Z for class of CRB	32
TRB2 (effectively)	Therapist Responds Effectively to CRB2 (improvements in session)	therapist responds effectively to in-session improvements	subscript as A,B, C, etc. or Z for class of CRB	35
TRB3 (effectively)	Therapist Responds Effectively to CRB3 (description of controlling variables)	therapist responds to the client describing how different controlling variables impact the client's behavior; therapist shapes or models CRB3 for client		38
RO1	Therapist Responds to Client's Discussion of Clinical Problems Outside the Therapeutic Relationship (to "outside CRB1s")	therapist comments on problem behaviors the client describes having engaged in outside of the therapy session	subscript as A,B, C, etc. or Z for class of "outside" CRB	40
RO2	Therapist Responds to Client's Discussion of Clinical Improvements Outside the Therapeutic Relationship ("to outside CRB2s")	therapist provides verbal reinforcement in response to the client describing improved behaviors outside of the therapy session	subscript as A,B, C, etc. or Z for class of "outside" CRB	41
TPR	Therapist Positive Session Progression	therapist engages in generally effective or facilitative behavior	**cannot** be coded using one of the specific FAP codes outlined above	43
M1	Therapist Misses/Does not Respond to CRB1	Therapist misses an opportunity to respond to a CRB1	subscript as A,B, C, etc. or Z for class of CRB	47
M2	Therapist Misses/Does not Respond to CRB2	Therapist fails to reinforce an instance of a client's CRB2 or a reasonable approximation of a CRB2	subscript as A,B, C, etc. or Z for class of CRB	49
M3	Therapist Misses/Does not Respond to CRB3	therapist misses an opportunity to respond to a client's description of important controlling variables or reasonable approximation		51

List of Therapist Codes Continued

IRB1	Ineffective Response to a CRB1	Therapist emits a contingent response that was ineffective in reducing the frequency of a CRB1	Subscript as I, II, III for type ineffective therapist behavior	53
IRB2	Ineffective Response to a CRB2	Therapist emits a contingent response that was ineffective in increasing the frequency of a CRB2	Subscript as I, II, III for type ineffective therapist behavior	55
IRB3	Ineffective Response to a CRB3	Therapist emits a contingent response that was ineffective in increasing the frequency of a CRB3	Subscript as I, II, III for type ineffective therapist behavior	56
IN	Generally Ineffective Therapist Responding	Therapist engages in generally ineffective behavior	**cannot** be coded using one of the specific FAP codes outlined above	57

The above list of Therapist codes should be used only as reference and reminders for code abbreviations and special notes for each code. When raters have questions about codes, they must consult the manual.

CLIENT BEHAVIOR CODES

1. **Clinically Relevant Behavior 1: Client Problems In-Session (CRB1)**

Definition: Code CRB1 when the client engages in problematic behavior in-session, as it occurs in the context of the therapeutic relationship. Do not code a CRB1 when the turn is problem behavior the client describes as it occurs in other relationships outside therapy, unless the function of that describing behavior operates as a CRB1 with the therapist.

CRB1s are defined based on their functional relation to the impact the behavior has on the therapist. CRB1s are not defined solely on their topographical features with one exception. The problem behavior needs to be in direct relation to the therapist, unless the case conceptualization specifies otherwise. The case conceptualization is what distinguishes the threshold for CRB1s and O1s.

Distinguish CRB1s based on the functional class of which they are examples and are denoted by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $CRB1_A$s may include behaviors such as avoidance of negative affect during session, and $CRB1_B$s may be defined as client behaviors that function to prevent the client from accessing important interpersonal social reinforcers from the therapist.

A list of CRB1s (e.g., $CRB1_{A,B,C}$s etc.) should be made based on the case conceptualization and raters should code each CRB1 consistent with the functional class of which each CRB1 is an instance.

Therefore, code a **CRB1** using the appropriate subscript for that CRB1 (e.g., **$CRB1_A$**) to distinguish the specific CRB1s as they occur.

In the event that both a $CRB1_A$ and a $CRB1_B$ occur in the same turn, indicate that both have occurred by making both subscripts, that is, $CRB1_{AB}$

A case conceptualization may not be complete, and a rater could note the occurrence of an additional class of CRB not provided. If a CRB occurs that is not indicated by the conceptualization, the rater should mark that code as $CRB1_Z$, where the Z subscript indicates that the CRB is not part of the conceptualization but is an important CRB, nonetheless. When a rater utilizes the $CRB1_Z$ code he or she should notify the project manager when the session is finished being coded. The manager will discuss the issue with the rater to determine the appropriateness of continued use of the code.

EXAMPLES

In this example, a brief case conceptualization is provided to illustrate why the codes would be applied. Readers should assume the client has presented for problems with interpersonal relationship difficulties and has specific problems with accepting that others care about her and inaccurately tacts how others are feeling towards her ($CRB1_A$). She also tends to underestimate her ability to be in a close relationship with someone, a discrimination deficit ($CRB1_B$). When these behaviors occur in the context of the therapy relationship, they function as CRB1s. CRB2s are the corresponding improvements in therapy with these problems. For example, when the client recognizes that the therapist cares about her, accurately identifies this, and she tacts how this makes her feel, this would be a $CRB2_A$. A $CRB3_A$ would be approximations at identifying the important controlling variables that impact how she is able to engage interpersonal relationships more effectively or those contingencies that make it difficult to do so.

The following are examples of a **CRB1** code:

The client engages in behavior with the therapist that is similar to those behaviors the client exhibits in other relationships outside therapy which cause problems in interpersonal relationships
T: How are you feeling right now? (ERB_1)
C: I feel like you don't care about me, that's how I feel. Like you don't care. ($CRB1_A$)

Client engages in behavior that severely limits the therapist's response options to the client
T: So, do you think you might be able to try that with me next time? (ERB_1)
C: Oh, I don't know. I'll never be good in relationships. I'm just no good at this at all. ($CRB1_B$)

MARGINAL EXAMPLE

T: Tell me more about what you're feeling. (ERB)
C: I hate talking about this, being in a relationship is really difficult. Talking about my feelings with you is difficult. (?)
T: Yeah, this is hard stuff, I know… (depends on code given to client behavior)

In the event that the client's behavior functions to distance the therapist from him or her, this client behavior would be coded a **$CRB1_A$**. However, if the client is merely pointing out that talking about emotions is difficult for him or her, it may be a **CRB3**. This would more likely be the case if the client was not able to identify or disclose that talking about emotions is difficult as indicated by the case conceptualization. In this example, based on the decision code hierarchy would assign the code CRB3 if he or she were unsure whether the behavior functioned as a CRB1 or CRB3.

COUNTER EXAMPLE

T: Tell me about your week, what's been going on. (TPR)
C: Well, John and I really got along great. We were able to talk about the rough spots we've been having and how to maybe have less of those. I was really able to, you know, open up to him and let him know how I have been feeling lately. I even talked about how I felt right then, like you've been telling me all this time. (O2)

In this example the client refers to the therapist but is clearly discussing issues in outside relationships. This turn would be coded as a client improvement outside of session (O2). Coders should not assign a code of CRB when the talk is focused on issues outside the therapeutic relationship unless the client behavior functions as a problem or improvement specific to the CRBs outlined by the case conceptualization (see description of codes O1 and O2 below).

2. **Clinically Relevant Behavior 2**: **Client Improvements In-Session (CRB2)**

Definition: Code CRB2s when the client engages in improved behavior in-session in the context of the therapeutic relationship. A CRB2 is not improved behavior that the client describes as it occurs in other relationships outside therapy, unless the function of that describing behavior operates as a CRB2 with the therapist.

CRB2s are defined based on their functional relation to the impact the behavior has on the therapist. CRB2s are not defined solely on their topographical features with one exception. The improved behavior needs to be in direct relation to the therapist, unless the case conceptualization specifies otherwise. The case conceptualization is what distinguishes the threshold for CRB2s and O2s.

Distinguish CRB2s based on the functional class of which they are examples and denote each by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $CRB2_A$s may include behaviors such as not engaging in avoidance of negative affect during session, and $CRB2_B$s may be defined as client behaviors that function to allow the client to access important interpersonal social reinforcers with the therapist.

A list of CRB2s (e.g., $CRB2_{A,B,C}$s etc.) should be made based on the case conceptualization and raters should code each CRB2 consistent with the functional class of which each CRB2 is an instance.

Therefore, a **CRB2** should be coded using the appropriate subscript for that CRB2 (e.g., $\mathbf{CRB2_A}$) to distinguish the specific CRB2s as they occur.

In the event that both a $CRB2_A$ and a $CRB2_B$ occur in the same turn, the coder should indicate that both have occurred by making both subscripts, that is, $CRB2_{AB}$

If a CRB occurs that is not indicated by the conceptualization, the rater should mark that code as $CRB2_Z$, where the Z subscript indicates that the CRB is not part of the conceptualization but is an important CRB, nonetheless.

EXAMPLES

The following are examples of a **CRB2** code :

The client engages in useful behavior, improving his or her ability to relate closely and interpersonally to the therapist and others
T: Tell me how you felt before coming in here today. (ERB_2)
C: Well, to be honest, I was nervous. Sometimes I feel worried about how things will go, but I am really glad I came in even though I was feeling that way. (CRB2)

The client engages in improved behavior in such a way as to be at least a successive approximation to a CRB2.
T: Tell me how you felt before coming in here today. (ERB_2)
C: Well, I don't know what to say. I felt, well, nervous I guess. (CRB2)

Both examples are difficult to code without a conceptualization. However, if the client were assumed to have difficulty sharing any negative affect in the context of the therapeutic relationship, the first would be a CRB2. The second is coded as a CRB2, as the client is tacting an emotional experience. These are examples and should not be treated as rules. Statements that might serve as CRB2s early in therapy would not necessarily be coded as such in later sessions as the threshold for approximations is higher as therapy nears its end.

MARGINAL/COUNTER EXAMPLE

T: Do you want to talk with me about something in particular? (ERB_1)
C: Well, I was wondering if there was anything you'd like to talk about? (see below)

This is another example where the conceptualization will determine the code for the client's behavior. In the event that the client is unable to assert his or her needs, the therapists has presented the client with an opportunity to do so, and the client has deferred to the wishes of the therapist; raters code this as a **CRB1** for the client behavior. However, if the client were unable to acknowledge the needs of others in a relationship, typically overlooking opportunities to have input about the needs of others, this would be an example of a **CRB2** by a client.
Again, the case conceptualization will determine the code for the client's behavior.

3. **Clinically Relevant Behaviors 3: Client Description of Important Functional Relationships among controllable variables** (**CRB3**)

Definition: Code client verbal behavior as CRB3 when the client describes how different controlling variables impact the client's behavior and when the client makes these functional descriptions. These descriptions of controlling variables are CRB3s whether the client describes in-session variables, or those contingencies outside the therapy setting. Code CRB3s when the client paraphrases the therapist's description and appears to alter the description in his or her own words. Also code CRB3s when the client makes an approximate functional description of important controlling variables as part of a three term contingency, but misses one of the terms in that relationship.

EXAMPLES

The following are examples of a **CRB3** code:

The client describes events in his or her life that lead to more successful interactions or satisfaction.
C: I get so worked up beginning relationships, I try to force my values on my partner before he is ready. The next thing I know, the relationship is over. (CRB3)
In this example the client has identified all of the variables, antecedent events, identified behaviors, and the consequences of that behavior (here, punishing effects).

The client could describe this functional relationship, but leave out a term and still receive a CRB3 code.
C: I get so worked up beginning relationships, I try to force my values on my partner before she is ready. (CRB3)
Note that while the client has left off the final term of this analysis he or she has identified an important controlling variable over his or her behavior that can be manipulated to produce a different outcome.

MARGINAL EXAMPLE

C: When I'm around women, I get so nervous, I just can't stay around long enough to have a conversation. (see below)

This is potentially an approximation of a CRB3 as it does not adequately describe the relationship between the client's behavior and variables that can be manipulated or otherwise altered. When coding sessions in linear order, the rater can accept this as a **CRB3** in that it is an approximation of a CRB3. However, future descriptions of variables by the client would have to be more elaborate.
When coding sessions out of order (as in most research using the manual), the coder will have to use his or her judgment as to whether the client behavior in question is elaborated enough to be considered an approximation of a **CRB3** at that time.

COUNTER EXAMPLE

C: I just can't be around men. (CRB1/CPR - see below)

In this example, the client's behavior is not elaborated enough to be considered an approximation of a description of controlling variables. The client is simply stating that as a fact. If the case conceptualization indicated that the client had difficulties with very contrasting or "black and white" styles of thinking, this would be a **CRB1**, as the impact it has on the therapist and client is to limit available reinforcers and other ways of engaging this problem.
If the rater does not determine this behavior to be a CRB1 nor a CRB3, the code that would be assigned is Client Positive Session Progression (CPR, see p. 26) because the client is relating difficulties he or she is having in relationships outside of therapy.

4. **Client Focus on the Therapeutic Relationship (CTR)**

Definition: Code CTRs when the client focuses on the therapeutic relationship (i.e., his or her feelings in the moment about the therapist) but when this behavior is ***not*** codeable as a specific FAP response (e.g., a CRB1, 2, or 3).
Code client behavior as CTR when the client continues focusing on therapeutic relationship including sharing the client's feelings in response to the therapist. This

code is also given when the client focuses on the discussion of in-session behaviors as pointed out by the therapist.
This code is also given when the client discusses with the therapist the purpose of focusing on the therapeutic relationship during treatment (i.e., a discussion about how FAP works).
Lastly, use this code to categorize client talk that talks about past therapy sessions when the function of those sessions isn't currently present.

EXAMPLES

The following are examples of a **client focus on the therapeutic relationship** code (**CTR**):

The client continues to describe how he or she feels about the therapist or the therapeutic relationship, but this behavior is ***not*** a CRB for the client
C: I like talking to you. You make me feel safe here. (CTR)

The client discusses with the therapist the fact that therapy is a relationship
T: Therapy is like other relationships you have, like we've been talking about, but it is not like others in some important ways. Can you think of some of those? (TTR)
C: Yeah. Well one is that I only see you for an hour a week and another is that there are limits about what you can say to other people about what goes on in here. (CTR)

MARGINAL EXAMPLE

C: Sometimes I get pretty upset with people. (CPR)
T: Do you get upset in here, with me? (ERB)
C: Yeah, I do. Sometimes I really do.

In this example, the first client behavior is talk about problems outside the relationship and would be coded as **CPR** (described below). The therapist then brings the focus of the discussion on the therapeutic relationship (**ERB**) and the client responds affirmatively. This could be coded as a **CRB2** only if the conceptualization indicated that any discussion by the client with the therapist about the therapeutic relationship and engaging the therapist honestly about such feelings were a clinical improvement. If the conceptualization did not indicate the client had any problem with doing this with the therapist, the code would then be a client focus on the therapeutic relationship (**CTR**).
This is another example of the importance of the case conceptualization determining the coded response.

COUNTER EXAMPLE

T: I'm glad you can share this with me. It makes me feel close to you when you do that. (TTR)
C: Yeah. You know, it feels good to finally open up about this to someone. (CRB3)

If the client engages in a clinically relevant behavior as determined by the case conceptualization, then the client behavior is coded as that type of **CRB**, in this case a CRB3. Again, this CRB is only coded as such if, based on the case conceptualization and the ongoing information provided in therapy, the behavior functions as a CRB.

5. **Client Discussion of Problems Outside Session (O1)**

Definition: Code client behavior using O1 when the client discusses or describes problem behaviors that have been the focus of treatment but that occur in other situations outside of session. These problems are not considered CRB1s because the behaviors are not problems evidenced in-session in the context of the therapeutic relationship. They should provide clear descriptions of the problem behavior or it is coded as CPR.

