The Realistic Accuracy Model

Tera D. Letzring

Idaho State University

David C. Funder

University of California, Riverside

Pre-copy edited draft


Acknowledgments: Work on the State and Trait Accuracy Model presented in this chapter was supported by the National Science Foundation Award No. 1551822 to Drs. Tera D. Letzring, Judith A. Hall, and Jeremy C. Biesanz.
Abstract

This chapter describes the Realistic Accuracy Model (RAM), starting with a history of its development. We then describe the four moderators of accuracy in personality judgment – good judge, good target, good trait, and good information – and how these moderators interact with each other. Next, we describe the four stages in the process of making accurate judgments, which are relevance, availability, detection, and utilization. Implications of the model for improving judgment accuracy and applications to judgments of states are then discussed. The chapter concludes with suggested directions for future research including judgments of other levels of personality besides traits, interactions between moderators, the development of judgmental ability, and the consequences of judgmental accuracy.

Keywords: judgment accuracy, judgmental ability, self-other agreement, realistic accuracy, relevance, availability, detection, utilization, moderators of accuracy
The Realistic Accuracy Model

Do personality traits exist and, more specifically, do people behave consistently enough across situations and over time to make trait descriptions meaningful and useful? Is it possible to make accurate judgments of the personality traits of others? What factors moderate how accurately people can make judgments of others’ traits? What is the process that needs to happen for accurate trait judgments to be possible?

The search for answers to these questions led to the formulation of the Realistic Accuracy Model, or RAM. David Funder (the second author of this chapter) first formally presented this model in writing in 1995 in a paper published in Psychological Review, although it was foreshadowed for several years in other research and thinking. The goal of this chapter is to describe this model, including its history and its descriptions of four moderators of accuracy and the four-step process of making accurate judgments. Then, we will discuss the implications of the model for improving judgment accuracy and applications to judgments of states, and point out some promising and important future research questions that remain unanswered.

History of the Realistic Accuracy Model

A first step in the development of RAM was taken when Funder (1980) examined relations between ratings people made of themselves and ratings other people made of them, which is referred to as *self-other agreement*. This research was grounded in attribution theory, self-concept theory, and personality theory, and addressed the question of “whether we appear to others as we appear to ourselves” (Funder, 1980, p. 473). More specifically, the research evaluated how people rated themselves and others on personality characteristics related to inner states and external behaviors. The theoretical perspective on which this work was based led to three predictions. First, how people see themselves would be similar to how others see them.
Second, people would be more likely to use the situation to explain their own behavior because the situation is more salient to them from their internal perspective, whereas people would be more likely to use personality traits to explain the behavior of others because the personality of others is more salient from the external observer’s perspective. And third, which trait was being judged would influence explanations of behavior such that traits related to inner states would be rated by the self as more characteristic, while traits related to external behaviors would be rated by others as more characteristic. These predictions were tested using self-reports of personality and other-reports from roommates or friends, and support was found for all three predictions. Importantly for personality theory, the article concluded that personality traits “do exist, at the very least as perceptions people have of themselves and share with the people around them” (p. 487, emphasis in original), and that perceptions of traits are based on patterns of behavior that are observed over time. This paper set the stage for continued research on agreement between self-ratings and other ratings of personality.

The need for a model of accurate personality judgment quickly became apparent (Funder, 1987). Accuracy research differed in important ways from other types of person perception research that were popular in the mid- to late-twentieth century, and a model would help guide theorizing and research on accuracy and distinguish it from similar areas. At the time, the most influential research was based on a cognitive approach and focused on describing the process of judgment in terms of how people combine information to make judgments about hypothetical targets. This came to be known as the social-cognitive approach. Much has been learned from research in this area, but this approach did not consider the accuracy of these judgments, as there was not a real person about whom judges could be accurate.
Not all researchers were satisfied with the social-cognitive approach. Its focus on the process of how judgments of others are made (Asch, 1946; Jones, 1985; Kelley, 1967) bypassed issues concerning the content and accuracy of such judgments in the real world (Funder, 1990). Furthermore, this approach overwhelmingly emphasized errors in judgment (Funder, 1987). Such errors were defined as any deviations from the judgment that would be expected if judges were to follow a normative model of “perfect” judgment, such as the laws of probability or logic (Nisbett & Ross, 1980; Tversky & Kahneman, 1983). Perhaps not too surprisingly, people do not always precisely follow such models and laws, and therefore often make judgments that differ from the prescribed outcomes (Nisbett & Ross, 1980, 1991; Shaklee, 1991). Within the social-cognitive approach, any such differences, regardless of their magnitude, are labeled as errors. However, such errors tell us little about the accuracy of judgments that are made in the real world about real people because formal normative models do not always correctly prescribe what the “correct” judgment should be in the real world1 (Ebbesen & Konecni, 1980; Funder, 1987, 1993, 1999; Ross, Nisbett, & Funder, 2007; Trope, 1986, 1989; Trope & Lieberman, 1993). However, this conventional view of error was so pervasive that many person perception and social judgment researchers came to doubt that it was possible for people to make judgments of others that attained even moderate levels of accuracy (Evans, 1984; Ross et al., 2007).

