
Special issue: New Developments in Psychological Wisdom Research

Measuring Wisdom: Existing Approaches, Continuing Challenges, and New Developments

Mag. Dr. Judith Glück

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

Address correspondence to: Mag. Dr. Judith Glück, Department of Psychology, Alpen-Adria Universität Klagenfurt, Universitätsstraße 65–67, 9020 Klagenfurt, Austria. E-mail: judith.glueck@aau.at

Received: July 17, 2017; Editorial Decision Date: October 13, 2017

Decision Editor: Bob G. Knight, PhD

Abstract

The question how wisdom can best be measured is still open to debate. Currently, there are two groups of wisdom measures: open-ended performance measures and self-report measures. This overview article describes the most popular current measures of wisdom: the Berlin Wisdom Paradigm, the Bremen Wisdom Paradigm, Grossmann's wise-reasoning approach, the Three-Dimensional Wisdom Scale, the Self-Assessed Wisdom Scale, and the Adult Self-Transcendence Inventory. It discusses the specific challenges of both open-ended and self-report approaches with respect to content validity, convergent and divergent validity, concurrent and discriminant validity, and ecological validity. Finally, promising new developments are outlined that may bridge the gap between wisdom as a competence and wisdom as an attitude and increase ecological validity by being more similar to real-life manifestations of wisdom. These new developments include autobiographical approaches and advice-giving paradigms.

Keywords: Wisdom, Measurement, Validity

Measuring Wisdom: New Developments and Continuing Challenges

When I tell people that I am a wisdom researcher, they usually first ask me what wisdom is—and then almost invariably the next question is, “But can you measure that?” How wisdom can be measured is indeed a complex question, and I do not think we have found a fully convincing answer yet. This article intends to review the current state of wisdom measurement, but also to stimulate new research that adds to our toolbox of wisdom measures.

Why do we even need to measure wisdom? Arguably, wisdom has long been studied in philosophy, theology, or the historical sciences without a need to assign numbers to individuals and run them through complex statistical analyses.

I believe that wisdom research benefits greatly from the use of qualitative methodologies (e.g., [Edmondson, 2005](#); Igarashi, Levenson, & Aldwin, in revision) or from combining qualitative and quantitative approaches (e.g., [DeMichelis, Ferrari, & Rozin, 2015](#); [Glück, Bluck, Baron, & McAdams, 2005](#); [König & Glück, 2013](#); [Weststrate & Glück, 2017](#)). However, reliable and valid measures of wisdom allow us to study complex psychological research questions in larger samples of individuals that could not otherwise be investigated. In addition, as I hope to show in the following, trying to capture an elusive concept like wisdom in a standardizable way is a rather fascinating creative endeavor in itself that can teach us a lot about the potentials and the limitations of psychological measurement in general.

The essential question of the ongoing discussion about how wisdom can best be measured (e.g., Ardel, 2004; Brienza, Kung, Santos, Bobocel, & Grossmann, 2017; Glück et al., 2013; Staudinger & Glück, 2011) is really about validity: how does wisdom manifest itself, and how can we access its manifestation in our typical measurement conditions? I first review current measures, then discuss validity challenges, and finally describe some promising new developments.

An Overview of Operationalizations of Wisdom

Arguably, there are almost as many definitions of wisdom in the field as there are wisdom researchers (for overviews, see Bangen, Meeks, & Jeste, 2013; Glück, 2013, 2016; Staudinger & Glück, 2011), and several of them have been operationalized into measures. Here, I focus on the most well-known and established measures, starting with the chronologically oldest one. The online supplement gives example items or vignettes and summarizes evidence on reliability and validity.

Performance Measures of Wisdom

The Berlin wisdom paradigm

The Berlin wisdom model was developed by Paul Baltes and his coworkers at the Max Planck Institute for Human Development in Berlin, including Ursula M. Staudinger, Jacqui Smith, and Ute Kunzmann (Baltes & Smith, 1990; Baltes & Staudinger, 2000). Based on earlier work on dialectical and post-Piagetian cognition and in line with other research in the 1980s, they defined wisdom as expert knowledge. The term “expert knowledge” usually refers to broad and deep knowledge and skill in a specific domain that is acquired through long-term deliberate practice (e.g., Ericsson, Krampe, & Tesch-Römer, 1993). The Berlin group argued that the subject matter of wisdom is the fundamental pragmatics of human life, including dealing with mortality, resolving moral dilemmas, or balancing intimacy and autonomy. Expert knowledge was usually measured using think-aloud methods. The *Berlin wisdom paradigm* (BWP) requires participants to think aloud about brief vignettes describing difficult fictitious life problems. An example is, “Somebody gets a phone call from a good friend. The friend says that he cannot go on any more, and that he has decided to commit suicide. What could one/the person consider and do in such a situation?” The Berlin group proposed five criteria for determining the level of wisdom in think-aloud transcripts. Typical for expert knowledge, a wise response shows high levels of factual knowledge (e.g., knowledge about why people might want to commit suicide) and procedural knowledge (e.g., knowledge about strategies to deal with the caller). In addition, it demonstrates life-span contextualism (discussing how life phases,

life situations, or historical and cultural settings can influence behavior), value relativism (awareness and acceptance of different values, beliefs, and priorities), and recognition and management of uncertainty (awareness of and ability to deal with life’s inherent unpredictability). Response transcripts are evaluated by two independent trained raters per criterion using seven-point scales, and the average across the five criteria is used as the wisdom score.

The Berlin wisdom paradigm is a prototypical measure of wisdom as a competence: an expert way of thinking about life problems. The question how real-life competences can best be measured dates back at least to McClelland (1973), and it will return in my discussion of ecological validity below. Generally, however, competences are measured by assessing performance, that is, by presenting participants with relevant problems and evaluating the quality of their solutions. The approaches reviewed here all use open-ended responses. Closed-response approaches are not available yet. Our own unsuccessful attempts at developing “wisdom tests” with predefined response alternatives suggest that many more people may pick the wisest solutions from among some alternatives than would actually be able to produce a wise solution by themselves (but see Mitchell, 2016; Sternberg, 2001). The following two approaches also measure wisdom as a competence.

