



Pharmacological Approaches to Stuttering Treatment: Reply to Meline and Harn (2008)

Purpose: To address questions raised by T. Meline and W. E. Harn (2008) in their critique of our previous article (A. K. Bothe, J. H. Davidow, R. E. Bramlett, D. M. Franic, & R. J. Ingham, 2006).

Method: Additional information is provided to address several issues raised by Meline and Harn.

Results and Conclusions: Our previous systematic review omitted 1 relevant article about the use of olanzapine in stuttering, but the minimal effectiveness and the known serious side effects of this drug limit the implications of this omission. While we do not agree with many of Meline and Harn's critiques of our review, we do agree with them that several larger points raise interesting questions about the structure, analysis, and usefulness of literature reviews in stuttering and in other areas. Fundamentally, we reassert our agreement with Meline and Harn that there is insufficient evidence to support the use of existing pharmacological agents in the treatment of stuttering.

Key Words: stuttering, olanzapine, systematic review

Meline and Harn (2008) raised several interesting issues about pharmacological treatments for stuttering, about our recent review of the literature in this area (Bothe, Davidow, Bramlett, Franic, & Ingham, 2006), and about reviews of the literature in general. We are pleased to have this opportunity to respond to, and expand on, their key points.

Olanzapine in the Treatment of Stuttering

Among the most straightforward of Meline and Harn's points is their first: that our review did not include or refer to Maguire et al.'s (2004) double-blind, placebo-controlled trial of the use of olanzapine for developmental stuttering. In the terms used in our review, Maguire et al.'s study met three out of five methodological criteria (random assignment to groups, pretreatment and posttreatment data, and reliability data, but no speech rate or speech naturalness data and no beyond-clinic data). It should have been included in our review as a relatively well-designed study and discussed in more detail in our article. Our only response to Meline and Harn on this point is simply to agree that we missed it (and, indeed, to wonder with them how we could possibly have missed it).

More interesting than this study's absence, however, are its implications. Maguire et al. (2004) concluded that olanzapine was "statistically superior to placebo" (p. 65) for three different dependent variables and described it as "useful" (p. 66) in the treatment of stuttering. In the terms

used in our review, however, Maguire et al.'s results met neither of our two outcomes criteria (stuttering, disfluency, or any such measure below 5%, or documented improvement in any nonspeech variable) at either of the two times we assessed (immediately posttreatment and 6 months post-treatment). Maguire et al.'s results also did not meet our secondary outcome criterion of a 50% reduction in stuttering (or any such variable), which we assessed in case the absolute 5% criterion could be viewed as too stringent.

Meline and Harn questioned whether Maguire et al.'s results might nevertheless be interpreted as positive. They mentioned the quality of life literature in this context, writing that "a standardized effect size of 0.50 (half a standard deviation) is equivalent to increasing the success rate from 25% to 75%" (p. 95). Equating standard deviations of test scores and effect sizes is problematic, however; the intended meaning within the quality of life literature was that 0.5 SD represents a change of 8% of the quality of life scale (Sloan, Symonds, Vargas-Chanes, & Fridley, 2003). Overall, in addressing our choice of outcomes criteria, Meline and Harn suggested that some improvements can be important even if they do not represent a complete cure, and of course we agree. The point of our review, however, was not to identify all treatments that might represent some improvement to someone, but to identify those treatments that met one set of reasonable, objective, and actually rather lenient outcomes criteria.

It is also important to note that Maguire et al. (2004, p. 66) described olanzapine as "well-tolerated." Meline and Harn, similarly, quoted both a well-known stuttering textbook (Guitar, 2006) and a new chapter from a book on stuttering (Ludlow, 2006) as suggesting that side effects are not problematic for olanzapine, "indicating that further research would be warranted" (Ludlow, 2006, p. 248; Meline & Harn, 2008, p. 93). The known side effects of olanzapine are serious and numerous, however, even if they have not been reported in the relatively small studies of its use for stuttering. It is critical for readers to be aware, for example, that a "Dear Doctor" letter about olanzapine (marketed by Eli Lilly as Zyprexa) was released on March 1, 2004. "Dear Doctor" letters are typically required by the U.S. Food and Drug Administration to be forwarded to physicians by manufacturers if there have been problems with a product. In this case, it was noted that "hyperglycemia, in some cases extreme and associated with ketoacidosis or hyperosmolar coma or death" (U.S. Food and Drug Administration, 2004, para. 2) had been reported as associated with use of Zyprexa (olanzapine). More recently, a *New York Times* article ("Lilly Settles," 2007) reported that the company has now settled lawsuits involving 18,000 individuals who claimed they had "developed diabetes or other disorders after taking Zyprexa"

(p. C1). This means that, along with earlier settlements, at least 28,500 individuals have received compensation from the manufacturer for claims that they were injured by Zyprexa.