The fundamental limitation of the social-cognitive approach is that accurate judgment requires more than simply not making errors. This limitation arises because research focusing on error begins, ironically, by assuming that judgment should be perfect, and then proceeds to catalog imperfections. However, it fails to answer the question of how judgments of any sort

---

1 This is because the laws of probability or logic are assessed by “coherence” criteria, whereas the accuracy of a judgment in relation to a real-life criterion is assessed by “correspondence” criteria (Hammond, 1996).
arise in the first place, because it fails to ask that question. To address the source of judgments – accurate and inaccurate both – requires an ecological approach that describes how judgments arise out of information that is available to the judge under realistic circumstances, and assesses the accuracy of these judgments in terms of correspondence with real-world criteria. Such correspondence is called realistic accuracy.

The study of realistic accuracy fits within an ecological approach that gives primary importance to the natural environment\(^2\) in which people behave and interact with others. According to the ecological approach, accurate perceptions are defined in terms of their adaptive functions, such as helping to navigate one’s social world and coordinate social interactions, and only make sense when we consider them in the context in which they were made and the purposes for which they are used (Neisser, 1988; Zebrowitz, 1990). In terms of the accuracy of trait judgments, an ecological approach assumes that traits actually exist, cues come from those traits, and people can, at least sometimes, detect those cue and use them to make accurate judgments (Funder & Sneed, 1993). Research within the tradition of the ecological approach has found substantial levels of agreement between judgments of traits and the traits themselves, and that judgments are more accurate when judges use valid cues more than invalid cues (Borkenau & Liebler, 1992, 1995; Funder & Sneed, 1993; see Ch. 13 by Breil et al. in this handbook). These findings led to the conclusion that people do have some wisdom when it comes to trait judgments (Funder & Sneed, 1993). Research within this approach has also led to ideas about how to train people to be more accurate, which could be done by helping them focus on and use valid cues while ignoring invalid cues (Borkenau & Liebler, 1992; Funder & Sneed, 1993; Ch. 4 by Hirschmüller, Breil, Nestler, & Back in this handbook).

\(^2\) In contrast to judgments about hypothetical stimuli or people.
The Realistic Accuracy Model was formulated as a corrective complement to error research, by focusing on how judgments could be accurate rather than inaccurate (Funder, 1987). The development of RAM was also motivated by Gordon Allport’s (1937) list of the most important problems that were related to judgments of personality that included “(1) the nature and reliability of first impressions, (2) the chief factors involved in judging, (3) the value of interviews, (4) the question whether ability to judge people is general or specific, [and] (5) the qualifications of a good judge” (p. 499).

The desire in developing RAM was to account for factors, or moderators, that are related to accuracy, and to describe the process that people use to make accurate inferences about personality (Funder, 1999; Funder & Sneed, 1993). The early stages of empirical research identified four moderators of accuracy, or factors that influence the level of accuracy, and later theoretical development proposed a 4-stage process of accurate judgment that sought to account for these moderators (Funder, 1995, 2012). RAM was also built upon Brunswik’s (1955) Lens Model of perceptual constancy, which describes how cues are related to an object and used to make judgments of that object (see Ch. 4 by Hirschmüller et al. in this handbook). In the lens model, there are two processes that are identified as cue validity and cue utilization. Cues are valid when they are related to the object being judged, and they are utilized when they are used to make judgments about the object. Accuracy will be higher when all valid cues are utilized, and lower when invalid cues are utilized or when valid cues are not utilized. In the case of RAM, the object being perceived was a personality trait of a real person. RAM further broke the Lens Model processes down into a 4-stage process for making accurate judgments of personality traits (see Figure 1).

**Moderators of Accuracy**
The first and perhaps most obvious advantage of focusing on moderators of accuracy, rather than simply on errors, is that the operation of moderator variables can explain both correct and incorrect judgments. The four moderators identified by Funder include categories of factors that are related to differing levels of accuracy. Some of the moderators were identified in earlier work (Allport, 1937; Funder, 1987, p. 87; 1990, p. 208; Funder & Dobroth, 1987) before all four were summarized in a review chapter (Funder, 1993). The four moderators are the good judge, good target, good trait, and good information. Each moderator is identified as “good” to emphasize a positive focus on accuracy (and thus suggest possible ways to improve accuracy; Funder, 1993, 1999), but as just noted, each moderator also implicitly identifies a low level of each variable as a possible source of lower accuracy.

