ELSEVIER
Journal of Fluency Disorders xxx (2005) xxx–xxx
Journal of FLUENCY DISORDERS

Evidence-based practice in stuttering: Some questions to consider

Nan Bernstein Ratner *

Department of Hearing and Speech Sciences, The University of Maryland, 0100 Lefrak Hall, College Park, MD 20742, USA

Received 30 March 2004; received in revised form 19 April 2005; accepted 22 April 2005

Abstract

A recent forum in *JFD* (28/3, 2003) evaluated the status of evidence-based practice in fluency disorders, and offered recommendations for improvement. This article re-evaluates the level of support available for some popular approaches to stuttering therapy and questions the relative value placed on some types of programs endorsed by the forum. Evidence-based practice is discussed within the context of emerging concerns over its application to non-medical interventions and suggestions are made for grounding fluency interventions by reference to empirically supported principles of change. A popular, evidence-based intervention for stuttering in young children (the Lidcombe program) is evaluated within the suggested parameters.

Educational objectives: After reading this article, the reader will be able to: (1) evaluate the status of evidence-based practice in fluency disorders; (2) list concerns about the impact of narrow interpretation of EBP on research and practice in the field of fluency disorders and other non-medical domains; (3) articulate the role of theory and basic research in selecting among and evaluating therapy approach options.

In a recent issue of *JFD* (28, 3) (hereafter, 28/3), a number of authors (Bothe, 2003; Finn, 2003; Ingham, 2003; Onslow, 2003) proposed certain standards to which stuttering treatment should be held and further suggested that only a limited number of therapy programs currently meet such evidence-based practice (EBP) standards. In

* Tel.: +1 301 405 4217.
E-mail addresses: nratner@hesp.umd.edu, nratner@aol.com (N.B. Ratner).

0094-730X/$ – see front matter © 2005 Published by Elsevier Inc.
doi:10.1016/j.jfludis.2005.04.002

the same series of articles, some authors appeared to imply that treatment programs not meeting the specified standards, and the clinicians who administer them, may in fact be engaging in less than ethical clinical practice, since they volitionally forgo a small set of "validated" techniques for those seemingly supported by a lesser evidence base. In this space, I would like to further discuss such implications as well as related, seemingly provocative issues. I will also address larger issues in evaluating the degree to which EBP is currently fully "ready for prime time" implementation in the field of fluency disorders. In doing so, I will frame my comments in the form of a series of questions that I think we need to ask and consider answering before applying some of the extended principles of EBP to the field of stuttering intervention.

On the surface, evidence-based practice is a noble concept and goal. Indeed, it would seem nonsensical to argue for therapeutic practice that is not based on some body of evidence. I will not, therefore, position myself as saying that evidence-based practice is wrong. At the same time, certain embodiments and extensions of evidence-based practice seem less obviously of value and may in fact pose difficulties for researchers, clinicians and their patients.

In framing this article partially as a response to the authors in 28/3, I prefer to start with areas of agreement. I wholeheartedly concur with the obvious need for practitioners to document the rationale for their selection of therapy approaches. As such, I also most wholeheartedly agree with the general consensus of the authors in the issue that we need far more research into therapeutic efficacy in stuttering treatment. Thus, I agree with Ingham (2003) that researchers need to develop more interest in therapy trials, and that our funding agencies, particularly the American National Institutes of Health, need to invest in them more aggressively. It is possible to conjecture that the lack of funded research in therapy efficacy is at least in part due to the relative paucity of such applications when weighed against the bulk of submissions that propose basic research questions. Despite these strong areas of agreement, however, my affinity for some, if not many, of the arguments raised in the issue begins to wane considerably, because they raise a number of vexing questions. I address what I view to be the most important of these questions in the remainder of this article. Among the issues that I will consider are:

1) The nature and scope of "evidence" and its relationship to clinical practice;
2) The limitations that may be associated with the use of a single framework to implement evidence-based practice across medicine and the many health-related professions;
3) The role of different types of evidence in determining the value of specific therapy approaches in stuttering;
4) The role of theory in evaluating treatment approaches;
5) Potential barriers to the gathering of clinical evidence and its implementation by practitioners; and finally,
6) Some logical "next steps" that will be required if practitioners and researchers are to bridge the perceived gaps between evidence and practice in stuttering treatment.

1. What is evidence?

Is evidence a "fuzzy category"? A "fuzzy" category in psychology or linguistics is one that seems to have an identity that can be agreed upon, but has features that are difficult to specify exactly (Rosch, 1973; Rosch & Mervis, 1975). Fuzziness has also been applied to the evaluation of data, as in fuzzy logic (Zadeh, 1965). In the past 30 years, "fuzziness" has spilled over to accounts of logical decision-making in the physical, biological and social sciences, indicating that one person's data may or may not be sufficient to be useful to another's evaluation of a set of facts or features.

In much the same way that semantics has argued about the boundaries that separate cups from bowls, it is not clear that any field can arrive at a perfect definition of evidence, or list of its features, although some will find it easier to define operationally, such as pharmaceutical interventions in medicine. Narrow criteria going into an evidence-based test of intervention effectiveness will improve the likelihood that professionals will agree on the value of the evidence that is produced. For example, if the question is whether a drug lowers blood pressure, there are only a few accepted measures of outcome.

In medicine, many questions which could appear simple are complicated by the "messiness" of the typical patient, who rarely presents with a single problem, or canonical features. Thus, the problem of treating the "whole person" appears to create a sense of fuzziness in dealing with evidence—to what extent will carefully gathered treatment evidence bear a reasonable relationship to the actual case under consideration? How are treatment outcomes in one domain related to the overall functioning of the individual in other domains? This fact has generated a certain degree of tension between certain groups in medicine regarding the importance of personal experience with well-defined individuals in one's application of evidentiary meta-analyses derived from large groups (Pope, 2003). There is a growing body of reports in which physicians, for example, argue that the complexity of individual profiles supercedes evidence from carefully controlled trials. Thus, there is growing documentation that some medical practitioners "resist" evidence, when, in fact, they appear to allocate value to different types of evidence in making individual clinical decisions. An important issue in such resistance is not so much a devaluation of new evidence, but personal experience with previous existing treatments that "work" for the clients one typically sees. This type of evidence can be called "anecdotal" as various authors in 28/3 note, but is powerful when the clinician is both gatherer and applier of the data, something quite different from anecdotal evidence associated with authority of peers or other respected professionals (so-called "authority-based" evidence; Onslow, 2003). One could say that there is evidence that one reads about and evidence that one encounters in everyday practice.

2. Evidence of *what*?

This may seem like a silly question, but if one wants to evaluate evidence, one must ask what concept, notion, or idea the evidence has been designed to inform. In fluency treatment, this is not necessarily straightforward, although some would like to limit the nature of the "primary" evidence of interest to measures of stuttering frequency as outcomes of certain interventions. Ingham (2003), for example, is quite open in her assessment that

observable behavior "trumps" less observable concomitant or predisposing components of the disorder, and that is a reasonable opinion to adopt. I would like to point out that the notion that one treats a condition only by addressing its observable symptoms is somewhat controversial, however. For example, going beyond the field of stuttering, it is not clear that current psychological theory uniformly endorses such an approach, which might presume, for example, that maladaptive classroom behavior is best treated symptomatically, rather than by searching for its root causes, motivators, or maintainers. Modern criminology, for instance, looks for the roots of delinquent behavior in a plethora of preceding and underlying factors, rather than merely focusing on deterrence (selective consequation of undesirable behaviors). As an example, in their large-scale meta-analysis, Petrosino, Turpin-Petrosino, and Buehler (2002), noted that cognitive–behavioral interventions are far more effective in preventing recidivism in juvenile delinquency than punitive system responses. In such models of treatment, frequency of incarceration is ultimately compelling evidence, but the treatment targets are not the same as the outcome targets. In such studies, evidence also comes in "stages," first those that show an immediate, targeted change, and then those that show achievement of the final, different goal. In our field, this would be similar to asking whether or not attitude change eventually induces fluency changes, even if this cannot be measured in a short-term study.

Similarly, stuttering frequency is certainly one logical outcome measure for documenting therapeutic effectiveness when treating people who stutter, but not all agree it should be the *only* one or the first one (St. Louis, in press). In particular, if a program diminishes percent stuttered syllables (%SS) by 50% but maladaptive attitudes are neither assessed at intake or therapy dismissal, or are shown to remain unchanged, how is such evidence to be evaluated or balanced with other observations of the same individual? At the same time, it is not clear that attitude change unaccompanied by any changes in stuttering frequency or quality can be considered evidence of fully successful therapy.