Distinguish O1s based on the functional class of which they are examples and denote each by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $O1_A$s may include behaviors such as engaging in avoidance of negative affect outside of session, and $O1_B$s may be defined as client behaviors that function to prevent the client from accessing important interpersonal social reinforcers with others outside therapy.

Therefore, **O1** should be coded using the appropriate subscript for that O1 (e.g., $O1_A$) to distinguish the specific O1s as they occur.

If an O1 occurs that is not indicated by the conceptualization, the rater should mark that code as $O1_Z$, where the Z subscript indicates that the "outside" CRB is not part of the conceptualization but is an important O1, nonetheless.

EXAMPLES

The following is an example of a **Client problems outside of session** code **(O1)**

C: I had a fight with my partner last week, I told him he could go to hell. I wasn't going to talk to him anymore. (O1)

In this example, the client's behavior is assumed to be problematic based on the case conceptualization (e.g., the client avoids interpersonal conflict by terminating the discussion. The behavior does not occur in the context of the therapeutic relationship and would be coded as a problem behavior outside of session.

MARGINAL EXAMPLE

C: I don't know, I just didn't want to tell her how I felt. I guess that's how I feel in here sometimes. (CTR - see below)

Here, the client is describing outside problems but has briefly linked those problems to the therapeutic relationship. Based on the decision hierarchy, the rater would assign the code of client focus on the therapeutic relationship (CTR). If the rater believed that the client's behavior was for any reason a clinically relevant behavior (i.e., CRB1 or CRB2), the rater would assign that code, provided the case conceptualization indicated the behavior could function in that way (i.e., as a problem behavior or as an improvement in-session).

COUNTER EXAMPLE

C: I realized that when I talk to my father like that, you know, being demanding, he just won't listen to me at all. (CRB3)

In this example, the client is specifying functional relationships between the variables of his behavior with respect to the consequences that behavior has. This is an approximation to a CRB3 and should be coded as such.

6. **Client Discussion of Improvements Outside Session (O2)**

Definition: Code client behavior as O2 when the client discusses or describes improvements that have been the focus of treatment but that occur in other situations outside of session. These improvements are not considered CRB2s because the behaviors are not improvements evidenced in-session in the context of the therapeutic relationship. They should provide clear descriptions of the improved behavior or it is coded as CPR.

Distinguish O2s based on the functional class of which they are examples and denote each by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $O2_A$s may include client description of behaviors such as not engaging in avoidance of negative affect with others outside session, and $O2_B$s may be defined as client behaviors that function to allow the client to access important interpersonal social reinforcers with others.

Therefore, a **O2** should be coded using the appropriate subscript for that "outside" CRB (e.g., **$O2_A$**) to distinguish the specific O2s as they occur.

If an O2 occurs that is not indicated by the conceptualization, the rater should mark that code as $O2_Z$, where the Z subscript indicates that the outside CRB is not part of the conceptualization but is an important O2, nonetheless.

EXAMPLES

The following is an example of a **Client improvements outside of session** code **(O2)**

C: I just faced my fear of confronting him, that he was wrong, and I told him how I felt. (O2)
T: I think that's terrific. You really said how you felt. (RO2)

Provided the case conceptualization indicated this, the client in this example is describing an improvement outside of session she displayed. Note that the behavior occurs in the context of a relationship, but it is not the therapeutic relationship.

MARGINAL EXAMPLE

C: I did what I do in here, I told him how I felt, and we had a long talk about our relationship. (O2 - see below)

Here, the client is describing outside improvements but briefly brings the focus on the therapeutic relationship, then takes it back out to the original problem. The rater could either code this as a client focus on the therapeutic relationship (CTR) or as an improvement in session (O2). If the coder is not sure which code to assign, the CTR takes priority. In this example, the client is only alluding to the therapeutic relationship and is not focusing on this issue as would qualify the CTR code. This example should employ the O2 code.

C: I talked to her and I really opened up. I felt, well, like I feel now talking to you, really safe. I know that opening up to you won't make you want to not see me anymore. I feel really safe with you now. (CRB2)
T: That's really great. (TRB2)

In this example the client is assumed to have difficulties with talking about trust and being open emotionally with the therapist. Here the client has begun with a description of an outside improvement (O2), but focused on feelings in the room with the therapist that function as a clinical improvement in session, a CRB2.

COUNTER EXAMPLE

C: I just faced my fear of confronting him, that he was wrong, and that he could go to hell. (O1)

This is an example where the client's behavior likely was a problem behavior outside of session. The code that should be assigned is O1. Although it begins as a potential improvement, and the client did say to the other person how she felt, the client was (we will assume for this example) not engaging in effective behavior outside session (coded as O2).

7. **Client Positive Session Progression** (**CPR**)

Definition: Code **CPR** when the client engages in generally on-task behavior that is facilitative to discussion and that **cannot** be coded using one of the specific FAP codes outlined. This code is provided when the content of the client's verbal behavior functions to keep the discussion focused on the task in-session without impacting the function of the therapist's behavior.

These can include: when the client discusses or describes problems as they occur in situations other than the therapeutic relationship; when the client clarifies or provides context to the therapist about the problems he or she is discussing; when the client and therapist discuss homework assignments (that are not CRBs on the part of the client).

The CPR code is ***not*** assigned when client behavior is considered a CRB and is not a clinical problem or improvement that occurs outside therapy (O1 & O2).

As with all other codes, this cannot be given based strictly on the topography of discussing outside issues. Discussing problems may be considered a CRB1, CRB2, O1, or O2 depending on the conceptualization. The code must be given based on the case conceptualization and on the context in which the behavior occurs.

In the event that a coded response contains both **CPR** and another specific FAP response, the specific FAP response is coded.

CONTINUATION: CPR sometimes functions to continue the speech of the therapist but does not directly impact the next therapist utterance. When this occurs, the therapist codes that precede and follow the CPR code would be identical (see example).

EXAMPLES

The following is an example of a continuation **CPR** code:
T: It seems when I mention your improvement that you become…(TRB1)
C: Anxious. I know. I know. (CPR)
T: And that this isn't what I intended at all. (TRB1)

The following are examples of a **CPR** code:

The client describes difficulties at work, home, school etc. that do not impact the therapeutic relationship directly, but warrant consideration and addressing by the therapist but that cannot be coded specifically as discussion about clinical improvements or problems occurring outside that have been a focus of treatment.
C: So how was your week? (TPR)
T: Well, my roommates got in a fight, and my car broke down, other than that, well, not much. (CPR)

Provided this behavior did not function as a CRB, and that the roommate fight or car problems have not been a focus of treatment, this client behavior is coded as discussion about problems.

Code as **CPR** any client behavior that functions to clarify or provide context for what a discussion about outside problems.
T: So this was your father that said that?
C: Yeah. OK. What you need to know about him is that he rarely interrupts my mother. When he does, it's like, everyone sit up and listen. This is going to be big. So that's really key, you know.

MARGINAL/COUNTER EXAMPLE

T: Were you able to complete your homework assignment this week? (TPR)
C: No, I didn't. My mom fell ill and I spent most of my time home caring for her. (02)

This is an example where the client's behavior would be coded as **O2** if the client's conceptualization included too much rule-governance. Because she let the contingencies in her life appropriately dictate her behavior, it would not be coded as CPR.

C: When my boss said to do that, well, I don't know, I just froze. I knew I shouldn't do it, it's not even my job, but I just couldn't tell him no. (O1)
T: So he really put you in a bad spot. But this is an issue we've been working on isn't it. What could you've said to him? (RO1)

In this example the client is discussing problems in asserting him or herself, an issue that has been a focus of therapy. The behavior is an outside problem and could be coded as **CPR**, but the description of the problem is sufficient and can be coded as an "outside" CRB1 – **O1**.

THERAPIST BEHAVIOR CODES

EFFECTIVE THERAPIST BEHAVIORS

1. **Therapist Focuses on the Therapeutic Relationship (TTR)**

Definition: This is the therapist's response to in-session client behavior such as the client's feelings in the moment about the therapist that is not codeable as a specific FAP response (e.g., response to a CRB1, 2, or 3).
To receive this code the therapist continues focusing on the therapeutic relationship including sharing the therapist's feelings in response to the client
This code is also given when the therapist describes the purpose of focusing on the therapeutic relationship during treatment (i.e., a discussion about how FAP works).
Lastly, use this code to categorize therapist talk that talks about past therapy sessions when the function of those sessions isn't currently present.

EXAMPLES

The following are examples of a **Therapist Focuses on the Therapeutic Relationship** code (**TTR**):

The therapist encourages the client to describe how he or she feels about the therapist
T: So can you tell me more about what you're feeling about me right now? (TTR)

The therapist discusses the fact that therapy is a relationship
T: The reason I'm asking you these questions, ones about how you feel in here right now, is that I think this relationship is an important one. It's not like other relationships you have in some important ways we can talk about, but it is a relationship even still. (TTR)

The therapist describes FAP and the importance of the therapeutic relationship in effecting change in client behavior
T: You know, I can best help you work on the problems you're having in other relationships by focusing on what goes on between you and me in here…(TTR)

The therapist discloses his or her feelings about the client that is not in response to a CRB1
T: And I think about you between sessions. I care about you and what happens to you, and I look forward to us meeting each week. (TTR)

The therapist notifies, clarifies, or makes a connection between how the client's in-session behavior is like his or her behavior out of session
C: We really got along great. I opened up to her and she really seemed to listen to me.
T: That's great. Do you feel like I listen to you, you know, so that you can open up in here? (TTR)

MARGINAL EXAMPLES

C: …So I told him that I think this is an important relationship…(O2)
T: Like this one. (TPR)
C: and he really didn't listen to me...

Here, the therapist attempted to point out that the client is describing a relationship that has features similar to those the client is describing in another relationship outside therapy. However, this is not a

discussion in any way about those similarities, and would not be rated using this code. This response would have to be rated as a **Therapist Positive Session Progression** (TPR) code because it fails to meet the criteria for the **TTR** code.

C: and he really didn't listen to me. That made me pretty upset. (CPR)
T: Do you ever feel like I don't listen to you in here? (TTR)

In this continued example, the therapist is now bringing the focus of the discussion to the therapeutic relationship and how the client feels in-session with the therapist. This example would be coded using the **TTR** code.

COUNTER EXAMPLE

C: It really feels good to finally open up about this to someone. (CRB2)
T: You know, it's good to hear you share that with me. (TRB2)

If the client engages in a CRB and the therapist responds to that by sharing how this behavior impacts the therapist, this is a Response to a CRB2 **(TRB2)**. Again, this response to a CRB is only given if, based on the case conceptualization and the ongoing information provided in therapy, the behavior functions as a CRB. This is not a **TTR** code because the specific response to the CRB by the client.

2. **Therapist Evokes a CRB (ERB)**

Definition: Code ERBs when a therapist attempts to evoke a clinically relevant behavior by the client, either CRB1, 2, or 3. However, the code is not dependent upon the client behavior that follows the ERB. For example, the therapist could attempt to evoke a CRB, but the client needs clarification of the question (CPR) or misunderstands the therapist's request and provides an answer that is unresponsive to the question (CPR). Each of these examples could be considered as possible CRB1s depending upon the case conceptualization (e.g. avoidance).

An ERB code is typically not given in response to a series of CRBs (coded as TRB1, 2, 3). The exception would be if the therapist clearly attempted to shift the focus of the session to a different response class and this shift would not be better coded as a TRB1, TRB2, or TRB3.

There are two common forms of ERBs.
1) The first is when the conversation is focussed on events outside of the current therapy session and the therapist shifts the conversation onto the client/therapist relationship in the current session. The therapist redirects the focus of the session from comments concerning past therapy sessions (CTR/TTR), instances of the client's outside problems/improvements (O1/O2), or progressive talk (CPR/TPR) to the therapeutic relationship, presumably to elicit in-session client behavior (CRBs) that the therapist might respond to.
2) The second common form is when the therapist clearly shifts the discussion from one client response class (as specified by the case conceptualization) to a different client response class during the course of FAP-specific responding (see example).

ERB3s are direct mands for behavior by the client to identify controlling variables. When CRB3s are being reinforced or modeled, the behavior is coded as TRB3.

Distinguish the type of ERB by the client's response that follows as a CRB1, CRB2, or CRB3 using a subscript for that CRB number, i.e., **ERB_1** or **ERB_2**.

EXAMPLES

TYPE 1 Example
C: You know, she made me feel, well, I'm not sure. (O1)
T: What are you feeling right now? (ERB)

Here, as with most ERB codes, the therapist could be evoking a CRB1 or CRB2 and possibly a CRB3. It is not known what behavior the therapist will elicit until the client responds. When the client responds, the rater should add the appropriate subscript to indicate whether the ERB was followed by a CRB1, 2, or 3. In the event that the client did not respond with a CRB, the coder does not need to indicate the subscript.

TYPE 2 Example
C: So I've been noticing that I've really come to depend on your support and caring. ($CRB2_A$)
T: Are there things you aren't happy with in here? (ERB_B)

Here, the conceptualization is vitally important as the distinction between classes A and B would have to be delineated. In this example, class "A" would be tacting reinforcers and class "B" would be tolerance of interpersonal conflict.

MARGINAL/COUNTER EXAMPLES

C: Oh, this is just stupid! (CRB1)
T: OK, why don't you try something different than that. (TRB1)

In this example, it will be assumed that the client's behavior is a CRB1. The therapist's response can be taken to be a response to a CRB1 (**TRB1**) as well as evoking a CRB2 (**ERB_2**) depending on what followed. The definition of the code, however, necessitates that the **TRB1** code be given in place of the **ERB** code when the therapist's response follows a client CRB.

C: I'm thinking I probably won't come back to therapy. (as CRB1)
T: I see. Tell me more about what you're feeling about therapy right now. (M1)

In this case the therapist is seeking more information about the client's behavior. It is difficult in just this interaction to determine whether the therapist is missing an opportunity to respond to the client's behavior (**M1**), or if he or she is attempting to determine the function of the client's behavior (coded as **TPR**). In this case, the therapist's response is coded as missing a CRB1 (**M1**), because the decision hierarchy indicates that a specific FAP behavior takes priority over a general response by the therapist.

3. **Therapist Responds (Effectively) to CRB1** (**TRB1**)

Definition: Code TRB1s when the therapist's response is to in-session client problem behavior (as defined in the case conceptualization).
This code is given when the therapist responds to the client when he or she engages in behavior that has been defined earlier by the therapist as functioning to interfere with or prevent more effective client behavior. The therapist does not need to

comment specifically on the function that the CRB has on him or her and may respond less explicitly or more naturally to the impact that behavior has on the therapist.

Even though the therapist will make some type of response to the CRB1 by the client, coders should not assume that the therapist has responded effectively to a CRB1 when it occurs.

Distinguish CRB1s based on the functional class of which they are examples and denote each by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $CRB1_A$s may include behaviors such as avoidance of negative affect during session, and $CRB1_B$s may be defined as client behaviors that function to prevent the client from accessing important interpersonal social reinforcers.

A list of CRB1s (e.g., $CRB1_{A,B,C}$s etc.) should be provided based on the case conceptualization and raters should code each CRB1 consistent with the functional class to which each CRB1 is an instance.

Therefore, a therapist's response to a CRB1 should be coded using the appropriate subscript for that CRB1 (e.g., **$TRB1_A$**) to distinguish the therapist's responses to specific CRB1s as they occur.

