Psychological Methods
2006, Vol. 11, No. 1, 123–125

1082-989X/06/$12.00 DOI: 10.1037/1082-989X.11.1.123

REPLY

Paper or Plastic Revisited: Let's Keep Them Both—Reply to Broderick and Stone (2006); Tennen, Affleck, Coyne, Larsen, and DeLongis (2006); and Takarangi, Garry, and Loftus (2006)

Niall Bolger
Columbia University

Patrick E. Shrout and Amie S. Green
New York University

Eshkol Rafaeli
Barnard College, Columbia University

Harry T. Reis
University of Rochester

The authors review commentaries by J. E. Broderick and A. A. Stone (2006); H. Tennen, G. Affleck, J. C. Coyne, R. J. Larsen, and A. DeLongis (2006); and M. K. T. Takarangi, M. Garry, and E. F. Loftus (2006) on their original article (A. S. Green, E. Rafaeli, N. Bolger, P. Shrout, & H. T. Reis, 2006). The authors were pleased to find more agreement than disagreement regarding the choice of methods for conducting diary studies. It is clear that continued critical evaluation of all diary methods, both paper and plastic, is warranted. However, on the basis of their initial findings, the authors conclude that paper diaries are still likely to have a valuable place in researchers' toolboxes.

Keywords: diary studies, experience sampling method, ecological momentary assessment

Our goal in reporting the three studies that compared data from paper and plastic diaries was to stimulate an open debate about the utility of these methods in current psychological research. We very much appreciate the contributions to this debate provided in the three commentaries in this issue (Broderick & Stone, 2006; Takarangi, Garry, & Loftus, 2006; Tennen, Affleck, Coyne, Larsen, & DeLongis, 2006). All three reminded readers of the important methodological issues—some obvious, some subtle—that need attention as researchers adapt diary methods to study psychological processes. Broderick and Stone (2006) provided a critical perspective on participant compliance and motivation in paper and electronic diary studies. Tennen et al. (2006) argued for an expanded view of diary collection methods, and they discussed the conditions under which paper and electronic diaries will yield diverging results. Finally, Takarangi et al. (2006) drew attention to memory processes that affect the validity of self-reports and apply to both paper and electronic data collection methods. It is noteworthy that none of the commentators disputed our basic empirical findings—namely, that the substantive data obtained by paper and electronic methods were very similar across our studies.

There was much to agree with across the commentaries, particularly the conclusion that more research is needed to understand how memory, compliance, motivation, and changing technology affect data quality. In the sections that follow, we highlight some areas of disagreement and some topics that deserve further consideration. We consider, in turn, issues raised by Broderick and Stone (2006), Tennen et al. (2006), and finally, Takarangi et al. (2006).

Broderick and Stone (2006)

Broderick and Stone faulted us for not reviewing the broader compliance literature, particularly the literature on medical compliance. However, in their own review of that literature, we believe their reading of the evidence on noncompliance with paper-and-pencil protocols was less than evenhanded. Consider their conclusion about Litt, Cooney, and Morse's (1998) study, which they cited as evidence of poor compliance with paper diaries. Fourteen out of 20 participants (70%) who were debriefed following the study

Niall Bolger, Department of Psychology, Columbia University; Patrick E. Shrout and Amie S. Green, Department of Psychology, New York University; Eshkol Rafaeli, Department of Psychology, Barnard College, Columbia University; Harry T. Reis, Department of Psychology, University of Rochester.

Correspondence concerning this article should be addressed to Niall Bolger, Department of Psychology, Columbia University, Schermerhorn Hall, New York, NY 10027. E-mail: bolger@psych.columbia.edu

admitted delaying between one and five of their recordings until later (with a median "closer to 4 than to 3"; D. M. Litt, personal communication, November 22, 2005). Participants had been signaled eight times per day and were considered delayed if their response was more than 5 min after the signal. Given that 14 participants admitted a median of 4 delays out of 8 signals and that 6 participants reported no delays, then their compliance rate would be 65%, a figure not very different from our reported Study 1 result (75%).[1] Furthermore, in Litt et al.'s study, a response was counted as delayed if it was as little as 5 min late; our Study 1 allowed for up to a 15-min delay, a criterion that many experience-sampling researchers consider reasonable (Delespaul, 1992). When one further considers that Litt et al.'s participants were recently discharged alcohol-dependent inpatients reporting on their cravings and relapses—a group that might well be motivated to noncompliance—and that we advocate dropping participants who delay all recordings (reported as 1 out of the debriefed 20 participants by Litt et al., 1998), the overall rate of compliance in Litt et al.'s study does not seem as poor or as potentially consequential as Broderick and Stone have suggested.

