

# Reflective Judgment: Theory and Research on the Development of Epistemic Assumptions Through Adulthood

Patricia M. King

*Center for the Study of Higher and Postsecondary Education  
University of Michigan*

Karen Strohm Kitchener

*School of Education  
University of Denver*

The reflective judgment model (RJM) describes the development of complex reasoning in late adolescents and adults, and how the epistemological assumptions people hold are related to the way they make judgments about controversial (ill-structured) issues. This article describes the theoretical assumptions that have guided the development of the RJM in the last 25 years, showing how these ideas influenced the development of assessment protocols and led to the selection of research strategies for theory validation purposes. Strategies discussed here include a series of longitudinal studies to validate the proposed developmental sequence, cross-sectional studies examining age/educational level differences, and studies of domain specificity. Suggestions for assessing and promoting reflective thinking based on these findings are also offered here.

Do the benefits of inoculating health care workers against smallpox outweigh the risks? Will a proposed urban growth policy protect farmland without sacrificing jobs? Is affirmative action an effective tool for promoting genuine access to higher education? Controversial problems such as these about which “reasonable people reasonably disagree” are called ill-structured problems (Churchman, 1971; Wood, 1983); they are characterized by two features: that they cannot be defined with a high degree of completeness, and that they cannot be solved with a high degree of certainty. In the last 25 years, we have investigated how late adolescents and adults come to understand and make judgments about these kinds of controversial problems. In examining the responses that hundreds of individuals across a wide range of age and educational levels have given to such questions, we have made three major observations: (a) there are striking differences in people’s underlying assumptions about knowledge, or epistemic assumptions; (b) these differences in assumptions are related to the way people make and justify their own

judgments about ill-structured problems; and (c) there is a developmental sequence in the patterns of responses and judgments about such problems. The reflective judgment model (RJM; King & Kitchener, 1994; K. S. Kitchener & King, 1981) provides a theoretical framework for understanding and organizing these observations.

In this article, we begin by presenting the original theoretical grounding and underlying assumptions that have guided the development of the RJM, as well as the influence of subsequent theoretical developments, and then show how these assumptions guided the development of assessment protocols and led to the selection of research strategies for theory validation purposes. The next major section of this article summarizes research on the RJM that tested the theoretical claims about the development of reflective judgment. Last, we explore some of the implications for practice and research based on the theoretical ideas and empirical findings presented here.

---

Requests for reprints should be sent to Patricia M. King, Center for the Study of Higher and Postsecondary Education, University of Michigan, 610 E. University, Room 2117, Ann Arbor, MI 48109–1259. E-mail: patking@umich.edu

## REFLECTIVE JUDGMENT MODEL

The RJM is a model of the development of reflective thinking from late adolescence through adulthood. This construct was

first defined by Dewey (1933), who argued that reflective judgments are initiated when an individual recognizes that there is controversy or doubt about a problem that cannot be answered by formal logic alone, and involve careful consideration of one's beliefs in light of supporting evidence. This kind of reasoning remains a central goal of education, especially higher education; this is evident in several recent national reports on undergraduate education, each of which reiterated the need for college graduates to think reflectively (American Association of Colleges and Universities [AAC&U], 2002; American Association of Higher Education, American College Personnel Association [ACPA], & National Association of Student Personnel Administrators, 1998; ACPA, 1994). The RJM describes a progression of seven major steps in the development of reflective thinking leading to the capacity to make reflective judgments; each step represents a qualitatively different epistemological perspective. In defining these perspectives, we use K. S. Kitchener's (1983) definition of epistemic cognition (as distinguished from cognition and metacognition), focusing on individuals' underlying assumptions about knowledge and how it is gained. For each step in this progression (which we call stages, as defined below) the RJM includes a description of individuals' views of knowledge and concepts of justification, showing the relationship between the epistemological assumptions people hold and the way they make judgments about controversial (ill-structured) issues. The model shows how the assumptions are interrelated and how they reflect an internal logic within each stage. (See Table 1.) In the following description, we offer a general overview of these seven stages and offer examples from Stage 4, which is characteristic of the reasoning of a majority of college students, and Stage 7, which is indicative of the kind of reasoning many colleges aspire to teach (AAC&U) and which has been associated with the kind of thinking skills adults need to function effectively in today's complex societies (Baxter Magolda, in press; Kegan, 1994). These examples illustrate how qualitatively different sets of epistemic assumptions are associated with distinctly different ways of justifying beliefs in adulthood. Because a detailed descriptions of each stage is available elsewhere (King & Kitchener, 1994), here we offer only a brief summary of the seven stages.

As an introduction to this developmental progression, consider the seven stages grouped into three levels: prereflective thinking (Stages 1–3), quasi-reflective thinking (Stages 4–5), and reflective thinking (Stages 6–7). We are aware that labels such as these (e.g., “absolutist”) risk being interpreted as overly simplistic reflections of complex epistemological perspectives and that similar terms are used to refer to quite different epistemologies (as documented by Hofer & Pintrich, 1997). For this reason, we use numbers to reflect the order in which the epistemological perspective typically emerges and offer these broader categories as a more general introduction to the model. Although clustering qualitatively different stages into levels reduces complexity

by collapsing stages within levels and highlighting similarities, this strategy also risks obscuring within-level differences. Thus, the summaries that follow should be read with the understanding that important differences exist between stages, both within and across levels.

A major hallmark of *prereflective thinking* is that knowledge is assumed to be certain, and accordingly, that single correct answers exist for all questions and may be known with absolute certainty, usually from authority figures. Well- and ill-structured problems are not differentiated, as all problems are assumed to be well structured. Further, those using prereflective assumptions do not use evidence to reason toward a conclusion, relying instead on a restatement of beliefs or on unsubstantiated personal opinions. With *quasi-reflective thinking* comes the recognition that uncertainty is a part of the knowing process, the ability to see knowledge as an abstraction, and the recognition that knowledge is constructed. This is a major advance, as it lays a foundation for the construction of beliefs that are internally derived, not simply accepted from others. Further, evidence is now understood as a key part of the knowing process, as it provides an alternative to dogmatic assertions that are characteristic of prereflective thinking. Those using quasi-reflective assumptions are aware that different approaches or perspectives on controversial issues rely on different types of evidence and different rules of evidence, and that factors like these contribute to different ways of framing issues. Here are two examples of Stage 4 reasoning:

#### Stage 4 Reasoning, Example 1

*Interviewer (I)*: “Can you say that one point of view is better and another worse?”

*Respondent (R)*: “No, I really can't on this issue. It depends on your beliefs since there is no way of proving either one.”

*I*: “Can you say that one is more accurate than the other?”

*R*: “No, I can't. I believe they're both the same as far as accuracy.”

#### Stage 4 Reasoning, Example 2

*I*: “Can you say one view of creation is right and one is wrong?”

*R*: “No, because no one can prove how the world was created or how man evolved. Scientists can get close to it—an actual answer. When it comes right down to it, as to the actual change, they don't know because they can't draw a straight relationship between apes and man. There isn't a straight relationship ...”