3. Can all fields use the same evidence-gathering framework?

Another way to phrase this question is, "how is evidence best produced?" One cannot argue the merits of progressively more demanding tests of treatment effectiveness in any discipline. However, requiring all evidence to fit within the medical community's current hierarchy of evidence types may not be possible for fields such as ours. Fields such as physical therapy, education and social work face difficulty in implementing components of Randomized Clinical Trials (RCTs) such as blinding, placebos, and other common constructs. They don't seem to be a natural fit for all therapies within the discipline (Swinkles, Albarran, Means, Mitchell, & Stewart, 2002), despite being hailed as a gold standard for determining therapeutic efficacy in medicine, and now, stuttering (Ingham, 2003). For example, placebo pills are clearly easier to administer than placebo behavioral treatments. "Washout" phases (to assess the effects of removing a treatment), so appropriate for short-duration medications, are not as easily evaluated when treatments involve client learning of techniques, concepts, reactions, etc. Once some treatments have been applied, they are not readily "un-applied." Thus, while many fields agree that evidence can be derived from RCTs (Feinstein & Horwitz, 1997; Sailor & Stowe, 2003) not all fields would agree that it

has to be, and in some cases, argue that it should not be. Even in medicine, Culpepper and Gilbert (1999, p. 831) note that

> . . . the emphasis on (EBT) . . . has raised the Randomised Clinical Trial (RCT) to a position above other forms of biomedical research. Yet, primary care research, with its emphasis on outcomes, natural history, human behavior, cultural issues, and questions related to care rather than cure, is of vital importance to an understanding of our patients, their problems, and the way we can best serve them. These issues are suited to qualitative research designs, descriptive work, and case-control or cohort designs; RCT designs rarely apply. That we neglect these critical research priorities in favor of (RCTs) . . . is tragic.

One apparent concern with EBP, as identified by Sailor and Stowe (2003), is that we cannot allow social policy to view program efficacy as either a race or as a construct only evaluable by application of a tightly specified subset of research methods. I will not bore readers with their detailed and hopelessly familiar analogy of the drunk seeking his keys only under the lamppost because "the light is better there," but will note their concern that in the messier, real-world contexts of behavioral change in childhood and beyond, we cannot limit ourselves only to those questions which lend themselves to randomized clinical trial designs; nor can we necessarily view an RCT within a discipline as superior to other forms of therapeutic evidence being gathered concurrently. Sailor and Stowe are adamant in their distaste for the RCT as the model for educational research, a close colleague of the work performed by speech language pathologists. Salkovskis (2002) echos their concern in the application of RCTs to cognitive–behavioral treatments. As noted earlier, it is not clear that RCTs will apply to situations where conditions cannot be counterbalanced, blinded or "withdrawn" to verify the impact of the intervention on the behavior of interest. Only operant or similar behavioral approaches to fluency treatment seemingly fit this model. Therefore, blanket acceptance that RCTs represent the gold standard for speech–language intervention may artificially constrain which approaches to intervention are even eligible for equal consideration. It is not clear that a fluency shaping or modification skill, once taught, can be untaught, or the components of cognitive therapy be undone.

4. What are the qualities of good evidence?

First, in choosing from the available data, publication in a peer-reviewed journal is often considered the default, as it should be. Our best evidence is to be found there, rather than in books or workshops, neither of which are held to rigorous peer-review standards Journals tend to be quite concerned by the design of trials that report therapeutically relevant evidence. To raise the bar, we could set minimum standards for the number of subjects required to provide evidence of therapeutic effectiveness, require thorough baselining, demand evidence of skill generalization, durability of results, etc., in order to publish any results; however, to date, we have not set such standards. And the lack of such standards is probably fortuitious given the fact that historically, such requirements would be likely to stultify progress. Witness the fact that many major discoveries in medicine, psychology

and other fields have begun with publication of individual case reports or case series that spurred further basic and applied research (Vandenbroucke, 2001).

Beyond the features of individual research reports, an important component of the value of evidence is its replication by sites not viewed as major "stakeholders," something recognized even in 1984 by Andrews, Craig, Feyer, Hoddinott, Howie, and Neilson in their review of stuttering therapies, and certainly held as important in other fields. Why is this relevant? In stuttering, of the roughly 50 published articles on the effectiveness of the behaviorally or operant-oriented Lidcombe program of intervention (Onslow, Packman, & Harrison, 2003), virtually all have program authors as article authors. As we improve the level of evidence-based practice in fluency disorders, we will clearly need more evidence from clearly independent sites (Lohr, DeMaio, & McGlynn, 2003) that have neither actual nor perceived personal, professional or financial interests in the outcomes.

It is interesting that, outside the field of speech–language pathology, other fields have moved from asking that therapies or interventions be grounded in evidence to asking deeper questions about the nature of the evidence itself, rather than the design of trials meant to produce evidence. For example, an important concept now operative both in medicine and psychology is that therapy programs should be "dismantle-able" into their component efficacious parts (Rosen & Davidson, 2003) to isolate "empirically supported principles of change" or enable understanding of the "causal processes and mechanisms" (Michie & Abraham, 2004). I will return to this issue later.

5. Is there evidence in stuttering treatment?

Yes, of many differing approaches and components. Ramig and Bennett (1997) and Culatta and Goldberg (1995) are among those who provide annotated summaries of the degree to which extant therapies are evidence-based. Surveys of practicing clinicians suggest that they employ a healthy mix of these evidence-supported approaches, including fluency shaping, attitude modification, and operant techniques. Cooper and Cooper (1996) reported that only 7% of practicing clinicians disagreed with the premise that a "combination of operant and self-evaluative-type therapies (sic) is probably the most efficacious type of therapy for stutterers," a finding that Cooper and Cooper (1996) believed to be "reassuring" (p. 132).

Thus, in some respect, much of what seems to be at stake in positioning "camps" in stuttering treatment is the relative focus on, contribution(s) and merits of potential therapy components involving stuttering modification, cognitive therapies or counseling in addition to negative contingencies, and/or fluency shaping components and/or GILCU (to paraphrase Ingham, 2003, pp. 203–204). Despite some expressed pessimism about the level of empirical support for many popular approaches to stuttering therapy (Ingham, 2003), one can find ready evidence for the value of many therapy components in the treatment of stuttering. This body of evidence includes documented support for the so-called "traditional" approaches (fluency shaping with or without components of stuttering modification; Baumeister, Caspar, & Herziger, 2003; Conture, 2001; Conture & Melnick, 1999; Gottwald & Starkweather, 1999; Runyan & Runyan, 1999). Within therapy programs that address the cognitive and affective concomitants of stuttering, there is evidence of the effective-

ness of interventions to improve self-efficacy and locus of control (I note that some of these are from outside the field of speech–language pathology, a concept I will return to later; Bandura & Locke, 2003; Graves, 2003; Multon, Brown, & Lent, 1991). Again, from outside the field of speech–language pathology, there is excellent evidence of the value of cognitive–behavioral therapy (CBT), an empirically supported treatment approach (Salkovskis, 2002), shown to be effective in treatment for Generalized Anxiety Disorder (Lohr et al., 2003). CBT has also been effective in reducing speech anxiety that not only impacts speakers' comfort in communication but can logically hinder everyday generalized application of other therapeutic strategies for more fluent speech (Cho, Smits, & Telch, 2004; Heimberg, 2002). Desensitization in stuttering treatment is also validated by analogy to other disorders (e.g., Fava et al., 2001). Studies of relapse, for example, demonstrate the effects of anxiety on skilled motor practice (Craig, 1998; Woodman & Hardy, 2003). Self-reports of successfully treated and self-treated adults (Finn, 1996; Yaruss et al., 2002), which tend to emphasize some degree of cognitive–behavioral readjustment, should not be dismissed as forms of evidence when they are congruent with other types of data, as they appear to be. There is a substantial difference between such survey results and testimonials that cannot be reconciled with the larger body of basic and applied science in a domain (e.g., the use of copper bracelets in healing cancer). Vos, Willems, and Houtepen (2004) note the value of patient testimonial in evaluating appropriate treatment options for conditions which at this point in time are "medically inexplicable," as stuttering would seem to be. Finally, there is a very large literature on the effectiveness of operant/behavioral treatments (some of which is summarized in the 28/3 series), which I am neither explicitly nor implicitly trivializing. I simply believe that data derived from such treatment approaches do not represent the only "decent" evidence out there. Furthermore, I believe that some existing treatment effectiveness data could benefit from better integration with evolving data and theories regarding the nature and etiology of stuttering, as I will discuss further below.

6. What will affect the "shape" of future evidence?

It seems reasonable to suggest that at least some of the information being gathered in "basic research" on stuttering will have to be integrated into the design and interpretation of efficacy research in fluency disorders. For example, most researchers now agree that approximately 80% of children who begin to stutter will spontaneously remit from symptoms, primarily within a window that is between one and two years post-onset (Curlee & Yairi, 1997). This relatively recent set of facts "ups the ante" for demonstrations of clinical effectiveness in preschool stuttering therapy because it leaves open the possibility that even very efficacious therapies look to be so because of sampling error. It is probable that we are looking for a small effect at play among the 20% of children who are not expected to improve simply through the "tincture of time." As such, failure to balance clinical populations varying in inherent risk for chronic stuttering may produce outcomes that unduly reflect either spontaneous recovery or speeding of the natural course of spontaneous recovery. Fixing a problem earlier than it might have resolved on its own, even if it is quite dramatically linked to treatment timing, is not trivial but is not the same thing as efficacy

in treatment of chronic stuttering. Specifically to the point, in preschoolers, even so-called treatment control groups, such as those reported by Harris, Onslow, Packman, Harrison, and Menzies (2002) show a tendency for both untreated and treated children to show similar improvement curves, although the treated children showed more improvement during the treatment period. Thus, a primary challenge for all interventions, particularly those involving preschool children who stutter, is to show that they are demonstrably better than delayed therapy or no therapy at all, something that would be particularly difficult for small scale studies to prove or disprove. To date, to this writer's knowledge, no large scale published studies have contrasted multiple types of intervention for early childhood stuttering, a critical step if we are to rule out what appears to be the case in contrasting legitimate forms of psychotherapy, that is, that most such psychotherapies are quite effective (Wampold et al., 1997). As we now gather information regarding predictors of childhood persistence and recovery (see Yairi & Ambrose, 2004), all treatment studies in children will need to address and report these patient variables. Likewise, it is rare to see information about adult clients that includes potential outcome-influencing data, such as concomitant problems (Watson et al., 1994), and the degree to which the perceived handicap of stuttering (Yaruss, 1998) may impact broad and lasting generalization of skills. This is, of course, not simply a problem in stuttering or speech–language pathology. There has been a recent cross-disciplinary sea change (in medicine, allied health, etc.) to include more patient-centered outcome measures and a broader model of impairment (Swinkles et al., 2002) in evaluating therapy effectiveness.

