Critique and Rebuttal

Crossmodal Matching: The Case for Developing and Employing a Valid and Feasible Approach to Equate Perceived Stimulus Intensities in Multimodal Research

Sara Lu Riggs, Clemson University, South Carolina, USA, and **Nadine Sarter**, University of Michigan, Ann Arbor, USA

Keywords: multimodality displays, displays and controls, tactile, touch, haptic, sensory and perceptual processes, multisensory integration, uninhabited aerial vehicles, human-robot interaction

We are grateful to the Editor of *Human Factors* for extending us this opportunity to highlight the need for developing and including a critical step in multimodal research: crossmodal matching. We would also like to thank the reviewer who wrote the critique of our article for his/her helpful and constructive feedback on our work. In the following sections, we will address the reviewer's main comments regarding the specifics of our approach to crossmodal matching. However, the higher-level goal for us is to use this published exchange to raise awareness of a shortcoming of most multimodal research to date and to trigger a discussion in the human factors community about appropriate and feasible crossmodal matching techniques.

The goal of crossmodal matching is to ensure that the perceived intensities of stimuli across two or more sensory channels are equivalent (Colman, 2008; Pitts, Riggs, & Sarter, 2016; Stevens, 1959). Crossmodal matching is necessary to account for the known large between-subject variability in stimulus perception and to avoid confounding modality with intensity when comparing performance with visual, auditory, and tactile stimuli (e.g., Gomes & Riggs, 2016; Pitts, Lu, & Sarter, 2013; Pitts & Sarter, 2014). A recent review of the multimodal literature in human factors and ergonomics journals revealed that only 2.7% of studies performed or reported crossmodal matching (Pitts et al., 2016). This raises questions about the validity of many findings in this important research area. When crossmodal matching was conducted, details about the specific method used are often not shared, and when the information is provided, it shows that most but not all studies based their approach on psychophysics research conducted by Stevens and colleagues in the 1950s and 1960s (e.g., Bond & Stevens, 1969; Stevens, 1959). The findings from our review highlight the need for developing a valid, agreed-upon method for matching signal intensity across sensory channels, especially for *applied* multimodal research.

The current study did employ a crossmodal matching technique that was based on the *method of adjustment* (Bond & Stevens, 1969). Specifically, six participants were asked to adjust the intensity of the variable cue (i.e., the stimulus that could be adjusted) in one modality until they felt that its intensity was equal to that of the reference cue (i.e., the stimulus that the participant was matching to) in another modality. For this study, participants were matching between vision and touch. The values obtained for each variable cue were then averaged across participants, and the mean values were used for the visual and tactile stimuli during the subsequent experiment. The reviewer points out that the mean values derived from the pilot sample may not have been perceived as equivalent by participants in the subsequent

Address correspondence to Sara Lu Riggs, Clemson University, 100 Freeman Hall, Clemson, SC 29634, USA; e-mail: sriggs@clemson.edu.

HUMAN FACTORS
Vol. 61, No. 1, February 2019, pp. 29–31
DOI: 10.1177/0018720818816439
Article reuse guidelines: sagepub.com/journals-permissions

TABLE 1: Intensity Levels for Visual and Tactile Cues

	Modality Condition	
Intensity Level	Visual	Tactile
Low (baseline)	Low-level brightness	No vibration
Medium (to direct attention to UAV sector of interest)	Medium-level brightness	Matched medium level of vibration intensity
High (occurred only during change trials which were the focus of our research)	High-level brightness	Matched high level of vibration intensity

Note. UAV = unmanned aerial vehicle.

experiment and that, instead, crossmodal matching should have been conducted for each participant before the start of each experimental session. We agree that, in principle, this would be the preferable approach, but we would like to highlight some problems with this technique. First, the crossmodal matching task in this study took each participant 15–25 min to complete. This means that, if all participants in an experiment perform crossmodal matching, the overall time required for this step will quickly become prohibitive. Also, keeping each participant engaged and mitigating the effects of stimulus fatigue are challenges we have faced while conducting crossmodal matching. Second, there is no consensus on the mapping of parameters across modalities. For example, frequency can be modulated for touch and audition, but what is the equivalent parameter in the visual channel? Ensuring crossmodal equivalence becomes inherently more complex with studies involving stimuli that embed more than one parameter in a sensory modality. The situation becomes even more problematic when a range of perceptual contexts and stimulus parameters are included, as suggested by the reviewer. Again, we do not, in principle, disagree with this suggestion; however, this will further increase the duration and complexity of crossmodal matching. We believe that the method used in the present study represents a tradeoff. While it can likely be improved, our approach allowed us to establish, in a reasonable amount of time, an approximation to signal equivalence.