In the event that both a $TRB1_A$ and a $TRB1_B$ occur in the same turn, the coder should indicate that both have occurred by making both subscripts, that is, $TRB1_{A,B}$

If a CRB occurs that is not indicated by the conceptualization, the rater should mark that code as $TRB1_Z$, where the Z subscript indicates that the CRB is not part of the conceptualization but is an important CRB, nonetheless.

EXAMPLES

The following are examples of a **Therapist Responds to CRB1** code **(TRB1)**:

The therapist conveys to the client how that response impacts the therapist (i.e., serves to distance the therapist from the client). The therapist shares his or her feelings about the way that an interaction is impacting him or her

C: Well I think this sucks, I think I've had as much as I can take from you. ($CRB1_A$)
T: This doesn't feel good to me—The way we're relating to each other here. ($TRB1_A$)
C: Oh, well, I guess I am kinda yelling at you when I'm really mad at Tom for what he did. I'm so angry. ($CRB2_A$)

The therapist tells the client directly that that response will not work to get the client what he or she needs from the therapist in this moment. The therapist asks the client what he or she expects will happen in response to what the client did and follows this with a comment about the impact the behavior actually has on the therapist.

C: I'm just a closed person. I'm defensive, and I've always been that way. There's nothing you can do about it so don't try. (CRB1)
T: OK, I want to tell you something. When you do that, tell me how you are, defensive, closed, and that you're unchangeable, I kind of sit here not knowing what to do. I don't feel like you even want me to try to help you. Is that what you want, for me to leave you alone? (TRB1)

The therapist responds naturally or subtly to the impact that the client's behavior has on him or her.
C: Well anyway, I think next week will be a better week, Sherry will be out of town, I'll see a film, there's a new one opening this week. It looks good. (CRB1)
T: Are you doing that thing again? - Distracting us from talking about how you feel after fighting with Sherry? (TRB1)

MARGINAL EXAMPLES

C: Oh, I don't know why you even care about me (to therapist) (CRB1)
T: What do you mean? (TPR)

Here, the therapist may be responding to the impact of the client's behavior, but he or she has not supplied a specific response to a CRB1. This would instead need to be coded as **Therapist Positive Session Progression** (**TPR**) because it does not meet any other criteria defined in this system, and it fits the criteria of clarification by the therapist for the **TPR** code.

If the therapist were to simply continue with:

C: Oh, I don't know why you even care about me (to therapist) (CRB1)
T: What? I really feel like you're discounting an important relationship to me, here. (TRB1)

Then the code would be a **Therapist Response to a CRB1** because the therapist is sharing how the client's behavior is impacting the therapist in a manner that is not functioning to provide the client with more opportunities for social reinforcement (provided that this is a CRB1 in this example).

COUNTER EXAMPLE

C: Let's just talk about something else (in response to the therapist's attempt to continue a difficult discussion). (CRB1)
T: I really appreciate you having stayed with this for a while. (M1)

In this example, the therapist missed an opportunity to respond to a **CRB1** by the client. This therapist response would be coded as a **Missed CRB1** (**M1**) due to the decision hierarchy defined below. In no way is this a contingent response to the CRB code because the therapist has not described or discussed how the client's problem behavior has impacted the therapist. (See below for a description of **Missed CRB1** code.)

4. **Therapist Responds (Effectively) to CRB2 (TRB2)**

Definition: Code TRB2s when the therapist responds effectively to in-session improvements in client behavior as defined in the case conceptualization.
Assign this code when the therapist responds to the client when he or she engages in more effective behavior in the therapeutic relationship. The therapist does not need to comment specifically on the function that the CRB has on him or her and may respond less explicitly or more naturally to the client impact that behavior has on the therapist.

The therapist should make some type of response to the CRB2 emitted by the client, but coders must not assume that the therapist has responded to a CRB2 effectively when it occurs.

Distinguish TRB2s in the same manner as TRB1s by categorizing each based on the functional class to which it belongs (e.g., $CRB1_{A,B,C}$s, etc.).

Therefore, a therapist's response to a CRB2 should be coded using the appropriate subscript for that CRB2 (e.g., **$TRB2_A$**) to distinguish the therapist's responses to specific CRB2s as they occur.

In the event that both a $TRB2_A$ and a $TRB2_B$ occur in the same turn, the coder should indicate that both have occurred by making both subscripts, that is, $TRB2_{A,B}$

If a CRB occurs that is not indicated by the conceptualization, the rater should mark that code as $TRB2_Z$, where the Z subscript indicates that the CRB is not part of the conceptualization but is an important CRB, nonetheless.

EXAMPLES

The following are examples of a **Therapist Responds to CRB2** code (**TRB2)**:

The therapist reinforces the client's behavior that may include not avoiding feelings, talk that is not superficial, staying with difficult discussions, etc. and reinforces specific approximations toward more effective in-session behavior.
C: This is just really hard to talk about. (CRB2)
T: Yeah, it is, but I'm glad you're talking about it with me today. (TRB2)

The therapist responds to the client by conveying the positive impact that the client's behavior had on the therapist in a more natural way.
T: That was neat that you were able to tell me that. That's just great. (TRB2)

The therapist conveys how this client behavior is more effective in helping client meet his or her needs and/or goals with respect to the goals of therapy.
C: OK, OK, I don't mean to yell at you, and I really do want you to help me with this stuff. Can you, can you help me figure out what to do? (CRB2)
T: I'd be glad to help you, and you know, when you ask me like that, telling me what's going on for you, it makes me lots more happy to do that with you. (TRB2)

The therapist does not need to comment specifically on the function that the CRB has on him or her and may respond less explicitly or more naturally to the client impact that behavior has on the therapist.
C: I really do care about you, you know. (CRB2)
T: Thanks, the same goes for me. (TRB2)

MARGINAL EXAMPLES

T: So, I'll see you next week? (TPR)
C: Actually, I'd like to change my appointment time. (CRB2 -see below)
T: That's no problem. Let's see what I can do. What time were you thinking of? (TRB2)

This is an example of the need to be clear on the case conceptualization for the particular client-therapist dyad being rated. In the event that the client has a history of being unable to assert his or her needs, even

over more minor issues in the therapeutic relationship, the client's behavior here would be understood as a CRB2. The therapist's response to this is a **TRB2** because he is making accommodations to the client given his or her request. If the therapist had responded by saying:

T: That's really not possible. Sorry about that. (M2 - see below)

This would be considered a missed opportunity to reinforce a CRB2 (see **Missed CRB2** description below), and would be coded accordingly.

If however, the therapist responded by saying:

T: That's really not possible. Sorry about that, but that was really great of you to ask me to do that given it's not easy for you to assert your needs sometimes. (TRB2 - see below)

This would be considered a TRB2 using this system. The two therapist behaviors that receive a rating of Responds to a CRB2 (the first and third) are different with regard to natural versus arbitrary responding by the therapist (e.g., Ferster, 1967).

COUNTER EXAMPLE

If the client has a history of making excessive demands of others in interpersonal relationships and this is an example of that behavior (CRB1), then the raters must code the client and therapist behavior accordingly. In the example above:

T: So, I'll see you next week? (TPR)
C: Actually, I'd like to change my appointment time. (CRB1 - see below)
T: That's no problem. Let's see what I can do. What time were you thinking of? (M1 - see below)

The therapist's response to the client's CRB1 (using the above conceptualization) would not be a TRB2, but would be **Missed CRB1** (**M1**, defined below). It is essential to note that this response is immediately tied to the way the client's behavior is defined based on the conceptualization. If the client emits excessive demands that function as CRB1s, this does not indicate that all demands the client engages in are excessive, some demands may be requests or reasonable demands. This is essential to remember, as all client behaviors are understood as they function in the therapeutic *relationship based on the case conceptualization.*

5. **Therapist Responds to CRB3 (TRB3)**

Definition: Code TRB3s when the therapist responds effectively to the client describing how different controlling variables impact the client's behavior. Additionally, this code captures when the therapist models this description for the client.
This code is given when the therapist responds to the client when he or she makes these functional descriptions. The code is also given when the therapist describes them for the client. The code is given if the client paraphrases the therapist's description and appears to alter the description in his or her own words, *and* the therapist responds to this client behavior with support or assistance in the analysis of the specific client behavior. This code can also be given when the therapist provides general descriptions of the importance of the client conducting basic behavioral analyses on his or her own, noticing how variables interact and examining functional relationships between variables.

EXAMPLES

The following are examples of a **Therapist Responds to CRB3** code (**TRB3)**:

The therapist reinforces the client's verbal behavior that includes a description of the controlling variables that likely bring about more effective client behavior.
C: The one thing I do know is that when I tell someone how I feel, you know, when I feel like I might really like them, that if I say it too strongly or too much, they tend to back away from me. That makes me feel pretty lousy. (CRB3)
T: It sounds like you've really picked out the important things going on there with respect to the things we talked about, what comes before and after the particular thing you do that you're interested in. (TRB3)

The therapist assists the client in his or her analysis by suggesting different controlling variables to attend to, or otherwise helps strengthen the client's own analysis of his or her behavior or models a CRB3 for the client, describing the important variables the client should consider.
C: I don't know, he just leaves. I say how I feel, and he leaves. (CRB3)
T: Well, tell me what's going on for you when you want to say how you're feeling, you know, what comes before you saying what you feel. Let's tease that apart. Are you feeling anxious when that happens? (TRB3)

MARGINAL/COUNTER EXAMPLES

C: I just can't be around other people. (CPR - see below)
T: That's really great that you can identify that. (IN)

While this may be an example of a very general approximation that will at some point lead to a CRB3, in its current state, the client's behavior is not a **CRB3**. Therefore, the therapist's response to that behavior cannot be considered a **TRB3** code. The client's behavior in this interaction, provided it is not a CRB1, would be coded as discussion about problems (**CPR**) as it fails to meet the criteria for any other code at this point. The therapist's response to the client discussing problems is rather poorly matched and would likely receive a Generally Ineffective Therapist behavior code (**IN**). If the client's behavior were a **CRB1**, the therapist clearly missed an opportunity to respond to that **CRB1** and the therapist turn would be coded as **M1**.

If the interaction had appeared as follows:

C: When I'm around other people, I get so nervous that I leave the room really fast, and I don't end up ever getting to know any of them. (CRB3)
T: That's really great that you can identify that. (TRB3)

This interaction would be coded as a **TRB3** because the client's behavior is an approximation to, if not a full **CRB3**.

6. **Therapist Responds to Client Problem Behaviors Outside Session (RO1)**

Definition: Code RO1 when the therapist comments on problem behaviors the client describes having engaged in outside of the therapy session. This code cannot be given when the therapist comments or responds to in-session behavior (TRB1 or IRB_1)

Distinguish RO1s based on the functional class of which they are examples and denote each by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $RO1_A$s may include responses to client behaviors such as engaging in avoidance of negative affect outside of session, and $RO1_B$s may be defined as therapist responses to client descriptions of behaviors that function to prevent the client from accessing important interpersonal social reinforcers with others outside therapy.

Therefore, a **RO1** should be coded using the appropriate subscript for that RO1 (e.g., $RO1_A$) to distinguish the specific RO1s as they occur.

If an RO1 occurs that is not indicated by the conceptualization, the rater should mark that code as $RO1_Z$, where the Z subscript indicates that the response to that "outside" CRB is not part of the conceptualization but is an important RO1, nonetheless.

EXAMPLES

The following are examples of a **Therapist Responds to Client Problem Behaviors Outside Session** code (**RO1**):

The therapist comments on how the client engages in problems outside of session that have been a focus of treatment.
C: So I told him if he didn't like it, he could go to hell. (O1)
T: Now, do you think that's the best thing you could have done, I mean, did everything turn out the way you'd hope when you told him that? (RO1)

MARGINAL/COUNTER EXAMPLES

C: They were all really ganging up on me, I just left, I couldn't take it when I was the focus of that kind of attention. (O1)
T: So, let's talk about what things were going on when you felt like you had to leave, what kind of things were going on in the environment, like we've been talking about, that made it so you couldn't tell them how you felt about what was going on. (ERB3)

In this example, the client was engaging in talk about problems outside of the therapeutic relationship (**O1**), and the therapist focused the talk on describing the functional relationships between variables, modeling CRB3 behavior (**ERB3**), even though the client did not appear to be engaging in this behavior.

7. **Therapist Provides Reinforcement for Client Improvements Outside Session** (**RO2**)

Definition: Code RO2 when the therapist provides verbal reinforcement in response to the client describing improved behaviors outside of the therapy session. This code cannot be given when the therapist comments or responds to in-session improved behavior (TRB2 or IRB_2).

Distinguish RO2s based on the functional class of which they are examples and denote each by a subscript letter (A, B, C, etc.) based on the case conceptualization. For instance, $RO2_A$s may include responses to client descriptions of behaviors such as not engaging in avoidance of negative affect outside of session, and $RO2_B$s may be defined as therapist responses to client

descriptions of behaviors that function to allow the client to access important interpersonal social reinforcers with others outside therapy.

Therefore, a **RO2** should be coded using the appropriate subscript for that RO2 (e.g., $RO2_A$) to distinguish the specific RO2s as they occur.

If an RO2 occurs that is not indicated by the conceptualization, the rater should mark that code as $RO2_Z$, where the Z subscript indicates that the response to that "outside" CRB is not part of the conceptualization but is an important RO2, nonetheless.

EXAMPLES

The following are examples of a **Therapist Provides Reinforcement for Client Improvements Outside Session** code (**RO2**):

The therapist provides verbal reinforcement (e.g., praise) to the client when he or she engages in more effective behaviors outside of the therapeutic relationship.
C: I was able to really assert myself with my partner when she told me I was being unreasonable. (O2)
T: That's great. Sounds like you were able to stand up for yourself there. (RO2)

MARGINAL/COUNTER EXAMPLES

C: I was able to really assert myself with my partner when she told me I was being unreasonable. (O2)
T: So you were able to assert yourself? (TPR)

In this example, the therapist has chosen to paraphrase the client's statement, possibly in order to encourage the client to describe more about the interaction, a therapist positive session progression behavior (**TPR**). It is possible, that this type of therapist response will serve to evoke a CRB if the client responds to the question as a lack of support or some other reason. If this occurs, the rater can change the code to **ERB**. In either case the therapist's response is not reinforcing the description of improvements outside the therapeutic relationship (**O2**).

8. **Therapist Positive Session Progression (TPR)**

Definition: Code TPR when the therapist engages in generally effective behavior that **cannot** be coded using one of the specific FAP codes outlined above.
This code is given when the therapist responds to or facilitates talk about problems the client has in other relationships or in other situations outside therapy including discussions about homework.
In the event that a coded response contains both **TPR** and another specific FAP response, the specific FAP response is coded.