In quasi-reflective reasoning, the link between gathering evidence and making a conclusion is tenuous; this link becomes explicit in reflective thinking, the third level of the RJM.

TABLE 1  
Summary of Reflective Judgment Stages

| Prereflective Thinking  | Quasireflective Thinking   | Reflective Thinking  |
|---|--|--|
| <p>Stage 1<br/><i>View of knowledge:</i> Knowledge is assumed to exist absolutely and concretely; it is not understood as an abstraction. It can be obtained with certainty by direct observation.</p> <p><i>Concept of justification:</i> Beliefs need no justification because there is assumed to be an absolute correspondence between what is believed to be true and what is true. Alternate beliefs are not perceived.</p> <p>“I know what I have seen.”</p>   | <p>Stage 4<br/><i>View of knowledge:</i> Knowledge is uncertain, and knowledge claims are idiosyncratic to the individual because situational variables (such as incorrect reporting of data, data lost over time, or disparities in access to information) dictate that knowing always involves an element of ambiguity.</p> <p><i>Concept of justification:</i> Beliefs are justified by giving reasons and using evidence, but the arguments and choice of evidence are idiosyncratic (e.g., choosing evidence that fits an established belief).</p> <p>“I would be more inclined to believe evolution if they had proof. It is just like the pyramids: I do not think we will ever know. Who are you going to ask? No one was there.”</p>              | <p>Stage 6<br/><i>View of knowledge:</i> Knowledge is constructed into individual conclusions about ill-structured problems on the basis of information from a variety of sources. Interpretations that are based on evaluations of evidence across contexts and on the evaluated opinions of reputable others can be known.</p> <p><i>Concept of justification:</i> Beliefs are justified by comparing evidence and opinion from different perspectives on an issue or across different contexts and by constructing solutions that are evaluated by criteria such as the weight of the evidence, the utility of the solution, or the pragmatic need for action.</p> <p>“It is very difficult in this life to be sure. There are degrees of sureness. You come to a point at which you are sure enough for a personal stance on the issue.”</p>   |
| <p>Stage 2<br/><i>View of knowledge:</i> Knowledge is assumed to be absolutely certain or certain but not immediately available. Knowledge can be obtained directly through the senses (as in direct observation) or via authority figures.</p> <p><i>Concept of justification:</i> Beliefs are unexamined and unjustified or justified by their correspondence with the beliefs of an authority figure (such as a teacher or parent). Most issues are assumed to have a right answer, so there is little or no conflict in making decisions about disputed issues.</p> <p>“If it is on the news, it has to be true.”</p> | <p>Stage 5<br/><i>View of knowledge:</i> Knowledge is contextual and subjective because it is filtered through a person’s perceptions and criteria for judgment. Only interpretations of evidence, events, or issues may be known.</p> <p><i>Concept of justification:</i> Beliefs are justified within a particular context by means of the rules of inquiry for that context and by context-specific interpretations of evidence. Specific beliefs are assumed to be context specific or are balanced against other interpretations, which complicates (and sometimes delays) conclusions.</p> <p>“People think differently and so they attack the problem differently. Other theories could be as true as my own, but based on different evidence.”</p> | <p>Stage 7<br/><i>View of knowledge:</i> Knowledge is the outcome of a process of reasonable inquiry in which solutions to ill-structured problems are constructed. The adequacy of those solutions is evaluated in terms of what is most reasonable or probable according to the current evidence, and it is reevaluated when relevant new evidence, perspectives, or tools of inquiry become available.</p> <p><i>Concept of justification:</i> Beliefs are justified probabilistically on the basis of a variety of interpretive considerations, such as the weight of the evidence, the explanatory value of the interpretations, the risk of erroneous conclusions, consequences of alternative judgments, and the interrelations of these factors. Conclusions are defended as representing the most complete, plausible, or compelling understanding of an issue on the basis of the available evidence.</p> <p>“One can judge an argument by how well thought-out the positions are, what kinds of reasoning and evidence are used to support it, and how consistent the way one argues on this topic is as compared with how one argues on other topics.”</p> |

(continued)

TABLE 1 (Continued)

| <i>Prereflective Thinking</i>  | <i>Quasireflective Thinking</i> | <i>Reflective Thinking</i> |
|--|---------------------------------|----------------------------|
| <p>Stage 3</p> <p><i>View of knowledge:</i> Knowledge is assumed to be absolutely certain or temporarily uncertain. In areas of temporary uncertainty, only personal beliefs can be known until absolute knowledge is obtained. In areas of absolute certainty, knowledge is obtained from authorities.</p> <p><i>Concept of justification:</i> In areas in which certain answers exist, beliefs are justified by reference to authorities' views. In areas in which answers do not exist, beliefs are defended as personal opinion because the link between evidence and beliefs is unclear.</p> <p><i>"When there is evidence that people can give to convince everybody one way or another, then it will be knowledge; until then, it is just a guess."</i></p> |                                 |                            |

*Note.* From *Developing Reflective Judgment* (p. 14–15), by P. M. King and K. S. Kitchener, 1994, San Francisco: Jossey-Bass. Copyright 1994 by John Wiley & Sons, Inc. Reprinted with permission.

*Reflective thinkers* consistently and comfortably use evidence and reason in support of their judgments. They argue that knowledge claims must be understood in relation to the context in which they were generated, but that they can be evaluated for their coherence and consistency with available information. Because new data or new perspectives may emerge as knowledge is constructed and reconstructed, individuals using assumptions of reflective thinking remain open to reevaluating their conclusions and knowledge claims.

#### Stage 7 Reasoning, Example 1

I: "Can you ever say you know for sure?"

R: "It's [the view that the Egyptians built the pyramids] very far along the continuum of what is probable."

I: "Can you say one is right and one is wrong?"

R: "Right and wrong are not comfortable categories to assign to this kind of item—more or less like or reasonable—more or less in keeping with what the facts seem to be."

#### Stage 7 Reasoning, Example 2

R: "It's my belief that you have to be very skeptical about what you read for popular consumption ... even for professional consumption."

I: "How do you ever know what to believe?"

R: "I read widely ... of many points of view. Partly [it's] reliance on people you think you can rely on, who seem to be reputable journalists, who make measured judgments. Then reading widely and estimating where the reputable people line up or where the weight of the evidence lies."

These examples illustrate the kinds of developmentally ordered differences in the way people reason about ill-structured problems that are described in the RJM.

### Theoretical Assumptions Underlying the RJM

We turn next to one of the questions that is the focus of this volume, the paradigmatic assumptions underlying each theory of personal epistemology. Because research on the RJM spans more than 25 years, we will introduce the theoretical assumptions from both historical and contemporaneous perspectives.

