This is an Accepted Manuscript of an article published by Taylor & Francis in the *Intl. Journal of Clinical and Experimental Hypnosis* in 2022, available online:
https://www.tandfonline.com/doi/full/10.1080/00207144.2022.2104648

For the final typeset article or reference to this article please note the following reference:
Intl. Journal of Clinical and Experimental Hypnosis, 70(3): 220-226, 2022

DOI: 10.1080/00207144.2022.2104648

The "Whack-a-Mole" Challenge of Hypnosis Research: A Commentary Regarding "Guidelines for the Assessment of Efficacy of Clinical Hypnosis Applications"

Michael D. Yapko, Ph.D.

ABSTRACT

In this short commentary, I have acknowledged the merits of trying to establish treatment guidelines for the use of hypnosis in treatment and applauded the efforts and intentions of the Task Force for Establishing Efficacy Standards for Clinical Hypnosis. I have identified a few of the complex issues in trying to promote guidelines for conducting research and clinical practice in the domain of hypnosis; these include the difficulties in defining hypnosis and hypnotically-based interventions, the divergent ways hypnosis is applied in actual practice by clinicians who rely on their own understandings and biases in designing and delivering hypnosis, and the inevitable variations in skill level across practitioners. To their credit, the Task Force has considered these and other practical issues in their approach to formulating guidelines.

"Despite the very devious and unscientific history of hypnotism, there is excellent reason to expect a decided change for the better."

Clark Hull (1933/2002; p. 21)

"Whack-a-Mole" is a colloquial expression that characterizes a situation as next-to-impossible to resolve because as soon as you successfully complete one aspect of it (i.e., whack one mole), another one pops up, frustrating any possibility of a satisfying full resolution. It seems

an apt metaphor for the difficulties faced in trying to resolve critically important questions about the merits of hypnosis in treating physiological and psychological problems.

CONTACT: Michael D. Yapko email: michael@yapko.com Address: Private Practice, P.O. Box 487, Fallbrook, CA. 92088-0487. USA

The attempts over the last century, beginning with the pioneering book by Clark Hull, *Hypnosis and Suggestibility: An Experimental Approach* (1933/2002), to establish solid scientific protocols for studying hypnosis, have been impressively successful…up to a point. There have been significant methodological advances, more sensitive statistical analyses, and more well-elaborated theoretical constructs that have advanced our understandings and appreciation of the clinical merits of hypnosis. Yet, hypnosis remains (and I believe will likely forever remain) enigmatic in certain respects. After all, how do we objectively and precisely define and measure the subjective experience of people across different populations, much less across different individuals, especially in response to variable approaches to treatment all under the umbrella of "hypnosis?" It is a formidable challenge, to say the least, and we have a century's worth of evidence that highlights how trying to pin down one aspect of applied hypnosis inevitably gives rise to the theoretical and methodological "pop-up" of alternative views and priorities.

Thus, the authors of this fine article deserve to be congratulated for being willing to take on the challenge of the "whack-a-mole" nature of evaluating hypnosis efficacy research and should be appreciated for their thoughtful and measured approach to this exceedingly complex endeavor. The statement they offer in their article that provides the framework for their efforts describes "a disconnect between the scientific literature and much of clinical practice" (p.1).

They stated, “Many of the specific applications of this treatment modality, even some of the ones that are widely used in clinical practice, have still not been investigated in research and therefore are not supported by scientific evidence” (p.1).

A primary reason given for the disconnect between research and clinical practice is the lack of widely accepted standards for establishing the efficacy of interventions involving hypnosis. Their well-defined goal is to establish such standards. To do so, they organized the Task Force for Establishing Efficacy Standards for Clinical Hypnosis, simply referred to as the Hypnosis Efficacy Task Force. They assert that with such guidelines, “it is reasonable to expect researchers to be able to formulate evidence-based recommendations about clinical applications of hypnosis” (p. 2).

The merits of providing guidelines for determining treatment efficacy are well articulated in this article, especially in the emphasis the authors place on helping provide “clear, evidence-based recommendations about which therapies to use” to healthcare providers, insurers, clinicians and patients (p. 2). How valuable such guidelines may actually be to any or all of these parties is simply assumed to be self-evident. But there is good reason to believe that not everyone will embrace the rationale for them and then adopt them. Some will believe, not incorrectly, that treatment recommendations, even when based on solid research evidence, may simply not be a “good fit” for some individuals. Thus, it may pose an additional burden to the Task Force to successfully “sell” their guidelines to their intended buyers.

Evidence-Based Practice? Or Practice-Based Evidence?