EXAMPLES

The following are examples of a **Therapist Positive Session Progression** code **(TPR)**:

The therapist clarifies what client has said.
T: Can you tell me more about what you're feeling now? (TPR)

Therapist goes over **client homework** and client does not emit CRB behavior requiring response by therapist, nor is the client describing clinical improvements or problems in situations outside session that should be followed-up by the therapist.
T: Did you run into any problems when you did the homework assignment we talked about? (TPR)

Therapist encourages or facilitates the client's continued discussion (note: one and two word utterances are not coded in this manual). Therapist uses "microskills" and techniques such as reflecting what client has stated, "mirroring" or paraphrasing client's previous statement, asking open ended questions to facilitate continued discussion.
C: I felt sad when he told me that. (O2)
T: It made you feel sad. (TPR)

The general rule for this code is that the therapist initiates or continues discussion with client about issues outside of therapy provided these behaviors *do not function* to miss a CRB1, 2, or 3 (i.e., not related to the therapeutic relationship or the impact the client has on the therapist). When the therapist responds to client problems outside of session, they are coded using the **RO1** code. When the therapist responds to improvements that occur outside of session and reinforces those, the behaviors are coded using the **RO2** code. *Therefore, not all outside of session talk is coded as TPR. TPR is reserved for behaviors that cannot be coded using other, more specific codes when the therapist is engaging in generally effective responding.*

MARGINAL/COUNTER EXAMPLES

C: When I'm around other people, I get so nervous that I leave the room really fast. (CRB3 - see below)
T: Tell me more about that. (TPR)

The client's behavior in this example could be considered an approximation to a CRB3. In this example the therapist is simply encouraging the client to continue with that approximation, but has not expressly responded to it. This would be coded as a **TPR**, not a **TRB3**. If the rater determined the therapist never responded to the CRB3, he or she could adjust the code to M3 to reflect this therapist error (provided no more than 3 therapist turns have passed).

C: I just can't do this job. (CRB1)
T: Sounds like that job is really hard for you to do right now. (M1 - see below)

This behavior could be considered a CRB1 or CRB2 depending on the client and the corresponding case conceptualization. In this example, assume that the client's behavior was a CRB1, the therapist missed the opportunity to respond to that behavior. In this case the behavior would be coded as a missed CRB1 (M1, described below).
(In the event the client behavior was not coded as a CRB1 or 2, the therapist's response would be coded as **TPR**.)

C: What do you think she meant by that? (CRB1 - see below)
T: Why don't you make a guess about what she meant first. (TRB1 - see below

This example is similar to that given above in that the therapist's response depends on the client's behavior as it is coded. If the client's behavior is coded simply as question asking and not as a specific CRB, then the therapist's behavior would be considered **TPR**. If the client's behavior was considered a CRB1, the therapist's response would be considered a TRB1 because the therapist has prevented the client's problem behavior (i.e., not answering his or her own question) from continuing to occur. If, on the other hand, the client's behavior was understood to be a CRB2 (because the conceptualization

suggests the client is unable/unwilling to ask others for their feedback about interpersonal behavior), then the therapist's behavior would be considered a missed opportunity to respond to a CRB2.

C: Tell me what to do. I just need someone to tell me what to do here. (CRB - see below)
T: It sounds like you need me to tell you what to do. (M1 or M2 - see below)

The therapist's response again is determined by the client's behavior. In this example it is likely the client's behavior, because it is a mand (or command) for a response by the therapist, the therapist's code would be in response to the CRB and would not be considered TPR. The therapist's response here is one of no action. If the client's behavior was either a CRB1 or 2, the therapist's behavior was not in direct response to that behavior and would likely be given a **Missed CRB1** or **2** code (depending on the conceptualization for the client) described below.

INEFFECTIVE THERAPIST BEHAVIORS

Codes 9, 10, and 11 are considered errors of omission as opposed to errors of commission by the therapist (i.e., the therapist fails to make a therapeutic response based on the client's previous behavior.)

9. **Therapist Does not Respond to/Misses CRB1 (M1)**

Definition: Code M1s when the therapist does not respond to or misses an opportunity to respond to a CRB1 (e.g., story-telling by client; changing topic; client avoidance behaviors). This code is also given when a therapist allows discussion to drift away from the therapeutic relationship and relevant issues. Here, the therapist is passively or actively strengthening the client's problematic behavior.

The focus of this code is that the therapist doesn't recognize the occurrence of a CRB1 or gives no overt impression of that recognition either naturally or arbitrarily.

This code is used anytime CRB1s occur and the therapist fails to address this as a problem behavior. There may be times where this miss is intentional. Regardless of the therapist's intent, if the therapist's commission or omission resulted in strengthening the client's problem behavior, it is coded M1. The lone exception is if the therapist makes clear that it is useful to shift the discussion away from the client's ineffective repertoire in session and this shift isn't arbitrary.

This is different than IRB1. IRB1s are when the therapist's behavior indicates that they recognize an occurrence of a CRB1, but their response isn't as effective as it could be.

Distinguish M1s in the same manner that CRB1s and TRB1s are categorized - based on the functional class to which each belongs (e.g., $CRB1_{A,B,C}$s, etc.), the therapist's behavior of not responding to or missing CRB1s should be coded based on the category of client CRB1.

Therefore, when a therapist does not respond to or misses a CRB1, that behavior should be coded using the appropriate subscript for that CRB1 (e.g., $\mathbf{M1_A}$) to distinguish the therapist's responses to specific CRB1s as they occur.

EXAMPLES

The following are examples of a **Therapist Does not Respond to/Missed CRB1** code **(M1)**:

The therapist misses an opportunity to respond to a client's problem behavior in-session. The therapist may respond with some other codeable response, but in the absence of a specifically effective FAP behavior the response is coded as **M1**.

C(angrily): I don't know why I even come to therapy with you. (CRB1)
T: Sounds like you're having a bad week. (M1)

Here, the therapist provides the client with general support and generally effective therapist responding codeable as Positive Session Progression (TPR). However, the client's behavior is a CRB1 (in this example), and should have been contingently responded to by the therapist. In this case, the code M1 (as a specific FAP behavior) takes priority over the code TPR (generally effective therapist responding).

M1 is also applied if the therapist inadvertently supports (provides verbal reinforcement for) client behavior that is a **CRB1** (for example, if the therapist believed the client's behavior to be a **CRB2**, but the rater coded the client behavior as a **CRB1**).

C: I hate talking about this stuff with you. I just hate this. (When this is a CRB1, **not** *a CRB2)*
T: O.K. What would you like to talk about, then? (M1)

Clearly, the therapist has not responded to the client's problem behavior, and instead supports it by providing the natural reinforcer for the client's mand (command) to change the subject.

MARGINAL EXAMPLES

If the therapist simply fails to respond to a **CRB1** in an apparent effort to extinguish the **CRB1** behavior, this will be difficult to code. It is more likely that a **M1** code will be given when it is clear by the rater's judgment that the therapist either should have responded to the client behavior when it occurred, or the therapist simply did not notice the presence of a **CRB1**.

C:. I hate talking about this stuff with you. I just hate this. (as CRB1)
T: Mmm, hmm. Why don't you go ahead and keep talking about it with me, though. (TRB1)

Again, in this example, the client's behavior is taken to be a **CRB1**. The therapist here has chosen not to respond to the client's behavior directly and instead briefly acknowledges it and asks the client to continue. This would result in a **TRB1** code.

COUNTER EXAMPLE

C: I'm thinking I probably won't come back to therapy. (as CRB1)
T: Where is this coming from? I don't understand this at all. (TRB1)

While the therapist is asking a question here, he or she is responding to the problem behavior by asking why the client is saying this in as much as it is having an odd impact on the therapist at the moment. In this case the therapist would be responding to the impact of the **CRB1** and the therapist behavior would be coded as a **TRB1**.

10. **Therapist Does not Respond to/Misses/Stops a CRB2 (M2)**

Definition: Code M2s when the therapist fails to effectively reinforce (naturally or arbitrarily) an instance of a client's CRB2 or a reasonable approximation of a CRB2. This code is also given when the therapist punishes a CRB2 (e.g., disclosing) as it occurs.

The focus of this code is that the therapist doesn't recognize the occurrence of a CRB2 or gives no overt impression of that recognition either naturally or arbitrarily.

This is different than IRB2. IRB2s are when the therapist's behavior indicates that they recognize an occurrence of a client improvement, make a response that is in the class of strengthening the client's improved behavior, but their response isn't the most effective in strengthening client responding.

Distinguish M2s in the same manner that CRB2s and TRB2s are categorized - based on the functional class to which each belongs (e.g., $CRB2_{A,B,C}$s, etc.), the therapist's behavior of not responding to or missing CRB2s should be coded based on the category of client CRB2.

Therefore, if a therapist does not respond to, misses, or stops a CRB2 that behavior should be coded using the appropriate subscript for that CRB2 (e.g., $\mathbf{M2_A}$) to distinguish the therapist's responses to specific CRB2s as they occur.

EXAMPLES

The following are examples of a **Does not Respond to/Misses/Stops a CRB2** code **(M2)**:

The therapist fails to respond to a CRB2 when one is present or while it is occurring.

C: I really need you to listen to me, I need your help. (Where the client requesting assistance is a CRB2).
T: Mm hm. So it sounds like you need my help. (M2)

In this example the therapist is repeating the client's request back to him or her, but, the client had engaged in a clinical improvement in-session and the therapist did not respond to it at all.

MARGINAL EXAMPLES

C: I really felt belittled in the argument I had with my wife. (as a disclosure, in this case a CRB2)
T: That sounds important, but we need to get back on track. (M2 - see below)

This is a difficult type of therapist response because the therapist is technically responding to the client's CRB2, but is quickly shifting topics and is stopping the CRB2 from continuing. In this case the therapist behavior under consideration is predominantly stopping the CRB2 (disclosing) from continuing in an effort to hold to a pre-established agenda. This therapist behavior would be coded as **M2**.

C: I really like talking to you. (CRB2)
T: Mmm hmm. I enjoy talking to you, too. I'm glad you said that. (TRB2)

Here, the therapist begins the interaction with an acknowledgment of what the client has said. The therapist continues with the natural reciprocation and a brief statement of how that impacted the therapist. The therapist's behavior in this case would be coded as **TRB2**. If the therapist had stopped at just the acknowledgment of "Mmm hmm" and the topic of conversation moved to a less relevant area for therapy, the response would be coded **M2**.

COUNTER EXAMPLES

C: [Client continues struggling with some difficult issue during session, experiencing emotional responses in front of the therapist—coded as a CRB2]
T: I really feel closer to you when you show me how you feel like this. (TRB2)

In this example, the therapist is clearly conveying to the client how his or her behavior is impacting the therapist and how that behavior functions to create interpersonal closeness. This would be coded as a **TRB2**.

C: [Client expresses continued emotion disproportionate to the circumstances in a way that functions to distance the client from the therapist—coded as a CRB1]
T: It's hard for me to be here with you right now, when you do this. (TRB1)

Because the client's behavior was a **CRB1**, the therapist's response was appropriate (a **TRB1**). ***If*** the client's behavior had been coded as a **CRB2**, the response to the client's behavior would have been coded as **M2** because the therapist clearly tried to stop that behavior as it occurred.

11. **Therapist Fails to Respond to a CRB3 (M3)**

Definition: Code M3 when the therapist misses an opportunity to respond (i.e., reinforce) or takes an opportunity to punish client responding when a client engages in a description of important controlling variables or a reasonable approximation thereof. This code can also be given when the therapist misses an opportunity to model CRB3 behavior.

The focus of this code is that the therapist doesn't recognize the occurrence of a CRB3 or gives no impression of that recognition either naturally or arbitrarily.

EXAMPLE

The following are examples of a **Therapist Fails to Respond to CRB3** code **(M3)**:

The therapist misses a CRB3 when it occurs or stops one while it is occurring.
C: Like in with Jim, I know, when I get anxious about talking to Jim about how I feel, if I just persist, I can tell him about that, and then I usually end up feeling better. (CRB3)
T: OK, what about in here? (M3)

The therapist in this example has pulled the focus of the discussion on the therapeutic relationship (**TTR**) which is a goal for therapy, but the therapist has missed an opportunity to reinforce the client's **CRB3** when it occurred. This therapist turn would be coded as **M3**.

MARGINAL/COUNTER EXAMPLES

C: When I'm around men, I get so nervous that I leave the room really fast. I never get the chance to really connect with anyone, much less get to know him or her. (CRB3)
T: Sounds like you shouldn't be around men (laughs). (M3 - see below)

While humor is an important part of therapy and a part of naturally responding to a client, in this case the client was engaging in or approximating a **CRB3**. The therapist did not reinforce this response by the

client and stopped it from continuing. This would be coded as **M3** for this therapist response. The therapist may go on in a few turns to support that CRB3 (**TRB3**), but that response would receive a separate code. If after watching the client's behavior after the therapist responds, the rater notices that this response actually functions to have the client continue engaging in CRB3 behavior, the rater could go back and change the code.

C: I wanted to tell you this. I realized that when I'm around men, I get so nervous that I leave the room really fast. (when coded as CRB1 - see below)
T: O.K., that sounds like some important stuff there, but you just changed topics on me. We were talking about how you feel when you see your ex-husband. (TRB1)

In this case, the therapist noticed that the client's approximation to a **CRB3** came as a **CRB1** in that the client changed topics in response to a difficult discussion. Here the client's behavior would be coded as **CRB3**, and the therapist' response as **M3**. If the rater coded the client's behavior as a **CRB1**, the therapist's behavior would be coded as **TRB1**.

12. **Ineffective Response to a CRB1 (IRB1)**

Definition: Code IRB1 when the therapist supplies a contingent response that is clearly less than ideal at achieving the therapist's apparent goal in the interaction (as judged by the coder or by viewing the client's reaction). This code can include examples of when the therapist supplies grossly non-contingent or artificial reinforcement (e.g., gratuitous verbal praise of client behavior), or when the therapist responds to a CRB1 with feedback that could not achieve the goal of reinforcing an improved response by the client.

EXAMPLES

The following are examples of a **Ineffective Response to a CRB1** code **(IRB1)**:

C: I really don't feel like you care for me. (CRB1)
T (condescendingly): There you go again. (IRB1)

Here, the therapist is not supplying feedback effectively that will provide the opportunity for the client to emit a more useful response. Instead the therapist's response is likely more punishing and may temporarily prevent client responding or result in an escalation of CRB1s. This therapist response would receive an **IRB1** code.

MARGINAL EXAMPLE

C: I think I hate coming to therapy. (CRB1)
T: Well, why don't you tell me what you mean by that? (TPR)

This is a more difficult interaction to code, and it would be useful to have access to a larger history of the therapist's behavior to determine whether this is a more extreme response and has a more punitive impact on the client than the therapist's typical responding. (Although, if a therapist were typically providing punitive responses to client behaviors, all such behaviors should receive this code.) However, in this example, the therapist's behavior is more consistent with clarifying the client's statement and would be coded as **TPR** (session progression). In the event that the client persisted in what could be coded in this

case as a **CRB1**, and the therapist continued simply clarifying or asking questions, the therapist's behavior would be coded as **M1** (missing a CRB1).

COUNTER EXAMPLE

C: I just hate everybody right now. Everybody. (CRB1)
T: So, when you say this kind of thing, it's really hard for me to be close to you, and really know how to be here for you. (TRB1)

This is an example of the therapist supplying difficult feedback to the client in a way that simply puts the client in contact with the problematic behavior he or she is displaying in-session. Although this feedback would be difficult for the client to hear, it does not possess the more punitive properties described by this code in the first example above. The interaction just provided would be coded as **TRB1**.

This is an important distinction, while contingent feedback would be difficult to hear for some clients, it is not the straightforwardness of the therapist's responding that will necessitate a **IRB** code. Instead, it is when this behavior is responded to in a manner that is functionally problematic for the therapist, that the **IRB1** code would be used.

13. **Ineffective Response to a CRB2 (IRB2)**

Definition: Code IRB2 when the therapist supplies a contingent response that is clearly ineffective at achieving the therapist's apparent goal in the interaction (as judged by the coder or by viewing the client's reaction). This code can include examples of when the therapist supplies grossly non-contingent or artificial reinforcement (e.g., gratuitous verbal reinforcement of client behavior), or when the therapist responds to a CRB2 with feedback that could not achieve the goal of reinforcing an improved response by the client.