7. Do the therapies themselves matter?

For some readers, this will seem like a foolish question, but it shouldn't be if we examine the process of evaluating therapy effectiveness in other fields. To date, we simply lack sufficient numbers of comparison studies to perform the required meta-analyses in fluency therapy, but in psychotherapy, specific techniques have been found to account for no more than 15% of variance in therapy outcomes (Chwalisz, 2001; Lambert & Barley, 2001). This makes behavior change therapy, like that used in psychology or stuttering, very different from the types of interventions typically studied in evidence-based medicine. It is not clear that the beneficial effects of drugs require any interpersonal skills on the part of those who prescribe or dispense them. A surgeon's technical prowess would seem to be much more important than other aspects of his or her involvement with the patient. In such fields, we might say that interpersonal skills don't matter much in one's treatment outcome—but what about the interpersonal talents of one's speech–language pathologist, teacher, or psychologist? Given a high degree of evidence that specific therapies cannot in fact be favored over one another in the treatment of certain disorders, it has already been recommended that clients should seek out specific therapists, rather than specific therapies (Ahn & Wampold, 2001). The notion that all therapies appear to produce winners doesn't negate the importance of asking and trying to answer this question (Wampold et al., 1997). Taken to extremes, there seem to be two diametrically opposed views of the value of the therapy approach itself in achieving outcomes. In the first, exemplified by the Ahn and Wampold (2001) recommendation, the most active component of some treatments is the

therapist. This notion may remind some of detailed observations made by Van Riper (1973) regarding the potent roles of clinical suggestion and persuasion in promoting historically popular approaches to stuttering therapy, and is certainly one basis for the placebo effect across health-related disciplines (e.g., Benson & Epstein, 1975).

In the opposing view, assuming that the therapist understands the precepts of the treatment, all treatments are the treatment: when the treatment isn't working well, it has to be the therapist's fidelity in applying the approach; he or she may need more guidance or experience. The approach itself cannot be at fault if it has been experimentally validated somehow. In fact, this is an observation made in the recent issue of *JFD* (Onslow, 2003), when variations in the rate of clinical progress are to be considered. Between such extremes, there is of course the thorny problem that not all treatments work for all individuals, something medicine knows well (e.g., paradoxical (opposite) response to an intervention or lack of response to medication). There is also, in any literature, a published tendency to "blame" patient or family compliance for lack of success in therapy, a problem sometimes disguised in the completion or drop-out rate.

8. Can treatments arise in the absence of theory?

Theory is basic to any discipline, even when we do not directly appeal to it in framing interventions or evaluating them. Sailor and Stowe (2003) note that the age-old opposing philosophies of positivism (empiricism) and constructivism (subjectivism) have not made "good bedfellows" in the behavioral sciences, and thus, sometimes underlying theories are not overtly addressed in framing social policy.

Can effective therapy arise in the absence of well-articulated theory that addresses the nature and etiology of stuttering? Asking this question is akin to asking "*should* and *why* does this therapy work?" in addition to "does this therapy work? In fact, specification of causal mechanism or mechanisms of action is a component of Phase I studies aimed at approving new pharmaceuticals (www.fda.gov). While there are those who would argue that some approaches to treatment render theory unnecessary (after Skinner, 1950), there would appear to be tangible consequences if a treatment does not appear consistent with practitioners' views on how behavior change is achieved. First, there is likely to be lack of "buy in." Clinicians seem to have more of an affinity for programs whose mechanisms they understand. In speech pathology, this may be a factor in the declining popularity and/or sole use of operant-based approaches in treating stuttering, which have been less than vocal in linking themselves to a theory of stuttering, especially recent biological models. The current tempering of enthusiasm for using or only using operant approaches in treating stuttering has been seen earlier, in the decline of popularity of similar programs in treating language problems in children (Kamhi, 1999).

Lipsey and Cordray (2000) summarize recent concerns from a number of fields that are currently evaluating EBP by referring to "grounded" or "program" theory. Many interventions may share characteristics that lead to successful outcomes (both active or non-specific factors; Lohr et al., 2003) and some components may mediate outcomes in ways not anticipated by program framers. Thus, it is critical to explain what program outcomes are or are not actually produced by unique program actions—in other words, determining the

causal mechanisms through which the program produces the outcomes of interest. Lohr et al. (2003) express concern that there is a tendency to replace analytical efficacy studies that test and examine putative causative agents of change with studies of comparative efficacy. Treatment outcome studies also face difficulty in equating any number of non-specific but potentially active outcome factors such as degree of attention, the personality and expertise of the clinician as well as his or her allegiance to the intervention being employed.

Thus, using the experimental method to validate interventions is of course extremely valuable, but has its limitations. As Lipsey and Cordray (2000) note, speaking about educational interventions, using the experimental method to ask whether an intervention is effective only provides a "yes" or "no" answer. In today's social climate, because we wish to optimize correct answers and improve on less successful attempts, pragmatic assessment of efficacy "puts a premium on *why* (italics added) an intervention is successful or ineffective . . . in short, the explanation of program effects, or lack thereof, requires some THEORY that can serve as a framework for organizing and interpreting information from both descriptive and experimental components of an outcome evaluation" (p. 358). The inverse, therapy with no apparent theoretical grounding, that attempts to specify precise content or form of intervention, represents a technical approach to behavior change that reduces the role of clinicians (and researchers) as decision-makers and innovators (Evans, 1996).

Thus, model building is important in the applied social sciences (Dishion & Patterson, 1999). Models provide a way of conceptualizing the processes that lead to the outcomes of interest in the intervention sciences. Good models are testable, falsifiable and revisable. They can be improved by adding or changing theoretical constructs or reconceptualizing relationships among model components. Eifert (1996) contends that, "a theory-driven hypothesis-testing approach to conducting therapy that is both flexible and patient-focused, is critical for the long-term survival and prosperity of behavior therapy in an era of managed health care." Lohr et al. (2003) put it more concisely: we must "do the right clinical things for the right theoretical reasons."

Lack of an underlying theory for a therapy approach also leads to an inability to refine initial proposals to be more efficacious or can lead to "reinventing the wheel rather than re-applying it" (Michie & Abraham, 2004). In other words, if we understand what parts of a therapy are achieving its desired goals, we can "fine-tune" the therapy to maximize "facilitating" components while reducing apparently non-effectual components that do not seem to exert much effect on the outcome. Unfortunately, in some cases, the cost of a disconnect between a therapy and its theory is a spate of initial successes that don't hold up under scrutiny because post hoc explanation of a potential relationship between a therapy and its mechanism of action turned out to be inaccurate (e.g., the embarrassing case of eye movement desensitization therapy (EMDR) in psychology; Follette & Beitz, 2003).

One way to choose among therapy approaches is to relate them theoretically and empirically to beliefs about what we think stuttering is, etiologically. Increasing amounts of basic research suggest that stuttering reflects underlying neurological disfunction, and interacts with speech and language planning demand (see Anderson & Conture, 2004; Buchel & Sommer, 2004; Conture, Zackheim, Anderson, & Pellowski, 2004; Cuadrado & Weber-Fox, 2003; Forster & Webster, 1991, 2001; Kleinow & Smith, 2000; Smith & Kelly, 1997 for representative but not exhaustive examples). As stuttering becomes chronic, there is also evidence that it is shaped by experience. But little empirical data exist to suggest that

stuttering is "learned" per se, and thus, it is not clear why operant approaches (the primary ones favored by authors in *JFD* 28/3 with the exception of Langevin & Kully, 2003) should work, for instance. Moreover, most of the therapy presentations and efficacy data reports for operant procedures do not attempt to ground the therapy in theoretical or evidence-based models of the underlying disorder. Years ago, Andrews et al. (1983) noted that:

> The purely operant model of stuttering . . . by implication involves some qualitative differences in the child-rearing practices of stutterers' parents . . . behavior that has been looked for and not found. The model . . . (also) predict(s) that stuttering behaves as a lawful operant. Even the evidence for this last proposition is somewhat equivocal . . . hedonically ambiguous stimuli such as verbal stimuli have been employed (in studies, and) . . . Clearly the data on the effect of punishment are unsatisfactory . . .. (p. 237)

Attanasio (1999) suggests that operant programs are atheoretical by nature and do not require grounding in models or theories, but others would vociferously disagree (Eifert, 1996; Lohr et al., 2003; Rosen & Davison, 2003; Salkovskis, 2002). In fact, there is real risk that treatments not based on any articulated theory may well "paint themselves into corners" when challenged by the need for refinement.