Grossmann’s conception of wise reasoning

Igor Grossmann defines wise reasoning as “the use of certain types of pragmatic reasoning to navigate important challenges of social life” (Grossmann et al., 2010, p 7246). Wise reasoning involves dialectical thinking and intellectual humility as manifested, for example, in taking different perspectives, recognizing the limitations of knowledge, making flexible predictions, and searching for compromise. To measure wisdom, Grossmann and colleagues present participants with difficult life problems concerning personal or larger-scale societal issues and ask them to discuss how the situations might unfold and why. Responses are collected in written or oral formats and evaluated by trained raters. Grossmann et al. (2010) showed that wise reasoning about group conflicts increases with age and wise reasoning about individual conflicts increases or stays stable. Experimentally, Kross and Grossmann (2012) demonstrated that wise reasoning increases when participants take a self-distanced perspective.

The Bremen wisdom paradigm

Ursula M. Staudinger (Mickler & Staudinger, 2008; Staudinger, Dörner, & Mickler, 2005; Staudinger & Glück, 2011) has argued that psychological definitions of wisdom can be grouped into two categories: general wisdom, concerning questions of human life in general, and personal wisdom, which concerns oneself and one’s own life. General wisdom may developmentally precede personal wisdom, as it is often difficult to apply insights to one’s own life. Mickler and Staudinger (2008) developed a performance measure

of personal wisdom, the Bremen wisdom paradigm (BrWP) that conceptually parallels the Berlin Wisdom Paradigm. The criteria for personal wisdom are rich self-knowledge (about one's strengths and weaknesses, priorities, and life meaning), heuristics of growth and self-regulation (knowing how to deal with challenges and develop positively), interrelating the self (being aware of one's contextual and social embeddedness), self-relativism (self-criticism and self-reflection balanced with self-esteem), and tolerance of ambiguity (recognizing and managing uncertainty and uncontrollability). To measure personal wisdom, participants are asked to think about themselves as a friend and answer questions concerning typical behaviors, dealing with difficult situations, strengths and weaknesses, and reasons for their own behavior. As in the BWP, two raters per criterion evaluate the response transcripts.

The three approaches reviewed above measure wisdom as a competence, that is, as the ability to find good solutions to a certain kind of problems: wise thinking is based on rich, experience-related knowledge about the complex questions of human life and involves distancing oneself from one's own perspective, acknowledgement of the variety of people's needs, values, and perspectives, and a broad and integrative view. Thus, the assumption is that wisdom can be measured by evaluating to what extent people's responses to open-ended problems display these and related characteristics. Developers of competence measures do not necessarily consider wisdom as an exclusively cognitive phenomenon (see, e.g., Baltes & Kunzmann, 2004), but they believe that competence is the core aspect of wisdom that measures should focus on. Other researchers, all studying personal wisdom, take a different approach.

Self-report Measures of Wisdom

Ardelt's three dimensions of wisdom

In 2004, Monika Ardeli published a critique of the Berlin wisdom paradigm arguing that the driving force of wisdom is not knowledge, but personality. She wrote that only experiential, internalized knowledge constitutes wisdom—that there are many things that everybody “knows” as they are part of our societal knowledge as proverbs or maxims, but to actually realize these insights in one's own life and to act according to them, one needs personal experience. In other words, Ardeli considers as wisdom only what Mickler and Staudinger called personal wisdom. In addition, Ardeli (2003, 2004) argued, the core of wisdom is not knowledge—although a wise person will certainly have knowledge—but a certain personality structure that leads people to gain experience-based insights and grow from them. She defined wisdom as a constellation of three personality dimensions. The reflective dimension is a willingness to take different perspectives. Reflective individuals will also try to see themselves from others' perspectives, which enable them to learn from mistakes and overcome defense mechanisms. The

cognitive dimension is a deep desire for understanding, for truth, even if the truth may compromise one's positive self-image. The affective dimension is defined as compassionate love for others—a caring concern for the needs and problems of other people. Only individuals high in all three dimensions are considered as wise.

Obviously, the idea of wisdom as a personality construct suggests a different measurement approach than wisdom as a competence. In line with typical personality measures, Ardeli (2003) developed the *Three-Dimensional Wisdom Scale* (3D-WS), a 39-item self-report scale that assesses participants' agreement to statements reflecting the three dimensions of wisdom, such as “Things often go wrong for me by no fault of my own” (reverse-coded) for the reflective dimension, “Sometimes I feel a real compassion for everyone” for the affective dimension, or “Ignorance is bliss” (reverse-coded) for the cognitive dimension. It is probably the most-used wisdom measure to date.

Several other authors have also defined wisdom as an attitude or a personality trait—a way of experiencing life and reflecting upon it. They do not deny that this attitude will bring about knowledge and competence, but their measurement approaches focus on the attitude, which they measure by self-report.

Webster's self-assessed wisdom scale

Webster (2003, 2007) defined wisdom as “the *competence* in, *intention* to, and *application* of, *critical* life experiences to facilitate the *optimal* development of *self* and *others*” (Webster, 2007, p.164; italics by original author). Based on a literature review, he proposed five components of wisdom: critical life experience (having had personal experiences that were complex and uncertain), openness (to different views, knowledge, strategies, and one's own inner experience), emotional regulation (sensitivity to complex feelings and being able to regulate them), reminiscence and reflectiveness (evaluation and integration of past experiences, applied to future problems), and humor (recognizing ironies and using humor for stress reduction and bonding). To measure wisdom, Webster (2003, 2007) developed the Self-Assessed Wisdom Scale (SAWS), which consists of 40 items (eight per component) such as “I have had to make many important life decisions” (critical life experience), “I'm very curious about other religious and/or philosophical belief systems” (openness), “I can regulate my emotions when the situation calls for it” (emotion regulation), “I often think about my personal past” (reminiscence and reflectiveness), or “I can chuckle at personal embarrassments” (humor).