Such issues are obviously serious and raise questions in our minds about describing this agent as “well tolerated.” While there may yet be situations in which olanzapine could be used, its known side effects must be considered in any calls for “further research.” We might suggest that any further research of its use in stuttering must, at a minimum, include repeated fasting and postprandial blood glucose testing; the blood testing reported by Maguire et al. was insufficient for an agent known to have hyperglycemic side effects (Cavalot et al., 2006; Kendall, 2005; Liebl, 2003). In addition, because this agent is an antipsychotic drug, further information is needed regarding its cognitive impact in nonschizophrenic patients before it could reasonably be recommended as a treatment for stuttering.

Responses to Points Raised by Meline and Harn

The bulk of Meline and Harn’s letter proceeded to critique the Bothe, Davidow, Bramlett, Franic, and Ingham (2006) article in several larger senses, or to review our review. Their efforts in this area raise several interesting issues.

Previous summaries. Meline and Harn noted, to begin with, that the review’s conclusions were more negative than the conclusions reached in some previous reviews of related literature. As they also noted, we did address this question in our original article, which also discussed part of the answer: It has become relatively well established that more stringently conducted clinical research often leads to more negative results than those obtained in smaller or less controlled studies. Because our review incorporated trial quality assessment, and based its results on those articles that met a trial quality criterion, its conclusions could be predicted to be more negative than the conclusions reached in less selective reviews. We also must point out that our conclusions did not differ from all previous conclusions. Meline and Harn quoted Ingham’s (1984, p. 413) suggestion, for example, that some pharmacological agents “are undoubtedly reliable ... for reducing state anxiety.” That particular quote was less relevant to the issues at hand, however, than the same chapter’s succinct and quite negative summary a few pages later: “For the present, it is safe to conclude there is little evidence that stuttering can be usefully treated by drug therapy” (Ingham, 1984, p. 424). In fact, if we have learned anything from the drug treatments of stuttering, especially investigations into the efficacy of anxiolytic drugs, it must be that stuttering is neither inextricably nor functionally related to anxiety. There is simply no evidence that anxiolytics have helped to eliminate or reduce stuttering.

Search strategies, inclusion criteria, and references. Meline and Harn also raised some specific questions about the databases and search strategies we had used, and they questioned the information we provided about the articles selected for review. The answer to most of their questions is simply that one must draw a line somewhere: We selected

1970 because we wanted to address stuttering treatments developed since the introduction of prolonged speech; we searched certain journals and used the key words we selected because they seemed a relevant and comprehensive subset (and we note that Meline and Harn suggested only the one article that we missed, an omission that was a human error, not a flaw in our selection of databases or a problem that would have been solved by more databases); and we chose, given the structure of our project and after consultation with the editors during the review process, to make the larger list of all the articles we had consulted available to interested readers upon request rather than taking the journal space to publish them. Meline and Harn are correct in noting that we did not describe in detail the process through which we removed inappropriate articles (e.g., those that were about pharmacological agents that caused stuttering-like speech) from the initial pool retrieved by our searches; that task, like the others in our project, was completed by multiple readers.

Meline and Harn also commented on the fact that our review included only published articles. They are correct in noting that some definitions of systematic reviews emphasize the importance of including unpublished material, but there is by no means universal agreement about this point. One important reason to seek unpublished literature is because of an accepted general tendency for positive results to be published and for negative results not to be. Drawing the conclusion from published literature that a treatment is effective may be problematic, therefore, if there is important information to the contrary in unpublished sources. This problem did not occur for the Bothe et al. review, given our conclusion of the relative ineffectiveness of the pharmacological agents.