**Developmental traditions.** The RJM evolved out of a careful examination of the few models of late adolescent and adult intellectual development that existed in the late 1970s. Our initial conceptualization (K. S. Kitchener & King, 1981) was grounded in the cognitive-developmental tradition of Piaget (1965; Piaget & Inhelder, 1969) and Kohlberg (1969). Other developmental theorists in this tradition whose work informed our early conceptualization of the RJM were Perry (1968, 1970), Broughton (1975, 1978), Loevinger (1976),

and Harvey, Hunt, and Schroder (1961). The cognitive-developmental tradition has much in common with more recent constructive-developmental perspectives (e.g., Fischer & Pruyne, 2002; Kegan, 1982, 1994). What these two approaches share are (a) the underlying assumption that meaning is constructed, (b) the emphasis on understanding how individuals make meaning of their experiences, and (c) the assumption that development (not just change) occurs as people interact with their environments. Another central defining feature is that patterns of meaning-making are described in developmental terms, that is, the frameworks people use for interpreting their experiences (e.g., categories and organizing principles) are described as becoming more complex, integrated, and complete over time. These changes do not occur automatically but rather through interaction with an environment that both challenges and supports growth.

However, our data led us to reject two well-known assumptions espoused by prominent theorists from this tradition. First, unlike Piaget, we do not assume that cognitive development is best measured by deductive reasoning, nor do we assume that it is complete with the emergence of formal operations at age 16 (indeed, our data show that this is not the case). And in contrast to Kohlberg, we do not claim cross-cultural universality, and we endorse Rest's (1979) concept of a complex rather than a simple stage model of development.

**Stage theory.** At the time the RJM was being developed in the 1970s, Rest (a faculty member who supervised our initial research) was also working within the cognitive-developmental tradition. As a researcher of moral development, he was beginning to raise questions about the adequacy of what he called the "simple stage" model being advanced by Kohlberg (1969), a critique Rest (1979) later published in his first book. We took the opportunity to ask similar questions based on our initial study (K. S. Kitchener & King, 1981), such as whether there was stage variability or consistency among an individual's responses. Our scoring procedures were intentionally designed to allow for this question to be tested (i.e., allowing raters to record multiple stages if several were apparent in a given interview protocol). We found that Rest's alternative, the "complex stage" model of development, provided a good explanatory framework for our data. That is, we observed that development in reasoning about ill-structured, controversial problems has stage-like properties, but not that it evolves in a lock-step, one-stage-at-a-time fashion. Hence, we refer to the major categories of thinking and interrelated clusters of assumptions as stages, but our use of this term is qualified, based on specific assumptions and definitions that fall outside more traditional usage. Below, we offer data illustrating development across stages that support this approach.

We acknowledge that stage models within the cognitive-developmental tradition have been criticized as providing inadequate conceptual frameworks for describing devel-



opment (e.g., Flavell, 1971). Two underlying assumptions about stages (traditionally defined) have drawn considerable criticism. The first is that individuals utilize only one organizing framework (stage) at a time and, therefore, that development from stage to stage is abrupt with any overlap between stages occurring briefly only during transitions between stages. At the time of stage consolidation, stage usage is assumed to peak at 100%, consistent with the common phrase used when referring to stage theories, being “in a stage.” The second criticism is that the stages constitute an invariant sequence that exists across all cultures. Kohlberg’s (1969, 1984) claim to the universality of his sequence of stages of moral development was based on his refutation of moral relativism as an inadequate philosophical framework (Kohlberg, 1991) and on cross-cultural studies indicating that the pattern of development he proposed was also apparent among individuals across several cultures. We do not make these claims. We do, however, support the other claims within this tradition (that meaning is constructed, that these constructions are developmentally ordered, and that development is the result of person-environment interactions).

Rest’s (1979) complex stage model better captures the nature of development of reflective judgment because it accounts for the observed patterns in data gathered using the Reflective Judgment Interview (RJI). For example, it is common to find an individual who relies heavily on Stage 4 assumptions while reasoning about a controversial problem, but who also makes statements that are consistent with Stage 3 and Stage 5 assumptions. By contrast, someone who relies heavily on Stage 2 assumptions rarely uses assumptions of any stage higher than Stage 3. As Rest noted, this approach suggests a “much messier and complicated picture of development” (p. 65) than does a simple stage approach.

Does the complex stage model proposed by Rest accurately capture reflective judgment data? We examined variability of scores of those in our 10-year longitudinal sample (described later in the review of RJM research) to answer this question. In only two cases were the RJI ratings limited to a single stage; in the vast majority of cases, a subdominant score was assigned, and this was almost always an adjacent stage. In a small proportion of cases, more than two stage scores were assigned. Wood (1997) examined the variability of RJI scores using data from 15 studies for which raw data were available ( $n = 1,995$  problem scores; reported in Wood, 1993). He constructed a “percent stage utilization score” based on all responses across the four problems; this score indicated the proportion of time each stage was assigned. He then calculated a series of spline regressions (Darlington, 1990), which predicted stage utilization on the basis of overall RJI score. (A graph of these may be found in King & Kitchener, 1994, Figure 6.2.) Here, development is pictured as a series of uneven, overlapping waves, where usage of given stage assumptions rises and falls in different proportions over time. As this figure shows, for those whose modal score was Stage 2, 70% of the ratings were for Stage 2, with

less than 20% at Stage 3. About two-thirds of the ratings were at Stage 3 for those with a modal Stage 3 rating, with the remainder fairly equally distributed between Stages 2 and 4. A similar pattern was obtained for those with a modal Stage 4 rating; here, the remaining ratings were split fairly equally between Stages 3 and 5. However, the shape of the “wave” was much flatter for Stage 5, with only about half of the ratings at Stage 5; the remainder were spread two stages higher and lower than the mode. The shape of the curve for Stage 6 was more similar to those for Stages 3 and 4. In other words, variability in reasoning across stages was the norm and not the exception in these ratings. No individuals evidenced non-adjacent utilization patterns (3/5, 4/6, etc.). This evidence is consistent with the assumptions of complex stage theory (Rest, 1979) and adds further evidence that characterizing individuals as being “in” or “at” a single stage is misleading. Based on these patterns, King, Kitchener, and Wood (1994) suggested that development in reflective thinking be characterized as

... waves across a mixture of stages, where the peak of a wave is the most commonly used set of assumptions. While there is still an observable pattern to the movement between stages, this developmental movement is better described as the changing shape of the wave rather than as a pattern of uniform steps interspersed with plateaus. (p. 140)

This shift from simple to complex stage theory represents a radical change in how development is conceptualized; indeed, it may be considered a change of paradigmatic proportions within stage theory.

*Skill theory.* The second theoretical model that has affected our thinking about RJM research is Fischer’s skill theory. Fischer and his colleagues (Fischer, 1980; Fischer, Bullcock, Rosenberg, & Raya, 1993; Fischer & Lamborn, 1989; K. S. Kitchener & Fischer, 1990) identified seven developmental levels that emerge between ages 2 and 30. These levels are divided into two overlapping tiers, the representational tier and the abstract tier. The focus of the representational tier is on individuals’ ability to manipulate concrete representations, objects, people, or events; the focus of the abstract tier is on individuals’ ability to integrate, manipulate, and reason using abstract concepts. This portion of skill theory has much in common with Kegan’s (1982, 1994) theory of the development of mature capacity toward self-authorship. The upper levels of Fischer’s model also have much in common with the RJM; in fact, the seven stages of the RJM can be readily mapped onto Representational Levels 1–4 and Abstract Levels 2–4 (Fischer & Pruyne, 2002; King, 1985; K. S. Kitchener, 2002; K. S. Kitchener & Fischer, 1990). Reflective thinking requires the ability to think abstractly, which explains the correspondence between the abstract levels of skill theory and Stages 4–7. K. S. Kitchener (2002) also suggested that skill theory provides a framework

for comparing the multiple models of folk epistemology (R. Kitchener, 2002) and personal epistemology, such as those in Hofer and Pintrich's (1997, 2002) comprehensive reviews.