Wisdom as self-transcendence

Michael R. Levenson defined wisdom as self-transcendence (Levenson, Jennings, Aldwin, & Shiraishi, 2005) based on contemplative traditions, conceptions of gerotranscendence (Tornstam, 1994), and a philosophical analysis of commonalities of wisdom across cultures

(Curnow, 1999). Self-transcendence is achieved through self-knowledge (awareness of the sources of one's self), detachment (awareness of the provisional nature of external sources of self), and integration (acceptance of all self-aspects). Self-transcendence (wisdom) is defined as independence from external self-definitions and the dissolution of rigid boundaries between the self and others. To measure wisdom, Levenson et al. (2005) devised the Adult Self-Transcendence Inventory (ASTI). The current, revised version of the ASTI consists of 35 items, 10 of which refer to alienation which is assumed to be the conceptual opposite of wisdom. Sample wisdom items include "My peace of mind is not easily upset" and "I feel that my individual life is part of a greater whole."

Brief wisdom screening scale

Using a purely empirical approach, we identified those 21 items from the three self-report scales described above that had the highest correlation with the common factor across the three scales (Glück et al., 2013). Thus, the BWSS is not based on any particular conception of wisdom and does not have any specific subcomponents. It is recommended as a screening measure for studies where wisdom is not the main variable of interest. Sample items include "I've learned valuable life lessons from others" (from the SAWS) and "I can accept the impermanence of things" (from the ASTI).

Other self-report wisdom scales include the Foundational Value Scale (Jason, Reichler, King, Madsen, Camacho, & Marchese, 2001) and the Wisdom Development Scale (Brown & Greene, 2006; Greene & Brown, 2009).

Two Approaches Tapping the Same Phenomenon?

In sum, while some definitions measure wisdom as a competence, others measure it as an attitude or personality trait, a way of experiencing and reflecting on life that includes a desire to achieve meaning and growth rather than closure and satisfaction (Weststrate & Glück, 2017), an open, compassionate stance toward others and a willingness to reflect deeply and self-critically. Wisdom as an attitude is perfectly compatible with wisdom as a competence, as it is likely to lead to the acquisition of broad and deep knowledge about oneself and human nature in general.

Thus, the two groups of wisdom measures may actually tap aspects of the same construct. In one of our studies, independent raters evaluated difficult-event interview transcripts from 94 participants (including 47 wisdom nominees; for details see Glück et al., 2013) with respect to the components of four different wisdom conceptions: the three-dimensional wisdom model, the Berlin wisdom paradigm, the Bremen wisdom paradigm, and the MORE Life Experience model (*see below*). The transcripts were also rated for wisdom by lay raters. The same participants completed the 3D-WS, ASTI, SAWS, and BWP. Table 1 shows the correlations between the ratings for the four

Table 1. Correlations Among Ratings of Interview Transcripts According to Four Conceptions of Wisdom

	BWP rating	BrWP rating	MORE rating	Lay rating
3D-WS Rating	.72**	.80**	.83**/.48** ^a	.63**
BWP Rating		.69**	.72**	.62**
BrWP Rating			.76**	.68**
MORE Rating				.77**

Note: $N = 94$ (for details about the sample, see Glück et al., 2013). Transcripts of autobiographical interviews about a difficult life event were rated on four-point scales for all subcomponents of the four wisdom conceptions. Two independent trained students rated each transcript. BWP = Berlin wisdom paradigm; BrWP = Bremen wisdom paradigm.

** $p < .01$.

^aCorrected for overlapping subcomponents (reflectivity, empathy removed from MORE score).

Table 2. Correlations Between Four Measures of Wisdom

	ASTI	SAWS	BWP
3D-WS	.58**	.26**	.25*
ASTI		.50**	.30**
SAWS			.25*

Note: ASTI = Adult Self-Transcendence Inventory; BWP = Berlin wisdom paradigm; SAWS = Self-Assessed Wisdom Scale.

* $p < .05$; ** $p < .01$. $N = 94$ (for details, see Glück et al., 2013).

conceptions, and Table 2 shows the correlations between the four actual measures. As the tables show, the correlations between the different ratings were all above .60, with an average of .72, suggesting some common phenomenon being tapped by all ratings. The correlations between the measures, however, were only in the .20–.30 range except for those of the ASTI with the 3D-WS (.58) and SAWS (.50). Thus, there seems to be far more common variance among conceptions than among measures of wisdom, suggesting a crucial role of variations in measurement methodology.

Assuming that wisdom includes a competence component and an attitude component, is it sufficient to measure only one of them or do we need to assess both? The "wisdom attitude" may not necessarily lead to wisdom if, for example, a person is intellectually unable to mentally represent complex issues. On the other hand, highly intelligent people may be able to "fake" wisdom in verbal responses. To answer the question on an empirical basis, we first have to devise valid measures of both components.

Evaluating the Validity of Wisdom Measures

How can we test whether a measure of wisdom indeed measures wisdom? All measures described above have been evaluated for validity (see online supplement), still, none seems completely satisfactory. Evaluating the validity of

wisdom measures involves specific challenges, as will be discussed in the following.

Content Validity

Content validity describes how well the items or problems in a measure represent the respective content domain. In constructing self-report measures, researchers usually start from a definition of the construct. Ardel (2003), for example, used definitions of her three wisdom dimensions to select 140 candidate items from existing personality measures and constructed 18 additional items. Five experts assigned the 158 candidate items to the three wisdom dimensions. Those 90 items for which at least four experts agreed were administered to participants, and statistical criteria (such as response range, skewness, and item-total correlations) were used to select the final 39 items of the 3D-WS. Other authors (Levenson et al., 2005; Webster, 2003) did not select their items from a pool—they just wrote them and used factor analysis to confirm the expected structure.

We believe that the importance of the first steps in the construction of a measure is often underestimated (Koller, Levenson, & Glück, 2017). The precision of the construct definitions used for item development or selection determines the representativeness of the items. If the items are clearly representative of the intended dimensions, empirical data are likely to confirm the expected structure. It is important, in our view, to select items carefully and to rely on external experts to ensure objectivity. Using an elaborate mixed-methods procedure for evaluating content validity, Koller et al. (2017) analyzed the 35-item version of the ASTI and identified five inter-related subdimensions.