EXAMPLES

The following are examples of an **Ineffective Response to a CRB2** code **(IRB2)**:

C: You know, I feel a lot closer to you. (CRB2)
T: That's great. You said that you feel a lot closer to me. This is a big step for you. (IRB2)

While this type of therapist response is attempting to reinforce a CRB2 (**TRB2**), the therapist is supplying very arbitrary reinforcement that serves largely to distract the client from hearing how that behavior impacted the therapist. This would receive an **IRB2** code.

MARGINAL EXAMPLE

C: I can't meet at our regular time next week. I'd like to meet later next week, though. (CRB2)
T: What is different today that you can ask that? (ERB3)

The therapist could have a myriad of responses that might be useful. Here, they are asking for controlling variables of the client's behavior. If it seemed out of place, it might be considered an **IRB2** or **IRB3**. However, if it works in the interaction to elicit statements of controlling variables, **ERB3** would be the appropriate code.

COUNTER EXAMPLE

C: Every time you ask me how I'm feeling, I panic inside. It's like I need to give you some right answer instead of how I may be feeling. (CRB3)
T: That's a very interesting connection. Could you tell me more about it? (TRB3)

This is an example of the therapist recognizing the client's tacting of the impact of the therapist as a CRB3 and making an attempt to encourage the client's further disclosure coded as a **TRB3**.

14. **Ineffective Response to a CRB3 (IRB3)**

Definition: Code IRB3 when the therapist supplies a response to the client's talk about controlling variables that serves to mitigate the effectiveness of the therapist. This can be in response to a client's CRB3 or could be a poor attempt at modeling controlling variable talk.

EXAMPLES

The following are examples of a **Ineffective Response to a CRB3** code **(IRB3)**:

C: This pattern just puzzles me. (CPR)
T: Well, it seems that there can be plenty of opportunities for certain things to come up which have a strange effect on people sometimes. (IRB3)

Here, the therapist fails to discuss manipulable variables. Instead, they say things with such generalities that no point of intervention can be determined and no useful is provided to or modeled for the client. It would be coded **IRB3**.

MARGINAL EXAMPLE

C: It makes me sad when my wife doesn't consider my opinions. (CRB3)
T: This sounds like what happens with us sometimes. (ERB1)

The client is articulating distinct controlling variables and responses. The therapist is attempting to move the focus of the conversation onto the relationship between the therapist and client. It would be coded **ERB1** if it was eliciting client problem behavior. You might consider **IRB3** if the therapist seems to be getting at the controlling variables of the client's feeling as though their feelings aren't considered by others, but does it too obscurely.

COUNTER EXAMPLE

C: How long should I wait until I decide that my boyfriend is a putz? (CRB1)
T: This sounds like you not wanting to make decisions again. (TRB1)

Here the therapist is recognizing the occurrence of a problematic client behavior that emerges at different points in therapy. The interaction just provided would be coded as **TRB1**.

15. **Generally Ineffective Therapist Responding (IN)**

Definition: Code IN when the therapist engages in generally ineffective behavior that ***cannot*** be coded using one of the specific FAP codes outlined above.
Therapist behaviors rated using this code also include when the therapist engages in more plainly topographical responding that does not achieve its desired function (e.g., rule-following on the part of the therapist, adhering to an agenda, despite the inappropriate conditions for following that rule - perhaps given by a supervisor).
In the event that a coded response contains both **IN** and another specific missed FAP response, the specific missed FAP response is coded (i.e., **M1**, **M2**, **M3**).

EXAMPLES

The following are examples of a **Generally Ineffective Therapist Responding** code **(IN)**:

The therapist engages in responding that is determined to be ineffective for reasons other than missing or not responding to a CRB1 (**M1**), missing or stopping a CRB2 (**M2**), or missing or stopping a CRB3 (**M3**).
C: I just can't be around other people. (CPR)
T: That's really great that you can identify that. (IN)

As discussed above, this is an odd therapeutic response to what could be later shaped into a **CRB3** by the client, but given the (assumed) context in therapy it is coded as **CPR**. This therapist response is coded as IN, as the therapist did not miss a **CRB3** (**M3**), and provided that the client's behavior did not function as a **CRB1**.

The therapist adheres to a prearranged agenda that prevents him or her from attending to in-session client behavior.
C: Well, my dog died, and I broke up with my girlfriend. (CPR)
T: OK, so today we're really going to focus on how you identify what goes on in the environment when you just, how did you put it, "shut down." (IN)

The therapist is moving the session toward discussing CRB3 related issues, but the client has listed several key issues that need to be addressed and discussed before going on, the TRB3 code here is not appropriate given the client's list of important events that should be dealt with in session.

This code is also given when the therapist allows the client to continue with Fill Talk () for too long and is also continuing in Fill Talk () and is not engaging in a productive treatment strategy during session (i.e., therapist is wasting session time)

MARGINAL/COUNTER EXAMPLES

C: I was wondering what you wanted to talk about today. (CRB1 - see below)
T: OK, I see. (M1)

In this example, the therapist has emitted a very brief response that appears to not be contingent on what the client has said. In the event that the conceptualization about the client indicates that the client's response is a **CRB1**, the therapist has responded by not reinforcing that **CRB1**, but has not responded to the impact of that comment, per se. An M1 code would be given because the therapist failed to comment on the impact of this problematic behavior.
If this is simply fill talk, the therapist has not reciprocated with continued fill talk, and if the tone of the therapist's response were difficult to determine or even less cordial, the code would be ineffective therapist responding (**IN**).

Ultimately, it will be the judgment of the rater that will determine the code. It is helpful in these instances to allow any knowledge of the case, the therapist, and the client inform the type of code given. The tone of the therapist's verbal behavior as well as any observable nonverbal behavior (in the event coding is based on videotaped material) could also influence the code. The general guideline to be taken from this example is to let the case conceptualization guide the type of response each has emitted.

MULTIPLE DECISION CODE HIERARCHY

When it appears that a turn contains more than one codeable response by either client or therapist, the following decision hierarchy must be consulted to determine what code should be applied over other possibilities.

CLIENT CODES:

CRB1, CRB2, CRB3 > CTR, O1, O2, CPR

Specific FAP-related client behaviors are coded over less specific (client focus on the therapeutic relationship) or non-FAP behavior (discussion of problems in outside relationships, discussion of improvements in outside relationships, discussion of other problems, question asking/clarification, client fill talk)

CRB2, CRB3 > CRB1

When a client engages in CRB2 or CRB3 behavior (i.e., more effective behaviors), these are coded over a CRB1 if it occurs in the same turn.

CRB2 > CRB3

If a CRB2 and a CRB3 occur in the same turn, a CRB2 (improvement in the context of the therapeutic relationship) is coded over a description of relevant controlling variables.

CTR > O1, O2, CPR

Focus on the therapeutic relationship (CTR) is coded over all other non-FAP specific client behavior

O1, O2 > CPR

Discussion of problems (O1) and improvements (O2) which occur outside of session, that have been a focus of treatment, are coded over general discussion of other problems (CPR), facilitative discussion by the client (CPR), and other non-FAP specific client behaviors

O2 > O1

Improvements that occur outside (O2) the therapeutic relationship are coded over problem behaviors outside the therapeutic relationship (O1)

THERAPIST CODES:

TRB1, TRB2, TRB3 > M1, M2, M3, IRB1, IRB2, IRB3

Specifically effective FAP behavior takes precedent over specifically ineffective FAP behavior

TTR, ERB, TRB1, TRB2, TRB3 > RO1, RO2, TPR

Specifically effective therapist behaviors (General Contingent reinforcement, Focus on therapeutic relationship, Responds to CRB1, Responds to CRB2, Responds to CRB3)

take priority over generally effective therapist behaviors (responds to outside problems, responds to outside improvements, positive therapy progression)

TRB1, TRB2, TRB3 > TTR, ERB

Specific contingent responding (Responds to CRB1, 2, 3, respectively) is coded over general FAP responding (General Contingent reinforcement, Focus on therapeutic relationship, evokes CRBs)

TRB2, TRB3 > TRB1

Responding to a CRB2 or CRB3 is coded over responding to a CRB1 if both behaviors occur in one turn

TRB2 > TRB3

Responding to a CRB2 is coded over responding to a CRB3 if both occur.

ERB > TTR

Evoking a CRB is coded over a therapist focusing on the therapeutic relationship because ERB assumes TTR

M1, M2, M3, IRB1, IRB2, IRB3 > TTR, RO1, RO2, TPR

Specifically ineffective therapist behaviors (Miss CRB1, Miss CRB2, Miss CRB3, ineffective contingent feedback) take priority over generally effective therapist responding codes (GCR, TTR, RO1, RO2, TPR)

M1, M2, M3, IRB1, IRB2, IRB3 > IN

Specifically ineffective therapist behaviors (Miss CRB1, Miss CRB2, Miss CRB3, ineffective contingent feedback) take priority over generally Ineffective Therapist Responding (IN)

IRB1, IRB2, IRB3 > M1, M2, M3

Specifically ineffective delivery of therapist responses takes priority over the failure to recognize the occurrence of therapeutically relevant client responding

IRB2 > IRB1, IRB3

Specifically ineffective delivery of therapist responses to a CRB2 is coded over specifically ineffective delivery of therapist responses to a CRB1 or CRB3 if both occur in the same turn.

IRB1 > IRB3

Specifically ineffective delivery of therapist responses to a CRB1 is coded over specifically ineffective delivery of therapist responses to a CRB3 in the unlikely event both occur.

M2 > M1, M3

Missing or failing to respond to a CRB2 is coded over missing a CRB1 or CRB3 if both occur in the same turn.

M1 > M3

Missing or failing to respond to a CRB1 is coded over failing to respond to a CRB3 in the unlikely event both occur.

References

Callaghan, G. M., Naugle, A. E., & Follette, W. C. (1996). Useful constructions of the client-therapist relationship. *Psychotherapy, 33*, 381-390.

Ferster, C. B. (1967). Arbitrary and natural reinforcement. *The Psychological Record, 22*, 1-16.

Follette, W. C., Naugle, A. E., & Callaghan, G. M. (1996). A radical behavioral understanding of the therapeutic relationship in effecting client change. *Behavior Therapy, 27,* 623-641.

Kohlenberg, R. J., & Tsai, M. (1991). *Functional analytic psychotherapy: Creating intense and curative therapeutic relationships*. New York: Plenum.

Author Contact Information:

Glenn M. Callaghan, PhD.
Department of Psychology
San Jose State University
San Jose CA 95192-0120
Phone: 408.924.5610
Email address for first author:
Glenn.Callaghan@sjsu.edu

William C. Follette, PhD.
Department of Psychology/298
University of Nevada, Reno
Reno, NV 89557-0062
775-682-8699
Follette@unr.edu

The Functional Analytic Psychotherapy Rating Scale (FAPRS): A Behavioral Psychotherapy Coding System

Glenn M. Callaghan William C. Follette, L. E. Ruckstuhl, Jr., & Peter J. N. Linnerooth

Abstract

Many researchers and clinicians believe that the therapeutic relationship is essential in bringing about clinical change. Empirical research to support this contention is scarce in part due to the difficulty of specifying and measuring theoretically derived mechanisms of change and the important dimensions of the client-therapist relationship. Functional Analytic Psychotherapy (FAP; Kohlenberg & Tsai, 1991) is a behavioral treatment that delineates how the therapeutic relationship brings about clinical change in clear and measurable terms. While initial research has been conducted to demonstrating the effectiveness of FAP with different populations, the purported mechanism of clinical change in FAP has not been sufficiently documented. This study describes the creation of a behavioral coding system (the Functional Analytic Psychotherapy Rating Scale; FAPRS) to identify and specify the components believed to be essential in bringing about client behavior change in FAP. Interobserver agreement values indicated moderate to high levels of reliability for the coding system. Implications for future tests of FAP's proposed mechanism of change and the validity of the coding system are discussed.
Keywords: Functional Assessment, Behavioral Therapy, Functional Analytic Psychotherapy, Behavioral Rating Scale, Behavioral Coding

---

Psychotherapy is a set of circumstances that can produce meaningful, lasting and positive behavior change. It is unclear, however, which aspects of many treatments actually cause change. Most therapies assume that the manipulation of certain critical variables in psychotherapy, called mechanisms of change, will reduce client problem behaviors and increase adaptive behaviors. Although mechanisms of change are assumed to operate in the day-to-day practice of psychotherapy, they are rarely studied empirically, and are often left either unspecified or defined in a way that makes detection and measurement of their operation difficult. The client-therapist relationship is one such potential mechanism of change in need of further specification and study.

Psychotherapy has been described as the "systematic use of a human relationship for therapeutic purposes" (Butler and Strupp, 1986, p. 36). Many other therapists and researchers also regard the therapeutic relationship as essential to the process of client change (Beck, Rush, Shaw, & Emery, 1979; Fiedler, 1950; Greenberg, 1994; Hentschel & Bijleveld, 1995; Horvath & Luborsky, 1993; Krasner, 1962, 1963; Lazarus, 1972; Morganstern, 1988; Rimm & Masters, 1974; Rogers, 1957; Rosenfarb, 1992; Schaap, Bennun, Schindler, & Hoodguin, 1993; Wright & Davis, 1994). There is an important gap, however, between assumptions about the importance of the therapeutic relationship and empirical documentation of the events that change client behavior.

To empirically establish a link between the therapeutic relationship and client behavior change, psychotherapy researchers must accomplish several tasks. First, a treatment must be identified and described. Second, the hypothesized mechanism of change for that treatment must be specified in a way that is clear and measurable. Third, the dimensions of behavior that are hypothesized to change following the manipulation of the mechanism must be defined. Fourth, a measurement system that allows reliable observation of behavior occurring within that treatment must be designed. Fifth, the validity of the hypothesized mechanism of change must be tested by empirically linking the manipulation of that

mechanism to client improvement. This improvement should be documented within the therapeutic relationship, and ultimately it should be shown to generalize to other relationships the client has outside of therapy.

This article presents a method for accomplishing the goals described above for Functional Analytic Psychotherapy (FAP; Kohlenberg & Tsai, 1991) an interpersonally oriented, behavioral psychotherapy. It defines FAP's mechanism of change (therapist contingent responding; discussed below) and the dimensions of client behavior that are impacted by the therapist's contingent responses in-session. It describes the development of and presents data on a reliable coding system for FAP (the Functional Analytic Psychotherapy Rating Scale; FAPRS; Callaghan, 1998). Finally, it presents initial data testing the validity of FAP's hypothesized mechanism of change.

The primary purpose of this study was to create and evaluate a coding system that could reliably demonstrate FAP's mechanism of change (and the client behaviors hypothesized to be impacted by it). In the next section, the important dimensions of the therapeutic relationship in FAP will be introduced and FAP's mechanism of change, therapist contingent responding, will be defined.

Functional Analytic Psychotherapy

Functional Analytic Psychotherapy (FAP; Kohlenberg & Tsai, 1991; see also Kohlenberg, Hayes & Tsai, 1993) is used primarily to treat psychological disorders that are interpersonally based, such as those with pervasive repertoire deficits including generalized anxiety disorder, dysthymia, social phobia, and the Axis-II personality disorders (American Psychiatric Association, 1994). FAP assumes that the interpersonal problems that clients have in relationships outside of therapy will also occur in the context of the therapeutic relationship. The FAP therapist focuses on creating clinical change in this context. The therapist can influence client problem behaviors and improvements because he or she has direct access to the contingencies that occur in-session. Not all behaviors occur in the therapeutic relationship exactly as they do outside of session. However, the FAP therapist attempts to understand these behaviors principally, or functionally, and looks for opportunities to respond to them in-session, rather than dealing exclusively with events that happened outside of therapy.