The good judge moderator reflects the idea that people differ in how accurately they tend to judge others and this is due to something about the judges themselves. In general, good judges have high levels of intelligence and cognitive ability, knowledge about how behaviors and personality are related, agreeableness and social skills, perspective taking and empathy, and good psychological adjustment (Christiansen, Wolcott-Burnam, Janovics, Burns, & Quirk, 2005; Colman, Letzring, & Biesanz, 2017; Human & Biesanz, 2011a; Human, Biesanz, Finseth, Pierce, & Le, 2014; Kolar, 1995; Letzring, 2008; see Ch. 6 by Colman in this handbook). The good target moderator reflects that people differ in how accurately they are judged by others, and therefore there is something about the targets themselves that tends to make it easier for them to be judged with relatively high levels of accuracy. In general, good targets have high levels of psychological adjustment, behavioral consistency, social skills, extraversion, and expressiveness (Ambady, Hallahan, & Rosenthal, 1995; Colvin, 1993; Human, Mignault, Rogers, & Biesanz, 2018; see Ch. 7 by Mignault & Human in this handbook). The good trait reflects that personality
characteristics differ in how accurately they tend to be judged, and therefore there is something about the traits that allows them to be judged with relatively high levels of accuracy across judges, targets, and situations. In general, good traits are easily observable in most situations (Funder & Dobroth, 1987; John & Robins, 1993; Paunonen & Kam, 2014; see Ch. 8 by Krzyzaniak & Letzring in this handbook). Finally, good information reflects that information differs in how useful it is for making accurate judgments, and therefore there is something about the information that makes it easier for judges to make more accurate judgments. This moderator has two aspects. Information quantity reflects that having more information tends to result in more accurate judgments than having less information, which is referred to as the acquaintanceship effect (Biesanz, West, & Millevoy, 2007; Blackman & Funder, 1998; Colvin & Funder, 1991; Funder & Colvin, 1988; Letzring, Wells, & Funder, 2006). The other aspect, information quality, reflects that some types of information are more useful for making more accurate judgments than other types of information. Highly useful information includes information about general and specific behaviors, and thoughts and feelings (Andersen, 1984; Letzring & Human, 2014; see Ch. 9 by Beer in this handbook). In this handbook, an entire chapter is devoted to describing and exploring each of these moderators, so more will not be said here about individual moderators. As a whole, research has supported the usefulness of these moderators for understanding how accurate judgments are made and how various factors are related to levels of accuracy.

**Interactions between Moderators**

In addition to these moderators, theoretical work on RAM (Funder, 1995) also proposed that each moderator could interact with every other moderator and used the following terms to describe these interactions: relationship (interaction between judge and target), expertise (judge
and trait), sensitivity (judge and information), palpability (target and trait), divulgence (target and information), and diagnosticity (trait and information). These terms have not come into common use to the same degree as the four basic moderators, and have not been tested as systematically, although some of the interactions have been examined empirically.

One moderator-interaction that has received a fair amount of research attention is the relationship moderator, or how something about the interaction between the judge and target could affect accuracy. The relationship moderator operates when something about the unique alignment between a target and a judge results in that judge being more accurate about the target than are any other judges and more accurate than they are when they are judging anyone else. For example, the quality of relationship between a husband and a wife could result in the husband judging his wife more accurately than anyone else judges her. His augmented ability to make an accurate judgment would not extend to other targets—just to his wife. Ch. 17 by Luo and Watson in this handbook addresses accuracy within romantic relationships and includes excellent examples of how this moderator-interaction has been examined. A meta-analysis by Connelly and Ones (2010) went a step farther and examined how accuracy for specific traits differed across types of relationships, and therefore examined a 3-way interaction between judge, target, and trait. Accuracy for judgments of extraversion differed little across relationships, but accuracy for judgments of emotional stability, openness, and agreeableness varied quite a bit across relationships. In particular, accuracy for these traits tended to be higher in more intimate relationships (i.e., family and friends) than less intimate relationships (i.e., work colleagues, incidental acquaintances, and strangers).

A few studies have examined expertise, which is the interaction between the judge and trait moderators, and have therefore considered how characteristics of judges are related to how
accurately they judge certain traits. One study examined how several positive characteristics of judges, including intrapersonal characteristics such as satisfaction with life and positive emotions, and interpersonal characteristics such as social network size and social intimacy, were related to accuracy of ratings of five broad traits that are commonly assessed in trait accuracy research\(^3\) (Letzring, 2015). Two types of accuracy were examined: normative accuracy, which is the level of similarity between judgments of targets and what the average person is like, and distinctive accuracy, which is the level of similarity between judgments of targets and how each target differs from the average person. Consistent with the expertise moderator-interaction, several positive characteristics of judges were positively related to normative accuracy for certain traits. For example, judges with larger social networks and higher levels of agreeableness and openness were more likely to judge targets on openness to experience in a way that was similar to the average person; and judges with higher levels of interpersonal control and support, satisfaction with life, and positive emotions were more likely to judge targets on conscientiousness in a way that was similar to the average person. However, these positive judge-characteristics were unrelated to distinctive accuracy. Another study of expertise found that judges with high social curiosity about “how other people behave, think, and feel” (Hartung & Renner, 2011, p. 796) made more accurate judgments of extraversion and openness following a 10-minute interaction with a stranger, than did judges with low social curiosity. An explanation for this finding was that judges with high social curiosity tended to detect and correctly use more cues to extraversion and openness in comparison to judges with low social curiosity. Finally, a question that is sometimes raised is whether being high on a trait makes a person an expert at

\(^3\) These are the Big Five traits of extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience from the Five Factor Model of personality (Goldberg, 1999; John et al., 2008; McCrae & Costa, 1999).
judging that trait in others. However, research to date has not found that judges who are higher on a certain trait typically make more accurate judgments of that same trait in others (Hartung & Renner, 2011; Letzring, 2015).