In addition, the conceptual breadth of a construct should be considered. For example, the subcomponents of the SAWS are more internally consistent than those of the 3D-WS because they are more narrowly focused on specific aspects. Narrow scales, however, may not adequately represent complex constructs like wisdom. The 3D-WS subdimensions and the ASTI are based on relatively broad definitions, which may reduce internal consistency but is perhaps more representative of wisdom. Thus, there may be a tradeoff between precise, narrow scales that assess small portions of a construct and broader, less precise, but more comprehensive scales.

Importantly, the content of a measure also determines its score distribution. Figure 1 displays the score distributions of the 3D-WS, SAWS, ASTI, and BWP (Glück et al., 2013). The distributions of the self-report measures are almost exclusively located in the upper half of the response scales, whereas the BWP scores are in the lower half. Thus, far more people describe themselves as highly wise in the self-report scales than are actually likely to be. The most plausible explanation for this is that most people are not particularly good at evaluating themselves concerning positive characteristics. Self-report measurement of concepts that include self-criticism creates an important paradox (Aldwin, 2009; Glück et al., 2013): relatively unwise people may describe themselves as wiser than wise people who are keenly aware of their limitations. In the 3D-WS, Ardel (2003) tried to ameliorate this problem by formulating most items negatively, assuming that a wise person would disagree with, for example, “Ignorance is bliss,” while an unwise person might agree. In the ASTI, some items circumvent the problem by making sense only to highly self-transcendent individuals. The item “Whatever I do to others, I do to myself,” for example, regularly confuses our students. Beyond correlations with social desirability scales, the effects of these strategies have not yet been empirically investigated.

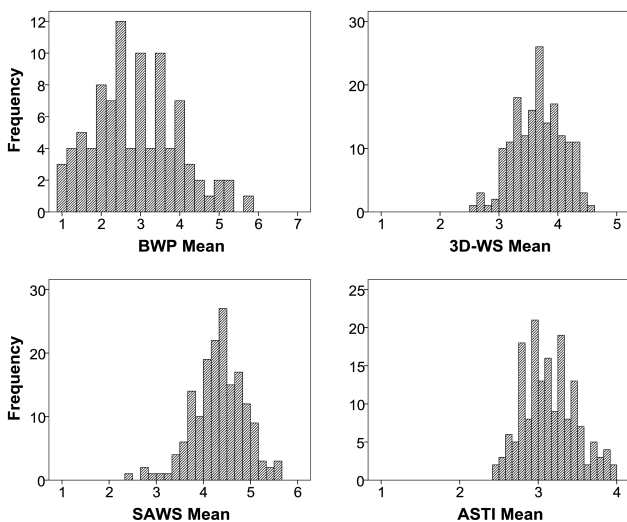


Figure 1. Score distributions of four wisdom measures. The horizontal axis reflects the response scale for each measure. Note: $N = 94$ for the BWP; $N = 170$ for the 3D-WS, ASTI, and SAWS. See Glück et al. (2013) for further information. ASTI = Adult Self-Transcendence Inventory; BWP = Berlin wisdom paradigm; SAWS = Self-Assessed Wisdom Scale.

Content validity of performance measures

Of course, the importance of content validity is not limited to self-report scales. Developers of performance measures also think carefully about the problems that they use to elicit wisdom. The vignettes of the Berlin wisdom paradigm were developed systematically to represent difficult problems of life review, life planning, and life management for different age groups (Smith & Baltes, 1990). As the BWP takes a lot of time and effort, however, most researchers use no more than three problems, and the ones listed in the online supplement are used most frequently (Glück & Baltes, 2006; Staudinger & Baltes, 1996). Specific tasks were developed for adolescents (Pasupathi, Staudinger, & Baltes, 2001). Grossmann’s problem vignettes vary somewhat across studies, but they generally focus on either societal or individual problems of high uncertainty and complexity (see online supplement). Mickler and Staudinger (2008) discussed several reasons for selecting friendship as the topic of the BrWP. Thus, the content of the existing performance measures is consistent with definitions of competence aspects of wisdom.

Details of wordings and instructions can be important as well. For example, the BWP suicide problem sometimes elicits very concrete, problem-focused responses from people with relevant experience, focusing on how to prevent the caller from killing him- or herself rather than on variations in contexts and values or on uncertainty, which leaves them with a low BWP score. It is important, therefore, to emphasize that participants should talk about what one *could*, rather than *should*, consider and do.

A more general problem with performance measures is that they may disregard affective aspects of wisdom. From my own experience with the BWP (Glück & Baltes, 2006), I clearly remember one interview where I distinctly felt that the participant, a man in his thirties, was saying things that sounded very wise, but was lacking something important. He seemed to be highly intelligent and to have an excellent idea of what we wanted to hear, but I was not convinced that he would act half as wisely faced with the same problem in real life. Generally, problem vignettes about fictitious persons may not include affective aspects of wisdom like compassion or emotion regulation, which are crucial to wisdom in real life. I will come back to this issue later.

Convergent and Divergent Validity

An important aspect of validity concerns relationships with other variables. As the table in the online supplement shows, most correlations are in line with what one would expect: wisdom is positively related to intelligence, openness to experience, well-being, and aspects of self-maturity. These relationships are again influenced by method variance: self-report scales have higher correlations with other self-report scales, whereas performance measures have higher correlations with open-ended measures. In addition, the specificity of the correlational patterns is somewhat limited, especially for the self-report measures: positively valued constructs generally tend to correlate in the .20–.30 range, with little differentiation between aspects that one would expect to be particularly close to wisdom and others that seem more distant. More specific, theory-based predictions about the size of specific correlations would be helpful. Divergent validity has hardly been explored except for neuroticism, which is negatively related to at least the ASTI and the SAWS. It might be worthwhile to think about other, preferably positively valued, constructs that should be unrelated or negatively related to wisdom.

It also seems important to go beyond correlations. Quadratic relationships have been found, for example, with fluid intelligence (Mickler & Staudinger, 2008) and extraversion (Staudinger, Maciel, Smith, & Baltes, 1998). The same might be true for many other constructs where a healthy balance is more typical for wisdom than any extreme. Other forms of relationships, such as a certain minimal level of intelligence as a prerequisite of wisdom but no correlation above that level, might also be worth testing.