Another important and influential aspect of Fischer's work is his assumption that no skills exist independent of the environment and that the skill levels a person demonstrates will vary depending on the conditions under which they are assessed. (Notably, the acknowledgement that performance varies with task demands is incompatible with the simple stage assumption that individuals are "in" one stage at a time.) Fischer and his colleagues (Fischer & Pipp, 1984; Lamborn & Fischer, 1988) posited that variability in individuals' responses across tasks reflects the degree of "contextual support" (e.g., memory prompts, feedback, opportunity to practice) available at the time of the assessment. He suggested that tasks that require performance without support elicit a person's "*functional level*" capacity, but that tasks that provide contextual support can elicit performance at levels that are closer to the upper limit of the person's cognitive capacity, called "*optimal level*." Contextual support can be provided by offering participants a high-level example of the skill, the opportunity to ask questions about the example, the chance to practice the skill in a variety of settings, and so on:

It is the emergence of this general capacity [for abstract thinking] that establishes an upper limit on the level of independent functioning an individual can potentially achieve in reflective thinking or other domains involving advanced abstract thinking. This upper limit of skill development is termed the *optimal level*. (Fischer & Pruyne, 2002, p. 169; italics in original)

The space between functional and optimal levels is called one's "developmental range" and reflects the range of skills that an individual can access and produce depending on the circumstances. That is, the nature of the person's experience—including the structure of the learning and assessment tasks—affects where within this developmental range a person's performance will fall. If courses and other opportunities for student learning do not provide contextual support for developing the skills associated with forming abstract concepts like reflective thinking (a criticism commonly levied at both schools and colleges), students will be more likely to perform at functional rather than optimal levels. Further, those who have access to higher levels of development would also have access to a larger repertoire of responses from which to choose, explaining Fischer's (1980; Kitchener & Fischer, 1990) hypothesis that optimal and functional level will diverge to a greater degree as the person approaches higher levels of development.

Fischer (1980; Kitchener & Fischer, 1990) also hypothesized that functional level performance would improve in a slow, steady fashion, resulting in a gradual, even slope if graphed over time. By contrast, he hypothesized that optimal level performance would be less even and instead be charac-

terized by spurts at given age levels, followed by plateaus between spurts. Thus, researchers would expect different developmental trajectories depending on whether their measures yield data on functional or optimal level performance.

The ability to operate at an optimal level is influenced not only by support and practice, but also by changes in brain activity and the reorganization of neural networks (Fischer & Pruyne, 2002; Fischer & Rose, 1994). The emergence of abstractions and reflective thinking appears to involve brain development that does not occur until late adolescence and early adulthood.

As this brief summary shows, skill theory provides an innovative approach to the study of human development in general and the development of reflective thinking (with its grounding in epistemic cognition) in particular. For example, the concept of developmental range provides an alternative way of addressing the question of being "in" or "at" a single stage on the RJM, and a way of targeting educational interventions to students' developmental levels. Further, its differentiation of functional and optimal level suggests the need to analyze measures of cognitive development or personal epistemology for degree of contextual support, and to develop measures that assess both levels. And although skill theory is certainly consistent with the person-environment interaction assumptions inherent in the cognitive-developmental paradigm, it specifies particular environmental variables (e.g., contextual support) that appear to affect how students learn to engage in the production of more advanced behaviors (here, reflective thinking).

### Measuring Reflective Judgment

Over the last 25 years, we have experimented with several assessment procedures to measure reflective judgment and its underlying epistemic assumptions. In order to illustrate the links between theoretical assumptions stemming from our research paradigm and our assessment approaches, we describe how the development of several assessment procedures was grounded in theoretical considerations.

*The Reflective Judgment Interview.* The RJI was initially designed to measure reflective thinking as described by the RJM and to inform theory development. We used an iterative process between theory development and assessment ("boot-strapping") for much of the first decade of research on the RJM, moving back and forth between theory development and validation efforts. The RJI uses a semistructured interview format to elicit responses from participants regarding how they reason about ill-structured problems. A trained and certified interviewer asks a series of predetermined but open-ended questions regarding their reasoning in order to get at their fundamental assumptions concerning knowledge and how it is gained. The original interview consisted of four controversial problems (the accuracy of news reporting, the creation of human beings, the safety of chemical additives to foods, and the building of the Egyptian pyramids). A di-

lemma on the safety of nuclear power was added for the 10-year longitudinal retest, and several discipline-specific dilemmas have been used in subsequent studies (business, chemistry, and psychology). The RJI also includes a standardized series of probe questions; each question is designed to elicit comments that reflect individuals' epistemic assumptions (specifically, their assumptions about knowledge, how it is gained, how they decide what to believe). Probe questions ask about the basis for their point of view, the certainty with which they hold that view, whether differing opinions on the topic are right or wrong or better or worse, and how it is possible people (including experts) disagree about the topic. The one-hour interview was designed to yield a picture of how people approach the task of knowing and making judgments about controversial intellectual issues by looking at ways they understand and make meaning of concepts such as evidence, differences of opinion, uncertainty, and interpretation. (For a detailed description of the RJI, see King & Kitchener, 1994, Chapter 5 and Resource A.)

The RJI is scored by trained and certified raters using the Reflective Judgment Scoring Rules (K. S. Kitchener & King, 1985). Consistent with the complex stage assumptions noted above, raters can assign three scores to each dilemma to reflect whatever characteristics of reasoning they observe in the interview; these typically range across two adjacent stages. The stage most clearly or frequently observed is coded first as the dominant stage, followed by the subdominant stage(s). Occasionally, one dilemma includes statements that reflect three different stages; this is rare, but recording all three is an option available to raters if they determine that this best captures the reasoning in the interview. The point here is that scoring is designed to reflect whatever stage-characteristic responses are evident in the transcript and not to assume a priori that consistency (or inconsistency) will be observed. Assigned scores are then weighted across dominant and subdominant stages, and an overall dilemma score is calculated. The training and certification programs for interviewers and raters were put into place to assure comparability across studies and researchers. (We have recently discontinued these programs in order to focus on the development of other measures.)

Fischer's (Fischer & Pruyene, 2002; Kitchener & Fischer, 1990) differentiation of functional and optimal levels of performance raised several questions for research on the RJM. In particular, it called for the consideration of the level of performance characterized by the RJI: because no contextual support is offered, the RJI may be considered a measure of functional level. As such, it may underestimate a person's capacity to engage in reflective thinking, yielding a score at the lower rather than the higher end of the individual's developmental range. Implications of this insight for research and teaching are explored later in the article.