Another moderator-interaction that has been empirically tested is diagnosticity, which is the interaction between the information and traits moderators, and can be examined when research on information quantity and quality considers accuracy for individual traits. There is evidence that some kinds of information are more useful for judging certain traits. For example, neuroticism is frequently found to be judged with low levels of accuracy, but judges were able to make relatively accurate judgments of neuroticism after observing targets in a socially stressful situation (i.e., being video-recorded while introducing one’s self to strangers; Hirschmüller, Egloff, Schmukle, Nestler, & Back, 2015). Judgments of neuroticism were also more accurate when judges had more information about specific behaviors of the targets, current events in targets’ lives, and when targets talked about their own personalities (Letzring & Human, 2014). Judgments of conscientiousness were more accurate when judges were given information about facts that differentiated targets from most other people than when they were given information about values that were important to targets (Beer & Brooks, 2011). Letzring and Human (2014) examined how quantities of several types of information were related to accuracy for specific traits, which yielded several interesting findings. For example, the amounts of information about targets’ relationships, past events, and targets’ own personalities were positively related to accuracy of judgments of agreeableness; while the amounts of information about targets’ thoughts and feelings, specific and general behaviors, and current and past events were positively related to accuracy of judgments of openness (Letzring & Human, 2014). Analyses have also
examined which cues are especially diagnostic for certain traits\(^4\) and have yielded many interesting findings. For example, in photographs, a healthy appearance was diagnostic for extraversion, emotional stability, likability, and self-esteem; and smiling was diagnostic for extraversion, agreeableness, likability, self-esteem, and religiosity (Naumann, Vazire, Rentfrow, & Gosling, 2009). Additionally, a fashionable and refined appearance, energetic and expressive facial expressions, posture, and body movements, and an energetic and cheerful voice were diagnostic for extraversion (Hirschmüller, Egloff, Nestler, & Back, 2013). For more information about diagnosticity, see Ch. 8 by Krzyzaniak and Letzring, Ch. 9 by Beer, and Ch. 14 by Wall and Campbell in this handbook.

There has been at least one study on palpability, which is the interaction between target and trait moderators such that some traits might be judged more accurately for some targets. Traits that were more central to a target’s self-concept were judged with higher levels of self-other agreement by close friends, but this finding only held when targets also thought the trait had little influence on how much others liked them (Koestner, Bernieri, & Zuckerman, 1994).

Overall, research has shown that in addition to the importance of the good judge, target, trait, and information moderators, how these moderators interact with each other can influence levels of accuracy. It is possible that inconsistent findings in research that focuses on individual moderators may be due to a failure to address how the moderators interact with each other. For example, there have been some inconsistencies in studies of characteristics of the good judge (Allik, de Vries, & Realo, 2016; Christiansen et al., 2005; Letzring, 2008), and this may be due

\[^4\] These analyses are referred to as Lens Model analyses because they are based on Brunswik’s Lens Model and have the goal of determining which cues are valid and utilized. See Ch. 4 by Hirschmüller, Breil, Nestler, & Back in this handbook for a complete description of the Lens Model.
to asking judges to rate different traits, basing judgments on different types and amounts of
information, and using targets that differ in important ways from each other. Examining how
accuracy relates to these interactions has the potential to shed much light on the complex process
of making accurate judgments.

**The Process of Making Accurate Judgments**

The other main contribution that resulted from the formulation of RAM was the
description of the process that needs to take place in order for accurate judgments to be possible.
Notice that theorizing about RAM does not assume that the process always happens, or even that
it often (or perhaps ever) happens; these are separate empirical issues. However, the theory does
assert that the process is *necessary* for accurate personality judgment - without this four-step
process, accurate judgment would simply be impossible. These four stages are relevance,
availability, detection, and utilization (Funder, 1995; 1999; 2012; see Figure 1).

**Relevance**

The first stage, *relevance*, indicates that there must be behavioral cues relevant to the trait
or characteristic being judged. For example, let’s say you are trying to assess a person’s level of
creativity. First, that person has to do something related to her level of creativity, such as having
a thought or idea or creating something. Success at this stage is primarily influenced by the target
because targets have to do things that are relevant to and consistent with their traits. Targets tend
to be judged more accurately when they are more extraverted, self-confident, emotionally stable,
psychologically adjusted, and behaviorally expressive; and less deceitful, self-defensive, and
likely to keep others at a distance (Colvin, 1993; Human & Biesanz, 2013; see also Ch. 7 by
Mignault & Human in this handbook). This set of characteristics of good targets supports the
importance of the first stage of RAM, because targets with these characteristics are likely to behave in ways that are consistent with their personalities, which would increase relevance.