Relations to demographic variables

How should a measure of wisdom be related to age? Different relationships have been found for different measures, including zero correlations for the BWP (Staudinger, 1999), the ASTI (Glück et al., 2013), and Grossmann's wise reasoning measure (Grossmann & Kross, 2014), negative correlations for the BrWP (Mickler & Staudinger, 2008) and the 3D-WS (Ardelt, 2003), a positive correlation for Grossmann's wise-reasoning measure (Grossmann et al., 2010), an inverse U-shaped relationship for the SAWS (Webster, Westerhof, & Bohlmeijer, 2012)—a pattern that may be more ubiquitous, but has not been tested very often (but see Ardel, in revision), and a U-shaped relationship for the Situated Wise Reasoning Scale (Brienza et al., 2017; see below). A linear correlation between age and wisdom seems unlikely: laypeople and wisdom researchers agree that wisdom may come with age (or rather, with life experience), but only in a few individuals who are able to reflect and integrate life experiences in wisdom-fostering ways (e.g., Ardel, 2004; Glück & Bluck, 2014; Staudinger, 1999). Thus, in samples with a broad age distribution, one would expect the wisest participants to be older than the rest, but not a general linear trend. Such a pattern has rarely been found, however (Glück et al., 2013).

An important validity issue concerns the relationship between wisdom and education. On the one hand, people striving for learning and growth are likely to also seek out formal education, suggesting a positive relationship. On the other hand, open-ended measures in particular might be biased favoring more educated participants. Correlations of the BWP and the 3D-WS with education are significant, but relatively low, suggesting no strong confounding effect (Glück et al., 2013).

Concurrent and Criterion Validity

Concurrent validity concerns relationships with other measures of the same construct. As shown in Tables 1 and 2, the correlations between different measures of wisdom are markedly lower than the correlations between different conceptions of wisdom rated for the same transcripts. Thus, again, measurement variance plays an important role.

Criterion validity is the correlation with a criterion variable representative of the construct. The main "criterion" that has been used in wisdom research is wisdom nomination. On average, wisdom nominees score higher than other participants in wisdom measures, but not quite as high as one might expect for truly wise individuals (Ardelt, 2003; Baltes, Staudinger, Maercker, & Smith, 1995; Glück et al., 2013). One lesson that we learned is that not everyone whom someone considers as wise really has wisdom—there is a wide range of reasons for nominations, which seem to have validity issues of their own. For example, people may view a relative stranger as wise who once told them something that changed their life for the better—but that may

have happened for many other reasons than the nominee's wisdom. Including wisdom nominees may be very useful to raise the average level of wisdom in a sample, but it does not guarantee individual wisdom.

It may also make sense to think about other criteria than nomination (I am grateful for this suggestion from one anonymous reviewer of this paper!). Philosophers have suggested to define wisdom by "knowing how to live a good life" (Grimm, 2015; Ryan, 2014). While they have wisely abstained from defining what a good life is, this could also be considered an empirical question that may open up an alternative way to look at wisdom. Another primary building block of lay and expert conceptions of wisdom is advice-giving; a wise person should be able to give good advice to a wide range of people facing a wide range of problems, a point that I will revisit in the next section.

Ecological Validity

Ecological validity concerns whether a research setting approximates the real-world situation that it refers to. It may be rather crucial in measuring wisdom because wisdom manifests itself most clearly in specific, rare situations. While wiser and less wise individuals may differ in many aspects of how they live their life, individual differences in wisdom are amplified in situations that are difficult, uncertain, personally relevant, and emotionally challenging (Glück & Bluck, 2014): while, for example, many people may be able to maintain their calm and take everyone's perspective when they are only slightly affected by a conflict between other people, far fewer can do so when they are personally involved in a conflict involving important domains of their life. This assumption is supported by the results of a study where we interviewed participants about situations where they thought they had done something wise. Of the narrated events, about 90% were coded as "fundamental", referring to, for example, complex life decisions, conflicts, or negative life events (Glück et al., 2005). Neither thinking about a difficult life problem as in typical performance measures nor describing one's own typical behavior in a self-report scale may optimally predict a person's behavior in a wisdom-requiring real-life situation. Emulating such emotionally challenging situations in laboratory situations, however, is difficult and ethically problematic. How can we devise measures that approximate actual real-life behavior more closely? Several interesting new measures have explicitly or implicitly attempted to increase ecological validity.

Promising New Approaches to Measuring Wisdom

Several research groups currently pursue new approaches that may move the assessment of wisdom closer to real life. Some focus on participants' autobiographical reflection, others use videos of real-life problems.

The MORE Wisdom Interview

In the MORE Life Experience Model, Susan Bluck and I proposed that wisdom-related knowledge develops through an interaction of life experiences with psychological resources (Glück & Bluck, 2014). Therefore, wisdom should manifest itself in how people reflect upon past experiences. To test this hypothesis, we interviewed participants about difficult events from their past, including a free narration and several questions about emotions, strategies, and lessons learned. The interview transcripts were rated for the wisdom resources by trained raters. Encouragingly, reliabilities were satisfactory and correlations to other measures of wisdom were in the usual range. However, two somewhat cautionary findings were, first, that the range of events that people narrated was so broad that it would seem better to ask for more specific events (e.g., relationship conflicts). Second, the correlations between wisdom scores for two different narratives were only around .30, suggesting that people who think very wisely about one life challenge may be rather unwise about another one. This finding is consistent with recent research from Igor Grossmann's group.

Grossmann's Situational Measures of Wise Reasoning

Grossmann, Gerlach, and Denissen (2016) asked participants to fill out online diaries over nine days. They completed self-report scales measuring wise reasoning with respect to the most difficult challenge of each day. The average correlation across days was only .20, again suggesting large situational variation of wisdom. This variation was not random: in situations where participants reasoned more wisely, they also reported higher emotional complexity, better emotion regulation, and more forgiveness. For an aggregated score across the 9 days, relationships to other variables were far weaker.