While questions in evidence-based practice can be phrased as *Yes/No* or *More/Less* (Does this treatment work? If so, which treatment is better than another?), there is an important role for questions that ask, "How" and "Why." We need to know *why* therapies succeed, so "lessons can be learned and improvements can be made . . . program stakeholders . . . want an explanation for the results . . ." (Lipsey & Cordray, 2000), not just the results themselves. Across other fields, there is growing concern that evidence-based practice should be superimposed on a coherent body of knowledge of pathophysiology and clinical observation, not the other way around; as van Weel and Knottnerus (1999) note, we need to test "key elements of the intervention as identified from a conceptual basis" (p. 918).

Thus, reasonable questions in selecting a preferred therapy approach are: How do we think therapy works, when it does seem to, and why wouldn't it be expected to work when it doesn't seem to? As Kamhi noted, discussing approaches to language remediation in children: "One reason not to use a . . . treatment approach is that we do not understand what it is about the treatment approach that makes it work" (Kamhi, 1999, p. 96). Onslow (2003) chastises clinicians for employing "assertion-based practice." However, it is possible that an educational emphasis on the growing body of basic research on the nature of stuttering has led clinicians to apply an educationally enculturated form of clinical bias in dealing with accumulated evidence of neurophysiological, physiological, neurolinguistic and cognitive factors in the disorder. Because so much research into developmental disorders now emphasizes "constitutional" factors in their etiology, I do not find it surprising that surveys show a recent slight downturn in interest in the use of purely operant programs (Crichton-Smith, Wright, & Stackhouse, 2003; Kuhr, 1994) by clinicians. The "traditional" approaches tend to emphasize components that include fluency shaping, stuttering modification, densensitization and establishment of self-efficacy and locus of control (LOC), each of which has therapy goals grounded in part in neurophysiological research. Thus, it is relatively easier to understand WHY these approaches should work, and it is merely a next step to further document their application to this population. There are also, contrary to the assertion of

some, already data to show that they work, as noted earlier. The fact that such components appear to have such wide-spread application (Cooper & Cooper, 1996) and that adults report relatively good success with them (Yaruss et al., 2002) is distinctly at odds with claims that they lack efficacy. The data are consistent to some degree with a claim that their current profile of efficacy data may not be as compelling when contrasted with a small but intense body of data on a small number of fairly singular approaches.

However, supportive data are there if one chooses to look for them, both in stuttering and related fields. For example, attitude change does in fact bear on the risk of relapse (Andrews & Craig, 1988; Craig, 1998). In social phobia, a meta-analysis of exposure therapy has confirmed its efficacy in desensitizing patients (Fava et al., 2001). This is not "assertion-based practice." To believe that clinicians prefer to utilize non-efficacious techniques even when they do not work for their clients requires a more misanthropic view of our professional peers than seems warranted, and threatens the type of collegial dialogue that best fosters cooperative design, collection and analysis of therapy outcome studies.

9. Will we need more than one theory? Will we need more than one therapy approach?

There is already some consensus that the theory of what causes stuttering is not sufficient to explain what stuttering becomes. In health professions that emphasize behavioral change, "no single theory dominates ... the most useful approach is to combine concepts from more than one theory to address a problem, and to bring these together in a comprehensive plan. ... no model or theory will get it right all the time, and in practice, often a single theory (approach) explains only a small amount of the variance in targeted behaviors" (Cockburn, 2004). As noted earlier, there is evidence that speech–language pathologists already believe this, and use a healthy mix of therapy approaches, including fluency shaping, attitude modification and operant. The concept that more than one fluency therapy may work equally well already appears documented for older children who stutter (Hancock et al., 1998), and future research may wish not only to enlarge the approaches that are compared, but seek predictors of which clients profit best from which therapies.

10. Applying evidence to theory: a case in point

A number of the contributors to the JFD symposium cite the Lidcombe program (Onslow et al., 2003) as one that has better efficacy data than popular alternatives for the treatment of stuttering in very young children. I would like to illustrate some of my concerns about the application of EBP by applying some of the concepts in this article to this value judgment.

First, as noted, Lidcombe is grounded as an operant program. As mentioned earlier, operant procedures have fallen out of favor in stuttering, as well as across many fields of human behavioral intervention. Indeed, such procedures appear to be primarily used with profoundly impaired children in the field of communication disorders and related educational and behavioral domains. In language, there is consensus that operant programs did not meet the clinical challenge, even in Pervasive Developmental Disorder, where general-

ization of results has not met expectations, even when trial by trial documentation appears promising (Delprato, 2001).

Moreover, in other fields, there is a trend to reconceptualize some older operant terminology in order to better explain effects of certain interventions. In particular, I find more recent, non-operant interpretations of concepts such as "reward" and "reinforcement," which I will enlarge on below. However, I find few parallel conceptualizations of extinction, a concept clearly at play in Lidcombe, where "correction" (earlier termed "punishment") is the presumed agent of stuttering reduction. Negative consequation of behavior in most current behavioral therapies implies some level of relative patient choice of action and definition of reward that can impact action, such as a broad group of maladaptive and habitual behaviors. Bandura and Locke (2003) reflect that, "If knowledge and skills could be acquired only by negative response feedback, human development would be greatly retarded . . .." Finally, as I note in more detail below, punishment of the stuttered moment (or time-out from positive reinforcement) does not easily map onto our understanding of the mechanisms that produce or prevent children's stuttered events.

Typically, a therapeutic program is most attractive to discerning, experienced clinicians if they extract principles that permit them to understand, adapt and trouble-shoot the program. In other words, if I or any other clinician knows why a component is integral to the program, if the child does not respond as predicted, we can fine-tune using similar principles. Right now, in operant programs such as Lidcombe, the contingency schedule appears to be the only manipulable component, which limits the degree to which one can fine-tune or "dismantle" (Rosen & Davidson, 2003) the program to investigate its mechanisms of action or refine its level of effectiveness for a particular client. Simply put, when we teach a skill, we can think of other ways to convey the information or skills we seek to instill. When we change behavior through contingencies, there is no teaching, only learning, of a sort. When learning does not occur, therapeutic course adjustments or options appear relatively restricted.

Still another way to endorse the outcomes of a program is to be able to interpret the program from other theoretical perspectives, and there is actually a wealth of ways to view a purportedly operant program such as Lidcombe. I will take this opportunity to speculate on only a few. For example, it is relatively easy to hypothesize that a potential mechanism of action in Lidcombe is manipulation of linguistic demand. Parents must design activities that progressively create a proportion of fluent and stuttered utterances and it is possible that they help children shape less ambitious language attempts as they gain fluency. Intake/outcome language measures for Lidcombe children (Bonelli, Dixon, Bernstein Ratner, & Onslow, 2000) suggest that spontaneous language tendencies of children who have completed the program have been modified to be more age-appropriate than they were when the child was first referred for intervention. This may have occurred if children associated correction of dysfluent utterances with parental disapproval of relatively long or complicated utterances. This does not mean that children 'lost language' skills as a result of the reinforcement schedule, but rather that they may have been encouraged to keep their language production attempts more comfortably within their fluency abilities (Bonelli et al., 2000). Such data are consistent with recent data reported by Watkins and Johnson (2004) that show spontaneous recovery to be linked to slowed pace of expressive language growth post-onset of stuttering symptoms, while persistence was linked to more developmentally advanced spontaneous language profiles.

It is *not* just a matter of semantics to ask if one can describe an operant program such as Lidcombe in alternative terms. Casting therapy components in a new light may allow us to test mechanisms of action. One can ask, for example, if Lidcombe "binds parental anxiety" (Bernstein Ratner & Guitar, in press) by making parents active participants in their children's therapy. Reductions in parental anxiety levels have been shown to improve some child therapy outcomes (Cobham, Dadds, & Spence, 1998). Similarly, because it acknowledges stuttering openly, does Lidcombe reduce *child* anxiety? This might predict that any structured therapy involving a high degree of parental involvement would yield positive outcomes, despite varying specific goals, a notion recently supported by data from Cook, Millard, and Nicholas (2004). Does Lidcombe wind up reinforcing other behaviors that influence the frequency of stuttering? (This is likely to be true of many if not most stuttering therapies.) One can even venture into less comfortable questions, such as, "does it simply drive some behaviors "underground"? (create covert stuttering)?

I ask such questions not to diminish the available data, which are truly impressive, but because to a certain degree, the data don't make a lot of sense when presented in a theoretical and scientific vacuum. For example, one of the most troubling "disconnects" in understanding the success of operant techniques with preschool stutterers is the finding that in adults, such therapies appear to work by calling upon extant fluency skills obtained through prior types of therapy (James, Ricciardelli, Rogers, & Hunter, 1989). Since Lidcombe toddlers have not ever been instructed in fluency-enhancing techniques, it is not clear what extant resources the negative feedback spurs into action.