Several other lines of reasoning also suggest an emphasis on published literature; it is available to all readers, for example, and it has already passed the test of peer review. The decision to concentrate on published literature also draws an objective line that avoids the complex problems inherent in deciding which unpublished literature to include and which to exclude, because as reasonable as the inclusion of some unpublished information might be, the decision to include all unpublished information about a given topic will usually be unreasonable. Biddle, Watson, Hooper, Lohr, and Sutton (2002), as one example, in a relevant recent review, initially concentrated on published peer-reviewed information only, but then made the decision to include published but not peer-reviewed test manuals as another source of information for some specific reasons. While other lines can be defensible and could certainly serve legitimate purposes, in summary, it is not an error to draw a line between published peer-reviewed research and other types of information.

Outcome criteria. Meline and Harn discussed the possibility that our use of a 5% criterion “biased” the conclusions drawn because it is an “imprecise measure of treatment success or failure” (p. 94). While we certainly agree wholeheartedly with the latter, we do not believe that selecting a binary outcome criterion, or using speech outcomes as one of two types of outcome criteria in a review of stuttering treatment, creates “bias” in a review. Quite the opposite: Such a method provides an objective (i.e., specifiable and repeatable) method through which a set of articles can be divided into those that meet a relevant

outcome criterion and those that do not. There are other ways to create reviews of the literature, but such methods as meta-analysis (combining results mathematically) or calculation of effect sizes (as Meline and Harn discussed) impose multiple restrictions on the articles that can be included (because of the need for individual data or variability data) in a manner that we believed to be too restrictive for our purposes in reviewing stuttering treatment research.

It is also important to note that Meline and Harn (2008, p. 94) summarized our review as having required “the frequency of ‘stuttering’ to be reduced to 5% or less.” This summary is misleading. We used a 5% criterion for any measure of disfluency, nonfluency, stuttering, or any reported variable; we investigated the alternative possibility that any such measure had been reduced by half if the absolute 5% criterion were not met; we identified all studies where pretreatment values had been below 5% to assess them separately; and we also looked for any documented pretreatment to posttreatment change in any social, emotional, or cognitive variable. We certainly agree with Kazdin’s (2001, p. 461) warning, quoted by Meline and Harn in this context, that “we cannot merely decide a criterion as a cutoff and assume those who have made the cutoff have improved.” Kazdin’s (2001) point is not that no criterion or cutoff should ever be used, however, only that one must also consider the impact of treatment on quality of life or patient preferences as well. In stuttering, one of the many possible variables that must be assessed in comparing treatment options is whether they tend to result in improved speech; our review did this. It is also reasonable to address other variables, and our review did that as well (as has some of our other work; see Bramlett, Bothe, & Franic, 2006; Franic & Bothe, 2008; Ingham & Cordes, 1997).

Sample size, external validity, and confounding variables. Finally, Meline and Harn mentioned several problems with many clinical studies that they believed the Bothe et al. reviews should have emphasized. The problems discussed in this section of Meline and Harn’s letter are indeed problems with many studies. They were not addressed in detail in our original articles in part for reasons of space and in part because we sought to use a relatively short list of extremely basic methodological criteria to assess all studies in an objective way.

Reviewing Reviews of Reviews of Reviews: Criteria for the Criteria for the Criteria

The most intriguing feature of Meline and Harn’s letter may be that it parallels so closely several of the complex issues we also faced in developing and reporting our systematic review articles. The essential goal of our project, which was a systematic review with trial quality assessment, was to identify those stuttering treatment research articles that had been done “well enough” that we believed they could reasonably serve as a body of literature to be summarized and potentially acted upon. Because there is no single definition of “well enough,” we turned to multiple previous sources and assembled five common methodological requirements, and we also compared the results from those criteria with the results obtained using two existing

sets of criteria (see Bothe, Davidow, Bramlett, & Ingham, 2006, esp. Appendixes A and B).