*The Prototypic Reflective Judgement Interview (PRJI).* As noted earlier, there are several points of correspondence between Fischer's skill theory and the RJM. In

order to evaluate whether the developmental patterns Fischer (1980) had predicted occurred for the development of reflective thinking as defined by the RJM (K. S. Kitchener & Fischer, 1990), a measure of optimal level performance was needed. Fischer argued that optimal measures required two qualities: first, there had to be an independent assessment for each step in the developmental sequence. Second, the research design had to vary relevant characteristics of the participants (especially age) as well as characteristics of the environment (e.g., the amount of environmental support provided for a response). Using these criteria, K. S. Kitchener, Lynch, Fischer, and Wood (1993) designed a new measure, the prototypic reflective judgment interview (PRJI) to assess reflective judgment under conditions of support and practice. The study measured both functional level (using the RJI) and optimal level (using the PRJI) to determine whether scores differed between the two measures of reflective thinking and whether there was evidence for age-related spurts and plateaus using the optimal level measure. Because of its theoretical significance, this study is summarized here.

To construct the PRJI, two problems from the RJI were selected, and stage-prototypic responses were written for reflective judgment Stages 2 through 7. These responses were based on answers given by people for the same problems used in the RJI; they were presented in order to give contextual support for high-level responses to the reflective judgment problems. Participants were first asked to complete the RJI, then to read one of the prototypic statements, and to respond to a series of questions that directed their attention to key elements of the statement. They were then asked to explain the prototypic statement in their own words; each answer was scored as a "hit" if it accurately paraphrased the statement and as a "miss" if it did not. This procedure was repeated for each reflective judgment stage and both of the problems. They were then given two prototypic statements addressing a different ill-structured problem; the statements were selected to correspond with the highest and second highest stages they had paraphrased in the interview. Participants were asked to think about these statements prior to the next testing, which took place two weeks later; this strategy provided contextual support in the form of exposure to and the opportunity to think about higher stage responses to ill-structured problems. The procedure was repeated at the next testing, fulfilling the practice component of contextual support.

Three findings from the study are relevant for the current discussion of how stage theory and skill theory as complementary theoretical models have informed research on the RJM. First, participants scored higher on the PRJI than on the RJI. This supports the idea that individuals are not "in" a stage but rather that they have access to several stages, and that this reflects the effects of contextual support and practice. In other words, contextual support appears to increase the individuals' access to higher stage functioning,



yielding more advanced levels of performance (here, higher reflective thinking scores). Second, there was an age-related ceiling in the PRJI even after practice, suggesting that optimal levels are age-related. This is consistent with the developmental trends observed in RJM research (described later) but offers new information about the nature and limits of age-related trends. Third, there was evidence of age-related developmental spurts on the PRJI at reflective judgment Stages 4 to 6. This is consistent with Fischer's hypothesis that the emergence of optimal levels is marked by spurts in performance and then plateaus. It also helps explain the growth in reflective thinking that has been observed with samples of traditional-age college students (whether this growth is consistent with collegiate goals is discussed later).

*The Reasoning About Current Issues Test (RCI).* Although the RJI and PRJI provided extremely rich information for theory development, their expense both in time and money was problematic in terms of conducting the kind of validation and application studies that were of interest to researchers and educators alike. In the process of developing an objectively scored measure of reflective thinking, we developed and tested several different approaches; these are described by Wood, Kitchener, and Jensen (2002). Here, we discuss the most recent measure, the Reasoning About Current Issues Test (RCI). This is an objectively scored instrument that was built on research using prior measures, but using a format that is amenable to large-scale administration. Because this is described elsewhere, the focus of this discussion is on ways our theoretical assumptions guided measurement development.

The RCI was modeled after Rest's Defining Issues Test (DIT) of moral judgment (Rest, 1979, 1986; Rest, Narvaez, Bebeau, & Thoma, 1999). The DIT has been found to be a highly reliable measure of moral judgment, able to detect developmental change over time, and sensitive to macrolevel changes in reasoning about social issues among adults. (For reviews of research using the DIT among college students, see King & Mayhew, 2002, in press.) In the RCI, respondents are asked to read a dilemma similar to those used in the RJI. In addition to the chemicals in foods dilemma, several others have been used that reflect contemporary issues (causes of alcoholism, workforce preparation, immigration policy, determinants of sexual orientation). The RCI first asks respondents to write a short statement describing their response in their own words. These written statements served to "prime the pump" by encouraging respondents to start thinking about their views on the given topics. Respondents are then asked to rate and rank in order a series of short statements to indicate the statements' similarity to the respondents' own views; each statement reflects the epistemic assumptions of one of the reflective judgment stages.

Each statement was based on responses made by respondents taking the RJI and modified from the statements developed for the PRJI. By merit of being a recognition task, this format provides contextual support (in contrast to the RJI). These brief statements are not written to capture an individual's whole network of underlying assumptions on which a judgment is based, nor to yield a nuanced articulation of how the individual approaches making judgments about controversial issues. Rather, it appears that responding to short items serves to activate the internal organizing schemas that individuals use to make judgments about the given issue, but without filling in the details about the specific rationale used and strategies employed, and without articulating the specific epistemic principles underlying the approach. (For a more detailed description of this rationale as applied to the DIT, see Rest et al., 1999.)

In addition, we wished to control for the possibility that the respondents would endorse statements that sounded impressive (e.g., that used sophisticated vocabulary) but that were not similar to the approach they used, or even to an approach they aspired to use. To address this concern, we created a series of statements that are grammatically correct but nonsensical. When these are selected, the responses for that problem are excluded from the analyses. The RCI score is calculated across all dilemma topics based on the statements most often ranked as similar to the participant's own view. Internal consistency reliabilities have been in the low to mid-.70s (depending on the sample). It takes 30 to 45 minutes to complete (Wood et al., 2002).

There are many trade-offs to be made when moving from a production task with open-ended questions (the RJI) to a recognition task where respondents are asked to choose from a limited set of predefined options (the RCI). Although both approaches are designed to tap into related skills required for the production of reflective thinking, the two are not simply different formats that yield comparable scores; rather, each serves a different purpose, makes different demands on respondents, and yields a different "snapshot" of the development of reflective thinking. Having participants evaluate statements provides more contextual support than responding to open-ended interview questions; therefore, we would anticipate that individuals would score higher on the RCI than the RJI, and this has been the case.

### Research on the RJM

In the last 25 years, we have learned a great deal about reflective thinking and how it develops. The centerpiece of our book, *Developing Reflective Judgment* (King & Kitchener, 1994), is a review of this research base. It reports both the results of our 10-year longitudinal study of the development of reflective judgment using three age/educational level cohorts ( $n = 80$  at Time 1), other longitudinal studies of 120 other respondents, as well as a review of cross-sectional studies in which more than 1,700 people

(high school students, college students, graduate students, and nonstudent adults) completed the RJI. Since the publication of that volume, Wood (1997) has completed a comprehensive secondary analysis of these data, and an updated literature review was published (King & Kitchener, 2002). Interested readers should consult these works for details. Here, we will summarize the major findings from this body of research, especially as they pertain to the focus of this special issue, how our theoretical framework has guided research on reflective judgment.