In addition to highlighting the situational variability of wisdom, these findings have implications for measurement. In an impressive multistudy paper, Brienza et al. (2017) introduced the Situated Wise Reasoning Scale (SWIS), a "hybrid" between self-report and autobiographical approaches to measuring wisdom. Participants are asked to recall a recent interpersonal conflict and answer a number of questions about the situation and their subjective experience, which serves to increase accuracy of their recall. Then, they fill out self-report items measuring to what extent they used aspects of wise reasoning (intellectual humility, recognition of a world in flux and change, appreciation of different perspectives, application of an outsider's vantage point, consideration of and search for compromise and conflict resolution) in dealing with the conflict. Brienza et al. (2017) showed that in contrast to global self-report measures of wisdom, SWIS scores are unrelated to various biases including social desirability. Thus, the SWIS is a promising new method to measure wise reasoning on a state level. Importantly, as Grossmann's group has repeatedly

demonstrated the variability of wisdom across situations, a valid overall assessment of wisdom would require collecting SWIS responses across several recalled situations. [Brienza et al. \(2017\)](#) suggested that two to five episodes would need to be sampled for sufficient reliability.

Wisdom as Guidance Provided to Others

A rather prototypical situation in laypeople's accounts of wisdom concerns guidance and advice given to others in difficult situations (e.g., [Montgomery, Barber, & McKee, 2002](#)). Some new approaches have tried to increase the extent of immersion of participants in paradigms that involve giving advice to others. Thomas and Kunzmann (2013) developed a measure of wisdom based on videos of real young couples discussing serious conflicts in their marriage. Out of 34 videos, three were selected that were judged as highly authentic, serious, and emotionally challenging. They were presented to participants together with standard BWP tasks. As in the BWP, participants were asked to think aloud about what the protagonists in the video could consider and do, and the responses were rated for the BWP criteria. Inter-rater reliabilities were acceptable and correlations with the BWP tasks were significant. Interestingly, the videos elicited higher levels of wisdom-related knowledge than vignette-based problems. Younger participants showed higher levels of wisdom than older adults concerning the marriage conflicts, which again suggest that wisdom may be somewhat context-specific.

The use of videos may provide a highly promising route to studying wisdom. Showing real people talking about real problems increases ecological validity, as it presumably engages participants emotionally much more than reading a short vignette. When the videos show real protagonists rather than actors, it might even be possible to collect the protagonists' evaluations of participants' responses. On the other hand, using other existing or specially developed film material might allow for specific manipulations of task content (e.g., [Kunzmann & Grünh, 2005](#); [Richter & Kunzmann, 2011](#)). It might also be possible to collect written responses, such as letters written to the video protagonists, which might reduce the effort of data collection.

[Hu, Ferrari, Wong, and Woodruff \(2017\)](#) introduced a novel advice-giving paradigm for measuring wisdom. They asked participants to imagine that a person they knew was faced with a specific problem and to talk into a camera as if they were giving advice to that person. I am not quite certain about the ecological validity of talking to a camera, but the general idea of inducing a second-person, rather than a third-person, perspective seems very promising. In experimental studies, Ethan Kross and Igor Grossmann have shown that taking a third-person rather than a first-person perspective, even if it is only in the way one thinks about a problem, can improve wise reasoning significantly ([Grossmann & Kross, 2014](#); [Kross & Grossmann, 2012](#)). A second-person perspective is somewhat in between, as

the participant is thinking about someone else's problem but imagining to be in personal contact with that person. It is an interesting question for future research how this perspective affects wisdom.

There is an important perspectival difference between advice-giving and autobiographical measures that reflects Staudinger's distinction between personal and general wisdom ([Staudinger et al., 2005](#)). Measures based on recall of autobiographical experiences, such as the MORE wisdom interview and the SWIS, look at participants' reflections of their own experiences, whereas measures involving advice-giving put the participant in an observer's perspective. New, ecologically valid approaches for measuring both of these forms of wisdom will allow us to investigate the empirical relationship between personal and general wisdom. Is it, for example, possible to be a very good advice-giver but, at the same time, very unwise with respect to one's own life?

Other Possibilities for New Measurement Approaches

There are several other routes that might be worth exploring. One concerns the use of informant perspectives. If wisdom is indeed in the eye of the beholder, we might learn a lot about a person's wisdom if we ask his or her friends, colleagues, and family. There are probably not very many people who are consistently viewed as wise by informants from different life domains, and studying them would seem very interesting. Beyond wisdom nominations, informant methods have hardly been used in wisdom research. Some exceptions include a study that compared self- and peer wisdom ratings of university faculty ([Redzanowski & Glück, 2013](#)) and studies that investigated the perspectives of wisdom nominators and nominees ([Baltes et al., 1995](#); [Krafcik, 2015](#)).

Another domain that has not yet been investigated is actual wise behavior. All our measures focus on verbal responses, but it is likely that just like morality, wisdom manifests itself in nonverbal intuitions as much as in verbal reflection (see, e.g., [Haidt, 2001](#)). What do wise people actually *do* when they are listening to a person in trouble or dealing with a serious conflict? While many aspects of wisdom, such as complex inner feelings and ultimate goals, may never be observable directly, it may be very interesting to look at how wisdom actually manifests itself in real-life situations.

New and interesting avenues for wisdom research may come from technological advances. For example, [Hu et al. \(2017\)](#) combined their new second-person paradigm with another innovation: participants' facial expressions, as recorded by the camera, were automatically analyzed for expressed emotions. While I have some doubts about the validity of the particular methodology used in that study, it may soon be possible to perform fine-grained analyses of a participant's facial and bodily expressions, speech, and behavior in general. In future decades, immersive virtual-reality approaches may allow us to engage participants in

life-like scenarios and record their behavior in real time.

Conclusions

The measurement of wisdom is not exactly easy, but a fascinating area of research. Currently, we have a relatively small toolbox of established methods that are very useful as long as we remain aware of their limitations. Several recent studies have triangulated different approaches by testing whether findings hold across both performance and self-report measures (Weststrate & Glück, 2017; Webster, Weststrate, Ferrari, Munroe, & Pierce, in press). In the future, new developments may lead to even more ecologically valid methods. Perhaps this overview will incite the creativity of some methodologists.