Meline and Harn, in a parallel fashion, sought to determine whether our review had been done “well enough.” Their efforts to define “well enough” led them to Oxman and Guyatt’s (1988, p. 698, Table I) “guidelines for assessing research reviews” and to Oxman et al.’s (1991, p. 92, Table 1) and Oxman and Guyatt’s (1991, p. 1272, Table 1) related “criteria for assessing the scientific quality of research overviews.” Their abstract stated that they had adopted “A. D. Oxman and G. H. Guyatt’s (1998) guidelines ... and A. D. Oxman and G. H. Guyatt’s (1991) criteria” for their review of our article. The structure of their letter was not based on Oxman and Guyatt’s lists, however; instead, Meline and Harn seem to have structured their letter around a Web site (Davies & Crombie, undated, cited by Meline & Harn at their p. 94) and around a previous Meline (2006) article. Their intent, as described in their abstract, to use a “7-point scale from *extensive flaws* on the high end to *minimal flaws* on the low end” (Meline & Harn, 2008, p. 93) also does not seem to appear explicitly in their letter, nor does such a scale appear anywhere in the Oxman and Guyatt (1988, 1991) or Oxman et al. (1991) articles.

The structure of their letter thus leads to some interesting observations. First, it is clearly possible to review a set of studies, or to review a review, using more than one set of criteria, and identifying or establishing those criteria is not necessarily straightforward. It is also possible, even given the same criteria, to reach different conclusions. As an example, we must respectfully disagree with Meline and Harn’s (2008, p. 94) conclusion that our review “lacked transparency” in a way that created “major flaws.” We specified the publication years, publication language, search strategy, and topics covered by our project; we published references to all the articles that met our methodological inclusion criteria, which were the articles our review discussed; we offered to make lists of all the articles we had initially reviewed available to interested readers (and we have in fact since done so); and we described in detail the methods and the results from the multiple reviewers who assessed each article using a written data-extraction instrument. That it would have been possible to select different publication years, different publication languages, or different search strategies is not a problem of the transparency of our review; it is simply a statement that other authors or researchers could have made other equally defensible decisions.

One problem with assessing a review of a review, however, is that these issues are in danger of becoming the old joke represented visually by a person looking at a picture of herself looking at a picture of herself looking at a picture of herself. The present authors enjoy these issues as academic exercises, and we suspect Meline and Harn do as well. In the particular case of pharmacological treatments for stuttering, we agree with Meline and Harn’s basic conclusion that “the existing evidence for the use of pharmacological agents with persons who stutter is insufficient to recommend them in practice” (p. 95). We cannot agree with some of the other implications in their final paragraph, such as that being “acceptable” to clients can be equated with “no undesirable side effects.” Undesirable side effects are common

and do not contraindicate the use of a drug, as long as the benefits are viewed as outweighing the risks.

Such minor differences aside, we appreciate Meline and Harn's careful reading of our articles. We also appreciate their help in bringing the problems of olanzapine to light for speech-language pathologists, and we appreciate the opportunity to address the larger issues raised by their questions. We also suspect that Meline and Harn would agree with us that all of us who have the luxury of expending energy on such exercises as reviewing reviews of reviews of reviews also have the responsibility to ensure that our efforts lead, somehow, to improved assessment, treatment, and management methods for the people living with the disorders we study. We hope that this exchange of letters might represent some small progress in that direction.

