JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 53, NO. 1, PP. 65–69 (2016)

Rejoinder

Utilizing Indicator-Based Methods: 'Measuring the Impact of a Science Center on its Community'

John H. Falk[1] and Mark D. Needham[2]

[1]*College of Education, Oregon State University, Corvallis, Oregon 97331*
[2]*Department of Forest Ecosystems and Society, Oregon State University, Corvallis, Oregon*

Received 21 June 2015; Accepted 26 June 2015

We have been generously provided an opportunity to write a rejoinder to Jensen and Lister's "Comments and Criticism" piece published in this issue. From the outset, we want to make it clear that we consider such dialogue healthy and constructive for the field. We appreciate the opportunity to clarify the points made by Jensen and Lister and thank this journal's editors for the opportunity to do so.

Jensen and Lister admirably summarized our original article (Falk & Needham, 2011). According to Jensen and Lister, the main issue at hand was our use of a conceptual marker as an "indicator" of science learning as a way to "circumvent the need to rely exclusively on self-report data" (Jensen & Lister, this issue). To clarify, we did not utilize this marker approach merely to circumvent self-report data. As stated in our article (Falk & Needham, 2011), our primary motivation for using this approach was because of challenges related to attribution, which self-report does not adequately address. Given the cumulative nature of learning, it is difficult for anyone to accurately determine exactly where or when they actually learned anything. As a result, we selected a single concept, homeostasis, that could be used as the learning equivalent of a radioactive tracer; "something that in and of itself may or may not be highly important, but which could be considered an indicator of something greater that was meaningful" (Falk & Needham, 2011, p. 3). Jensen and Lister raised three basic concerns with our approach: (a) the validity of our baseline dataset, (b) the reliability of our procedures for coding this marker across the three samples, and (c) that results do not support our conclusions. We address each of these concerns in this rejoinder.

Jensen and Lister questioned the validity of our use of 7% as a baseline estimate of the Los Angeles (L.A.) public's knowledge of homeostasis in 1998. As Jensen and Lister themselves summarized in their commentary and we also acknowledged in our own article, this figure was derived using a different methodology than was utilized in the subsequent two samples; 2000 and 2009 data were derived from random telephone surveys of the L.A. population, whereas 1998 data were from a random sampling of the first month's visitors to the newly re-opened California

Correspondence to: John H. Falk; E-mail: falkj@science.oregonstate.edu
DOI 10.1002/tea.21269
Published online 25 August 2015 in Wiley Online Library (wileyonlinelibrary.com).

Science Center (CSC) containing the exhibition related to homeostasis. However, as we originally stated and as Jensen and Lister reiterate, this 7% was a conservative estimate of public knowledge of homeostasis because, if anything, the understanding of science among the L.A. general public was almost certainly lower than that of those who self-selected to visit a science center. As suggested by our own data (Falk & Needham, 2011) and that of others (e.g., Falk *et al*., in press; Friedman, 2008; NRC, 2009), those who self-select to visit free-choice learning institutions such as the CSC tend to be more predisposed toward, interested in, and knowledgeable of science related topics than is the general public. Thus, it is highly unlikely that more than 7% of the L.A. population would have been able to accurately define this homeostasis marker if asked at the time. Jensen and Lister claimed that "no valid baseline measurement was developed," but we were forthright in our original article that "no true baseline was collected . . . but a useable baseline was created" (Falk & Needham, 2011, p. 8). Although our 1998 estimate is indeed not a true baseline, using this 7% figure represented a higher than would be expected bar to overcome. Although we obviously cannot definitively prove this was the exact baseline percentage, this figure appears to fit the expected trend line in growth in the knowledge of the concept of homeostasis among the L.A. population, assuming that changes in understanding of this concept demonstrated by research at the CSC did in fact make a contribution over time to the public's understanding of this one, somewhat obscure concept. As reported in our original article, the proportion of individuals in L.A. who were correctly able to define this concept increased to approximately 10% in 2000 and 20% in 2009.

