NONLOCALITY, INTENTION, AND OBSERVER EFFECTS IN HEALING STUDIES: LAYING A FOUNDATION FOR THE FUTURE

Stephan A. Schwartz[1#] and Larry Dossey, MD[2]

All research domains are based upon epistemological assumptions. Periodic reassessment of these assumptions is crucial because they influence how we interpret experimental outcomes. Perhaps nowhere is this reassessment needed more than in the study of prayer and intention experiments. For if positive results from this field of research are sustained, the reality of nonlocal consciousness must be confronted. This paper explores the current status of healing and intention research, citing a number of major studies and using the "Study of the Therapeutic Effects of Intercessory Prayer (STEP) in Cardiac Bypass Surgery Patients: A Multicenter Randomized Trial of Uncertainty and Certainty of Receiving Intercessory Prayer" as a case study of this line of research. The paper argues that the dose-dependent model typical of drug trials, and adopted for use in the STEP and other studies, is not the optimal model for intention-healing research, and critiques this approach in detail, citing apposite research from which we draw our recommendations and conclusions. The paper suggests that the usual assumptions concerning blindness and randomization that prevail in studies using the pharmacological model must be reappraised. Experimental data suggest that a nonlocal relationship exists among the various individuals participating in a study, one which needs to be understood and taken seriously. We argue that it is important to account for and understand the role of both local and nonlocal observer effects, since both can significantly affect outcome. Research is presented from an array of disciplines to support why the authors feel these issues of linkage, belief, and intention are so important to a successful, accurate, and meaningful study outcome. Finally, the paper offers suggestions for new lines of research and new protocol designs that address these observer-effect issues, particularly the nonlocal aspects. The paper finally suggests that if these effects occur in intention studies, they must necessarily exist in all studies, although in pharmacological studies they are often overshadowed by the power of chemical and biological agents.

Key words: Intention, prayer, healing, observer effect, consciousness, placebo, nonlocal consciousness

(Explore 2010; 6:295-307.

INTRODUCTION

Epistemologically periodically reassessing what constitutes good research is crucial because this process gives us the ability to distinguish justified belief from opinion. Perhaps nowhere is this distinction of greater significance than in the study of prayer and intention experiments. For if the positive results from this field of research are sustained, the reality of nonlocal consciousness must be taken seriously. This paper explores the current status of healing and intention research, cites a number of major studies, and uses the "Study of the Therapeutic Effects of Intercessory Prayer (STEP) in Cardiac Bypass Surgery Patients: A Multicenter Randomized Trial of Uncertainty and Certainty of Receiving Intercessory Prayer" conducted by Herbert Benson et al[1] as a case study of this line of research. In April 2006, researchers from Harvard Medical School published this long-awaited study in the *American Heart Journal.*[1] The $2.4 million study was funded in large part by the John Templeton Foundation, which promotes the study of the intersection of religion and science. Its publication grabbed headlines across America for two main reasons: it originated from Harvard Medical School, and it had an unexpected result—patients who were prayed for, and knew they would be prayed for, fared the worst of the three intervention groups.

Because of its venue, its level of funding, and the media attention the study has engendered, we have selected this study as a case demonstrating issues common to much of this research field, and we use it to explore those issues. In the process, we also examine attitudes that prevail in the arguments of both proponents and skeptics of prayer and intention research. What we seek is a discussion on the basic assumptions implicit, but usually unacknowledged, in these studies, and a reappraisal of the design parameters upon which prayer and intention studies have been grounded. In our view, STEP is a noble failure; noble because it was done with integrity, on the basis of imperfect understanding, and because its failure has much to teach us.

STEP BACKGROUND

The STEP experiment involved 1,802 patients undergoing coronary artery bypass surgery at six US hospitals.[1] These 1,802 patients were then assigned to one of three subpopulations. Therapeutic intention in the form of prayer was provided by one Protestant and two Catholic groups, whose members were told to pray for a quick recovery with no complications. They were provided only with the first name and the initial of the last name

1 Senior Samueli Fellow for Brain, Mind and Healing of the Samueli Institute, columnist for *EXPLORE*, and editor of the SchwartzReport available at: www.schwartzreport.net
2 Executive Editor, *EXPLORE*

 # Corresponding Author. Address: P.O. Box 905 Langley, WA 98260
e-mail: saschwartz@earthlink.net

ISSN 1550-8307/$36.00

doi:10.1016/j.explore.2010.06.011

of the prayer participants–"John D." Prayers were initiated on the eve or the day of surgery and continued for two weeks thereafter.