Further, because the program is intrinsically linked to reinforcement of fluent speech, it is not trivial to note that the mechanisms by which verbal praise maximizes desired outcomes in children is still poorly understood (Henderlong & Lepper, 2002), but currently, praise is rarely viewed in simple operant terms. Research suggests that the effects of praise vary considerably depending upon the content of the message, the context in which it is delivered, and the way in which it is interpreted by the child. Praise tends to "bond" children to an activity insofar as it increases a child's motivation to continue participation in an activity. It also is clearly associated with increases in self-efficacy, which maximizes locus of control and coping behavior. All have been linked to fluency improvement in other models of therapeutic intervention.

Again, rephrasing the mechanism for change is more than mere semantics. Crucial to questions about the relatively higher success rate of Lidcombe with younger children, it is notable that preschool children do not clearly distinguish between implications of praise and corrections that target effort versus ability, while older children may make more complicated assumptions about what is being reinforced and penalized in feedback from adults (Henderlong & Lepper, 2002). Because Henderlong and Lepper review a large literature linking the presumed sincerity and affective value of praise to its effects, it is also interesting to speculate that Lidcombe might actually be much less effective if administered by clinicians, rather than parents. Such a notion places a new spin on the therapy design, which seems to have originally involved parents as a way to achieve efficiency in session programming (using parents as therapist surrogates in an environment where therapists are of limited availability). We may be able to predict more and less successful mixes of target populations and therapy providers, or adjust program components by age if we

evaluate these potential agents of behavioral change in designing and monitoring program outcomes.

If a program such as Lidcombe doesn't recruit some of the mechanisms I have discussed, its efficacious basis is unclear (although I am truly impressed by the efficacy data) since critically, as previously noted, operant programs in adults seem to work best when clients already have techniques to apply to managing fluency (James et al., 1989). I am certainly amenable to trying it with particular children, and using it if it yields results within the promised time frame (a limited number of sessions). But I am less than sanguine, at this state in our knowledge, about setting up this or any other intervention as the paragon of therapeutic virtue to which all other programs must be compared unfavorably.

As with many other fluency therapy programs in print, there are some other questions about Lidcombe that we need to answer: The majority of data points represent unknowns, other than overt fluency behaviors going in and coming out; thus, we don't know risk factors such as family history, concurrent language abilities, the child's emotional and cognitive responses to the stuttering problem, the responses of the parents, etc. The recent report from Harrison, Onslow, and Menzies (2004) is a promising beginning but still quite underspecified. Such concerns reflect earlier discussion about what forms of data are considered relevant in the evidence-gathering process. Additionally, with all that we now know about prognostic indicators (Yairi, Ambrose, Paden, & Throneburg, 1996), we have a paucity of information on the developmental profiles of Lidcombe-treated children (and most other young children enrolled in therapy protocols).

In sum, I wish to be very clear in my appreciation of the Lidcombe program and its efficacy data. The program;s authors are to be respected and commended for providing us with an evidence-based treatment option for young children, as well as for heightening our level of appreciation for rigorous treatment outcomes research. I do believe that the program works for the majority of children who have been enrolled thus far. But I do not understand why it works, or why it should work, and would appreciate answers to these questions before reading more reports telling me that it does.

11. What is the alternative to EBP?

If there is an alternative to EBP, it is certainly not the endorsement of non-efficacious treatments. However, this is often how the debate is framed, particularly in the many disciplines now entering the EBP fray (Levant, 2004). In psychology, there is already concern that, "controversy seems to stem from the attempts of some clinical scientists to dominate the discourse on acceptable practice and impose very narrow views of both science and practice" on the field. To some, this might not seem like such a bad thing, but one should be cautious. In psychology, outcomes of Task Forces assigned to evaluate EBPs tended to over represent those approaches that involved highly specified, manual-guided treatments that permit few clinician-selected adjustments. Conflicts among practitioners regarding "best practices" also provided grist for managed care and insurance companies to restrict coverage for those seeking services, on the grounds that particular therapies had less supportive data or conflicting data (Levant, 2004). This was not the intent of the movement, but it was its outcome.

12. How do we proceed best in this climate?

A first step might be by cooperating to gather more data. And diversity of approaches is probably a good thing, until proven otherwise. Everyone benefits if there is more than one way to "skin a cat". A dynamic, ever-developing discipline should not be reduced to a single manual of practice, because clients inherently differ in their symptoms and needs. Thus, we need flexibility in planning as well as conducting therapy, rather than strict adherence to a "one-size-fits" all model.

Second, it would appear to be prudent not to start turf wars or philosophical disputes. While interesting (or frustrating) to academics and practitioners alike, a field that seems to be populated by combatants rather than colleagues is likely to suffer casualties in the bid for research funding, since reviewers are more likely to provide support for treatments that appear to be supported by consensus.

It would appear to be imperative to obtain more empirical data regarding the effectiveness of the practices actually employed in the field. The components of stuttering modification therapy are particularly under-researched, which is somewhat surprising given the personal therapy histories of some of the major researcher/clinicians in our field. In stuttering, contrary to some views of the situation, we don't just have "authority-based practice" (Onslow, 2003). Many of our most visible clinician-researchers who are identified with certain therapy approaches are themselves "walking graduates" of the program orientations that they endorse—that makes us somewhat unique, and I tend as a fellow professional to give credence to therapies that by report helped to give my colleagues fluency. However, more systematic data are clearly needed.

As we do all this, we should derive comfort from the fact that some of our efficacy data come to us by close analogy from psychology and other fields (e.g., CBT). The success of CBT in treating a wide variety of behavioral disorders, including communicative anxiety, should be viewed as endorsement of the wisdom that they can be used to combat speaking fears and anxieties in stuttering. For example, we don't need to test aspirin in all conditions that produce fever, if we know it reduces the primary symptom.

13. How do we assure clinicians "get" the evidence?

Just as our discipline is experiencing the EBP movement later than some other fields, there are some questions we have not yet begun to ask ourselves that will be very important. I do not mean to scare anyone, but the gap between the evidence and actual practice is a big problem for the medical community these days—just because you produce evidence doesn't mean that it gets used (Davis, Ciurea, Flanagan, & Perrier, 2004). Although no formal studies have been made of the degree to which practicing SLPs keep up with, as well as employ, newer or better supported treatment strategies, there is some awareness practitioners in many fields are unlikely to read their professional journals, and are more likely to "Google" the problem (Nail-Chiwetalu & Bernstein Ratner, 2003, submitted) than consult the appropriate primary literature.

It seems safe to suggest that we need a sea-change in SLPs' valuation of the need to keep up with the published literature. The new continuing education requirement for the

CCC's alone won't do it; journal self-study is one of the least pursued options, and peer review of workshop contents is more or less a formality. In medicine, there is depressing evidence that continuing medical education has not changed physicians' behaviors (Davis et al., 2004). So, a reasonable question to ask is: How do we convince practitioners to gather and use evidence in support of their practice?

One way to increase use of "newly evidenced" therapies is clearly philosophical "buy in." In addition to what we see in our own literature and that in psychology, medicine faces similar hurdles. For example, as an extreme example, many physicians are not persuaded by studies of "distant healing," regardless of findings and quality of research design (Astin, Harkness, & Ernst, 2000), because they cannot reconcile such findings with their existing belief systems. We do have some philosophical differences among each other in terms of what we believe stuttering is, as well as how theories and the evidence that supports them should integrate with therapeutic approaches.

In such cases, simply "dumping the data" is not going to do the job, as Ingham (2003) and Bothe (2003) observe in *JFD* 28/3. This may be especially problematic in the USA, where training in fluency treatment is much more heterogeneous than in Australia, for example (Attanasio, Onslow, & Menzies, 1996), and requires buy-in of diversely educated professors and clinical supervisors.

14. Caution: there is no finish line ahead!

EBP is not a race, where the earliest finisher ends the discussion, particularly when the evidence base is very immature. Levels of data throughout the profession (e.g., the recent Cochrane report for speech/language therapy in children, Law, Garrett, & Nye, 2003) are currently low and need to be thoughtfully built, not just on data, but on scientifically and theoretically sound principles. Psychology had an interesting recent "run in" with this problem. As noted earlier, recent analyses of eye movement desensitization and reprogramming (EMDR) therapy have detected a troubling flaw: recent research clearly indicates that in this hugely "successful" treatment, the putative underlying construct upon which the therapy was built is completely dissociated with the outcome (Follette & Beitz, 2003; Wampold, 2001), and that outcomes may reflect a complicated mix of non-specific factors. Thus, Chwalisz (2001) warns that an over-interpretation of EBT might result in the endorsement of dubious treatments that produce superficially compelling results, but do not utilize conceptually and empirically sound principles of change.