Supplementary Material

Supplementary data is available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

This study was supported by Austrian Science Fund (P21011, P25425; PI: Judith Glück).

Conflict of Interest

None reported.

References

- Aldwin, C. M. (2009). Gender and wisdom: A brief overview. *Research in Human Development, 6*, 1–8. doi:10.1080/15427600902779347
- Ardelt, M. (2003). Development and empirical assessment of a three-dimensional wisdom scale. *Research on Aging, 25*, 275–324. doi:10.1177/0164027503025003004
- Ardelt, M. (2004). Wisdom as expert knowledge system: A critical review of a contemporary operationalization of an ancient concept. *Human Development, 47*, 257–285. doi:10.1159/000079154
- Ardelt, M. (in revision). *The relation between age and three-dimensional wisdom: Variations by wisdom dimensions and education*. Under revision for this volume. Unpublished manuscript.
- Baltes, P. B., & Staudinger, U. M. (2000). Wisdom. A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist, 55*, 122–136. doi:10.1037/0003-066x.55.1.122
- Baltes, P. B., & Kunzmann, U. (2004). The two faces of wisdom: Wisdom as a general theory of knowledge and judgment about excellence in mind and virtue vs. wisdom as everyday realization in people and products. *Human Development, 47*, 290–299. doi:10.1159/000079156
- Baltes, P. B., & Smith, J. (1990). Toward a psychology of wisdom and its ontogenesis. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 87–120). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9781139173704.006
- Baltes, P. B., Staudinger, U. M., Maercker, A., & Smith, J. (1995). People nominated as wise: A comparative study of wisdom-related knowledge. *Psychology and Aging, 10*, 155–166. doi:10.1037//0882-7974.10.2.155
- Bangen, K. J., Meeks, T. W., & Jeste, D. V. (2013). Defining and assessing wisdom: A review of the literature. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 21*, 1254–1266. doi:10.1016/j.jagp.2012.11.020
- Brienza, J., Kung, F., Santos, H. C., Bobocel, D. R., & Grossmann, I. (2017). Wisdom, bias, and balance: Toward a process-sensitive measurement of wisdom-related cognition. *Journal of Personality and Social Psychology*. doi:10.1037/pspp0000171
- Brown, S. C., & Greene, J. A. (2006). The Wisdom Development Scale: Translating the conceptual to the concrete. *Journal of College Student Development, 47*, 1–19. doi:10.1353/csd.2006.0002
- Curnow, T. (1999). *Wisdom, intuition, and ethics*. Aldershot, UK: Ashgate.
- DeMichelis, C., Ferrari, M., Rozin, T., & Stern, B. (2015). Teaching for wisdom in an intergenerational high-school-English class. *Educational Gerontology, 41*, 551–566. doi:10.1080/03601277.2014.994355
- Edmondson, R. (2005). Wisdom in later life: Ethnographic approaches. *Ageing & Society, 25*, 339–356. doi:10.1017/S0144686X04003320
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406. doi:10.1037/0033-295X.100.3.363
- Glück, J., & Baltes, P. B. (2006). Using the concept of wisdom to enhance the expression of wisdom knowledge: Not the philosopher's dream but differential effects of developmental preparedness. *Psychology and Aging, 21*, 679–690. doi:10.1037/0882-7974.21.4.679
- Glück, J., & Bluck, S. (2014). The MORE Life Experience Model: A theory of the development of personal wisdom. In M. Ferrari & N. Weststrate (Eds.), *The scientific study of personal wisdom* (pp. 75–98). New York: Springer. doi:10.1007/978-90-481-9231-1_4
- Glück, J. (2013). Wisdom. In D. S. Dunn (Ed.), *Oxford bibliographies in psychology*. New York: Oxford University Press. <http://www.oxfordbibliographies.com/view/document/obo-9780199828340/obo-9780199828340-0061.xml?rskey=iUBqE7&result=81&q=>
- Glück, J. (2016). Wisdom. In N. Pachana (Ed.), *Encyclopedia of geropsychology*. New York: Springer. doi:10.1007/978-981-287-080-3_155-1
- Glück, J., Bluck, S., Baron, J., & McAdams, D. (2005). The wisdom of experience: Autobiographical narratives across adulthood. *International Journal of Behavioral Development, 29*, 197–208. doi:10.1177/01650250444000504
- Glück, J., König, S., Naschenweng, K., Redzanowski, U., Dorner, L., Straßer, I., & Wiedermann, W. (2013). How to measure wisdom:

- Content, reliability, and validity of five measures. *Frontiers in Psychology*, 4, 405. doi:10.3389/fpsyg.2013.00405
- Greene, J. A., & Brown, S. C. (2009). The wisdom development scale: Further validity investigations. *International Journal of Aging & Human Development*, 68, 289–320. doi:10.2190/AG.68.4.b
- Grimm, S. R. (2015). Wisdom. *Australasian Journal of Philosophy*, 93, 139–154. doi:10.1080/00048402.2014.937449
- Grossmann, I. (2017). Wisdom in context. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 12, 233–257. doi:10.1177/1745691616672066
- Grossmann, I., & Kross, E. (2014). Exploring Solomon's paradox: Self-distancing eliminates the self-other asymmetry in wise reasoning about close relationships in younger and older adults. *Psychological Science*, 25, 1571–1580. doi:10.1177/0956797614535400
- Grossmann, I., Gerlach, T. M., & Denissen, J. J. (2016). Wise reasoning in the face of everyday life challenges. *Social Psychological and Personality Science*, 7, 611–622. doi:10.1177/1948550616652206
- Grossmann, I., Na, J., Varnum, M. E. W., Park, D. C., Kitayama, S., & Nisbett, R. E. (2010). Reasoning about social conflicts improves into old age. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 107, 7246–7250. doi:10.1073/pnas.1001715107
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, 108, 814–834. doi:10.1037//0033-295x.108.4.814
- Hu, C. S., Ferrari, M., Wang, Q., & Woodruff, E. (2017). Thin-slice measurement of wisdom. *Frontiers in Psychology*, 8, 1378. doi:10.3389/fpsyg.2017.01378
- Igarashi, H., Levenson, M. R., & Aldwin, C. (in revision). The development of wisdom: A social ecological approach. Unpublished manuscript.
- Jason, L.A., Reichler, A., King, C., Madsen, D., Camacho, J., Marchese, W. (2001). The measurement of wisdom: A preliminary effort. *Journal of Community Psychology*, 29, 585–598. doi:10.1002/jcop.1037
- Koller, I., Levenson, M. R., & Glück, J. (2017). What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Frontiers in Psychology*, 8, 126. doi:10.3389/fpsyg.2017.00126
- König, S., & Glück, J. (2014). "Gratitude is with me all the time": How gratitude relates to wisdom. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69, 655–666. doi:10.1093/geronb/gbt123
- Krafcik, D. (2015). Words from the wise: Exploring the lives, qualities, and opinions of wisdom exemplars. *Integral Review: A Transdisciplinary & Transcultural Journal for New Thought, Research, & Praxis*, 11, 7–35.
- Kross, E., & Grossmann, I. (2012). Boosting wisdom: Distance from the self enhances wise reasoning, attitudes, and behavior. *Journal of Experimental Psychology: General*, 141, 43–48. doi:10.1037/a0024158
- Kunzmann, U., & Grünh, D. (2005). Age differences in emotional reactivity: the sample case of sadness. *Psychology and Aging*, 20, 47–59. doi:10.1037/0882-7974.20.1.47
- Levenson, R., Jennings, P.A., Aldwin, C., & Shiraishi, R.W. (2005). Self-transcendence, conceptualization and measurement. *International Journal of Aging & Human Development*, 60, 127–143. doi:10.2190/XRXXM-FYRA-7U0X-GRC0
- McClelland, D. C. (1973). Testing for competence rather than for "intelligence". *American Psychologist*, 28, 1–14. doi:10.1037/h0034092
- Mickler, C., & Staudinger, U. M. (2008). Personal wisdom: Validation and age-related differences of a performance measure. *Psychology and Aging*, 23, 787–799. doi:10.1037/a0013928
- Mitchell, L. (2016). *An exploration of wisdom in mid- to older-aged adults: An Australian context*. Australia: Unpublished dissertation, University of Queensland.
- Montgomery, A., Barber, C., & McKee, P. (2002). A phenomenological study of wisdom in later life. *International Journal of Aging & Human Development*, 54, 139–157. doi:10.2190/28E3-ANPT-UEED-92WF
- Pasupathi, M., Staudinger, U. M., & Baltes, P. B. (2001). Seeds of wisdom: Adolescents' knowledge and judgment about difficult life problems. *Developmental Psychology*, 37, 351–361. doi:10.1037//0012-1649.37.3.351
- Redzanowski, U., & Glück, J. (2013). Who knows who is wise? Self and peer ratings of wisdom. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 68, 391–394. doi:10.1093/geronb/gbs079
- Richter, D., & Kunzmann, U. (2011). Age differences in three facets of empathy: performance-based evidence. *Psychology and Aging*, 26, 60–70. doi:10.1037/a0021138
- Ryan, S. (2014). Wisdom. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Open-access publication. <https://plato.stanford.edu/archives/win2014/entries/wisdom/>.
- Smith, J., & Baltes, P. B. (1990). Wisdom-related knowledge: Age/cohort differences in response to life-planning problems. *Developmental Psychology*, 26, 494–505. doi:10.1037/0012-1649.26.3.494
- Staudinger, U. M. (1999). Older and wiser? Integrating results on the relationship between age and wisdom-related performance. *International Journal of Behavioral Development*, 23, 641–664. doi:10.1080/016502599383739
- Staudinger, U. M., & Baltes, P. B. (1996). Interactive minds: A facilitative setting for wisdom-related performance? *Journal of Personality and Social Psychology*, 71, 746–762. doi:10.1037/0022-3514.71.4.746
- Staudinger, U. M., Dörner, J., & Mickler, C. (2005). Wisdom and personality. In R. J. Sternberg & J. Jordan (Eds.), *A handbook of wisdom: Psychological perspectives* (pp. 191–219). New York: Cambridge University Press. doi:10.1017/CBO9780511610486.009
- Staudinger, U. M., Maciel, A. G., Smith, J., & Baltes, P. B. (1998). What predicts wisdom-related knowledge? A first look at personality, intelligence, and facilitative experiential contexts. *European Journal of Personality*, 12, 1–17. doi:10.1002/(SICI)1099-0984(199801/02)12:1<1::AID-PER285>3.0.CO;2-9
- Staudinger, U. M., & Glück, J. (2011). Psychological wisdom research: Commonalities and differences in a growing field. *Annual Review of Psychology*, 62, 215–241. doi:10.1146/annurev.psych.121208.131659

- Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist*, *36*, 227–245. doi:10.1207/S15326985EP3604_2
- Thomas, S., & Kunzmann, U. (2013). Age differences in wisdom-related knowledge: Does the age relevance of the task matter? *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *69*, 897–905. doi:10.1093/geronb/gbt076
- Tornstam, L. (1994). Gero-transcendence: A theoretical and empirical exploration. In L. E. Thomas & S. A. Eisenhandler (Eds.), *Aging and the religious dimension* (pp. 203–225). London: Auburn House.
- Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development*, *10*, 13–22. doi:10.1023/A:1020782619051
- Webster, J. D. (2007). Measuring the character strength of wisdom. *International Journal of Aging & Human Development*, *65*, 163–183. doi:10.2190/AG.65.2.d
- Webster, J. D., Westerhof, G. J., & Bohlmeijer, E. T. (2012). Wisdom and mental health across the lifespan. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *69*, 209–218. doi:10.1093/geronb/gbs121
- Webster, J. W., Weststrate, N. M., Ferrari, M., Munroe, M., & Pierce, T. W. (in press). Wisdom and meaning in emerging adulthood. *Emerging Adulthood*. doi:10.1177/2167696817707662
- Weststrate, N. M., & Glück, J. (2017). Hard-earned wisdom: Exploratory processing of difficult life experience is positively associated with wisdom. *Developmental Psychology*, *53*, 800–814. doi:10.1037/dev0000286