Although researchers have only just begun to demonstrate the effectiveness of FAP in treating different problems (Gaynor, 2002; Kohlenberg, et al., 2004), treatments for depression (Kohlenberg, Kanter, Bolling, Parker, & Tsai, 2002), adolescent behavioral problems (Gaynor & Lawrence, 2002), personality disorders (Callaghan, Summers, & Weidman, 2003), and the use of FAP as an adjunctive treatment for post traumatic stress disorder (Prins & Callaghan, 2003) have already been undertaken. Data from these studies indicate mounting evidence for the effectiveness of this treatment with diverse populations and a strong need to attempt to document the mechanism of change in FAP.

Mechanism of Change

The conceptual literature on FAP has identified therapist contingent responding, a specifiable and measurable aspect of the client-therapist relationship, as FAP's mechanism of client change. Therapist contingent responses occur contingent on clients' in-session, clinically relevant behaviors (Callaghan, 1996; Callaghan, Naugle, & Follette, 1996; Follette, Naugle, & Callaghan, 1996; Kohlenberg & Tsai, 1995). Clinically relevant behaviors are clients' specified problem behaviors, targets for improvement, and descriptions of functional relationships between their own behavior and its antecedents and consequences and are described in detail below.

Clinically Relevant Behaviors

In order to understand what maintains the problems that brought the client to therapy and to plan an effective treatment, FAP therapists conceptualize clients' interpersonal behavioral repertoires in the context of the developing therapeutic relationship. FAP therapists look for and respond to three types of clinically relevant behaviors (CRBs) that occur in therapy. In general, clinically relevant behaviors can be described as behaviors that occur within the context of the therapeutic relationship that have an identifiable (desirable or undesirable) function on the therapist.

Clinically Relevant Behavior 1. Client problem behaviors, CRB1s, are the responses that clients emit in-session that contribute to the client's problems in living outside of treatment. CRB1s are often behavioral excesses that aversively stimulate persons in the client's life, but may also be behavioral or motivational deficits that negatively impact the client's social relationships. For example, a FAP client may present for therapy complaining that he or she would like to develop closer, more meaningful interpersonal relationships, but fears doing so because he or she may be rejected and feel hurt. A CRB1 for such a client would be any escape or avoidance behavior in the therapeutic relationship that prevents the client from developing a better interpersonal relationship with the therapist.

Clinically Relevant Behavior 2. Client improvements, CRB2s, are behaviors that occur in the context of the therapeutic relationship that are effective for the client given his or her goals for therapy. Initially, many FAP clients have repertoire or motivational deficits in areas important to good interpersonal relationships. The FAP therapist's job is to help the client construct an effective repertoire starting with the client's existing skills (e.g., Goldiamond, 1974). Continuing the above example, any client behavior that serves to create a meaningful, close relationship with the therapist, such as appropriate disclosure or displaying affect that feels "risky" to the client, would be identified as a CRB2 for the client.

Clinically Relevant Behavior 3. Client verbal descriptions of controlling variables, or CRB3s, are statements of the functional relationship between the client's behavior and the social-environmental antecedents and consequences of that response. Continuing with the above example, the client might say to the therapist, "The way that you talk to me shows that you care, even when you don't like what I've just done or said. This tells me I can take risks, or even fail miserably. You won't stop caring about me, even though I sometimes do things that are insensitive to you. That's why I like talking to you." Here, the client has emitted a CRB3 because he or she has described the function of an antecedent, the therapist's history of responding in difficult discussions, on his or her own current, private evaluation of the therapeutic relationship. The client has specified the conditions that make it more likely he or she will engage in a particular response with the therapist.

Functional case conceptualization. CRBs are defined according to their function on the therapist, and are formally organized in a functional case conceptualization based on behavioral principles. The case conceptualization describes the current interpersonal repertoire of the client as it is emitted within the therapeutic relationship. The conceptualization is used to guide an idiographic assessment of client behavior and corresponding intervention strategies. Functional case conceptualizations are iterative. They are updated as new client target behaviors are identified and as behavioral improvements are established.

The FAP coding system to be described below was designed to recognize the function of the behavior of both members of the therapeutic dyad. It was also designed to incorporate important contextual variables, and to draw functional distinctions among therapists' responses to client CRBs. In the

following section we will expand on our definition of therapist contingent responding and reveal some of the therapeutic process of FAP by presenting how FAP therapists decrease problem behaviors and shape or maintain improvements.

In-Session Contingent Responding in FAP

Therapist contingent responding is the contingent application of the therapist's natural (Ferster, 1967, 1972) reaction to the client's behavior as it occurs in session. Contingent responses are delivered based on how the client's behavior actually impacts the therapist (negatively or positively), not simply on the appearance of the behavior. When a FAP therapist detects the occurrence of CRBs in session, he or she seeks to make an immediate, contingent, and appropriate therapeutic response in order to impact the behavior as it occurs. This may include responses that appropriately decrease the frequency or intensity of, or extinguish altogether, problem behaviors. These responses are then combined with therapist behaviors that evoke and reinforce client improvements. The FAP therapist's goal is to affect the frequency and shape the function of the client's in-session behavior in ways that will improve both the therapeutic relationship and generalize to serve the client's goals in outside relationships.

Decreasing problem behaviors. To decrease problematic in session behavior (CRB1), FAP therapists can ignore the maladaptive response, in which case the absence of reinforcement leads to extinction of the problem behavior, or the response can be contingently punished. Examples of contingent punishers include explicitly stating the negative impact that the client's behavior has on the therapist in that moment during session. In the coding system to be described, such a response (an explicit therapist behavior that functions to decrease the future probability of CRB1) is called a TCRB1.

Shaping improvements. Extinction or punishment of maladaptive behaviors is never used alone. Following their contingent response to a CRB1, FAP therapists immediately attempt to construct adaptive, alternative behaviors (CRB2s) that will be more useful for the client (e.g., Hawkins, 1986). Immediately evoking and reinforcing a client CRB2 are particularly salient in the moments following the therapist's response to a CRB1. Thus, FAP therapists constantly seek to detect or evoke and then differentially reinforce adaptive alternative client behaviors. Consistent, differential reinforcement of alternatives to CRB1s allows the FAP therapist to shape functional improvements (CRB2s) in the client's repertoire, rather than just decrease the frequency of symptoms or maladaptive social behaviors.

Drawing functional connections. In the moments following a CRB1, the therapist may also choose to model a description of controlling variables (CRB3) for the client's behavior by pointing out the relationship between maladaptive behaviors in session, their function on the therapist, and their relation to aversive consequences that the client has experienced in other relationships. Client behaviors that evoke therapist escape, avoidance or countercontrol are likely to have the same function on significant others in the client's natural social environment. Although FAP therapists take as their first priority the direct training of improvements (evoking and reinforcing CRB2), drawing functional connections between client behavior in- and out of session is an important element of FAP.

FAP does not attend solely to client problem behaviors, psychological symptoms and distress in therapy sessions. FAP's focus on the development of a prosocial client repertoire marks it as a constructionist approach to therapy. FAP builds on existing client repertoires to enhance functioning, and FAP therapists explicitly conceptualize client strengths throughout treatment. Evoking and differentially reinforcing approximations toward more effective responding enhances the client's interpersonal repertoire and the therapeutic relationship. This two-pronged approach to therapist contingent responding is FAP's mechanism of client behavior change.

Initial Research on the Mechanism of Change in FAP

The purpose of this study was to begin the empirical documentation of FAP's mechanism of change. The study was thus designed to accomplish two goals: (1) to create a coding system capable of documenting both client and therapist behavior in FAP sessions and; (2) to demonstrate that observers using the coding system can reliably track the occurrence of client and therapist behavior in session. Most importantly this study sought to reliably document the occurrence of CRBs and therapist contingent responses to CRBs. Because reliability is a necessary precursor to validity in any measurement system, the data presented in this study represent a crucial first step toward conducting tests of the model of change proposed for FAP and the implementation of efficacy and effectiveness trials. Because there was no coding manual for FAP prior to this research, the development of the Functional Analytic Psychotherapy Rating Scale (FAPRS) will be described. Data demonstrating the reliability of the FAPRS system, using measures of intra- and interobserver agreement, will then be presented.

Functional Coding with the FAPRS

The FAPRS codes the function of client or therapist behavior rather than the form. Raters using the FAPRS must recognize the function of in session behavior by its effect on the other member of the client-therapist dyad, rather than simply noting the occurrence of a particular topography. As a simple example, two clients might ask their therapist "What did you do over the weekend?" Even though the same question is asked, it could not be coded topographically because an identical question would have a different function depending upon the conceptualization of the client asking it. For a client who has had difficulty showing caring within interpersonal relationships, asking such a question would be an improvement (CRB2). In another instance, however, the question may be representative of a client problem (CRB1) that inhibits the development of the therapeutic relationship and client progress by allowing the client to forestall, avoid, or escape discussing an important event that occurred during the time between sessions with the therapist. When using the FAPRS a client's responses is coded based on the function the behavior has on the therapist, given the client's specific case conceptualization.

As with the client's behavior, the therapist's responses are coded based on the function or impact the behavior has on the client. The therapist's responses are not coded based on how the behavior appears (e.g., making an "I-statement"), nor are they coded based on what the rater believes the therapist's intentions were when making the response. Effectiveness is determined by the function of each therapist response on the client given the case conceptualization. Therapist responses that reinforce or shape the quality of a client's ability to show caring or willingness to share reports of important in-session events would be coded as effective responses by the FAPRS. Missing a client's approximation to positive change or allowing a client to avoid sharing important information would both be coded as ineffective therapist responses.

RESEARCH DESIGN AND METHODS

The FAPRS Coding Manual

Development of Observational Codes and the FAPRS Manual

Reliable assessment of both members of a therapeutic dyad's behavior is a necessary foundation for future studies designed to determine the empirical relationship between specific types of therapist responding and client behavior change. It is also necessary to document FAP's mechanism of change. The FAPRS thus includes both client and therapist behavior codes. The FAPRS was created by extending a set of codes designed to identify therapist behavior employed in a previous empirical study (Callaghan & Follette, 1999).

Instructions for raters. The Functional Analytic Psychotherapy Rating Scale (FAPRS) employs the speaking "turn" (also called a "floor change") as its unit of analysis. Raters are instructed in procedures for turn-by-turn coding of sessions using session transcripts and videotapes. Raters are allowed to apply only one code to each turn and are provided with a multiple code decision hierarchy for use in situations where more than one codeable event occurs within a turn. The FAPRS gives raters examples, counter examples, and marginal examples for each code. Raters are instructed to code every turn based on the approximate function of the behavior in that turn given the case conceptualization for each client.

Client codes. The client codes used for this study are presented in Table 1. This set of codes indicates whether the client is engaging in the FAP-specific, clinically relevant behaviors (CRB1, CRB2, or CRB3) versus behaviors likely to occur in other therapies (CTR, O1, O2, CPR). For example, if the client talks about the therapeutic relationship, but this behavior is not a specific problem or improvement (CRB), it is coded using the Client Focus on the Therapeutic Relationship (CTR) code. When clients emit responses that are not improvements or problem behaviors (in or outside of session) but are part of general therapy

Table 1. List of Client Codes in FAPRS

Code	Full Name of Code	Brief Description
CRB1	Clinically Relevant Behavior 1 (problems in session)	client engages in problematic behavior in-session in the context of the therapeutic relationship
CRB2	Clinically Relevant Behavior 2 (improvements in session)	client engages in improved behavior in-session in the context of the therapeutic relationship
CRB3	Clinically Relevant Behavior 3 (description of important controlling variables)	client describes how different controlling variables impact the client's behavior and when the client makes these functional descriptions
CTR	Client Focus on the Therapeutic Relationship	client focuses on the therapeutic relationship
O1	Discussion of Clinical Problems Outside the Therapeutic Relationship ("outside CRB1s")	client discusses or describes problem behaviors that have been the focus of treatment but that occur in other situations outside of session
O2	Discussion of Clinical Improvements Outside the Therapeutic Relationship ("outside CRB2s")	client discusses or describes improvements that have been the focus of treatment but that occur in other situations outside of session
CPR	Client Positive Session Progression	client discusses or describes problems as they occur in situations other than the therapeutic relationship, or clarifies or provides context about problems; client engages in generally on-task behavior that is facilitative to discussion

progression, the turn is coded as "Client Session Progression" (CPR). When the client tells the therapist about an occurrence of a clinically relevant behavior outside of the therapy session (i.e. behavior that is a target for therapeutic change), it is coded as an "Outside CRB1" (O1) or "Outside CRB2" (O2). O1s and O2s are not CRBs because they do not occur within the therapeutic relationship.

FAP assumes that problems and improvements occurring between the client and therapist are directly relevant to client problems and improvements in relationships outside of therapy. O1s and O2s are important to document because their occurrence indicates the extent to which the client's behavior targeted for change in session generalizes outside the therapeutic relationship.

Therapist codes. The therapist codes are presented in Table 2. The FAP-specific therapist behavior codes (TCRB1, -2 and -3, ECRB, M1, -2, and -3) are designed to document therapists' contingent responses to clients' in-session, clinically relevant behaviors. TCRB1, TCRB2 and TCRB3 indicate effective therapist responses to client in-session CRBs 1, 2 and 3 respectively. Ineffective therapist behaviors are noted with the codes M1, M2 and M3, indicating that the therapist has "missed" a CRB occurring in-session. Misses may be due to an inability to discriminate or failure to respond to a CRB. Therapist behavior that brings a

Table 2. ***List of Therapist Codes in FAPRS***

Code	Full Name of Code	Brief Description
TTR	Therapist Focus on Therapeutic Relationship	therapist continues focus on therapeutic relationship including sharing the therapist's feelings in response to the client
ECRB	Therapist Evokes a CRB	therapist evokes a clinically relevant behavior by the client, either CRB1, 2, or 3
TCRB1	Therapist Responds Effectively to CRB1	therapist's response is to in-session client problem behavior
TCRB2	Therapist Responds Effectively to CRB2	therapist responds effectively to in-session improvements
TCRB3	Therapist Responds Effectively to CRB3	therapist responds to the client describing how different controlling variables impact the client's behavior; therapist shapes or models CRB3 for client
RO1	Therapist Responds to Client's Discussion of Clinical Problems Outside the Therapeutic Relationship	therapist comments on problem behaviors the client describes having engaged in outside of the therapy session that are a focus of therapy
RO2	Therapist Responds to Client's Discussion of Clinical Improvements Outside the Therapeutic Relationship	therapist provides verbal reinforcement in response to the client describing improved behaviors outside of the therapy session that are a focus of therapy
TPR	Therapist Positive Session Progression	therapist engages in generally effective or facilitative behavior
M1	Therapist Misses/Does not Respond to CRB1	therapist does not respond to or misses an opportunity to respond to a CRB1
M2	Therapist Misses/Does not Respond to CRB2	therapist fails to reinforce an instance of a client's CRB2 or a reasonable approximation of a CRB2
M3	Therapist Misses/Does not Respond to CRB3	therapist misses an opportunity to respond to a client's description of important controlling variables or reasonable approximation
IN	Generally Ineffective	therapist engages in generally ineffective behavior

	Therapist Responding	

CRB into the session is noted with the code, Evokes CRB (ECRB). The codes RO1 and RO2 denote effective responses to client descriptions of problem behaviors or improvements occurring outside of the therapeutic relationship. TPR indicates general therapist progression. TTR denotes clinician focus on the therapeutic relationship that is not in response to a targeted CRB, such as describing the process of FAP to the client. Ineffective contingent responding (ICR) and generally ineffective therapist behavior (IN) codes capture therapist behavior that is not beneficial (for example, inappropriately arbitrary reinforcers or punishers) to the client given the goals of treatment using FAP.