**Availability**

The second stage, *availability*, indicates that the cues have to be available in the external environment. If the cues are thoughts or ideas, then the target could talk or write about these in order to make the cues available. If the cues are the product of creating something, then the product has to be available to others. Success at the second stage is affected by an interaction between the target and judge, because relevant behaviors only became available in contexts that the two people share. In different contexts, such as family, work, or school, different behaviors are available for observation and therefore helpful for making accurate judgments of different traits. The accuracy that a judge attains will depend crucially upon the degree to which behaviors relevant to the trait being judged are available in the contexts that the judge and target share. In addition, judges can also influence how many cues targets make available, possibly by being more agreeable, communal or interpersonally oriented, sympathetic, and warm; displaying interest in the targets; and being less domineering, avoidant, narcissistic, anxious, and condescending (Christiansen et al., 2005; Letzring, 2008; Vogt & Colvin, 2003; see Ch. 6 by Colman in this handbook).

In the era of social media, “shared” contexts may also include Facebook, Twitter, Instagram, and other electronically-mediated interactions. Potentially useful information about personality is available in these contexts and includes cues such as the words people use, number of friends, content of pictures and status updates or posts, and use of emoticons (Gosling et al., 2011; Hall, Pennington, & Lueders, 2014; Park et al., 2015; Qiu, Lin, Ramsay, & Yang, 2012; Wall, Kaye, & Malone, 2016). Furthermore, there is evidence that people can make accurate
judgments of some personality traits based on viewing social media (Hall et al., 2014; Qiu et al., 2012; Wall et al., 2016; see also Ch. 14 by Wall & Campbell in this handbook). Given the increasing use of online social media and the diversity of ways to use social media, accuracy of judging personality traits based on viewing social media and interacting in an online environment is a promising topic for future research.

People sometimes confuse the relevance and availability stages, but they are distinct and both important. There can be many cues that are relevant to a trait, but they have to also be available to contribute to accuracy. Likewise, there can be many cues that are available, but the cues have to be relevant to the trait being judged to contribute to accuracy.

Detection

The third stage is detection, or noticing the relevant and available personality cues. The success of this stage is primarily influenced by the judge. Returning to the example of judging creativity, for this stage to be successful, the judge has to detect the cues to creativity. This means that the judge has to pay attention to the target. If the target is talking about her idea, then the judge has to be listening to what the target is saying; if the target wrote about her idea, then the judge has to read and understand what was written; if the target created something, then the judge has to look at and examine this product. A judge who is inattentive, distracted, uninterested, or even sensory impaired, will be less likely to detect the information necessary for accurate judgment.

This stage is typically thought of as being under the control of the judge, as it is up to the judge to pay attention to the target and detect the cues that are available. Evidence that supports the importance of the detection stage includes positive correlations between accuracy and engaging in more eye contact, seeming more interested in the targets, and reporting higher levels
of perspective-taking and empathy (Colman et al., 2017; Letzring, 2008). Additionally, there may be attributes or behaviors of targets that encourage judges to pay better attention to the targets and therefore detect more cues. For example, target judgability is positively related to social status, physical attractiveness, and attempting to make a good impression (Human & Biesanz, 2013; Human, Biesanz, Parisotto, & Dunn, 2012; Lorenzo, Biesanz, & Human, 2010). Additionally, there is evidence that judges attend more to high social status and physically attractive targets who present themselves favorably (Human et al., 2012; Langlois et al., 2000; Ratcliff, Hugenberg, Shriver, & Bernstein, 2011), but not all of this research looks at whether these attributes of targets are also related to how accurately they are judged. Therefore, it would be informative to examine whether targets who are perceived as being more interesting, possibly due to have higher status or being physically attractive, are also judged more accurately, as this may reflect better attention and detection by judges.

**Utilization**

Finally, the judge has to correctly *utilize* the cues to make an accurate judgment. Utilization includes several more specific abilities, including being able to determine which cues are relevant to the trait being judged, giving appropriate weights to various cues, combining the cues with each other, and accounting for other characteristics, such as those of the situation, that may influence behavior. Returning to our example of judging creativity, the judge has to determine which cues are relevant to creativity, properly weight these cues and combine them with each other, and consider how the situation may influence creativity in order to make an accurate judgment.

Research supports the importance of the utilization stage in that characteristics of judges that would be likely to influence the success of this stages are related to accuracy. These
characteristics include general mental ability, knowing how personality is related to behavior,\textsuperscript{5} and making complex explanations for behavior\textsuperscript{6} (Christiansen et al., 2005; Krzyzaniak, 2018).

**Necessity of Each Stage and Multiplicativeness of the Stages**

Within RAM, these stages are all described as necessary, which means that if any stage is unsuccessful, an accurate judgment is not possible (Funder, 1995). If the target does not do anything relevant to creativity, then the process cannot even be started. If cues are not made available, then there is nothing for the judge to detect. If the judge does not detect the available cues, then there is nothing to be used to make a judgment. And if the judge does not correctly use the detected cue, or uses the wrong cues, then the judgment will not be accurate.