Broderick and Stone also questioned whether participant motivation affects compliance, either with medical treatment regimens or with experience-sampling data-collection protocols. However, there is growing evidence in the health literature that the nature of the patient–medical provider relationship does indeed influence compliance. For example, autonomy-supportive interventions by physicians and clinical practitioners have been demonstrated to produce significant increases in self-managed glycemic control among patients with type 2 diabetes over 1 year and to contribute to smoking cessation (Williams, Gagné, Ryan, & Deci, 2002; Williams, McGregor, Zeldman, Freedman, & Deci, 2004).

More generally, in a review of this literature, Mead and Bower (2000) concluded that patient-centered communication "is regarded as having value . . . largely in terms of mediating positive outcomes from management decisions. For example, a friendly and sympathetic manner may increase the likelihood of patient adherence to treatment" (p. 1090). In other words, studies of medical care have shown that patient-centered communication and participatory decision making contribute to patient trust, and trust is surely indispensable to compliance (and for that matter, data quality) in any diary study, be it paper or electronic.

We have emphasized the importance of motivating diary participants both to be compliant and to report honestly any deviations of diary entries from the diary protocols. Maintaining an open, trusting, and collaborative relationship with participants was central in all three of our studies. Broderick and Stone are correct, however, in noting that we did not assess participant motivation and therefore cannot claim that our participants were more motivated to comply than were participants in other studies. This is clearly an area in which more systematic studies are needed.

Although we agree that the medical compliance literature can be informative about the issues raised in our target article, we are less convinced than Broderick and Stone about how generalizable that literature is to compliance in general. In addition to the special issues of patient–physician relationships just discussed, the medical compliance literature confounds influences such as unpleasant side effects of medication, social pressure to appear compliant, beliefs about one's own health, habit strength, and self-presentation concerns with simple inclination or disinclination to provide diary reports of behaviors such as moods and social interactions. For researchers who are considering diary studies in the context of treatment studies, the medical compliance literature provides an important call to be skeptical about respondent reports. For researchers who study daily processes, such as mood and relationship quality, in randomly sampled persons or couples, the medical literature seems to us to be less relevant than Broderick and Stone would think.

Tennen et al. (2006)

Tennen and his colleagues identified limitations of our article, and some of their points get no argument from us. Two of our three studies were based on reanalyses of existing diary data, and as Tennen et al. have pointed out, the studies did not use diary methodology that would be considered state of the art in 2006. These studies were not intended to be the last word on the benefits of the current diary technology, but rather they were intended to reopen discussion of whether paper diaries were in essence worthless and therefore should be dropped from the research toolbox. We are in agreement that further studies of other versions of paper and plastic methodology as well as other tools fast approaching on the research horizon (e.g., cell phones, pocket digital audio recording devices) are needed.

Like Broderick and Stone (2006), Tennen et al. directed readers' attention to the medical literature, and they reviewed several studies in that literature that directly compare diary methods. We thank them for calling our attention to Weiler's (2005) review. This review seems to underscore our conclusion that paper diaries may often produce data that are comparable to those obtained using more modern methods but that such comparability cannot be taken for

[1] In his personal communication of November 22, 2005, D. M. Litt speculated that the compliance rate was closer to 60% because some participants might have a skewed number of delayed responses. Even this worst-case scenario is not very different from the 75% result in Study 1, especially when one takes into account sampling variation that can occur when only 20 participants are studied.

granted. It also suggests to us that comparability is a matter of degree more than it is a dichotomy.