Because the level of evidence in our field is relatively low when compared to medicine or even psychology, it seems less than appropriate to chastise those who use alternative approaches "early" in the game. It is my personal impression that it has conventionally been easier for "behavioral" interventions to amass efficacy data in many fields. This concern is not unique to stuttering: in medicine, some approaches that take too narrow a view of the problem space result in "overcontrolled" studies. These in turn can lead to "hyperclaiming" (Rosenthal, 1994), or claims of a research team that a proposed intervention is likely to achieve goals that it will not when applied more broadly (or that it is the single acceptable approach to a problem). Faster accumulation of out-

come data can obviously lead some programs to document given levels of efficacy before others. This is particularly true when the current climate in EBP is partnered with a proportion of current stuttering research that emerges from national health service-sponsored research programs, or a professional environment in which most practitioners are trained in fairly uniform methods, such as Australia (Attanasio et al., 1996). Above all, at this stage in our development we must not confuse the concept "currently without substantial evidence" with "without substantial value" (Kerridge, Lowe, & Henry, 1998).

American researchers and clinicians, the apparent object of concern by authors of 28/3, face a number of interesting obstacles to rapid treatment efficacy research production, including more diverse training backgrounds (contrary to the opinions of Attanasio et al. (1996), not all represent the Iowa school via tangible training pedigrees), greater geographical distance among clinical programs, more diversely trained local practitioners to work with and internal competition for public health service grant support. Inequalities in the ease and pace with which therapy outcomes research can reach publication can be particularly problematic, because as Wampold et al. (1997) note, in outcomes research, "to receive a prize that is meaningful, the competitors must have a level playing field." This is also troublesome if one conceptualizes (wrongly) EBT to be some type of race, rather than an ongoing dialogue among researchers and clinicians.

It is important to realize that documenting that a program doesn't work (doesn't produce results) is different from having no evidence that it does work. "In EBP, there will always be interventions for which no evidence is (yet) available, but that is no reason to withhold the intervention. Where research provides unequivocal evidence of the absence of effects, the situation is much clearer" (van Weel & Knottnerus, 1999). Therefore, it is not yet appropriate to ask why clinicians use therapies having "less evidence" of effectiveness than proposed alternatives, as the authors of contributions to the *JFD* symposium suggest.

It will also be important to evaluate (or debunk) each opposing approach to practice on its own terms, rather than only as a synchronic comparison of recently complied efficacy data, since more than one approach is likely to eventually be validated, as noted earlier. Few competing approaches to a problem emerge and achieve equivalent levels of empirical support simultaneously, especially in a sea change of practice guidelines such as the new orientation toward EBP. Immediately below, I will argue that one inevitable outcome to the immediate and wholesale adoption of an EBP approach to treatment selection will be the uneven pace of comparably gathered efficacy data collection across treatment approaches, and across outcome criteria selected for analysis.

15. "Branding" of therapy approaches

Lidcombe joins a few other recently developed approaches for treatment of communication disorder that are either already proprietary or "branded" through allegiance to a therapist training regimen—FastForWord is a possible parallel for the treatment of SLI, as is Lee Silverman Voice Therapy. In evaluating comparable therapies that are "marketed" through training or franchising in other fields, Beutler (2000a, 2000b) and Rosen and

Davison (2003) express concern that we should not confuse empirically supported treatments with *empirically-supported principles of change*. An important result of a shift in focus to principles or mechanisms of change would be seen in training of future clinicians: since 'cookbooks' for all potential treatment regimens cannot be provided in advance, clinicians in training would need to "learn principles of change and the range of their applications" (Rosen & Davidson, 2003). As noted earlier, among the skills that Rosen & Davidson would teach is how to "dismantle" therapies to identify their process mechanisms. Only by teaching students such concepts, rather than therapies per se do we allow the field to advance in a scientific manner, because "principles of behavior change, after all, cannot be trademarked, for they belong to science."

A final concern that is likely to be mitigated by the passage of time is that Lidcombe in particular can show published data virtually exclusively from one research team (even "replications" in Great Britain involved a high degree of participation and authorship by the Sydney team). As Lohr et al. (2003) note, a treatment seeking to be considered "well-established" must additionally supported by research conducted by two independent sets of researchers. While the numbers of published articles on Lidcombe are currently large and impressive, they are still representative of a class of findings that Andrews et al. (1983) termed Class C (findings replicated by a single research center with no conflicting results). Without wide-spread open consideration of the possible effective approaches to stuttering treatment, legitimate treatment comparison studies will be impossible because of the well-known allegiance effect—clinicians with truly vested interests in the outcome of particular approaches have remarkable effects on outcomes. Some of the opinions expressed about alternative therapy approaches in 28/3 would seem to be potential markers of experimenter bias in testing the very question under consideration (Wampold, Ahn, & Coleman, 2001).

16. Some concluding thoughts

Bertrand Russell once observed that "The most savage controversies are those about matters as to which there is no good evidence either way." This might be an apt way of describing what we know about stuttering therapy at this time when all is said, done and weighed. Debate is fun for some of us, but:

> To practitioners, the continued jousting among researchers may seem tedious. When scientists debate clinical practice issues, there is often no middle ground. Scientists need to convince clinicians that their view is the only one that has both theoretical coherence and empirical evidence to support it. This often leads to the use of emotionally charged rhetoric and exaggerated claims of the dangers of using the wrong treatment approach and the benefits of using the right one. (Kamhi, 1999; p. 95)

We will probably not be able to change some of the inevitable differences in how scientists and clinicians view the major issues in evidence-based practice. What we need to do is find some basic concepts to agree on. Among them might first be an agreement that we need to keep working on the problem. It might be good to work together, since sample sizes required to show program component functionality may be enormous (Harrison et al., 2004). Second,

it seems clear that we—scientific clinicians and clinical scientists alike—should keep better data. Further, we need to increase the rate at which successful therapists publish the data. Throughout all this, we must remember to report individual differences even as we evaluate the group data. Ellis Weismer (1991) long ago advised advocates of opposing theoretical models of child language remediation to turn their attention to the variables that might predict which programs are relatively more efficient in producing specific desired outcomes, which may differ from child to child, rather than defend the premise that one type of approach was globally superior to any others. This is sound advice for the field of stuttering as well.

We need to join together to increase interest in therapy trials, both from researchers' and funders' perspectives. We need to investigate how we can increase partnering between the actual therapy providers and the publishing "researchers," who are not typically working together in the USA, unlike the case in Australia, for example. I believe that we need to show respect for the values of current providers (e.g., minimizing remarks suggesting that some seem to be endorsing 'an aluminum standard'; Ingham, 2003). These service providers, I believe, want the best for their clients and employ therapy approaches guided by this set of values. And I think we should let the data *and theory* guide further research and therapy but not dictate it, *yet*.

A real concern raised by Sailor and Stowe (2003) is that of viewing the recent popularity of EBP, especially as defined by the RCT, as an endgame for current policy. They evaluate the current status of educational research in ways that some of the authors of JFD 28/3 might appreciate: too often "exploratory or descriptive, because the literature is too thin or even nonexistent with regard to many of the issues being investigated . . . still in its adolescence compared with the fully matured science of medicine." They conclude, "It needs time and the opportunity to grow." The same is true in our field.

Suffice it to say, now is the time for conciliation and progress, not platforming. We need to stay, as much as possible, collegial. In a small field trying to enter the fray of EBP, some of the arguments raised in *JFD* 28/3 threaten to weaken rather than strengthen the field. In psychology, where the EBP debate is rampant, advocates of various positions are throwing large bodies of evidence at one another to argue margins of relative efficacy. I personally fear a disciplinary debate where one side claims the other brings virtually nothing to the table (e.g., one forum contributor advised readers to "quit reading" after consulting the preface to ASHA's *Guidelines for Stuttering Treatment*). I also fear that our field may well inherit a wasteland if it polarizes attempts to improve the evidence basis of practice for the field rather than seek to define a set of recommended practices at this point in time.

In this writer's opinion, the tone of some past debates is hardly likely to win over the non-believers (cf., Ingham & Cordes, 1999). Rather than castigate researchers and clinicians who do not universally and exclusively endorse some of the approaches recommended by authors in 28/3, we should be proud of those who do not leap to administer treatments inconsistent with the emerging scientific understanding of the basis of stuttering. To me, it implies that such individuals are engaged in reflective thinking and would, therefore, be receptive to strengthened accounts of the causative mechanisms by which supposedly effective and efficacious programs achieve their ends.

CONTINUING EDUCATION

QUESTIONS

1. What does the author conclude about the evidence-base for "traditional" approaches to stuttering therapy?
   a. They are poorly supported by the available data
   b. They are more strongly supported than operant options
   c. They are supported by data from fluency intervention and other allied disciplines, such as clinical psychology
   d. The evidence base is strong, but the basic research is not consistent with it
2. The author suggests that future therapy outcomes research needs to address which of the following factors that may influence treatment response?
   a. Spontaneous recovery
   b. Family history of chronicity
   c. Associated/comorbid conditions
   d. Perceived emotional and/or cognitive impact of the stuttering on everyday functioning
   e. All of the above
3. The author provides a number of reasons why theory and program dismantling are important in developing and evaluating treatment approaches. Which is NOT one of these reasons?
   a. Therapies that cannot be reconciled with theory have no value and should not be used
   b. Theory assists us in "fine-tuning" therapy approaches to maximize active components and minimize non-active components
   c. Theory may assist in encouraging practitioner "buy-in" in adopting new treatments
   d. Theory may help us distinguish between outcomes that are program-specific as opposed to clinician-specific
   e. Theory may help us to adapt programs to specific client needs
4. Which of the following is NOT one of the author's concerns in evaluating the evidence-base for a treatment approach?
   a. The degree to which it can be reconciled with existing data regarding the underlying nature of stuttering and characteristics of its development
   b. Whether or not a treatment approach has been evaluated by multiple, unaffiliated researchers/clinicians
   c. Whether or not a therapy approach can be dismantled into its active components
   d. The relative number of studies that have appeared to date documenting the program's effectiveness when compared to other programs
   e. All of the above
5. Which of the following statements best summarizes the author's reaction to the forum in 28/3?
   a. She fully agrees with it, and offers additional support for the viewpoints expressed in the issue
   b. She believes that there is an evidence-base for many approaches to stuttering intervention beyond those highlighted in the forum

c. She believes that most current disagreements about appropriate treatment for stuttering will be answered by larger and additional RCTs
d. She believes that no current therapies can be linked to theories about the nature of stuttering
e. She believes that no adequate treatment approaches for stuttering will be developed until we have a better understanding of the etiology of the disorder than we do know

References

Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, *48*, 251–257.