The three groups consisted of the following (the group names are our designation):

- Group A: 604 patients who were told they might or might not be prayed for, and were; of this group, 52% had postsurgical complications
- Group B: 597 patients who were told they might or might not be prayed for, and were not; among this group, 51% had postsurgical complications
- Group C: 601 patients who were told they would be prayed for, and were; among this group, 59% had postsurgical compli-
  cations

To many skeptics in both media and science, it was this last result that was the headline of the study, suggesting that prayers for the sick might actually be harmful.

How can these results be understood? To begin with, the differences between the two blind groups, those who were told they might or might not be recipients, one of which eventually was prayed for, whereas the other was not, are nonsignificant. *The only significant outcome in the study is between those who were blind and those who were not* ($P = .003$; $z = 2.8$). Yet an attempt to analyze this result is almost wholly absent in both the published report and much of the commentary about it.[2-5]

We are disturbed by the fact that not only skeptics, but the researchers themselves, turned a blind eye to this challenging result, and we are not alone. Duke University Medical Center cardiologist Mitchell W. Krucoff, and his research partners, Suzanne W. Crater and Kerry L. Lee, explain this carefully in their article accompanying the publication of the STEP study in the *American Heart Journal*.[6] They say,

> [T]he most striking element of the STEP report is in the interpretation of the study results showing significantly worsened outcomes in one of the experimental arms . . . . [T]he investigators take an almost casual approach toward any explanation, stating only that it 'may have been a chance finding.' It is rather unusual to attribute a statistically significant result in the primary end point of a prospective, multicenter randomized trial to 'chance.'
>
> "In fact, such attribution is antithetical to the very definition of what error and statistical certainty imply: that the worse outcomes are almost certainly related to the therapy and not the play of chance. If the results had shown benefit rather than harm, would we have read the investigators' conclusion that this effect 'may have been a chance finding,' with absolutely no other comments, insight, or even speculation?"[6]

OBSERVER-EXPECTANCY EFFECTS, BOTH LOCAL AND NONLOCAL

The randomized, double-blind clinical trial is widely considered the gold standard of judging the efficacy of any therapy. If a study is adequately randomized and blinded, it is assumed that the effects of belief, intention, and conviction of subjects and researchers are bypassed. Applying this logic to prayer experiments, it is assumed that what an experimenter privately thinks about the intercession is irrelevant. Yet experimental results suggest it is not that simple.

Rather, data suggest that intention, belief, attitude, and expectancy, on the part of everyone involved with a study, expressed both locally and nonlocally, can be determining variables. Chemist Douglas Dean and parapsychologist Karlis Osis showed that different experimenters carrying out the same experiment got different results.[7] Psychologists Gertrude Schmeidler and Michaeleen Maher made videos of well-known researchers conducting experiments and then played them for students with the volume turned so low as to be inaudible.[8] The students were asked to describe the researchers, assigning them words like "friendly" or "cold." Estimates were then made as to how experiments conducted by these researchers would turn out. Those with "cold" type responses were estimated to have respondents who produced lower scores; the converse was true for researchers described as "friendly." The actual results of the experiments were then compiled. Those with "cold" type adjectives did in fact have informants who scored lower.[8]

Perhaps the starkest example, however, showing the observer latency effect of belief is an experiment series done by psychologist Richard Wiseman, a leading denier of nonlocal consciousness, and anthropologist Marilyn Schlitz, a researcher long associated with successful studies exploring whether an individual knows through some kind of linkage that they are being stared at, even by a person at a distance looking at their image on a closed-circuit television.[9] The measurement for this effect being galvanic skin response. Schlitz had earlier worked with psychologist William Braud, and carried out a series of studies demonstrating this effect.[10]

Wiseman sought to replicate these studies and made three attempts, all unsuccessful. Schiltz then proposed that she and Wiseman do a new series, a kind of hyper-replication *using his same laboratory, the same protocol, the same participant pool.* Once again with Schlitz as the principal investigator, the study was successful. Once again, the participants being stared at showed significant physiological response that was absent when they were not being focused on.[9] Wiseman then ran the same study again, without success, confirming his passionate negative expectation.

Hazelrigg et al[11] examined "personality moderators of experimenter expectancy effects" and focused on five, looking at them from the perspective of both researcher and participant. They reported, "Experimenters with stronger interpersonal control orientations, more positively evaluated interpersonal interaction styles, and greater ability to encode nonverbal messages are believed to be more likely to produce expectancy bias."[11] They also looked at subjects with greater need for social approval and greater nonverbal decoding ability, and hypothesized that such individuals would be more susceptible to bias.

They reported two "moderators" mattered: "the experimenter control orientation and subject need for social approval hypotheses. There was also evidence for a boomerang effect–subjects low in need for social approval gave ratings opposite to the experimenter's outcome expectancy. Finally, effects appeared