Anne K. Bothe
Duska M. Franic
The University of Georgia, Athens

Roger J. Ingham
University of California, Santa Barbara

Jason H. Davidow
Hofstra University, Hempstead, NY

References

- Biddle, A., Watson, L., Hooper, C., Lohr, K. N., & Sutton, S. F.** (2002). *Criteria for determining disability in speech-language disorders* (Evidence Report/Technology Assessment No. 52; Prepared by the University of North Carolina Evidence-Based Practice Center under Contract No. 290-97-0011; AHRQ Publication No. 02-E010). Rockville, MD: Agency for Healthcare Research and Quality.
- Bothe, A. K., Davidow, J. H., Bramlett, R. E., Franic, D. M., & Ingham, R. J.** (2006). Stuttering treatment research 1970–2005: II. Systematic review incorporating trial quality assessment of pharmacological approaches. *American Journal of Speech-Language Pathology, 15*, 342–352.
- Bothe, A. K., Davidow, J. H., Bramlett, R. E., & Ingham, R. J.** (2006). Stuttering treatment research 1970–2005: I. Systematic review incorporating trial quality assessment of behavioral, cognitive, and related approaches. *American Journal of Speech-Language Pathology, 15*, 321–341.
- Bramlett, R. E., Bothe, A. K., & Franic, D. F.** (2006). Using preference based measures to assess quality of life in stuttering. *Journal of Speech, Language, and Hearing Research, 49*, 381–394.
- Cavalot, F., Petrelli, A., Traversa, M., Bonomo, K., Fiora, E., Conti, M., et al.** (2006). Postprandial blood glucose is a stronger predictor of cardiovascular events than fasting blood glucose in type 2 diabetes mellitus, particularly in women: Lessons from the San Luigi Gonzaga Diabetes Study. *The Journal of Clinical Endocrinology and Metabolism, 91*, 813–819.
- Franic, D. M., & Bothe, A. K.** (2008). Psychometric evaluation of condition-specific instruments used to assess health-related quality of life, attitudes, and related constructs in stuttering. *American Journal of Speech-Language Pathology, 17*, 60–80.
- Guitar, B.** (2006). *Stuttering: An integrated approach to its nature and treatment* (3rd ed.). Philadelphia: Lippincott, Williams & Wilkins.
- Ingham, R. J.** (1984). *Stuttering and behavior therapy: Current status and experimental foundations*. San Diego, CA: College-Hill.
- Ingham, R. J., & Cordes, A. K.** (1997). Self-measurement and evaluating stuttering treatment efficacy. In R. F. Curlee & G. S. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (Rev. ed., pp. 413–437). Boston: Allyn & Bacon.
- Kazdin, A. E.** (2001). Almost clinically significant ($p < .10$): Current measures may only approach clinical significance. *Clinical Psychology: Science and Practice, 8*, 455–462.
- Kendall, D. M.** (2005). Postprandial blood glucose in the management of Type 2 Diabetes: The emerging role of incretin mimetics. *Medscape Diabetes & Endocrinology, 7*(2), Retrieved August 30, 2007, from www.medscape.com/viewarticle/515693.
- Liebl, A.** (2003). Postprandial blood glucose as the driving force in pathogenesis of type 2 diabetes. *Medizinische Klinik, 98*(Supp. 1), 7–11.
- Lilly settles with 18,000 over Zyprexa.** (2007, January 7). *New York Times*, C1.
- Ludlow, C. L.** (2006). Neuropharmacology of stuttering: Concepts and current findings. In N. B. Ratner & J. Tetnowski (Eds.), *Current issues in stuttering research and practice* (pp. 239–254). Mahwah, NJ: Erlbaum.
- Maguire, G. A., Riley, G. D., Franklin, D. L., Maguire, M. E., Nguyen, C. T., & Brojeni, P. H.** (2004). Olanzapine in the treatment of developmental stuttering: A double-blind, placebo-controlled trial. *Annals of Clinical Psychiatry, 16*, 63–67.
- Meline, T.** (2006). Selecting studies for systematic review: Inclusion and exclusion criteria. *Contemporary Issues in Communication Science and Disorders, 33*, 21–27.
- Meline, T., & Harn, W. E.** (2008). Comments on Bothe, Davidow, Bramlett, Franic, and Ingham (2006). *American Journal of Speech-Language Pathology, 17*, 93–97.
- Oxman, A. D., & Guyatt, G. H.** (1988). Guidelines for reading literature reviews. *Canadian Medical Association Journal, 138*, 697–703.
- Oxman, A. D., & Guyatt, G. H.** (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology, 44*, 1271–1278.
- Oxman, A. D., Guyatt, G. H., Singer, J., Goldsmith, C. H., Hutchison, B. G., Milner, R. A., & Streiner, D. L.** (1991). Agreement among reviewers of review articles. *Journal of Clinical Epidemiology, 44*, 91–98.
- Sloan, J., Symonds, T., Vargas-Chanes, D., & Fridley, B.** (2003). Practical guidelines for assessing the clinical significance of health-related quality of life changes in clinical trials. *Drug Information Journal, 37*, 23–31.
- U.S. Food and Drug Administration.** (2004, March 1). *Safety data on Zyprexa® (olanzapine)—hyperglycemia and diabetes* [Dear Doctor letter]. Retrieved January 8, 2007, from the U.S. Food and Drug Administration MedWatch Web site: www.fda.gov/medwatch/SAFETY/2004/zyprexa.htm.

Received September 11, 2007

Accepted October 31, 2007

DOI: 10.1044/1058-0360(2008/010)

Contact author: Anne K. Bothe, Department of Communication Sciences and Special Education, 516 Aderhold Hall, The University of Georgia, Athens, GA 30602.
E-mail: abothe@uga.edu.