Some of the other criticisms of our article seem more a matter of emphasis than of substantive disagreements. Tennen et al. are inclined to believe that paper diaries can work for end-of-the-day designs but not for within-day diary studies. We understand their reasoning, but we are not convinced that paper diaries will never be useful in within-day diary studies. This is among the questions that remain to be studied definitively.

Takarangi et al. (2006)

Takarangi et al. considered compliance from the perspective of memory research. They discussed prospective memory processes that are used by respondents to remember to complete a diary entry on time. Better memory (or technological aids that enhance such memory) will help compliance. Thus, a time-based study requiring multiple responses each day could benefit from some external request to remember. As Takarangi et al. noted, signals of this sort will come from an electronic diary's alarm, "whereas a paper diary leaves them to their own personal strategies for remembering" (p. 120). Although we agree that alarm cues are an inherent benefit to electronic diary formats, it is important to recognize that alarm cues can be used in paper diary studies as well through the use of beepers, cell phones, or preprogrammed wristwatches. Note that all of these memory aids have been used successfully in paper diary studies.

Takarangi et al. also discussed retrospective memory processes, and this discussion cuts closer to the raison d'être of diary studies—their ability to capture life as it is lived. Takarangi and her colleagues cited abundant evidence of memory distortions that should be of concern to researchers who rely on self-reports in general and apply equally to paper and plastic methods. Their commentary, along with Robinson and Clore's (2002) accessibility model, can help guide researchers in deciding which kind of memory (episodic, experiential, and contextual, or semantic, conceptual, and decontextualized) is of greater interest to them; if it is the former, then diaries do offer an advantage over other forms of data collection. The equivalence of results found in our target studies suggests that the narrower question of "what kind of diaries" is of secondary importance.

Conclusion

In conclusion, we hope that it is clear that we welcome—and are enthusiastic users of—the new electronic methods of diary data collection that are now available to researchers. These methods can serve only to increase the value of intensive measurement designs for understanding psychological processes. The availability of these methods, however, should not lead researchers to feel compelled to abandon more traditional paper diary methods or lead grant and journal reviewers to dismiss past findings based on paper diaries. By demonstrating that paper diary data are indistinguishable from electronic diary data in three studies, our modest goal has been to ensure that the range of methodologies available to researchers should be widening rather than contracting.

References

Broderick, J. E., & Stone, A. A. (2006). Paper and electronic diaries: Too early for conclusions on compliance rates and their effects—Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods, 11,* 106–111.

Delespaul, P. A. E. G. (1992). Technical note: Devices and time-sampling procedures. In M. W. de Vries (Ed.), *The experience of psychopathology: Investigating mental disorders in their natural settings* (pp. 363–373). Cambridge, England: Cambridge University Press.

Green, A. S., Rafaeli, E., Bolger, N., Shrout, P. E., & Reis, H. T. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological Methods, 11,* 87–105.

Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. *Health Psychology, 17,* 48–52.

Mead, N., & Bower, P. (2000). Patient-centredness: A conceptual framework and review of the empirical literature. *Social Science and Medicine, 51,* 1087–1110.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128,* 934–960.

Takarangi, M. K. T., Garry, M., & Loftus, E. F. (2006). Dear diary, is plastic better than paper? I can't remember: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods, 11,* 119–122.

Tennen, H., Affleck, G., Coyne, J. C., Larsen, R. J., & DeLongis, A. (2006). Paper and plastic in daily diary research: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods, 11,* 112–118.

Weiler, J. M. (2005, April). *Comparison of controlled clinical trials of paper versus ePRO or ePRO versus ePRO.* Paper presented at the Drug Information Association Symposium, Arlington, VA.

Williams, G. C., Gagné, M., Ryan, R. M., & Deci, E. L. (2002). Facilitating autonomous motivation for smoking cessation. *Health Psychology, 21,* 40–50.

Williams, G. C., McGregor, H. A., Zeldman, A., Freedman, Z. R., & Deci, E. L. (2004). Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. *Health Psychology, 23,* 58–66.

Received October 18, 2005
Revision received December 5, 2005
Accepted December 5, 2005 ■