The gap between research findings and clinical applications has existed since the earliest days of modern practice. Researchers are usually not also clinicians and clinicians are generally not also researchers. They each interact with populations that differ, often dramatically, in their

motivations to engage with whoever is conducting the hypnosis session provided to them. Beyond motivational differences of the person being hypnotized, there are substantial differences in the way a hypnosis session is conceived, structured, and stylistically delivered. The difficulties in bridging these potential gaps are well-represented in the conversations I had with both Ernest Hilgard, Ph.D., and André Weitzenhoffer, Ph.D., widely regarded as two of the premier hypnosis researchers whose profound influence on the field cannot be overstated. The conversations were about the measurement of the difficult-to-define trait they called 'hypnotizability" and reveal how divided even they were on some of the most important issues researchers and clinicians face.

When discussing with Hilgard the apparent reluctance of clinicians to use the Stanford Hypnotic Susceptibility Scales (SHSS) he and Weitzenhoffer spent many years developing and refining, Hilgard said, "I think that if I take just a sociological or political point of view, a scale is really very useful, and is more useful for clinicians than they like to admit" (personal communication, August,1988, cited in Yapko, 2019, p.174). Weitzenhoffer felt similarly, initially, and in our personal discussions often wondered aloud why clinicians didn't embrace the use of the Stanford Scales in their practices. That feeling changed markedly for Weitzenhoffer, though, when he retired from his research career in the Stanford lab. He moved to a small town where he decided to open a private clinical practice and use his extensive knowledge of hypnosis in what was for him an entirely new context. He thereby made the transition from researcher to clinician.

Soon after he made this change in his professional focus, Weitzenhoffer shared with me his realization that the well-researched scales he had spent years developing just didn't seem to have a place in clinical practice. He said, "Unless a patient is taking part in a study that calls for

such an assessment, doing one probably has no clinical value…I don't see the use of the scales in the clinical setting as desirable" (personal communication, December, 1988, cited in Yapko, 2019, p.36). In his second edition of his hypnosis textbook, *The Practice of Hypnotism* (2000), Weitzenhoffer made these same statements publicly, in writing. He said, in essence, if you want to know whether someone can be hypnotized and to what extent they may respond, then do hypnosis with them.

The Hypnosis Efficacy Task Force addressed this concern well in guideline number four when they suggested that the assessment of hypnotizability is encouraged "but is not required to establish the efficacy of a hypnosis-based treatment" (p. 4). This is the heart of the questions posed in the heading at the top of this section: do we strive to have a research base that informs an evidence-based clinical practice, or do we strive to have clinicians respond to the unique attributes of individual clients and provide practice-based evidence?

I think the Hypnosis Efficacy Task Force did an excellent job of anticipating and responding to these important questions. They acknowledged the inherent difficulties in trying to define what constitutes the practice of hypnosis as well as what defines hypnotizability. They stated, "it is not clear from these [efficacy grading] systems whether and how to take into consideration the hypnotizability of participants in the trials and which studies can be taken into consideration in the efficacy assessment, when there are so many different intervention variants" (p. 2). They specifically and insightfully mention expectancy as an issue in attempting to design double-blind, placebo-controlled studies because they are the "gold standard" for demonstrating efficacy since not being blinded, i.e., having positive expectations for a treatment result, is so foundational to hypnotic interventions. In fact, hypnosis has even been described as a "non-

deceptive placebo" in the clinical context, an acknowledgment of how important expectations may be to treatment response (Kirsch, 1994).

There are many procedure-based related questions to consider in evaluating efficacy: Will only scripted or standardized approaches be evaluated or will spontaneous approaches also be included? What about direct versus indirect and metaphorical approaches? Will past-oriented approaches be distinguished from present and future-oriented approaches? What about emotion-focused versus cognitive approaches? How will the quality of the therapeutic alliance or "attunement" between therapist and client be evaluated? When the Task Force acknowledged there are many intervention variants, these are presumably just a few of them that will likely complicate their efforts.

In a thought-provoking article in a special issue of the *IJCEH* on "Hypnosis and Treating Depression" (April-June, 2010), Barbara McCann and Sara Landes highlighted some of the limitations of trying to employ the traditional "gold standard" research methods, namely the randomized controlled trials (RCTs). They wrote, "One particularly important concern is whether results from RCTs can be generalized to actual clinical practice" (p. 152). Further concerns they raise include problems defining inclusion and exclusion criteria, the requirements of standardization of treatments generally in the form of treatment manuals or, in the case of hypnosis, scripted approaches, and "the potentially erroneous assumption that progress in therapy is linear" (p. 152).

Again, the Hypnosis Efficacy Task Force anticipated and addressed this concern when stating in guideline number five that "Blinding of the participants/patients and the interventionists to group allocation is aspirational but is not required to establish efficacy of a hypnosis-based treatment" (p. 4).

Is Hypnosis a Tool or a Therapy?

How one answers the above question, whether a researcher or clinician, is a powerful determinant of how hypnosis will be applied and its effects measured. Many professionals employ the term "hypnotherapy" as the defining label of their use of hypnosis, thereby suggesting that the hypnosis itself is the primary therapeutic factor. Other professionals prefer the term "clinical hypnosis" and frame hypnosis as a tool of treatment, a vehicle for delivering ideas and perspectives and mobilizing personal resources that are embedded in a larger therapeutic framework (such as cognitive-behavioral or emotion-focused therapies).