*Validating the developmental sequence.* RJM was proposed as a model of reflective thinking in the cognitive-developmental tradition, where the major claim is that the stages constitute a developmental sequence. Documenting the existence of this sequence and validating the model requires longitudinal data. Toward this end, we conducted a 10-year longitudinal study using three age/educational cohorts ( $n = 80$  at Time 1,  $n = 53$  at Time 4). At Time 1, the three gender-balanced cohorts included high school juniors, college juniors, and third-year doctoral students; the younger two groups were matched to the doctoral students on gender and academic aptitude (based on scores from the Minnesota Scholastic Aptitude Test). This was designed as a check of the competing hypothesis that obtained cohort differences on the RJI (e.g., if graduate students scored higher than college students) could be attributed to differences in aptitude. Even with this control, age and educational level remained confounded in this study. By the time of the last testing, all but one of the high school cohort had completed a bachelor's degree, and about half of the college cohort had completed post-baccalaureate degrees. This yielded a well-educated sample and served as a leveling factor for the age/educational level confounding at Time 1.

Mean RJI scores were significantly different between groups at Time 1 (1977), with the doctoral students scoring the highest ( $M = 5.67$ ), followed by the college students ( $M = 3.76$ ), and the high school students ( $M = 2.77$ ). The RJI mean score increased consistently for the high school and college student groups at each subsequent testing (1979, 1983, and 1987). Over the 10-year period, the former high school students' RJI scores increased over 2.5 stages to 5.29, the former college students' scores rose an average of 1.29 stages to 5.05, and the mean scores of the former doctoral students increased an average of .54, to 6.21. The overall rate of increase (less than two stages in two years) suggests that reflective thinking evolves slowly and steadily, even among those engaged in postsecondary education.

To determine whether a developmental sequence existed, we carefully examined the pattern of scores over time—and even reprinted the list of individual RJI scores in their entirety for other researchers to examine (King & Kitchener, 1994, Table B6.2). We found that during the 10 years of this study, the use of higher stage reasoning increased, and

the use of lower stage reasoning decreased. This is evident in the individual examples reported by King and Kitchener (1994). For example, listed below are RJI scores for three individuals (a high school student, #105; a college student, #417; and a doctoral student, #310) at Times 1 (1977), 2 (1979), 3 (1983), and 4 (1987). The mean score is given first, followed by the mode (dominant/subdominant stage observed):

105, T1: 3.17; 3(4); T2: 3.96; 4(3); T3: 5.67; 6(5); T4: 6.59; 7(6)

417, T1: 3.92; 4(3); T2: 3.63; 4(3); T3: 4.54; 4(5); T4: 4.96; 5(6/4)

310, T1: 5.29; 5(6); T2: 5.63; 6(5); T3: 5.71 6(5); T4: 6.02; 6(7/5)

Both types of analyses offer strong support for the claim that the posited reflective judgment stages form a developmental sequence.

Similar patterns of change were obtained in six other longitudinal studies involving an additional 180 individuals who took the RJI (Brabeck & Wood, 1990; Polkosnik & Winston, 1989; Sakalys, 1984; Schmidt, 1985; Van Tine, 1990; Welfel & Davison, 1986) and ranging in duration from 3 months to 4 years. The most noteworthy finding among these studies is that the pervasive pattern is one of growth or stability. As King and Kitchener (1994) reported, "in every sample tested, the scores either stayed the same or increased over time. Further, with two exceptions, the mean score increased significantly for all groups tested at 1- to 4-year intervals" (p. 156). The amount of change was smallest in studies of short duration (3–4 months); significant increases were consistently observed in studies of at least a year's duration. In studies reporting incidence of regressions (Brabeck & Wood; King & Kitchener; Sakalys; Schmidt; Welfel & Davison), 0–16% of the mean scores declined between testings, while 84–100% of the mean scores either remained consistent or increased. This suggests that change in reflective thinking over time is better reflected as stability or development rather than decline, and that earlier stage assumptions are rarely used once they are replaced with more advanced assumptions.

Similarly, longitudinal data based on RCI scores obtained at the beginning and end of the first year of college yielded significant increases of about one third of a standard deviation, with comparable gain scores by gender and ethnicity (K. S. Kitchener, Wood, & Jensen, 2003). These freshmen were tested again as sophomores, and a sample of juniors was retested as seniors; RCI scores again increased significantly over time. These findings provide additional evidence that the RJM describes a developmental sequence. However, the growth correlation coefficient was significantly and negatively correlated with scores at Time 1 for the entire sample, by class and by gender: Those who had the lowest scores at Time 1 gained the most, and those who entered with the high-

est scores gained the least. The pattern was similar for a small subset of the participants who also completed the RJI (K. S. Kitchener et al., 2003).

*Differences by age/educational level.* Another desirable characteristic of a model of reflective thinking is that it can detect predictable changes in thinking across educational levels (e.g., that graduate students score higher than undergraduate students). Over two dozen cross-sectional studies have been used to examine educational level differences in reflective judgment; these studies include samples of high school students, traditional- and nontraditional-age college students, graduate students, and nonstudent adults. Questions related to educational level differences have been of particular interest to those interested in using reflective thinking as a college outcomes variable. Because promoting intellectual development (and especially skills associated with complex reasoning) is a common goal of higher education, studies documenting complex reasoning among college students have been of interest among many higher education researchers. As these studies have been summarized elsewhere, we present only a brief review of these findings.

King et al. (1994) reviewed 25 studies in which more than 1,500 respondents from across the United States took the RJI. Student RJI scores increased slowly but steadily across educational levels, from high school ( $M = 3.2$ ) to the first year of college ( $M = 3.6$ ) to the senior year of college ( $M = 4.0$ ) to early graduate study ( $M = 4.6$ ) to advanced doctoral study ( $M = 5.3$ ). The average RJI scores for nonstudent adults with and without college degrees were 4.3 and 3.6, respectively. The high school students consistently evidenced the assumptions associated with prereflective thinking, such as making decisions on the basis of beliefs that are not subject to evaluation, especially when a conclusion was consistent with what they wanted to believe. Among the college samples, the shift to Stage 4 reasoning indicates that the students had accepted uncertainty as part of the knowing process and were using evidence more consistently to make judgments. Kroll (1992) eloquently captured the shift from prereflective to quasi-reflective thinking as the movement from “ignorant certainty” (the dogmatic assertions characteristic of prereflective thinking) to “intelligent confusion” (acknowledging what you don’t know, and why). Although this represents an important step toward reflective thinking, it is not the kind of thinking that is consistent with intended college outcomes (Brabeck, 1983; King, 1992). Only advanced doctoral students consistently used the assumptions of reflective thinking.

