Who Knows Who Is Wise? Self and Peer Ratings of Wisdom

Uwe Redzanowski and Judith Glück

Department of Psychology, Project Wisdom, Alpen-Adria Universität Klagenfurt, Austria.

Objectives. This study investigated relationships between three measures of wisdom: self-ratings, peer ratings, and a self-report scale. We expected to find a zero or negative correlation between the self-rating and the average peer rating and low positive correlations of both to the self-report scale. We also tested whether there would be more convergence among measures in the top scorers.

Method. A total of 179 members of 17 university departments were rated by their department colleagues with respect to wisdom; about half of them also rated themselves and filled out a wisdom self-report scale (Three-Dimensional Wisdom Scale [3D-WS]; Ardelt, M. (2003)).

Results. There were no significant relationships among the three measures of wisdom, neither in the total sample nor among the top scorers.

Discussion. Depending on the focus of research, peer ratings may be a more suitable measure of wisdom than self-report scales or performance measures.

Key Words: Measurement—Peer rating—Self-rating—Three-Dimensional Wisdom Scale—Wisdom.

For wisdom to appear it takes at least two persons: the wise person and another who identifies him or her as wise. “I am wise” is not a well-formulated sentence in the cultural grammar of wisdom.

(Assmann, 1994, p. 222)

How should we measure wisdom? Wisdom is a complex construct that includes aspects of competence (knowledge, experience, problem-solving skills), reflection (self-evaluation, self-criticism), and self- and other-directed affect (emotion regulation, compassion)—three domains that imply somewhat different measurement approaches (Ardelt, 2003; Staudinger & Glück, 2011; Glück, König, Naschenweng, Redzanowski, Dorner, & Straßer, submitted). Although a number of wisdom measures are available, none of them are optimally suited to cover all three domains.

The two most common ways of measuring wisdom are self-report scales and performance assessments. A number of self-report measures are available (e.g., Ardelt, 2003; Greene & Brown, 2009; Levenson, Jennings, Aldwin, & Shiraishi, 2005; Webster, 2003, 2007), but there are at least two reasons to question their validity. First, most people are not particularly good at judging their own competences (Freund & Kasten, 2012), which may compromise self-ratings of competence aspects of wisdom. Second, the reflective aspect of wisdom may cause a paradox: if a self-critical attitude is a part of wisdom, wise persons—knowing how much they do not know—might actually describe themselves as less wise than unwise persons (Aldwin, 2009; Assmann, 1994). On the other hand, one could argue that truly wise individuals should also be aware of their own wisdom. Even if this is the case, it seems likely that some relatively unwise people, lacking self-reflection, also rate themselves as highly wise. It seems difficult to distinguish these two groups on a self-report basis.

The available performance measures of wisdom, on the other hand (Baltes & Staudinger, 2000; Mickler & Staudinger, 2008), require large effort for data collection, transcription, and coding and may not cover the affective aspects of wisdom adequately. Participants‘ responses may reflect their cognitive competence to think theoretically about life problems, but not their ability and willingness to actually deal with them well in real life (Ardelt, 2004; Glück, Bluck, Baron, & McAdams, 2005).

Empirically, the correlations among measures of wisdom tend to be low (Ardelt, 2011; Glück et al., submitted), which reflects differences in researchers‘ definitions and operationalizations of wisdom. Thus, empirical findings may depend on what measure of wisdom is used. Perhaps one way out of this unsatisfying situation could be the use of peer ratings: If several friends or colleagues of a person agree that he or she is highly wise, this judgment is likely to have some validity. Such peer ratings would obviously be based on the informants‘ personal ideas of what wisdom is, rather than on a researcher‘s theory. As a number of studies have shown, however, laypeople have relatively consistent ideas about the main components of wisdom (overview in Bluck & Glück, 2005).

In this study, we collected within-department peer ratings as well as self-ratings from the members of 17 departments of...
our university and presented them with a self-report wisdom scale (Ardelt, 2003). The main research questions were, first, how are peer ratings, self-ratings, and the scale score interrelated? We expected to find a zero or negative correlation between the self-ratings and peer ratings, based on the assumption that truly wise individuals would be viewed as wise by their peers but not by themselves, whereas people who would view themselves as highly wise would not be judged as wise by their peers. For the self-report wisdom scale, we expected low but positive correlations with both the self- and peer ratings because the scale covered competence-related aspects, reflection-related, and affect-related aspects of wisdom.

One open question is whether wisdom should be considered a continuous variable, such as intelligence or extraversion, or as something more like a clinical “syndrome,” a constellation of characteristics that produces a qualitatively different way of thinking and acting. Thus, peer ratings might be a relatively unreliable indicator of wisdom in the general population but highly consistent when it comes to those few individuals who are truly wise. Therefore, we analyzed the top scorers in each measure separately.

Method
Participants were 179 faculty members (65.9% men, 65.4% junior faculty) of 17 departments of University of Klagenfurt. Their academic fields included the humanities, social sciences, economics, and technical sciences. The number of participants per department ranged from 3 to 19. Peer ratings were collected for 169 participants, 79 participants provided self-ratings, and 93 filled out the self-report scale.