Hypnosis is obviously subjectively conceived and practiced; what one person does with hypnosis can be quite far removed from the way another applies hypnosis even in response to the same types of presenting issues. Likewise, how one person thinks about the nature of hypnosis and hypnotic phenomena can be quite different from the perspectives of another. How will differences in conceptual framework and practical applications of hypnosis be bridged by the suggested guidelines? On this point, the Hypnosis Efficacy Task Force is not as clear. The hope that strong empirical support for one approach over another will change peoples' ideas or methods is certainly optimistic. Only time will tell if it's realistic.

Empirically-Validate Therapies or *Therapists*?

Should we strive to empirically validate *therapies,* or should we strive to empirically validate *therapists*? As a corollary question, is it the therapy that is effective or is it the clinician applying the therapy skillfully based on relevant experience and good judgment? Clearly, not all clinicians possess the same level of skill, knowledge, and judgment. How will the differences in clinical skill across therapists be addressed? How will treatment fare when a lesser-skilled therapist applies an empirically supported therapy?

Considering the hard lessons learned from the so-called "Repressed Memory Controversy" in the 1990s and the fractures it revealed in how hypnosis was being misapplied even by experienced practitioners who should have known better, the words of Ernest Hilgard were prescient. When encouraging clinicians to be at least a little more research-minded in their work, he said, "So if there's any message I have [for therapists], it is to not get scientistic; you don't have to have analysis of variance and become a slave to statistics, [but at least have] just garden variety statistics. Here's half a dozen people that have these same symptoms. They've been treated in three different ways. Why were the different ways chosen? …Put some little design into it" (personal communication, August, 1988, cited in Yapko, 2019, p. 175). McCann and Landes echo Hilgard's suggestion for case studies and single-subject designs and further suggest that "benchmarking" is a viable research tool as well (2010). If the goal is to raise the standard of care by well-intentioned practitioners, a primary goal of the Hypnosis Efficacy Task Force, their suggested guidelines may go a long way in actualizing that goal.

Grading the GRADE

The Hypnosis Efficacy Task Force chose to endorse the GRADE guidelines for assessing efficacy of hypnotic interventions. Their twofold rationale for doing so (see pages 5 and 6) was clearly well thought out and should be supported, in my opinion. They acknowledge that it isn't by any means a perfect system for assessing hypnosis efficacy research but is currently the best option available. The fact that it is said to be "the most accepted clinical recommendation system in medical research" (p. 6) is both a strength and weakness. As is pointed out, the strength is that this is a system familiar to and respected by medical professionals such that any recommendations made may be seen as more credible. Paradoxically, this is also the weakness: While it is true that hypnosis is routinely applied in medical contexts, its use is even greater in

psychotherapeutic and other contexts where there are many more potential confounding variables and progress is often not measured linearly. To apply medical standards of quality of evidence to non-medical problems and subjectively conceived treatments will likely limit the value of the GRADE system. To their credit, however, the Hypnosis Efficacy Task Force has anticipated this as well and acknowledges the reality that testing the efficacy of a hypnotic intervention for, say, anxiety, cannot be the same as testing the efficacy of a specific medication or surgical procedure.

Conclusion

In this short commentary, I have identified a few of the issues in trying to promote guidelines for conducting research and clinical practice in the domain of hypnosis. The Hypnosis Efficacy Task Force has done an excellent job of defining what many of the complexities are in doing hypnosis research compared to research in areas that are more objectively conceived and then applied therapeutically. Their choice of the GRADE system, well-described in the article, as the foundation for assessing the level of treatment efficacy and offering clinical recommendations isn't perfect but still makes good sense since it is the most flexible of the systems evaluated for the unique needs of conducting hypnosis efficacy research.

The ten recommendations the Task Force have articulated are eminently sensible and well-considered. I agree with and support them in entirety. I want to express my thanks to the Hypnosis Efficacy Task Force for taking on this important issue that will undoubtedly have a great deal of impact on the future of the field. I also want to offer my thanks to the *IJCEH* Editor, Gary Elkins, Ph.D., for inviting me to contribute this commentary.

References

Hull, C. (1933/2002). *Hypnosis and Suggestibility: An Experimental Approach*. New York: Williston, VT.: Crown House Publishing Ltd.

Kirsch, I. (1994). Clinical hypnosis as a nondeceptive placebo: Empirically derived techniques. *American Journal of Clinical Hypnosis, 37*, 95-106.

McCann, B. and Landes, S. (2010). Hypnosis in the treatment of depression: Considerations in research designs and methods. *International Journal of Clinical and Experimental Hypnosis, 58*(2), 147-164.

Weitzenhoffer, A. (2000). *The Practice of Hypnotism* (2nd edition). New York: John Wiley & Sons.

Yapko, M. (2019). *Trancework: An Introduction to the Practice of Clinical Hypnosis* (5th edition). New York: Routledge.