A similar pattern of findings was reported based on studies that used the RCI. Wood, Kitchener, and Jensen (2003) conducted a meta-analysis based on all available studies using RCI data; this yielded a sample of 8,537 students who were enrolled in college, graduate, and professional programs at seven different colleges or universities. They found significant differences by educational level, even after controlling for academic aptitude and prior academic achieve-

ment. Graduate students scored significantly higher than did medical students, who scored significantly higher than did undergraduate students ( $p < .001$ ). Among the undergraduate students, significant differences were found between early level college students (freshmen and sophomores) and more advanced students (seniors). Thus, the educational level differences in reflective thinking that were found using the RJI were also found using the RCI; however, scores by educational level were about one stage higher on the RCI than on the RJI.

Data from cross-sectional studies showing upward trends in reflective judgment scores across age/educational levels offer corroborating evidence that the RJM describes a developmental sequence. In addition, this collection of studies (especially those that controlled for age) offers evidence that development in reflective thinking is associated with participation in educational programs.

*Domain specificity.* Do individuals reason using similar sets of epistemic assumptions across domains? That is, do respondents score similarly or differently when reasoning about controversies of different content? We have analyzed score variability using several indices. Internal consistency, as measured by coefficient alpha, has been high, with the median scores in the low .80s (King & Kitchener, 1994). Inter-dilemma correlations have been lower, varying with the heterogeneity of scores in the sample, typically in the mid .40s. King, Kitchener, Wood, and Davison (1989) examined individual modal RJI scores and found that the modal score was consistent across dilemmas 75% of the time. However, Wood et al. (2003) reported a significant main effect for dilemma topic using the RCI, as well as an interaction of topic by education level. Students in all four collegiate class levels (freshman through senior) plus graduate students tended to score higher on the two psychology dilemmas (origins of alcoholism and of homosexuality) than on the other three (artificial sweeteners, curricular reform, and immigration policy); however the class difference was accounted for by the higher scores of the seniors and graduate students. That is, the magnitude of the dilemma differences was more pronounced for the more advanced students. Whether this is an artifact of sampling (e.g., representation of behavioral science majors among the seniors) is not known. Interestingly, the scores within the two sets of dilemmas (psychology and nonpsychology) were quite similar.

These findings suggest that there is a relatively high rate of consistency in people’s use of epistemic assumptions when reasoning about ill-structured problems. This could be because the RJM describes development in molar rather than fine-grained terms, and therefore is less sensitive to differences in dilemma content. Alternatively, it may be that epistemic assumptions themselves provide a guiding framework for making interpretive judgments that individuals use across a variety of problems such as those measured by the RJI and RCI.

Another way to consider questions related to domain specificity is to look at whether people reason similarly in terms of reflective thinking as compared with how they reason about issues in other areas. Data from several studies that have examined this question (reviewed in King & Kitchener, 1994) strongly suggest that development in reflective judgment is related to but distinct from development in other aspects of cognitive development (verbal aptitude, formal operations, academic ability, critical thinking) and from moral and identity development, and strongly predictive of tolerance for diversity (Guthrie, King, & Palmer, 2000).

### IMPLICATIONS FOR PRACTICE AND RESEARCH

How can educators apply their understanding of the nature of the development of reflective thinking as described by the RJM to educational practice? The theory and research presented here offer many possibilities for answering this question. First, the strong effects associated with education offer a hopeful sign that the educational experiences for many students are effective in promoting growth toward reflective thinking. However, the nature of these practices remains largely unexplored, and there is considerable concern (e.g., Baxter Magolda, *in press*; King, 1992) that the observed reflective thinking skills are not as developed as those called for in the national reports mentioned at the beginning of this article. Nor are they at the level consistent with college goals for students, nor at the level associated with the complex issues and decisions college students will face upon graduation, whether as employees, citizens, consumers, or parents.

Second, consider the consistent finding that development in reflective thinking appears to unfold in a slow, steady manner following the sequence of stages outlined in the RJM. Without data on specific educational experiences affecting this growth curve, it is reasonable to assume that theoretically grounded interventions would yield increases in performance, but probably not in dramatic proportions. Given that stage assumptions are organizing categories for viewing knowledge and knowing, and given that each stage is more like a molar than a fine-grained unit of analysis for development, slow, steady progress is a more reasonable expectation; after all, each stage is a dramatic shift in world view and one's role as a knower. We have offered a number of suggestions elsewhere for promoting reflective thinking (King, 1992, 2000; King & Kitchener, 1994, 2002; K. S. Kitchener et al., 2003; Wood & Lynch, 1998). These range from intentionally incorporating ill-structured problems into the curriculum to improving discipline-specific contextual support, to structuring opportunities for practice and feedback to stimulate optimal level thinking. In each of these practices, students are encouraged to examine their assumptions, gather and interrogate the available evidence from multiple perspec-

tives, and be responsible for offering their own conclusions of the evidence.

Third, it is noteworthy that virtually all the studies that comprise the database for the RJM have measured functional level, not optimal level. In only one study (Kitchener et al., 1993) did the measure of reflective judgment offer contextual support, and probably not at a level that would elicit performance at the upper reaches of a participant's developmental range. According to skill theory, functional level measures offer a low estimate of individuals' ability to engage in reflective thinking. If the average educational level scores are low estimates, then the concerns indicating deficits in student performance around reflective thinking may be overstated. Kroll (1992) also discouraged educators from directing their efforts toward a student's average performance; instead, he encouraged teachers to focus on the leading edge of development, which would be at a higher level within the student's developmental range. Similarly, the finding of differences in performance with and without contextual support suggests that educators should be encouraged to evaluate the amount and kind of contextual support they offer when assessing reflective thinking, for example, in student papers.

K. S. Kitchener et al. (2003) provided new information on the role of student involvement in an assortment of campus activities in promoting reflective thinking. In addition to the RCI, they also administered the College Student Experiences Questionnaire (CSEQ; Pace, 1990), which asks students to indicate on a 4-point scale how frequently they participated in particular collegiate activities. Findings from the freshman sample highlight the complexity of the relationship between participation in college activities and epistemological thinking. Predictably, those who entered college with higher reflective judgment scores also graduated with higher scores. For freshmen, the relationship between Time 1 and Time 2 scores are consistent with expectations about students with higher and lower reflective scores: Those who entered with higher scores endorsed an appreciation for challenging courses, a willingness to work harder in classes, and a commitment to thinking through ideas themselves. They expressed enthusiasm for being in college and indicated an appreciation for the scientific method and further growth in understanding of science. By contrast, the amount of growth in RCI scores for freshmen was almost always negatively correlated with the educational college activities on the CSEQ, including almost all items having to do with seeking out experiences that were different from prior experiences, or seeking out others who were different from themselves. That is, those who relied on prereflective assumptions were less open to experiences that involved talking about different points of view or interacting with others who are different. These students simply may not seek out these experiences as frequently as students who enter with more advanced epistemological assumptions. Studies such as this that link types of collegiate experiences to patterns of college stu-



dent growth would be particularly helpful in advancing our understanding of the mechanisms of development in reflective thinking. Baxter Magolda's (1999, 2001) pedagogical framework for promoting development offers a promising conceptual tool for designing interventions to promote not only reflective thinking but also advanced capacities in identity and interpersonal domains (for examples in higher education contexts, see Baxter Magolda & King, in press).