The questionnaire had three parts. Participants were first asked to list characteristics that they associated with wise individuals, largely to get them to think about wisdom before responding to the other parts. Analyses of these lists are not reported here because the content categories were unrelated to all other variables. The second part consisted of one or more perforated A4-sized sheets that could be torn into 10 strips per page. Each strip contained the name of one member of the participant’s department and a 160-mm line with the end points “not at all wise” and “absolutely wise.” Participants were asked to mark one point on the line for each colleague, separate the strips, and throw them into an urn. This procedure had the disadvantage that we could not relate participants’ self-rating to their ratings of their peers, but it was necessary to ensure complete anonymity of the peer ratings. Participants then rated their own wisdom on the same 160-mm scale, and filled out the German version of Ardelt’s Three-Dimensional Wisdom Scale (3D-WS; Ardelt, 2003; German version provided by M. Ardelt). This scale consists of 39 statements measuring the cognitive (e.g., “I am not interested in a problem if I think there is no solution”; 14 items), reflective (e.g., “I always try to look at all sides of a problem”; 12 items), and affective dimension (e.g., “Sometimes I feel real compassion for all people” (13 items) of wisdom.

Results
Cronbach’s alphas were .74, .75, and .49, for the cognitive, reflective, and affective dimension, and .76 for the total 3D-WS. Thus, the affective dimension was less reliable than in other studies (Ardelt, 2003, 2011), perhaps because of
the German version or the academic sample. We also tried to estimate reliability for the peer ratings, which was difficult because the number of ratings per participant varied from 1 to 19. When only three ratings were analyzed, which were available for 158 participants, alpha was .52. For eight ratings (83 participants), alpha was .73, and for 18 ratings (12 participants), it was .83. We consider reliability of the peer ratings as comparable to wisdom self-report scales.

Table 1 shows the correlations among the measures. Significant correlations in the .30 range were only found among the 3D-WS dimensions. The correlations between self-ratings, peer ratings, and the 3D-WS were all insignificant, close to zero, and significantly below .40. The two highest correlations were between the self-rating and the average peer rating (r = .16, p = .19), and between the self-rating and the 3D-WS reflective dimension (r = .17, p = .14). The correlations remained insignificant when we included only participants for whom eight or more peer ratings were available, and when we divided the sample into humanities and social sciences versus engineering and mathematics.

As expected, professors judged themselves as wiser (M = 104.6, SD = 16.3) and were judged as wiser by others (M = 92.80, SD = 24.6) than nonprofessors (self-rating: M = 82.1, SD = 32.4; t[77] = 3.168, p = .002; peer rating: M = 79.3, SD = 20.0; t[167] = 3.831, p < .001). Also, the correlation of .16 between the self-rating and the average peer rating became almost zero when seniority was controlled (r partial = .05, p = .69); thus, it was largely caused by the general tendency to rate senior faculty higher.

Even if the different measures were not significantly correlated in the total sample, we assumed that there might be more agreement for the wisest participants. A scatterplot of the means and standard deviations of the peer ratings showed the highest variability around the scale midpoint; thus, peer agreement was markedly higher for individuals rated as particularly low or high in wisdom than near the scale midpoint. Therefore, we analyzed the top 20% scorers in each measure in more detail. As Figure 2 shows, however, none of the three top 20% groups was significantly above the scale midpoint in the two other measures. Only seven participants were in the top 20% in two measures, and none was in the top 20% in all three measures.

**Discussion**

This study found no statistical relationship between self-ratings, other-ratings, and self-report scale scores of wisdom; in fact, they might as well be measuring different constructs. Now, which of the three approaches is the most valid measure of wisdom? The data do not allow us any direct answer to this question. Conceptually, self-ratings of wisdom are obviously the most dubious approach, as it is unlikely that most individuals who describe themselves as highly wise have the self-reflective capacities that wisdom implies (Aldwin, 2009; Assmann, 1994). Self-report measures have the advantages that they are theoretically grounded and easy to administer. However, they may be partly compromised by the self-reflection problem and the general problems of measuring competences by self-report (Aldwin, 2009). In addition, correlations among self-report measures of wisdom are generally low (Glück et al., submitted), illustrating the large differences among researchers in their conceptions and operationalizations of wisdom.

Are peer ratings a promising alternative? Obviously, peers differ both in their understanding of wisdom and in their experience with a person, and all kinds of relational issues may confound their ratings. In fact, judging other people’s wisdom may require some wisdom in the informants, as an anonymous reviewer of this article noted. Still, our findings suggest peer ratings at least as a possible additional indicator of wisdom, and for some studies, as a possible alternative to more direct measures of wisdom. In fact, wisdom as a construct may generally be more “in the eye of the beholder” than the carrier (Assmann, 1994; Staudinger & Glück, 2011). Although it may sound paradoxical to have a self-reflective competence judged by others than the self, it may be easier to rate others on a consistent metric
than to judge oneself in comparison to others. It is important, however, to decide who the informants should be on a sound theoretical basis. A recent suggestion in the wisdom literature (Staudinger, Dörner, & Mickler, 2005; Staudinger & Glück, 2011) is to distinguish general wisdom (wisdom concerning life in general) as opposed to personal wisdom (wisdom about one’s own life). Informants from a professional context, such as the peers investigated in this study, may be more competent to judge a person’s general wisdom, whereas family members and close friends may be better at judging their personal wisdom. Thus, part of the null findings in this study may be due to the fact that the 3D-WS is a measure of personal rather than general wisdom.

To conclude, we believe that wisdom research would profit from a broader range of methods for measuring wisdom. One such avenue could be the use of peer ratings; others include new experimental measures (Grossmann, Na, Varnum, Park, Kitayama, & Nisbett, 2010; Kross & Grossmann, 2012) as well as new performance measures (Mickler & Staudinger, 2008).

Funding
This work was supported by the Austrian Research Fund (P21011-G14) and the John Templeton Foundation through the University of Chicago’s Defining Wisdom Initiative (JTF grant number 11976).

Correspondence
Correspondence should be addressed to Judith Glück, PhD, Department of Psychology, Alpen-Adria-Universität Klagenfurt, Universitätsstraße 65–67, 9020 Klagenfurt, Austria. E-mail: judith.glueck@aau.at.

References