Another issue the reviewer highlights is that, in the present study, the amount of information available for making intensity discrimination judgments may have been different in each modality. It is true that, in the visual channel, discrimination could be based on relative brightness of the active feed compared with the surrounding inactive feeds, as well as the successive brightness changes in the active feed. In the tactile condition, on the other hand, only successive changes in intensity for the active tactor were available to participants. It is difficult to see how this issue could be addressed. Making the two conditions equivalent would require the continuous activation of all tactors throughout the experiment, which would create problems due to sensory desensitization in the tactile channel (Caldwell, Lawther, & Wardle, 1996). It is also important to keep in mind that, for change blindness (the main focus of this study), the critical discrimination judgment in both modalities was based on successive changes between the medium and high intensity level of the visual and tactile cues. The increase in intensity from low (baseline; where the visual channel provides more information) to medium was introduced solely for the purpose of attention capture and guidance (Table 1).

In conclusion, we agree with the reviewer that our approach to crossmodal matching could be improved and that including some crossmodal matching technique is necessary but not sufficient; however, by including this step, the reported research goes beyond the vast majority of (published) studies on multimodal information processing and display design to date. As pointed out by the reviewer, even in isolation, our unimodal results make a valuable contribution. We are thankful for the opportunity to share these findings with the readers of *Human*

Factors. We hope that this open exchange will lead the human factors community to work toward the development of a crossmodal matching technique that is both feasible and valid and that will become a required step in multimodal studies. We encourage our colleagues to work toward the development of a crossmodal matching technique that considers the following:

- *Individual differences.* Namely, minimize within-subject variability but account for between-subject variability in perceived differences between modalities.
- *Differences between modalities*. Ensure that differences between sensory channels are accounted for—namely the parameters associated with each sensory channel—but balancing what is appropriate for the nature of the applied work.
- *Various contexts and tasks types*. Ideally, the method should be able to be adopted in various domains, but also in more controlled laboratory settings to match between modalities. The method should be generalizable across various domains, environments, tasks, contexts, and stimuli.
- *Feasibility*. It is critical to develop a method that researchers, display designers, and users will want to adopt. To this end, factors to consider include whether the method can keep the users engaged and how much time it adds to a study or the tailoring of a specific display to a customer.

In conclusion, crossmodal matching is critical for ensuring not only the validity of research findings, but also the robustness of multimodal interfaces that increasingly find their way into a wide range of data-rich domains (Sarter, 2006).

ORCID ID

Sara Lu Riggs https://orcid.org/0000-0002-0112-9469

REFERENCES

Bond, B., & Stevens, S. S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds. *Perception & Psychophysics*, *6*(6), 337–339.

Caldwell, D. G., Lawther, S., & Wardle, A. (1996). Multi-modal cutaneous tactile feedback. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, *2*, 465–472.

Colman, A. M. (2008). *A dictionary of psychology* (3rd ed.). Oxford, UK: Oxford University Press.

Gomes, K., & Riggs, S. L. (2016). Crossmodal matching: A comparison of two methods. In *Proceedings of the Human Factors and Ergonomics Society 60th Annual Meeting* (pp. 1595–1599). Santa Monica, CA: Human Factors and Ergonomics Society.

Pitts, B. J., Lu, S. A., & Sarter, N. B. (2013). Cross-modal matching: The development and evaluation of a new technique. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (pp. 1760–1764). Santa Monica, CA: Human Factors and Ergonomics Society.

Pitts, B., Riggs, S. L., & Sarter, N. (2016). Crossmodal matching: A critical but neglected step in multimodal research. *IEEE Transactions on Human-Machine Systems*, *46*(3), 445–450.

Pitts, B. J., & Sarter, N. B. (2014). Crossmodal matching: Validation of a more reliable technique. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 1751–1755). Santa Monica, CA: Human Factors and Ergonomics Society.

Sarter, N. B. (2006). Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics*, *36*(5), 439–445.

Stevens, S. S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology*, *57*(4), 201–209.

Sara Lu Riggs is an assistant professor in the Department of Industrial Engineering at Clemson University. She earned her PhD in industrial and operations engineering from the University of Michigan in 2014.

Nadine Sarter is a professor in the Department of Industrial and Operations Engineering – Center for Ergonomics at the University of Michigan. She received her PhD in industrial and systems engineering from Ohio State University in 1994.

Date received: November 6, 2018
Date accepted: November 9, 2018