Reliability Statistic

Interobserver agreement was determined for all reliability analyses using Cohen's (1960) kappa. The kappa statistic is a measure of inter-rater reliability that improves on percent agreement by correcting for agreement that occurs by chance due to high base rates of some codes compared to others. Fleiss's (1981) guidelines for interpreting kappa were used to judge obtained coefficients as "poor" (less than .40), "fair" (.40 to .60), "good" (.60 to .75), or "excellent" (above .75) for all analyses presented below. A kappa value of .60 or higher (at least "good") was chosen to define the minimum acceptable level of rater competence using the FAPRS.

Participants

In the following discussion participants will be named and numbered to differentiate the Clients (1, 2, 3) Therapists (1, 2) and Raters (1-6). Client-therapist pairs will be referred to as "Dyads" and labeled alphanumerically (A, B, C). The term "Cohort" will refer to the groups (1, 2, 3) of raters trained together. Cohorts are distinguished because not all raters were trained at the same time.

Clients

The 3 clients in this study were adult women with long standing problems across multiple interpersonal relationships. The clients sought psychotherapy services in university clinics and provided informed consent for psychological treatment and for the use of videotapes and other session materials for research.

Therapists

Therapist 1 was one of the co-originators of FAP and was considered an expert in conducting the treatment. Therapist 1 conducted FAP with 2 of the 3 clients (Client 1 and Client 2). Therapist 2 was a doctoral student in clinical psychology with 3 years of experience in FAP and conducted treatment with the third client (Client 3).

Raters

Rater selection. All raters (1-6) are numbered in descending order according to their experience level both in conducting FAP and in using the coding manual. Only raters familiar with the basic principles of FAP were selected to code tapes for the study. This decision was based on the assumption that the functional nature of the coding system requires raters to understand the behavior analytic distinction between how client or therapist responses function as opposed to the superficial appearance, or topography, they may have.

Criterion rater. Rater 1 was an expert FAP therapist and developed the FAPRS. Rater 1 was thus designated the "criterion" rater for this study. The performance of Raters 2-6 was compared to Rater 1's in some of the reliability analyses to assess their expertise in the use of the FAPRS.

Rater training. All raters received 70-100 hours of training. Training began with individual study of Kohlenberg and Tsai's (1991) text, followed by lectures and discussions on the principles of FAP and behavior analysis in a small group format. Raters then read the FAPRS manual and were instructed in the definition and use of its codes in a small group lecture and discussion format. Raters viewed segments of a FAP therapy session previously coded by Rater 1 and discussed the rationale for the codes applied to each client or therapist turn. In the next stage of training, raters were taught to code videotape segments taken from Dyad A. In each trial the raters observed, coded, and then discussed in small groups each code they applied as well as the rationale for their choice. The number of turns coded prior to discussion was progressively increased across the training from one (turn-by-turn discussion and instruction) to independent coding of an entire therapy session. Across all stages of training, informal quizzes were given to the raters to test their knowledge of the coding manual, their ability to discriminate between codes, and their understanding of the client's functional case conceptualization prior to each training session. No data from these training and discussion sessions are included in the reliability analyses presented in this study.

Rater training cohorts. Raters 2-6 were trained in three "cohorts." Rater 1 trained Rater 2 in the first Cohort. Raters 1 and 2 jointly trained Cohort 2 (Raters 3 and 4). Rater 2 alone trained Cohort 3 (Raters 5 and 6). The goal of using two different trainers was to ensure that effective use of the FAPRS was not dependent upon being trained by the FAPRS' author and to demonstrate the transfer of FAPRS instruction across trainers. This training approach provided a "layering" of training and allowed differences in coding skills to be examined within and across rater training cohorts.

Criterion test. At the conclusion of training, each rater was asked to independently code a "criterion session" tape. All raters coded the same criterion session tape. The criterion session was used to determine each rater's accuracy in meeting FAPRS training goals by comparing his or her performance to Rater 1. If a rater reached the overall (entire session) kappa value set at .60, he or she was allowed to begin independently coding tapes included in the data analyses.

Selection of criterion session. The session used for the criterion test was selected because it demonstrated as many of the FAPRS codes as possible within a single therapy session. This session was coded in its entirety. Raters were not allowed to code with others present, nor were they allowed to discuss their coding of this session.

Apparatus

Videotaped Stimulus Materials

Videotape recordings of FAP treatment sessions were examined for this study. The therapy sessions occurred in a room with two comfortable chairs. The video camera was positioned behind the therapist and was focused on the client.

Selection of videotaped sessions to be coded. A sample of sessions from the entire course of FAP treatment was selected in order to rate approximately every fourth therapy session. When a particular session was of unacceptable quality (e.g., the sound was inaudible or the video was non-viewable), the session occurring one week prior to or after the targeted session was chosen at random.

This substitution was made if the camera's view of the client was obscured, or when technical problems led to poor audio or video quality.

Transcripts

A person not involved in coding the tapes transcribed each of the sessions used in the analyses. Transcripts provided a close approximation of a verbatim record of each client and therapist "turn" in the treatment session. Transcripts were completed using Mergenthaller and Stinson's (1992) standards for transcription with modifications as described below.

Embedding of turns. Standards that were modified included "embedding" one and two word utterances in the previous speaker's turn. For example, the client might report, "I became extremely angry with my spouse this week," and the therapist simply responded, "Really?" then the turn would be embedded in the previous (client's) turn and would not be coded. One- and two-word utterances were embedded because they often function only to facilitate the previous speaker's verbal behavior. Utterances were embedded only when they were clearly not functional units, such as the answer to a question or a verbal reinforcer or punisher of the previous speaker's behavior. Typical utterances there were not functional units and were embedded included turns such as, "Uhh-huh" or "Umm-hmmm" and "I see" or "OK."

Procedures

Coding Procedures

For all reliability analyses, sessions were coded in a randomized order. No identifying information was included on any transcript that would indicate the order in which sessions progressed, or the point in therapy (e.g. early, middle, late) that was depicted for each dyad. The primary advantage to randomizing session order is that it prevents rater bias due to raters' assumptions about how the therapist or client's behavior should change over time.

The decision to code sessions in a randomized order was conservative, as it would tend to lower estimates of the reliability of the FAPRS. Lacking any information as to the point at which the session occurred during the course of treatment forced raters to code client and therapist behaviors based only upon their apparent function at that moment of the session. When coding sessions out of order, it is difficult for raters to determine whether a particular client behavior functioned as an improvement or problem. Raters could not invoke rules developed on the basis of previously coded sessions because the rater's past experience might have included interactions from later sessions in therapy. Although this was a methodologically sound decision given the development of the FAPRS to prevent bias in coding, this approach made some sessions difficult to code.

Codes not included in the analysis. The M3 (therapist misses/does not respond to a CRB3) code was never used and the ICR (therapist ineffective contingent responding and IN (Generally Ineffective Therapist Responding) codes occurred only a few times. Thus, no measure of reliability was calculable for these codes.

Embedding check. Approximately 25 percent of the Dyad A tapes coded for the study were selected and checked for appropriateness of embedding by Raters 1 and 2. A list of embedding rules was established prior to this check and was used to monitor whether phrases were embedded consistently across sessions. A kappa value of K = .91 between raters 1 and 2 demonstrated excellent agreement (Fleiss, 1981) between raters 1 and 2 for embedding one and two word facilitative utterances.

FAP Adherence Check

Transcripts from Dyad A were also coded using the Therapist In-vivo Strategies Scale (THISS; Parker, Bolling, & Kohlenberg, 1996) to demonstrate that these sessions were indeed FAP session and not cognitive therapy. The THISS system is a topographical measure of adherence to FAP strategies. It is not a functional measure of FAP competence and was not designed to measure the functional mechanism of change discussed above. The adherence check was performed, before testing the reliability of the FAPRS, to show that FAP therapist behavior in-session can be distinguished from therapist in session behavior in Cognitive Therapy (CT). The adherence check examined therapist use of Cognitive Therapy (CT) and/or Functional Analytic Psychotherapy (FAP) strategies within each session to ensure that FAP sessions contained behaviors specific to FAP, and more therapist strategies specific to FAP than to CT.

Reliability Drift Checks

Reliability decay, or "drift" occurs when a given rater's reliability to a criterion rater decreases over time (Bakeman & Gottman, 1986). This issue is especially important when a series of interactions are coded, such as the long series of Dyad A tapes coded by Rater 2. To assess the potential for drift using the FAPRS, interobserver agreement "drift checks" between Raters 1 and 2 were performed using coded sessions from Dyad A. Insufficient agreement was defined as overall kappa values below K = .5 ("fair" agreement; Fleiss, 1981). Failure to maintain criterion at any point would lead to "booster" training sessions until the rater obtained sufficient agreement values. Rater 1 checked 25% of the Dyad A tapes coded by rater 2 for decay.

RESULTS

Adherence Check Using the THISS Scale

Ratings for FAP-related strategies as measured by the THISS scale were significantly higher (m = 1.83 ; sd = .60) than for CT strategies (m = 1.38 ; sd = .20), $t(15) = -2.45$, $p = .027$, across the 16 Dyad A videotapes used in this study. The mean interobserver reliability for the THISS scale was calculated with Shrout and Fleiss' model (1979) for averaged ratings across a group, yielding consistently high rates of reliability between raters (mean ICC = .94, range = .88 to .98).

Reliability of the FAPRS

Four types of reliability assessment were conducted on the FAPRS coding system. All reliability data for this study are presented in the form of kappa coefficient (Cohen, 1960) values obtained between pairs of raters. The "generalized" kappa formula (which yields a single coefficient for three or more raters) was not used in this study. The authors decided to present only kappas obtained between pairs of raters because the generalized kappa is difficult to interpret. The generalized formula for kappa was judged inappropriate for this study because it cannot be directly compared to kappas obtained between pairs of raters. Because kappa values cannot be averaged across sessions, dyads, or raters, "overall" kappa values were used in this study. The term "overall" kappa indicates that all turns from multiple therapy sessions for a dyad were entered into a single, large kappa matrix for each rater.

Agreement with the criterion rater. Table 3 presents the kappa coefficients obtained between each pair of raters and between each rater and the criterion rater (Rater 1) for the criterion session. All 5 raters reached or exceeded the criterion of k = .60 with the criterion rater. Two of these 5 raters reached excellent agreement with the criterion rater. The remaining 3 raters achieved a good level of agreement (Fleiss, 1981).

Table 3. ***One Session Reliability to Criterion Tests for Dyad A.***

	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5
Rater 2	.83				
Rater 3	.64	.56			
Rater 4	.65	.61	.79		
Rater 5	.74	.77	.54	.56	
Rater 6	.62	.62	.48	.49	.71

Agreement between all other rater pairs. Six of the 10 possible pairs of raters (not including the criterion rater) reached or exceeded the criterion kappa of .60 with one another. Two rater pairs obtained excellent agreement. Good agreement was obtained by 3 of 10 rater pairs. Fair agreement was obtained by the remaining 5 rater pairs.

Interrater reliability across multiple sessions. Because kappa coefficients cannot be averaged across dyads, a separate matrix of interobserver agreement values is presented for Dyads A and B, and presented textually for Dyad C below. Table 4 presents the overall kappa coefficients obtained between Raters 2, 3, and 4 across three sessions of Dyad A. Table 5 presents the kappa coefficients obtained between Raters 1, 3 and 4 across three sessions of Dyad B.

Table 4. ***Interrater Agreement Matrix for Multiple Sessions of Dyad A.***

	Rater 2	Rater 3
Rater 3	.60	
Rater 4	.65	.70

Table 5. ***Interrater Agreement Matrix for Multiple Sessions of Dyad B.***

	Rater 1	Rater 2	Rater 3
Rater 2	.77		
Rater 3	.70	.75	
Rater 4	.58	.56	.61

The kappa coefficient obtained between Raters 1 and 3 across three sessions of Dyad C was .61. Of the 10 possible rater pairs in these analyses, 2 achieved excellent agreement, 7 achieved good agreement, and 1 achieved fair agreement.

Intrarater reliability. To test the intrarater reliability of the FAPRS coding system, Rater 1 coded three sessions of Dyad A at two different points in time. The overall kappa coefficient obtained by Rater 1 across an 18-month code-recode interval was .91.

Reliability drift checks. Table 6 presents the kappa values obtained between Raters 1 and 2 for each of 7 reliability drift check sessions of Dyad A, and the overall kappa coefficient across the 7 sessions. None of the sessions ever decreased below a "good" level of agreement.

Table 6. ***Kappa Values for 7 Drift-check Sessions of Dyad A.***

Session	1	2	3	4	5	6	7	Overall
Kappa	.75	.87	.68	.64	.64	.60	.61	.72

DISCUSSION

Reliability of the Coding Manual

Interrater agreement within dyad. The kappa coefficients between each rater and the criterion rater presented in Table 3 suggest that raters with varying levels of experience can reach at least a good level of agreement with the author of the manual, an expert FAP therapist. The kappas obtained between all possible rater pairs indicate that FAP raters can be expected to agree with each other as well as with a criterion rater. This suggests that all raters learned a set of skills for coding FAP, rather than just the ability to code reliably with the person who trained them.

Table 3 indicates that there may have been a rater training cohort effect. Raters 3 and 4 showed modally good agreement with Raters 1 and 2 who trained them, and excellent agreement with each other. Raters 4 and 5 showed good to excellent agreement with Rater 2 who trained them, and good agreement with one another. When compared to one another according to training cohort, however, Raters 3 and 4 and Raters 5 and 6 uniformly achieved only fair agreement. Interrater agreement is somewhat higher within than between cohorts, and some rater pairs agree better than others. It is possible that raters, across training cohorts, were making different types of coding decisions and errors. All raters, however, remained within an acceptable range of kappa values given the difficulties of functional coding and the fact that the FAPRS is still in development. The possible cohort effect will be addressed in future studies by using a standardized, computer-based training curriculum for training new raters and for retraining experienced raters to a common criterion (Ruckstuhl, Callaghan, & Follette, 2001).

Interrater agreement across multiple sessions. The across-sessions analyses in Tables 4 and 5 indicate that good levels of agreement can be consistently attained using the FAPRS across different sessions for the same client-therapist dyad. The data indicate that raters can sustain these good to excellent levels of agreement across several sessions. These data also show that raters can achieve high levels of agreement for different client-therapists dyads; that is, rater agreement levels generalize across dyads when coding different clients and different therapists.

Assessment of the reliability of the FAPRS for measuring client and therapist behaviors across multiple dyads is challenging and important. The FAPRS requires raters to discriminate and correctly categorize both client and therapist behaviors based upon the approximate function that each dyad member's behavior has on the other person. This can be a difficult task, as each dyad has a topographically different means of creating a common function within the therapeutic relationship. Topography can also change within a particular dyad across time. While two FAP therapists might identify "trusting the therapist" or "feeling close to the therapist" as CRB2s in the functional case conceptualization for two different clients, the way that each client's CRB2 looks is typically quite different, making coding across clients difficult.

The fact that kappa values from Dyads B (Table 5) and C (kappa = .61) matched or exceeded those of Dyad A (Table 4) suggests that using sessions from Dyad A to train the raters did not confound the results for the reliability analyses. Future studies of the FAPRS will train raters with therapy examples from multiple dyads, and to the greatest extent possible, use training materials taken from dyads not included in the data set of the study for which raters are being trained. Given the status of FAP as a relatively new psychotherapy practiced by a limited number of clinicians, it was necessary to train raters using certain dyads. It is encouraging that interobserver agreement rates in this study did not appear to be bound by rater training that used one of the client-therapist dyads included the primary data analysis.