The stages are also multiplicative, meaning that the level of accuracy of the judgment could theoretically be determined by multiplying the levels of success for each stage. However, success is not measured in a way that is objective and quantitative enough to actually assign a number to each stage, so levels of accuracy are not really determined in this way. Rather, this is a conceptual idea and demonstrates that accuracy is influenced by the level of success at each stage. Another implication of the multiplicative nature of the stages is that 100% accuracy is unlikely, because this is only possible with 100% success at each stage of the model. However, much research supports the conclusion that judges can often make fairly accurate judgments of the personality traits of others (Brown & Bernieri, 2017; Colman et al., 2017; Connelly & Ones, 2010), and therefore the level of success for each of the stages must usually be quite high.

The processes of RAM are often described in terms of making a judgment for a particular trait based on a particular cue, such as making a judgment of creativity based on hearing a target

\textsuperscript{5} Which is known as dispositional intelligence.

\textsuperscript{6} Which is known as attributional complexity.
describe one idea she has had. However, this is a simplification designed to identify the core process of accurate judgment. In real life settings, people are more likely to make judgments of many traits at the same time and to base these judgments on many cues they have detected, often from more than one interaction with the target. Furthermore, when people interact in groups, they may be detecting cues and making judgments of several targets simultaneously. It is easy to see how the process could get highly complex very quickly, as judges have to deal with multiple cues and decide which cues are relevant to which traits, and remember which cues came from each target. Furthermore, the situation in which the interactions take place can also change the meaning of the cues. For example, a target laughing at a party could mean something different and be utilized quite differently than a target laughing while playing a violent video game. When all of this is considered, it is quite impressive that people are able to make reasonably accurate judgments of others.

**Implications for Improving Accuracy of Judgments**

The moderators and stages of RAM provide insight into ways that the accuracy of judgments could be improved. Accuracy will be highest when judgments are made by good judges who have access to a large quantity of high quality information about good targets and are judging good traits. But what if someone is not naturally a good judge, or wants to make an accurate judgment of target or trait that is typically difficult to judge, or only has access to limited information? What should be done to increase the accuracy of those kinds of judgments? People who are not naturally good judges could learn more about how traits are related to behavior, be careful to pay attention to the targets of judgment so they can detect more cues, try to be agreeable and socially skilled so targets will feel comfortable revealing relevant information about themselves, and learn to think about what targets are thinking and feeling.
When making judgments of difficult targets, judges could gather additional information by spending more time with the targets in a variety of situations, or asking targets relevant questions in an attempt to compensate for targets with low expressiveness or consistency. When making judgments of traits that are typically more difficult to judge, judges could create situations or ask questions that are likely to elicit cues that are relevant to those traits. When only limited information is available, increasing the relevance of the information, again by creating situations or asking questions, could be especially useful.

Different situations elicit cues relevant to different traits, and therefore affect the accuracy of judgments made on the basis of observations of people in those situations (Funder, 2016; Rauthmann, Sherman & Funder, 2015). Interacting with a person in an unstructured situation where behavior is relatively free to vary yields more informative behavioral information and more accurate personality judgments than interacting in a structured situation where behavior is more constrained (Blackman, 2002; Letzring et al., 2006). For example, in order to get to know someone in a first-date situation, it may be preferable to do something like going on a hike rather than to a movie because behavior is much more likely to be informative about personality in the less-constrained hiking situation. Moreover, certain traits may become visible under particular circumstances – recall the research, cited earlier in this chapter, that neuroticism becomes more evident when a person is under stress (Hirschmüller et al., 2015). Thus, one route to more accurate judgment is simply the selection of the situations within which one interacts with or observes a person.

Some researchers have designed programs to train people to make more accurate judgments of others, and there is evidence that some of these programs are effective (see Ch. 21 by Blanch-Hartigan & Hill Cummings in this handbook). In particular, programs that include
practice and feedback, or instruction, practice, and feedback, tend to be more effective than programs that have only instruction, only practice, or only instruction and practice (Blanch-Hartigan, Andrzejewski, & Hill, 2012). Additionally, programs that train people individually or in small groups tend to be more effective than programs that use large groups (Blanch-Hartigan et al., 2012).

**Application to Judgments of States**

RAM was designed to describe the process of making accuracy judgments about personality traits, but it can also be applied to state-level constructs, or what is happening in the moment. States that are commonly judged in research include emotions, thoughts, truthfulness vs. deception, and physical pain (Bond & DePaulo, 2006; Hall, Andrzejewski, & Yopchick, 2009; Ickes, 2016; Ruben, van Osch, & Blanch-Hartigan, 2015). Judgments of these domains are likely to follow the same process as trait judgments, because in order to accurately judge what someone is feeling, thinking, or experiencing, judges would need to detect and utilize relevant and available cues. Research of accuracy of state judgments based on the RAM framework would test the generalizability of RAM and provide another theoretical foundation for research in state domains.