Anderson, J., & Conture, E. (2004). Sentence-structure priming in young children who do and do not stutter. *Journal of Speech, Language, and Hearing Research*, *47*(3), 552–571.

Andrews, G., & Craig, A. (1988). Prediction of outcome after treatment for stuttering. *The British Journal of Psychiatry*, *153*, 236–240.

Andrews, G., Craig, A., Feyer, A. M., Hoddinott, S., Howie, P., & Neilson, M. (1982). Stuttering: A review of research findings and theories circa. *Journal of Speech and Hearing Disorders*, *48*, 226–246.

Astin, J., Harkness, E., & Ernst, E. (2000). The efficacy of 'distant healing': A systematic review of randomized trials. *Annals of Internal Medicine*, *32*, 903–910.

Attanasio, J. (1999). Treatment of stuttering: Some reflections. In M. Onslow & A. Packman (Eds.), *The handbook of early stuttering intervention* (pp. 189–204). San Diego: Singular.

Attanasio, J., Onslow, M., & Menzies, R. (1996). Australian and United States perspectives on stuttering in preschool children. *Australian Journal of Communication Disorders*, *24*, 55–61.

Bandura, A., & Locke, E. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, *88*, 87–99.

Baumeister, H., Caspar, F., & Herziger, F. (2003). Treatment outcome study of the stuttering therapy summer camp 2000 for children and adolescents. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, *53*, 455–463 [English abstract from EBSCO research databases].

Benson, H., & Epstein, M. (1975). The placebo effect—A neglected asset in the care of patients. *Journal of American Medical Association*, *232*, 1225–1227.

Bernstein Ratner, N., & Guitar, B. (in press). Treatment of very early stuttering and parent-administered therapy: The state of the art. In N. Bernstein Ratner & Tetnowski (Eds.), *Stuttering research and practice*: *Volume* 2: *Contemporary Issues and Approaches*. Mahwah, NJ: Erlbaum.

Beutler, L. (2000a). Empirically based decision making in clinical practice. *Prevention and treatment* (Vol. 3) [Article 27]. Retrieved from the World Wide Web at www.journals.apa.org on December 26, 2004.

Beutler, L. E. (2000). David and Goliath: When psychotherapy research meets health care delivery systems. *American Psychologist*, *55*, 997–1007.

Bonelli, P., Dixon, M., Ratner, N. B., & Onslow, M. (2000). Child and parent speech and language following the Lidcombe Programme of early stuttering intervention. *Clinical Linguistics and Phonetics*, *14*(6), 427–446.

Bothe, A. K. (2003). Evidence-based treatment of stuttering: V. The art of clinical practice and the future of clinical research. *Journal of Fluency Disorders*, *28*(3), 247–258.

Buchel, C., & Sommer, M. (2004, February 17). What causes stuttering? *PLoS Biology*, *2*(2) [On-line]. e46. doi: 10.1371/journal.pbio.0020046.

Cho, Y., Smits, J., & Telch, M. (2004). The Speech Anxiety Thoughts Inventory: Scale development and preliminary psychometric data. *Behavior Research and Therapy*, *42*, 13–25.

Chwalisz, K. (2001). A common factors revolution: Let's not "Cut off our discipline's nose to spite its face". *Journal of Consulting Psychology*, *48*, 262–267.

Cobham, V., Dadds, M. R., & Spence, S. H. (1998). The role of parental anxiety in the treatment of childhood anxiety. *Journal of Consulting and Clinical Psychology*, *66*, 893–905.

Cockburn, J. (2004). Adaption of evidence into practice—can change be sustainable? *Medical Journal of Australia*, *180*, 66–67.

Conture, E. (2001). *Stuttering: Its nature, diagnosis, and treatment*. Needham Heights, MA: Allyn & Bacon.
Conture, E., & Melnick, K. (1999). The parent-child group approach to stuttering in preschool children. In M. Onslow & A. Packman (Eds.), *The handbook of early stuttering intervention*. San Diego: Singular.
Conture, E., Zackheim, C., Anderson, J., & Pellowski, M. (2004). Linguistic processes and childhood stuttering: Many's a slip between intention and lip. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 253–282). Nijmegen: Nijmegen University Press.
Cook, F., Millard, S., & Nicholas, A. (2004). *A long-term efficacy study of parent-child interaction therapy*. American Speech, Language Hearing Association annual convention, Philadelphia.
Cooper, E., & Cooper, C. (1996). Clinician attitudes towards stuttering: Two decades of change. *Journal of Fluency Disorders*, *21*, 119–135.
Craig, A. (1998). Relapse following treatment for stuttering: A critical review and correlative data. *Journal of Fluency Disorders*, *23*, 1–30.
Craig, A., Hancock, K., Chang, E., McCready, C., Shepley, A., McCaul, A., et al. (1996). A controlled clinical trial for stuttering in persons aged 9–14 years. *Journal of Speech and Hearing Research*, *39*, 808–826.
Crichton-Smith, I., Wright, J., & Stackhouse, J. (2003). Attitudes of speech and language therapists towards stammering: 1985 and 2000. *International Journal of Communication Disorders*, *38*, 213–234.
Cuadrado, E., & Weber-Fox, C. (2003). Atypical syntactic processing in individuals who stutter: Evidence from event-related brain potentials and behavioral measures. *Journal of Speech, Language and Hearing Research*, *46*, 960–976.
Culatta, R., & Goldberg, S. (1995). *Stuttering therapy: An integrated approach to theory and practice*. Boston: Allyn & Bacon.
Culpepper, L., & Gilbert, T. (1999). Evidence and ethics. *Lancet*, *353*, 829–831.
Curlee, R., & Yairi, E. (1997). Early intervention with early childhood stuttering: A critical examination of the data. *American Journal of Speech-Language Pathology*, *6*, 8–18.
Davis, D., Ciurea, I., Flanagan, T., & Perrier, L. (2004). Solving the information overload problem: A letter from Canada. *Medical Journal of Australia*, *180*, 68–71.
Delprato, D. J. (2001). Comparisons of discrete-trial and normalized behavioral language intervention for young children with autism. *Journal of Autism and Developmental Disorders*, *31*(3), 315–325.
Dishion, T. J., & Patterson, G. R. (1999). Model building in developmental psychpathology: A pragmatic approach to understanding and intervention. *Journal of Clinical Child Psychology*, *28*, 502–512.
Eifert, G. (1996). More theory-driven and less diagnosis-based behavior therapy. *Journal of Behaviour Therapy and Experimental Psychiatry*, *27*, 75–86.
Ellis Weismer, S. (1991). Child language intervention: Research issues on the horizon. In J. Miller (Ed.), *Research on child language disorders* (pp. 233–241). Austin, TX: Pro-Ed.
Evans, I. (1996). Individualizing therapy, customizing clinical science. *Journal of Behavior Therapy and Experimental Psychiatry*, *27*, 99–105.
Fava, G., Grandi, S., Rafanelli, C., Ruini, C., Conti, S., & Belluardo, P. (2001). Long-term outcome of social phobia treated by exposure. *Psychological Medicine*, *31*, 899–905.
Feinstein, A., & Horwitz, R. (1997). Problems in the 'evidence' of 'evidence-based medicine'. *American Journal of Medicine*, *103*, 529–535.
Finn, P. (1996). Establishing the validity of recovery from stuttering without formal treatment. *Journal of Speech, Language and Hearing Research*, *39*, 1171–1181.
Finn, P. (2003). Evidence-based treatment of stuttering: II. Clinical significance of behavioral stuttering treatments. *Journal of Fluency Disorders*, *28*(3), 209–218.
Follette, W., & Beitz, K. (2003). Adding a more rigorous scientific agenda to the empirically supported treatment movement. *Behavior Modification*, *27*, 369–386.
Forster, D. C., & Webster, W. G. (1991). Concurrent task interference in stutterers: Dissociating hemispheric specialization and activation. *The Canadian Journal of Psychology*, *45*, 321–335.
Forster, D. C., & Webster, W. G. (2001). Speech-motor control and interhemispheric relations in recovered and persistent stuttering. *Developmental Neuropsychology*, *19*, 125–145.
Gottwald, S., & Starkweather, C. W. (1999). Stuttering prevention and early intervention: A multi-process approach. In M. Onslow & A. Packman (Eds.), *The handbook of early stuttering intervention*. San Diego: Singular.