An arguably more substantial concern raised by Jensen and Lister is the reliability of the way that understanding of homeostasis was measured. Here, we need to offer a *mea culpa*, as we regrettably did not sufficiently detail our methodology in the article (Falk & Needham, 2011), nor reference the technical report in which the full methodology was described (Falk & Amin, 1999). That said, the absence of a full discussion in the paper did not mean we failed to apply good research technique in practice. Accordingly, we present that methodology here. The first step was to ensure validity of measurement and to accomplish this we engaged a team of five subject-matter experts (all human physiology professionals); each of whom was tasked with developing a scoring rubric for defining homeostasis. From their transcripts, a single rubric was developed around the concept of homeostasis. This rubric was then shown to all five experts who agreed that it was an acceptable definition—"homeostasis is the balance or equilibrium that organisms or cells strive to maintain." To ensure reliability, two new and different subject-matter experts were identified who, along with the first author, were shown the expert's rubric and asked if they had any clarification questions. These three individuals were then presented with a randomly selected subset of one-third of the 1998 survey responses to the question "could you please tell me what you think homeostasis is" and asked to independently categorize responses as either "correct" or "incorrect" based on the expert rubric developed in step one. Results were compared among these three individuals, any initial disagreements in scoring were discussed, and a collective agreement was reached on how to slightly amend the scoring rubric. Each of these three individuals then independently scored a second third of responses using the slightly revised rubric. Inter-coder reliability was 95%. All subsequent data (e.g., 1998, 2000, 2009) were then scored by the first author using the same rubric. Although we did not report all of these methodological details in Falk and Needham (2011), and thus Jensen and Lister had grounds for questioning the reliability of our approach, we believe these additional details address the concerns raised.

Jensen and Lister's final concern was that results of the homeostasis marker do not support our claims that some changes in understanding of this concept among the L.A. public can be attributed to visiting the CSC. In particular, they pointed out that a larger percentage of the 10% of the L.A. public who could correctly define homeostasis in 2000 had visited the CSC (75%) compared to

2009 when 61% of the 20% of correct respondents had visited the CSC.[1] Further analysis, however, suggested that although this difference was statistically significant ($\chi^2 = 4.92, p = .026$), the phi effect size (φ) was only .13. Using guidelines from Cohen (1988) and Vaske (2008) for interpreting effect sizes, the magnitude of this difference was "small" or "minimal," respectively. In fact, these discrepancies between years in the number of respondents who defined homeostasis correctly but had not visited CSC versus the number who defined it correctly and had visited CSC represented a difference of less than a dozen individuals, or just slightly more than 1% of the total sample; well within the margin of error. The bottom line is that perhaps this difference could slightly undermine the strength in the basic argument that we make, but given the minimal effect size and small number of people associated with this difference between years, we are inclined to believe that this inconsistency is more likely to be a minor blip in the data, and the observed growth trend in understanding of this single concept amongst the L.A. public (7% in 1998, 10% in 2000, 20% in 2009) if not wholly, is likely largely attributable to visiting the CSC.[2] Importantly, as initially reported, individuals in both 2000 and 2009 who were able to correctly define this homeostasis marker were statistically more likely to have visited the CSC.[3]

Taken together, we refute Jensen and Lister's assertion that our research is methodologically flawed because we did not follow established research procedures. Although we acknowledge that our final conclusions can be debated (as would be true in any study), the basis of this debate should not be flaws in methodology; these we believe were sound. Of course when all is said and done, the larger issue relates to whether it is a reasonable idea to use a conceptual marker (e.g., homeostasis in the current case) as an indicator of the effects of an educational intervention. We remain unwavering in our belief that use of such a tool, in tandem with other research approaches, is indeed worthwhile. As previously stated, the inherently incremental and distributed nature of science learning makes attribution of learning from a single institution or event extremely challenging. Most individuals develop science understanding, as well as science interests and identities, through an accumulation of experiences from various sources at different times (e.g., Barron, 2006; Ito et al., 2013; Lemke, Lecusay, Cole, & Michalchik, 2012; NRC, 2009; OECD, 2012; Renninger & Riley, 2013; Stocklmayer, Rennie, & Gilbert, 2010). As shown by our data and discussed in our original article (Falk & Needham, 2011), even individuals who ostensibly benefited from exposure to the homeostasis "lesson" presented at the CSC were still likely to attribute their initial learning to the place where they first encountered the concept (i.e., school). So, is the sole use of such indicators sufficient to demonstrate the impact of an institution such as a science center? The answer, of course, is no. Does use of such indicators enrich our understanding of how and what people likely learn in free-choice settings? The answer, we believe, is unequivocally yes. Collectively, the several types of data we collected—self-report, factual knowledge questions and indicator—provided a rich snapshot of the likely effects that visit experiences at the CSC had on the general public's science interest and learning.