If judgments of states follow the same process as judgments of traits, then the first stage would be the existence of cues relevant to the state of interest, and the second stage would be making these cues available to judges. These cues are often nonverbal and include facial expressions, body posture, and qualities of the voice such as rate of speech; although targets could also express emotions, thoughts, or inner experiences verbally (Buck, 1976; DePaulo, Stone, & Lassiter, 1985; Elfenbein & Eisenkraft, 2010; Elfenbein et al., 2010). Then, these cues would be detected and utilized by the judge to make a judgment of the state of interest.
The State and Trait Accuracy Model (STAM) was proposed as a way to integrate the processes of making accurate judgments about states (with a focus on emotional states to begin with) and personality traits (Hall, Gunnery, Letzring, Carney, & Colvin, 2017; Letzring & Funder, 2018). Essentially, the process described in RAM was proposed to happen twice: first for the judgment of emotions, and then for the judgment of traits. Judgments of emotions would be used as cues when making the trait judgments. The two main predictions derived from STAM are that 1) accuracy of judgments of emotions and personality traits are positively correlated when the emotions that targets experience are related to their personalities, and 2) accurate judgments of emotional states cause more accurate judgments of personality traits when judges do not have pre-existing information about targets. An initial test of this model revealed that accuracy of judgments of distressed emotions, including fear and negative affect, was positively related to accuracy of judgments of neuroticism, and accuracy of judgments of positive affect was positively related to accuracy of judgments of extraversion (Hall et al., 2017). Furthermore, the distinctive accuracy of judgments of emotional states was positively related to the distinctive accuracy of judgments of personality traits, and this was most pronounced for judgments of positive and negative affect and extraversion, and for positive affect and conscientiousness (Letzring, Biesanz, Hall, McDonald, & Krzyzaniak, 2018). Furthermore, the causal direction predicted by STAM was supported in a study that revealed that judges who were given false information about targets’ emotional states made less accurate judgments of personality traits than judges who were given no information or true information about emotions (Letzring, Biesanz, Hall, Krzyzaniak, & McDonald, 2019). The STAM is an example of how to combine what is known about making accurate judgments in two domains, and research testing this kind
of model can lead to advances in both research areas and to a deeper understanding of how accurate judgments can be achieved in multiple domains.

An important difference for judgments of traits vs. states is that traits are assumed to be relatively consistent across situations and over time, whereas states are constantly changing. To make trait judgments, judges could rely heavily on previously-detected cues and incorporate this information with new cues. On the other hand, for judgments of states, previous cues may be less useful than cues that happen in the same moment as when judgments are made. However, knowledge about previous states may help judges detect cues that are relevant to current states and to utilize these cues correctly, which would increase state accuracy. For example, knowing that someone felt proud of a previous accomplishment could help judges look for signs of pride and accurately interpret the response to a current accomplishment, or knowing that a target was often able to deceive others in the past may be useful for determining whether a person is currently being deceptive. This is especially likely to be true when targets are being observed in situations that are similar to those observed in the past, as they may be likely to have similar thoughts, emotions, and experiences when they encounter those situations again.

**Future Directions**

**Other Levels of Personality**

The Realistic Accuracy Model was originally conceptualized as a way to understand the process of making accurate judgments of broad personality traits that are relatively consistent across situations and over time, but it is possible to apply the model to judgments of other levels of personality such as personal concerns and identity (McAdams, 1995). Personal concerns differ importantly from traits because they take the context or situation into account; therefore personal concerns are expected to differ across time, situations, and roles that people have (such as friend
or mother). Personal concerns are often related to motivation, development, or strategies for achieving one’s goals. Examining the accuracy of judgments of personal concerns would be more complex than examining the accuracy of trait judgments because the situation in which the judgments are made would need to be accounted for, possibly by asking judges to describe targets in specific situations or roles and comparing the ratings to assessments of personal concerns in those same situations or roles.

An even more complex level of personality is identity, or one’s definition of the self in terms of one’s “overall unity and purpose” (McAdams, 1995, p. 381) that is usually captured in the answer to the question, “Who am I?” Having a clear identity is associated with having unity, purpose, and meaning in life (McAdams, 1995). Examining accuracy for judgments of identity would be quite complex, given that identity continually evolves over the lifespan and is typically assessed through the use of narrative descriptions of people’s life stories (Adler et al., 2017; McAdams, 1995; 2012). The narrative descriptions are coded to identify elements that include themes, images, integrative meaning such as how one’s view of the self or world has changed in response to the environment, and structural elements such as coherence and complexity. Judges would have to rate targets’ identities, and then the level of correspondence between the target’s identity and the judge’s ratings would have to be determined. One way this could be accomplished would be to have judges rate targets on the elements that have been coded from the narratives, and then compare these judgments to the codings. Assessing accuracy of judgments of identity and personal concerns would allow researchers to examine the extent to which accuracy is possible at these deeper levels of personality and how the moderators of accuracy function within these levels.
When assessing accuracy of levels of personality other than broad-level and decontextualized personality traits, researchers will need to think about and examine how the situation in which targets are observed and judgments are made influences accuracy. It is likely that accuracy for different types of personal concerns will vary across situations and roles that people have. For example, achievement motivation may be easier to judge in academic or work settings than in social settings. The situation may be less influential for judgments of identity, as this should be consistent across situations. The key for revealing identity would be to have targets discuss their life story and how they see themselves, which is more likely to happen in some situations than others (e.g., within intimate relationships rather than casual or professional relationships).