The data provided for the intrarater reliability check suggest good test-retest reliability for the FAPRS. These data should be interpreted cautiously. The intrarater kappa coefficient represents the coding of one rater (the criterion rater) examining only one client-therapist pair. These data are presented here because they suggest that raters may be able to use the FAPRS consistently across time. This is a promising finding given the length (18 months) of the code-recode interval used and the challenges imposed on the rater by the complexity and subtlety of the behaviors coded.

The data in Table 6 for the drift check analysis indicate that two raters can use the FAPRS, across multiple sessions with good overall reliability. Overall, Rater 2 was able to sustain sufficient agreement with the criterion rater across the sessions used in the data analyses. The session-by-session kappa values indicate that reliability was subject to slight decay. From these tables, we conclude that ongoing training in conjunction with reliability checks may be necessary, and that future studies of the FAPRS should continue to include both session-by-session and overall kappas when examining rater drift.

Testing the Hypothesized Mechanism of Change

The reliability data presented above suggest that the FAPRS is reliable and that an empirical examination of the therapeutic relationship and FAP's proposed mechanism of change is now possible. In this section, a brief example of using the data from the FAPRS to test the hypothesized mechanism of clinical change will be given.

Brief description of lag-sequential analysis. Sequential analysis uniquely captures what occurs in complex interactions (e.g., Bakeman & Gottman, 1986; Gottman & Bakeman, 1979; Gottman & Roy, 1990) and is especially well suited to examining what occurs when two people interact in the context of therapy. Lag-sequential analysis compares the conditional probability of a targeted event, given the occurrence of some other event, to the simple probability of that targeted event. To test whether the mechanism of FAP occurs as hypothesized, researchers can use data from sessions coded with the FAPRS to determine whether the therapist responds contingently to in-session client behaviors. Answering this question helps determine whether the therapist is engaging in the behavior necessary to bring about change. For example, given a CRB2 by the client, we can determine whether the therapist provides a contingent response following that CRB2 (TCRB2) more frequently than he or she emits a TCRB2 independently of the client engaging in improved behavior. A *z*-score and its associated probability are calculated for these comparisons to determine whether these two are statistically significantly different. Interpretations of a significant *z*-score are based on the assumption of the null hypothesis (that there is no sequential relationship in the data).

Analyses. The data used for these sequential analyses consist of 1,788 turns from 31 sessions of Dyad A coded by Rater 1. Transitional probabilities used in the sequential analyses and corresponding *z*-scores were calculated using Allison and Liker's (1982) equations for computing lag sequential analyses. The use of *z*-scores controls for the base rate, or unconditional probability, of the target code. *Z*-scores demonstrate the degree to which the target code is contingent on the occurrence of the given, or specified, code. *Z*-scores above 1.96 are considered significant at the alpha = .05 level and indicate that the target code occurs more often relative to the given event than would be expected, given the unconditional probability of the target event overall. Negative *z*-score values indicate that the target code occurred, statistically, at a level less likely than expected.

Lag-sequential analysis of coded client and therapist behaviors tested the hypothesis that the therapist actually engages in contingent responding; that is, the therapist engages in the behavior necessary

to effect client change. Data for this analysis are presented in Table 7. Table 7 shows the *z*-scores for the conditional probabilities for lag 1 of the therapist's responses to client problem behaviors in-session (CRB1), client improvements in-session (CRB2), and client identification of relevant controlling variables (CRB3) across all 31 sessions. Transitional probabilities are presented for the therapist missing CRB1s and CRB2s (M1 and M2). The table also indicates the therapist's positive session progression behavior code (TPR) in response to the three antecedent CRBs.

Table 7. ***Results for Sequential Analysis for Therapist Behavior.***
***overall *z*-score for 5165 events, significant at $p < .05$.**

	Therapist Behavior					
	TCRB1	TCRB2	TCRB3	M1	M2	TPR
Antecedent						
CRB1	53.81*	-2.87*	4.70*	15.98*	-.65	-9.16*
CRB2	-2.91*	57.88*	-.38	-.66	15.84	-5.21*
CRB3	-1.37*	-.79	14.66*	-.31	-.18	-.78*

These data demonstrate that at lag 1, the therapist was more likely to respond to a CRB1 than engage in other behavior. While the therapist also missed some CRB1s a test performed on the difference between the *z*-scores (calculated using $Z = (Z_1 - Z_2)/\sqrt{2}$; Rosenthal & Rosnow, 1991) for these two codes revealed that TCRB1 occurred more often than M1 ($Z = 26.75$, $p < .05$). The therapist was also more likely to respond to a CRB2 than engage in any other behavior, and a test between *z*-scores for TCRB2 and M2 was statistically significant ($Z = 29.73$, $p < .05$). The therapist was again more likely to respond to a CRB3 than engage in any other behavior when a CRB3 occurred. The negative *z*-score value for TCRB2 given a CRB1 indicates the therapist was unlikely to respond to the CRB1 behavior as though it were a client improvement. In the same manner, the therapist was statistically less likely to respond to CRB2s with a TCRB1, indicating the therapist did not respond to CRB2s as if they were client problem behaviors.

The lag sequential and *z*-score comparison results indicate that the therapist responded contingently and effectively to CRBs as required by the hypothesized model. The therapist was more likely to respond to the impact of the client's behavior appropriately (e.g., emit a TCRB1 when the client engages in a CBR1) than engage in either an incongruous or ineffective contingent response (e.g., emit a TCBR2 in response to a CRB2) or engage in general therapeutic progression responding (non FAP-specific behavior). This finding provides initial evidence for the presence of the mediating mechanism of FAP and the use of sequential analysis as an investigative method for researching mechanisms of change.

Relating mechanism of change to outcome. To relate this responding to changes in client behavior, future studies of FAP's mechanism of change can compute the simple probabilities with which clients engage in CRB1s, 2s and 3s can be examined across a data set. Data can be divided into thirds of treatment, identified as Early, Middle, and Late thirds, and the combined simple probabilities for CRB1s, 2s, and 3s, can be calculated for each third of the treatment by summing the total events for each code and dividing by the number of client events for that third of treatment. Chi-square analyses can then be used to demonstrate the changes in proportions (simple probabilities) of CRBs during each third to test the hypothesis that problematic behaviors (CRB1s) will decrease and improved behaviors (CRB2s and CRB3s) will increase over the course of treatment.

Implications for Future Research

In an ongoing effort to improve the training of FAP therapists, the coding system described in this study is being used to train therapists in the theory and practice of FAP. The FAPRS is being used to

teach therapists to accurately identify their own, as well as clients' behaviors in the sessions they conduct. This teaching method provides a technology to empirically document the change in therapists' behaviors over the course of their training and their corresponding effectiveness in bringing about a change in their clients' behavior. The reliability analyses presented in this study are only a first step toward the refinement of observational measures for FAP and eventual studies of the validity of FAP's proposed mechanism of change. Reliability studies are a necessary stage, however, and these initial data are encouraging.

Future studies of the FAPRS must address several weaknesses, including the potential for training cohort effects and possible confounds when training and testing of the FAPRS uses the same dyads. It is likely that a diversity of behavioral exemplars will be required for raters to become well trained in a functional coding system such as the FAPRS. Coders must learn to more quickly and accurately extract information about behavioral function regardless of topography. The development of standardized, computer driven training systems and expansion in the number and diversity of FAP dyads available for study are key steps toward addressing the weaknesses of this study that are being pursued currently.

Conclusion

This analysis of the Functional Analytic Psychotherapy Rating Scale provides evidence that the functional variables hypothesized to be important to FAP's mechanism of change can be reliably observed and documented. The FAPRS coding system will eventually allow clinicians and researchers to test hypotheses about the therapeutic relationship and its role in effecting clinical change in a manner similar to the example provided above. Other psychotherapies may be able to use the methodology described in this study to specify and measure important elements of the therapeutic relationship and their influence on client behavior change. Although this research is in its early phases, it lays the groundwork to explain variance that remains unaccounted for in traditional psychotherapy outcome studies, namely, "What occurs in-session that helps clients improve?"

References

Allison, P. D., & Liker, J. K. (1982). Analyzing sequential categorical data on a dyadic interaction: A comment on Gottman. *Psychological Bulletin, 91*, 393-403.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*, (fourth ed.). Washington, DC: American Psychiatric Association.

Bakeman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford.

Butler, S. F., & Strupp, H. H. (1986). Specific and nonspecific factors in psychotherapy,: A problematic paradigm for psychotherapy research. *Psychotherapy, 23*, 30-40.

Callaghan, G. M. (1996). The clinical utility of client dream reports from a radical behavioral perspective. *the Behavior Therapist, 19*, 49-52.

Callaghan, G. M. (1998). Development of a coding system for functional analytical psychotherapy for the analysis of the components effecting clinical change. *Dissertation Abstracts International, 59* (09), 5073B. (UMI No. 9907753).

Callaghan, G. M. & Follette, W. C. (1999). *In-vivo contingent shaping as an adjunct method of supervision: A single-case report*. Unpublished manuscript.

Callaghan, G. M., Naugle, A. E., & Follette, W. C. (1996). Useful constructions of the client-therapist relationship. *Psychotherapy, 33*, 381-390.

Callaghan, G. M., Summers, C. J., & Weidman, M. (2003). The treatment of histrionic and narcissistic personality disorder behaviors: A single-subject demonstration of clinical effectiveness using functional analytic psychotherapy. *Journal of Contemporary Psychotherapy, 33*, 321–339.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement. 20,* 37-46.

Ferster, C. B. (1967). Arbitrary and natural reinforcement. *The Psychological Record, 22*, 1-16.

Ferster, C. B. (1972). Clinical reinforcement. *Seminars in Psychiatry, 4,* 101-111.

Fiedler, F. A. (1950). The concept of the ideal therapeutic relationship. *Journal of Consulting Psychology, 14*, 239-245.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.

Follette, W. C., Naugle, A. E., & Callaghan, G. M. (1996). A radical behavioral understanding of the therapeutic relationship in effecting client change. *Behavior Therapy, 27,* 623-641.

Gaynor, S. T. (2002). Getting ahead of the data: Not all threats are equal. *Behavior Therapist, 25*, 137–139.

Gaynor, S. T., & Lawrence, P. S. (2002). Complementing CBT with learning through in vivo experience (LIVE): Conceptual analysis, treatment description, and feasibility study. *Behavioural & Cognitive Psychotherapy, 30*, 79–101.

Goldiamond, I. (1974). Toward a constructional approach to social problems: Ethical and constitutional issues raised by applied behavior analysis. *Behaviorism, 2*, 1-85.

Gottman, J. M., & Bakeman, R. (1979). The sequential analysis of observational data. In M. E. Lamb, S. J. Soumi, & G. R. Stephenson (Eds.), *Social interaction analysis: Methodological issues.* Madison, WI: University of Wisconsin Press.

Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis: A guide for behavioral researchers.* New York: Cambridge Press.

Greenberg, L. S. (1994). What is "Real" in the relationship? Comment on Gelso and Carter (1994). *Journal of Counseling Psychology, 41*, 307-309.

Hawkins, R. P. (1986). Selection of target behaviors. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 331-385). New York: Guilford.

Hentschel, U. & Bijleveld, C. C. J. H. (1995). It takes two to do therapy: On differential aspects in the formation of therapeutic alliance. *Psychotherapy Research, 5*, 22-32.

Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 61*, 561-573.

Kohlenberg, R. J., Hayes, S. C., & Tsai, M. (1993). Radical behavioral psychotherapy: Two contemporary examples. *Clinical Psychology Review, 13*, 579-592.

Kohlenberg, R. J., & Tsai, M. (1991). *Functional analytic psychotherapy: Creating intense and curative therapeutic relationships*. New York: Plenum.

Kohlenberg, R. J., & Tsai, M. (1995). I speak therefore I am: A behavioral approach to understanding problems of the self. *the Behavior Therapist, 18*, 113-116.

Kohlenber, R. J., Kanter, J. W., Bolling, M., Parker, C., & Tsai, M. (2002). Enhancing cognitive therapy for depression with functional analytic psychotherapy: Treatment guidelines and empirical findings. *Cognitive and Behavioral Practice, 9*, 213–229.

Kohlenber, R. J., Kanter, J. W., Bolling, M., Wexner, R., Parker, C., & Tsai, M. (2004). Functional analytic psychotherapy, cognitive therapy, and acceptance. In S. C. Hayes, V. M. Follette, & M. Linehan (Eds.), *The new behavior therapies: Expanding the cognitive behavioral tradition*. New York: Guilford Press.

Krasner, L. (1962). The therapist as a social reinforcement machine. In H. H. Strupp & L. Luborsky (Eds.), *Research in psychotherapy*. (Vol. 2, pp. 61-94). Washington, DC: American Psychological Association.

Krasner, L. (1963). Reinforcement, verbal behavior, and psychotherapy. *American Journal of Orthopsychiatry, 33*, 601-613.

Lazarus, A. A. (1972). *Clinical behavior therapy*. New York: Brunner/Mazel.

Mergenthaller, E. & Stinson, C. H. (1992). Psychotherapy transcription standards. *Psychotherapy Research, 2*, 125-142.

Morganstern, K. P (1988). Behavioral interviewing. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook*. (3rd ed., pp. 86-118). Boston, MA: Allyn and Bacon.

Parker, C. P., Bolling, M., & Kohlenberg, R. J. (1996). *Therapist In-vivo strategy scale. Unpublished manuscript.* University of Washington, Seattle.

Prins, A., & Callaghan, G. M. (2002, November). *Functional analytic psychotherapy as an adjunctive treatment for a client meeting criteria for PTSD*. Poster session presented at the 36th Annual Meeting of the Association for the Advancement of Behavior Therapy, Reno, NV.

Rimm, D. C., & Masters, J. C. (1974). *Behavior therapy: Techniques and empirical findings*. New York: Academic Press.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic and personality change. *Journal of Consulting Psychology, 21*, 95-103.

Rosenfarb, I. S. (1992). A behavior analytic interpretation of the therapeutic relationship. *The Psychological Record, 42*, 341-354.

Rosenthal, R. & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Ruckstuhl, L. E., Callaghan, G. M., & Follette, W. C. (2001, May). *Training more effective therapists using the Functional Analytic Psychotherapy Rating Scale-II*. Paper presented at the 81st Annual Meeting of the Western Psychological Association, Maui, HI.

Schaap, C., Bennun, I., Schindler, L., & Hoodguin, K. (1993). *The therapeutic relationship in behavioural psychotherapy*. New York: Wiley.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Wright, J. H., & Davis, D. (1994). The therapeutic relationship in cognitive-behavioral therapy: Patient perceptions and therapist responses. *Cognitive and Behavioral Practice, 1*, 25-45.

Author Contact Information:

Glenn M. Callaghan, PhD
Department of Psychology
San Jose State University
San Jose CA 95192-0120
Phone: 408.924.5610
Glenn.Callaghan@sjsu.edu

William C. Follette, PhD
Department of Psychology/298
University of Nevada, Reno
Reno, NV 89557-0062
775-682-8699
Follette@unr.edu

L. E. Ruckstuhl, Jr., Ph.D.
Department of Psychology/296
University of Nevada, Reno
Reno, NV 89557-0062
775-784-1126
sonnyr@charter.net

Peter J. N. Linnerooth, Ph.D.
Department of Psychology/296
University of Nevada, Reno
Reno, NV 89557-0062
775-784-1126
peterl@unr.edu