That is by no means to suggest that other methodological approaches for addressing this topic are not only possible but might even under certain circumstances be preferable, including the repeated measures approach that Jensen and Lister advocated (i.e., panel data targeting the same individuals pre and post visit). In fact, we as well as others have used this approach in other studies (e.g., Falk & Storksdieck, 2012; Ito et al., 2013; Leinhardt, Crowley & Knudson, 2002; Lemke et al., 2012). However make no mistake, this approach too is subject to methodological and theoretical issues. Panel samples may become less representative over time as the population changes and as panel members drop out (Groves, 1989; Taplan, 2005). Panel designs may also be prone to certain forms of measurement error, such as "conditioning" and "seam" bias (cf., Groves, 1989; Lavrakas, 2008). Also relevant in this particular context is the potential problems associated with imposing highly contrived experimental designs on individuals participating in a free-choice learning experience (Falk et al., in press; Lemke et al., 2012; NRC, 2009). The reality is that any effort designed to investigate something as complex as the long-term impact of a particular science experience, regardless of setting, will face methodological challenges. In part, this is because all methodologies have both benefits and constraints, independent of the rigor with which they are applied. The indicator methodology described in our article (Falk & Needham, 2011) is no exception. In the final analysis, all researchers need to be ever vigilant for lapses in reliability and validity, but as a community we will make little intellectual progress if we attempt to restrict our investigations to a singular vision of research purity.

Notes

[1]A fourth, 2015 data point has just been collected since the publication of (Falk & Needham, 2011) using the exact same methodology as 2000 and 2009. Seventeen years after opening, correct responses to the question about homeostasis continued to climb, from 20% in 2009 to 36% of L.A. adults in 2015. This rise parallels a second long-term trend, growth in the percent of L.A. adult residents having visited CSC at least once, rising from 40% in 2009 to 67% in 2015. Also, in 2015 the percent of those in L.A. able to correctly answer this question who had previously visited the CSC was 76% of correct respondents; statistically identical to the percent in 2000 (Falk, J.H., Pattison, S., Livingston, K., Meier, D., Bibas, D., Fifield, S. & Martin, L. (in prep.). Contributions of science centers in Los Angeles, Philadelphia and Phoenix to public science understanding, interest and engagement).

[2]The new data collected in 2015 seem to support this conclusion.

[3]Also the case in 2015.

References

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. Human Development, 49, 153–224.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.

Falk, J. H., & Amin, R. (1999). California Science Center: BodyWorks summative evaluation. Technical Report Annapolis, MD: Institute for Learning Innovation.

Falk, J. H., & Needham, M. D. (2011). Measuring the impact of a science center on its community. Journal of Research in Science Teaching, 48, 1–12.

Falk, J. H., Dierking, L. D., Swanger, L., Staus, N., Back, M., Barriault, C., . . . Verheyden, P. (in press) Role of science centers in supporting adult science literacy: An international, cross-institutional study. Science Education.

Falk, J. H., & Storksdieck, M. (2010). Science learning in a leisure setting. Journal of Research in Science Teaching, 47(2), 194–212.

A. Friedman, (Ed.) Framework for evaluating impacts of informal science education projects. Technical Report. Arlington, VA: National Science Foundation. (2008).

Groves, R. (1989). Survey costs and survey errors. New York: John Wiley.

Ito, M., Baumer, S., Bittanti, M., Boyd, D., Cody, R., Herr-Stephenson, B., . . . Tripp, L. (2013). Hanging out, messing around, and geeking out: Kids living and learning with new media. Cambridge, MA: MIT Press.

Jensen, E., & Lister, T. J. P. (2015). Evaluating indicator-based methods of 'Measuring long-term impacts of a science center on its community. Journal of Research in Science Teaching, this issue.

P. J. Lavrakas (Ed.). (2008). Encyclopedia of survey research methods. Thousand Oaks, CA: Sage.

Leinhardt, G., Crowley, K., & Knutson, K. (2002). Learning conversations in museums. Mahwah, NJ: Lawrence Erlbaum.

Lemke, J. L., Lecusay, R., Cole, M., & Michalchik, V. (2012). Documenting and assessing learning in informal and media-rich environments. Boston, MA: MIT Press.

National Research Council (NRC) (2009). Learning science in informal environments: People, places, and pursuits. Committee on Learning Science in Informal Environments. In: P. Bell, B. Lewenstein, A.

Shouse, & M. Feder (Eds). Board on Science Education, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Organization for Economic Cooperation and Development (OECD). (2012). PISA in Focus 18: Are students more engaged when schools offer extracurricular activities? Paris: OECD.

Renninger, K. A., & Riley, K. R. (2013). Interest, cognition and case of L- and science. In S. Kreitler (Ed.), Cognition and motivation: Forging an interdisciplinary perspective (pp. 352-382). Cambridge, MA: Cambridge University Press.

Stocklmayer, S. M., Rennie, L. J., & Gilbert, J. K. (2010). The roles of the formal and informal sectors in the provision of effective science education. Studies in Science Education, 46, 1–44.

Taplan, S. (2005). Methodological design issues in longitudinal studies of children and young people in out-of-home care: A literature review. Technical Report. NSW Centre for Parenting & Research. Sydney, Australia: NSW Department of Community Services.

Vaske, J. J. (2008). Survey research and analysis: Applications in parks, recreation and human dimensions. State College, PA: Venture.