

2005

## The Value of Pooling “Naive” Expertise

Marc S. Schulz  
mschulz@brynmawr.edu

Robert J. Waldinger

[Let us know how access to this document benefits you.](#)

Follow this and additional works at: [http://repository.brynmawr.edu/psych\\_pubs](http://repository.brynmawr.edu/psych_pubs)

 Part of the [Psychology Commons](#)

---

### Custom Citation

Schulz, M.S., & Waldinger, R.J. (2005). The value of pooling "naïve" expertise. *American Psychologist* 60, 656-657.

This paper is posted at Scholarship, Research, and Creative Work at Bryn Mawr College. [http://repository.brynmawr.edu/psych\\_pubs/22](http://repository.brynmawr.edu/psych_pubs/22)

For more information, please contact [repository@brynmawr.edu](mailto:repository@brynmawr.edu).

Schulz, M.S., & Waldinger, R.J. (2005). The value of pooling "naïve" expertise. *American Psychologist*, 60, 656-657.

### **The Value of Pooling “Naive” Expertise**

Marc S. Schulz, Bryn Mawr College  
Robert J. Waldinger, Harvard Medical School

We read with great interest the article by Westen and Weinberger (2004) exploring the benefits and limitations of clinical observation and judgment. They make a persuasive case for using statistical methods to aggregate clinicians’ observations in clinical and personality research. Just as increasing items on an assessment instrument can improve reliability, statistically aggregating the judgments of multiple observers is likely to yield a more reliable judgment than that of a single observer. The advantages of pooling multiple observations, however, are not limited to research in clinical and personality psychology. Westen and Weinberger identify two categories of informants – clinicians and participants – but these categories could be expanded to include other observers who might have particular expertise or experience related to the phenomenon of interest. The type of expert best suited to provide observations depends on the type of expertise required. For reports about subjective psychological or physical experience, the individual him- or herself is likely to be most expert. For assessment of clinically observable phenomena, clinicians may indeed bring value-added expertise. There are some domains, however, in which those with greatest expertise are neither specially trained observers nor self-reporters, but rather lay observers who have a native or learned ability to detect complicated social or psychological phenomena and make subtle discriminations. This type of expertise is often thought of as intuitive because it utilizes implicit knowledge that is not always accessible

to conscious awareness or capable of being fully articulated. One way to harness this intuitive expertise effectively is pooling the judgments of multiple lay observers.

Using lay expertise in observational research runs counter to the prevailing belief that manualized coding systems are required to obtain valid, reliable and reproducible results. To be sure, manualized approaches have contributed greatly to research in psychology. However, some phenomena do not lend themselves easily to codification and rule-driven observation. For example, emotions, power dynamics, validation, and empathy have all been shown to be of great relevance to interpersonal functioning but are challenging to operationalize. Yet, they are phenomena that individuals routinely evaluate in their everyday interactions with others. When using manual-based systems, coders may find that the rules and procedures designed to facilitate rating of these constructs do not fully utilize, and even interfere with, the natural decoding expertise that they have developed. Needing, for example, to process all the information in a manual and simultaneously attend to the behavioral segment being coded may overload the cognitive capacities of many individuals and interfere with their ability to utilize important cues.

Consider the example of assessing emotion. Because it is socially adaptive to do so, most human beings develop a sophisticated and almost instantaneous ability to recognize others' emotions. We use complex internal algorithms to synthesize disparate cues about emotion from body language, facial expression, vocal qualities, speech content and social context. Of course, individual differences in this proficiency exist. Just as aggregating judgments across clinicians removes idiosyncratic biases and yields a more reliable composite judgment, aggregating the individual judgments of multiple untrained judges yields similar benefits.

Our research has led us to believe that lay observers' intuitive judgments about emotions may in fact capture important information that is lost when coders depend on more commonly

used manualized approaches such as the Specific Affect Coding System (SPAFF, Gottman, McCoy, Coan, & Collier, 1996) and the Facial Action Coding System (FACS, Ekman & Friesen, 1978). In our laboratory, we experimented for a year with manualized approaches to coding emotion and then decided to explore the utility of what is often (we believe erroneously) termed “naïve” coding. Inspired by the work of Rosenthal and colleagues on coding non-verbal behaviors (e.g., Ambady & Rosenthal, 1993), we provided untrained, unmarried college-age individuals a list of 16 emotion descriptors without instructing them on how to interpret these descriptors or on markers of specific emotions. We asked them to watch videotapes of 10-minute couple interactions and to code the emotional expression of each partner. For each 30-second segment of the interaction, coders rated the intensity with which each of the 16 emotions was displayed. Despite the fact that each rater coded in slightly idiosyncratic ways, we obtained good reliabilities on factor-analytically derived scales by combining the ratings of 5 to 6 coders. The aggregated ratings of these “naïve” coders were linked in expected ways with concurrent marital satisfaction and independent assessments of marital adjustment. More importantly, they predicted with more than 80% accuracy a real-world consequence of great significance: whether couples in committed relationships would remain together after 5 years (Waldinger, Schulz, Hauser, Allen, & Crowell, 2004). When we compared emotion ratings of couple interactions using this approach with ratings of the same interactions by experienced coders using the manualized SPAFF, we found a high degree of correlation between the two sets of ratings (Waldinger et al., 2004). We also found evidence that the untrained raters were more sensitive to subtle or low intensity expression of emotions.

Our experience with using lay expertise in assessing emotion is consistent with the research cited by Westen and Weinberger (2004, p. 599) suggesting that “with increasing experience in a

given domain, people are typically able to make more subtle discriminations, process information more efficiently, and automatize procedures that initially required conscious attention and hence consumed working memory resources.” In order to succeed in the social world, individuals must learn to make judgments about many complex psychological constructs, such as emotion, that are complicated to define but are nevertheless meaningful and commonly observed phenomena. They must make these judgments quickly, often while simultaneously engaging in other cognitively demanding activities. Empirical research indicates that people can learn to recognize complex connections in the social world without having an explicit awareness of, or ability to, articulate this implicit knowledge (Westen & Weinberger, 2004). In our zealotry to reduce variability among coders and to make our methods more exportable to other investigators, researchers risk losing the intuitive expertise that people naturally develop in making judgments about the world.

Correspondence concerning this comment should be addressed to Marc S. Schulz, Department of Psychology, Bryn Mawr College, 101 North Merion Avenue, Bryn Mawr, PA 19010. E-mail: mschulz@brynmawr.edu

### References

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*, 431-441.
- Ekman, P. and W. V. Friesen (1978). *The Facial Action Coding System (FACS)*. Palo Alto, CA, Consulting Psychologists Press.
- Gottman, J., McCoy, K., Coan, J., & Collier, H. (1996). The specific affect coding system (SPAFF). In J. Gottman (Ed.), *What predicts divorce? The measures* (pp. 1-169). Hillsdale, NJ: Erlbaum.
- Waldinger, R. J., Schulz, M.S., Hauser, S.T., Allen, J.P., & Crowell, J.A. (2004). Reading others' emotions: The role of intuitive judgments in predicting marital satisfaction, quality, and stability. *Journal of Family Psychology 18*, 58-71.
- Westen, D. & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist, 59*, 595-613.