**Interactions between Moderators of Accuracy**

A second future direction that is likely to yield useful results is a more systematic evaluation of how interactions between moderators influence accuracy. Some work has already been done in this area, as described previously in the section on Moderators of Accuracy, but not all interactions have been carefully tested. Examining interactions may resolve some of the inconsistencies that have been found in previous work and deepen our understanding of the factors that influence and are related to accuracy.

**Development of Judgmental Ability**

An aspect of judgment accuracy that has received little if any research attention is how the ability develops over time. It is likely that it is not possible to judge others accurately until a theory of mind has developed, which involves developing an understanding of other people as mental beings and acquiring the ability to attribute mental stages such as desires, beliefs, and knowledge to the self and others. Theory of mind typically develops by around four or five years
of age (Astington & Edward, 2010; Wellman, 2017). Such an ability is likely needed to understand that other people differ from one’s self and to make accurate judgments of others. One study examined the development of theory of mind and the ability to identify emotions based on vignettes, and found evidence that theory of mind develops before children are able to accurately identify emotions in others, but also that children with more fully developed theories of mind were able to more accurately identify emotions (Brown, Thibodeau, Pierucci, & Gilpin, 2015). A longitudinal study that examines theory of mind and accuracy of trait judgments across the lifespan, from preschool children to adults, would begin to shed light on the development of judgmental ability. A longitudinal study that tracked the development of these and other related abilities and perceptions would be even more informative. These kinds of studies could also examine how moderators function at different ages and developmental stages of judges and targets.

**Consequences of Judgmental Accuracy**

Finally, more research that examines the consequences of judgmental accuracy is needed. This could be particularly important for relationships, as knowing when and under what conditions accuracy is associated with good vs. bad relationship outcomes could point to ways to increase overall relationship quality and satisfaction, as well as relationship maintenance. For example, do more accurate judgments of potential or actual relationship partners enhance the probability of having or maintaining a good relationship with that person? The answer, as always, is probably “it depends.” Additional research would help us determine what factors and conditions lead to increases or decreases in accuracy, and how accuracy is related to outcomes such as relationship quality and maintenance. In some situations, seeing only the best in a person could lead to enhanced relationships; whereas in other situations it may be more beneficial to see
relationship partners as they truly are and to be aware of what makes them different from other people. Research on empathic accuracy, or knowing what someone is thinking and feeling, has shown that empathic accuracy has complex relationships with relationship satisfaction, feelings of relationship closeness, and rates of relationship dissolution (Sened et al., 2017; Simpson, Ickes, & Grich, 1999; Simpson, Oriña, & Ickes, 2007). It would be useful to examine how accuracy of judgments of personality traits relates to or predicts these types of relationship variables. Additionally, being oblivious to incompatible traits or character flaws could lead to unsatisfactory or even disastrous outcomes such as partner violence (i.e., Clements, Holtzworth-Munroe, Schweinle, & Ickes, 2007). Some cross-sectional research has found that normative accuracy based on viewing videos of targets is related to relationship variables such as perceptions of interpersonal support and social intimacy (Letzring, 2015), but longitudinal studies that examine how accuracy predicts these kinds of outcomes at later points in time would provide important insight about the consequences of accuracy for interpersonal relationships.

Other relationships that may be affected by accuracy of personality judgments could be within contexts such as relationships at work between colleagues or between managers and subordinates, relationships within educational settings such as between teachers and students, and relationships within athletics such as between coaches and athletes. Accurate judgments in these contexts could lead to more satisfying working, learning, and playing conditions, and to higher levels of success for companies, students, and teams. A good deal remains to be learned about the consequences of accuracy.

**Conclusion**

The Realistic Accuracy Model is based on the assumption that personality traits exist and that people behave relatively consistently across situations and over time. The ability of people
to make accurate judgments of others provides evidence for these assumptions. Substantial research shows that people can make accurate judgments of the personality traits of others, although levels of accuracy differ across judges, targets, traits, and quantity and quality of information. Continuing to increase our understanding of the process, moderators, and consequences of accuracy will allow researchers to design even more effective training programs to increase levels of accuracy for judging others and for increasing the accuracy with which one’s self is judged. Increased accuracy is likely to lead to better decisions in multiple realms of life and to many favorable outcomes.
References


http://dx.doi.org/10.1016/j.jrp.2016.05.013


doi:10.1016/j.jrp.2011.01.001


Krzyzaniak, S. L. (2018). Personality judgment accuracy and the role of physical fitness, cognitive functioning, and psychological well-being. (Unpublished master’s thesis), Idaho State University, Pocatello, ID.


doi:10.1016/j.jrp.2007.12.003


doi:10.1016/j.jrp.2014.05.001

Accuracy of Trait and State Judgments. Poster to be presented at the annual meeting of the Society for Personality and Social Psychology, Portland, OR.


Figure 1. Overlap between the Lens Model and the Realistic Accuracy Model. The oval represents the lens that the cues pass through in the process of cue utilization. Cues 1 and 2 are both valid and utilized, Cue 3 is not valid but is utilized, and Cue 4 is valid but not utilized. Cue validity corresponds with the relevance stage, RAM adds the stipulation that the cues must be available, and cue utilization is broken down into the stages of detection and utilization.