Graves, K. (2003). Social cognitive theory and cancer patients' quality of life: A meta-analysis of psychosocial intervention components. *Health Psychology*, *22*, 210–219.
Hancock, K., Craig, A., McCready, C., McCaul, A., Costello, D., Campbell, K., et al. (1998). Two- to six-year controlled trial stuttering outcomes for children and adolescents. *Journal of Speech, Language and Hearing Research*, *41*, 1242–1253.
Harris, V., Onslow, M., Packman, A., Harrison, E., & Menzies, R. (2002). An experimental investigation of the impact of the Lidcombe Program on early stuttering. *Journal of Fluency Disorders*, *27*, 203–214.
Harrison, E., Onslow, M., & Menzies, R. (2004). Dismantling the Lidcombe Program of early stuttering intervention. *International Journal of Language and Communication Disorders*, *39*, 257–268.
Heimberg, R. (2002). Cognitive–behavior therapy for social anxiety disorder: Current status and future directions. *Biological Psychiatry*, *51*, 101–108.
Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, *128*, 774–795.
Ingham, J. C. (2003). Evidence-based treatment of stuttering: I. Definition and application. *Journal of Fluency Disorders*, *28*(3), 197–207.
Ingham, R. J., & Cordes, A. K. (1999). On watching a discipline shoot itself in the foot: Observations on current trends in stuttering treatment research. In N. B. Ratner & C. Healey (Eds.), *Stuttering research and practice: Bridging the gap* (pp. 211–230). Mahwah, NJ: Lawrence Erlbaum.
James, J., Ricciardelli, L., Rogers, P., & Hunter, C. (1989). A preliminary analysis of the ameliorative effects of time-out from speaking on stuttering. *Journal of Speech, Language and Hearing Research*, *32*, 604–610.
Kamhi, A. (1999). To use of not to use: Factors that influence the selection of new treatment approaches. *Language Speech and Hearing Services in Schools*, *30*, 92–98.
Kerridge, I., Lowe, M., & Henry, D. (1998). Ethics and evidence-based medicine. *British Medical Journal*, *316*, 1151–1153.
Kleinow, J., & Smith, A. (2000). Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. *Journal of Speech, Language and Hearing Research*, *43*, 548–560.
Kuhr, A. (1994). The rise and fall of operant programs for the treatment of stammering. *Folia Phoniatrica Logopedics*, *46*(5), 232–240.
Lambert, M., & Barley, D. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory/research/practice/training*, *38*, 357–361.
Langevin, M., & Kully, D. (2003). Evidence-based treatment of stuttering: III. Evidence-based practice in a clincial setting. *Journal of Fluency Disorders*, *28*(3), 219–236.
Law, J., Garrett, Z., & Nye, C. (2003). Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database Systemic Reviews*, *3*, CD004110.
Levant, R. (2004). The empirically validated treatments movement: A practitioner/educator perspective. *Clinical Psychology: Science and Practice*, *11*, 219–224.
Lipsey, M., & Cordray, D. (2000). Evaluation methods for social intervention. *Annual Review of Psychology*, *51*, 345–375.
Lohr, J., DeMaio, C., & McGlynn, F. (2003). Specific and nonspecific treatment factors in the experimental analysis of behavioral treatment efficacy. *Behavior Modification*, *27*, 322–368.
Michie, S., & Abraham, C. (2004). Interventions to change health behaviors: Evidence-based or evidence-inspired? *Psychology and Health*, *19*, 29–49.
Multon, K., Brown, S., & Lent, R. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*.
Nail-Chiwetalu, B., & Bernstein Ratner, N. (2003, April). *Fostering information literacy competency in communication sciences and disorders*. Council of Academic Programs in Communication Sciences and Disorders, Albuquerque. Full text available at http://www.capcsd.org/proceedings/2003/talks/chiwetalu2003.pdf.
Nail-Chiwetalu, B., & Bernstein Ratner, N. Information literacy for speech-language pathologists: A key to evidence-based practice. Submitted for publication.
Onslow, M. (2003). Evidence-based treatment of stuttering: IV. Empowerment through evidence-based treatment practices. *Journal of Fluency Disorders*, *28*(3), 237–245.
Onslow, M., Packman, A., & Harrison, E. (2003). *The Lidcombe Program of early stuttering intervention: A clinician's guide*. Austin, TX: Pro-Ed.

Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2002). "Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews*, (4).

Pope, C. (2003). Resisting evidence: The study of evidence-based practice as a contemporary social movement. *Health: An interdisciplinary journal for the study of health, illness and medicine*, *7*, 267–283.

Ramig, P., & Bennett, E. (1997). Clinical management of children: Direct management strategies. In R. Curlee & G. Siegel (Eds.), *The nature and treatment of stuttering: New directions* (2nd ed.). Boston: Allyn & Bacon.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). NY: Academic Press.

Rosch, E., & Mervis, C. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Rosen, G., & Davison, G. (2003). Psychology should list empirically supported principles of change (ESPs) and not credential trademarked therapies or other treatment packages. *Behavior Modification*, *27*, 300–312.

Rosenthal, T. (1994). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science*, *5*, 127–134.

Runyan, C., & Runyan, S. (1999). The fluency rules program. In M. Onslow & A. Packman (Eds.), *The handbook of early stuttering intervention*. San Diego: Singular.

Sailor, W., & Stowe, M. (2003). The relationship of inquiry to public policy. *Research and Practice for Persons with Severe Disabilities*, *28*, 148–152.

Salkovskis, P. (2002). Empirically-grounded clinical interventions: Cognitive–behavioral therapy progresses through a multi-dimensional approach to clinical science. *Behavioral and Cognitive Psychology*, *30*, 3–9.

Skinner, B. (1950). Are theories of learning necessary? *Psychological Review*, *57*, 193–216.

Smith, A., & Kelly, E. (1997). Stuttering: A dynamic, multifactorial model. In R. Curlee & G. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (pp. 204–217). Boston: Allyn & Bacon.

St. Louis, K. (in press). Measurement options in fluency disorders. In N. Ratner & J. Tetnowski (Eds.), *Stuttering research and practice II: Contemporary issues and approaches*. Mahwah, NJ: Erlbaum.

Swinkles, A., Albarran, J., Means, R., Mitchell, T., & Stewart, M. (2002). Evidence-based practice in health and social care: Where are we now? *Journal of Interprofessional Care*, *16*, 335–347.

Vandenbroucke, J. (2001). In defense of case reports and case series. *Annals of Internal Medicine*, *134*, 330–334.

Van Riper, C. (1973). *The treatment of stuttering*. Englewood Cliffs, NJ: Prentice Hall.

van Weel, C., & Knottnerus, J. (1999). Evidence-based interventions and comprehensive treatment. *The Lancet*, *353*, 916–919.

Vos, R., Willems, D., & Houtepen, R. (2004). Coordinating the norms and values of medical research, medical practice and patient worlds—the ethics of evidence-based medicine in orphaned fields of medicine. *Journal of Medical Ethics*, *30*, 166–170.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods and findings*. Hillsdale, NJ: Laurence Erlbaum Associates.

Wampold, B., Ahn, H., & Coleman, H. (2001). Medical model as metaphor: Old habits die hard. *Journal of Counseling Psychology*, *48*(3), 268–273.

Wampold, B., Mondin, G., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes". *Psychological Bulletin*, *122*, 203–215.

Watkins, R., & Johnson, B. (2004). Language abilities in children who stutter: Toward improved research and clinical applications. *Language, Speech and Hearing Services in Schools*, *35*, 82–89.

Watson, B., Freeman, F., Devous, M., Chapman, S., Finitzo, T., & Pool, K. (1994). Linguistic performance and regional cerebral blood flow in persons who stutter. *Journal of Speech and Hearing Research*, *37*, 1221–1228.

Woodman, T., & Hardy, L. (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of Sports Sciences*, *21*, 443–457.

Yairi, E., & Ambrose, N. (2004). *Early childhood stuttering*. Pro-Ed: Austin, TX.

Yairi, E., Ambrose, N. G., Paden, E. P., & Throneburg, R. N. (1996). Predictive factors of persistence and recovery: Pathways of childhood stuttering. *Journal of Communication Disorders*, *29*, 51–77.

Yaruss, J. S. (1998). Describing the consequences of disorders: Stuttering and the international classification of impairments, disabilities and handicaps. *Journal of Speech, Language and Hearing Research*, *41*, 249–257.

Yaruss, J. S., Quesal, R., Reeves, L., Molt, L., Kluetz, B., Caruso, A., et al. (2002). Speech therapy and support group experiences of people who participate in the National Stuttering Association. *Journal of Fluency Disorders, 27*, 115–135.
Zadeh, L. (1965). Fuzzy sets. *Information Control, 8*, 338–353.

Nan Bernstein Ratner is Professor and Chairman, Department of Hearing and Speech Sciences, University of Maryland at College Park. An ASHA Fellow and Board Recognized Fluency Specialist with degrees in child development, speech-language pathology and psycholinguistics, she is the author of numerous publications that address child language acquisition and developmental language and fluency disorders.