Educators have often reported that they are puzzled by how students defend their beliefs—for example, why some reduce complex controversies to simple, black-and-white terms, and why others are so appreciative of the value of multiple perspectives that they are unable to make their own judgments. Our hope here is that educators can better interpret their observations about student behaviors by understanding how such behaviors are grounded in their epistemic assumptions, and how these assumptions about knowledge and how it is gained are related to the ways students justify their own judgments about controversial issues.

## REFERENCES

- American Association of Colleges and Universities (2002). *Greater expectations: A new vision for learning as a nation goes to college*. National Panel Report. Washington, DC: Author.
- American Association of Higher Education, American College Personnel Association, National Association of Student Personnel Administrators (1998). *Powerful partnerships*. Washington, DC: Author.
- American College Personnel Association (1994). *The student learning imperative*. Washington, DC: Author.
- Baxter Magolda, M. B. (1999). *Creating contexts for learning and self-authorship: Constructive-developmental pedagogy*. Nashville, TN: Vanderbilt University Press.
- Baxter Magolda, M. B. (2001). *Making their own way: Narratives for transforming higher education to promote self-development*. Sterling, VA: Stylus Publishing.
- Baxter Magolda, M. B. (in press). Self-authorship as the common goal of 21st century education. In M. B. Baxter & P. M. King (Eds.), *Learning partnerships: Educating for self-authorship*. Sterling, VA: Stylus Publications.
- Baxter Magolda, M. B., & King, P. M. (Eds.). (in press). *Learning partnerships: Educating for self-authorship*. Sterling, VA: Stylus Publications.
- Brabeck, M. (1983) Critical thinking skills and reflective judgment development: Redefining the aims of higher education. *Journal of Applied Developmental Psychology*, 4(1), 23–34.
- Brabeck, M., & Wood, P. K. (1990). Cross-sectional and longitudinal evidence for differences between well-structured and ill-structured problem solving abilities. In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, R. A. Grotzer, & J. D. Sinnott (Eds.), *Adult development 2: Models and methods in the study of adolescent and adult thought* (pp. 133–146). New York: Praeger.
- Broughton, M. (1975). *The development of natural epistemology in years 11 to 16*. Unpublished doctoral dissertation, Graduate School of Education, Harvard University, Cambridge, MA.
- Broughton, M. (1978). Development of concepts of self, mind, reality, and knowledge. In W. Damon (Ed.), *Social cognition. New directions for child development* (Vol. 1, pp. 75–100). San Francisco: Jossey-Bass.
- Churchman, C. W. (1971). *The design of inquiring systems: Basic concepts of systems and organization*. New York: Basic Books.
- Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill.
- Dewey, J. (1933). *How we think: A restatement of the relations of reflective thinking to the educative process*. Lexington, MA: Heath.
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477–531.
- Fischer, K. W., Bullock, D., Rosenberg, E. J., & Raya, P. (1993). The dynamics of competence: How context contributes directly to skill. In R. Wozniak & K. W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments* (pp. 93–117). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Fischer, K. W., & Lamborn, S. D. (1989). Mechanisms of variation developmental levels: Cognitive and emotional transitions during adolescence. In A. de Ribaupierre (Ed.), *Transitions mechanisms in child development: The longitudinal perspective* (pp. 33–670). Cambridge, England: Cambridge University Press.
- Fischer, K. W., & Pipp, S. L. (1984). Processes of cognitive development: Optimal level and skill acquisition. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 45–80). New York: Freeman.
- Fischer, K. W., & Pruyne, E. (2002). Reflective thinking in adulthood: Emergence, development, and variation. In J. Demick & C. Andreoletti (Eds.), *Handbook of adult development* (pp. 169–198). New York: Plenum.
- Fischer, K. W., & Rose, S. P. (1994). Dynamic development of coordination of components of brain and behavior. In G. Dawson & K. W. Fischer (Eds.), *Human behavior and the developing brain* (pp. 2–66). New York: Guilford.
- Flavell, J. (1971). Stage related properties of cognitive development. *Cognitive Psychology*, 2, 421–453.
- Guthrie, V. L., King, P. M., & Palmer, C. J. (2000). Cognitive capabilities underlying tolerance for diversity among college students. Manuscript submitted for publication.
- Harvey, L. J., Hunt, D., & Schroder, H. M. (1961). *Conceptual systems and personality organization*. New York: Wiley.
- Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88–140.
- Hofer, B. K., & Pintrich, P. R. (2002). Personal epistemology: The psychology of beliefs about knowledge and knowing. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kegan, R. (1982). *The evolving self*. Cambridge, MA: Harvard University Press.
- Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Cambridge, MA: Harvard University Press.
- King, P. M. (1985). Choice-making in young adulthood: A developmental double-bind. *Counseling and Human Development*, 18(3), 1–12.
- King, P. M. (1992). "How do we know? Why do we believe?" Learning to make reflective judgments. *Liberal Education*, 78(1), 2–9.
- King, P. M. (2000). Learning to make reflective judgments. In M. B. Baxter Magolda (Ed.), *Linking student development, learning, and teaching: New directions for teaching and learning* (Vol. 82, pp. 15–26). San Francisco: Jossey-Bass.
- King, P. M., & Kitchener, K. S. (Eds.). (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.
- King, P. M., & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 37–61). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- King, P. M., Kitchener, K. S., & Wood, P. K. (1994). Research on the reflective judgment model. *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.



- King, P. M., Kitchener, K. S., Wood, P. K., & Davison, M. L. (1989). Relationships across developmental domains: A longitudinal study of intellectual, moral and ego development. In M. L. Commons, J. D. Sinnott, F. A. Richards, & C. Armon (Eds.), *Adult development: Vol. 1. Comparisons and applications of developmental models* (pp. 57–72). New York: Praeger.
- King, P. M., & Mayhew, M. J. (2002). Moral judgement development in higher education: Insights from the Defining Issues Test. *Journal of Moral Education, 33*, 247–270.
- King, P. M., & Mayhew, M. J. (in press). Theory and research on the development of moral reasoning among college students. *Handbook of Higher Education*.
- Kitchener, K. S. (1983). Cognition, metacognition and epistemic cognition: A three-level model of cognitive processing. *Human Development, 4*, 222–232.
- Kitchener, K. S. (2002). Skills, tasks, and definitions: Discrepancies in the understanding and data on the development of folk epistemology. *New Ideas in Psychology, 20*, 309–328.
- Kitchener, K. S., & Fischer, K. (1990). A skill approach to the development of reflective thinking. In D. Kuhn (Ed.), *Contributions to human development: Developmental perspectives on teaching and learning thinking skills* (Vol. 21, pp. 48–62). Basel, Switzerland: Karger.
- Kitchener, K. S., & King, P. M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology, 2*(2), 89–116.
- Kitchener, K. S., & King, P. M. (1985). *The reflective judgment scoring manual*. Unpublished manuscript.
- Kitchener, K. S., Lynch, C. L., Fischer, K. W., & Wood, P. K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. *Developmental Psychology, 29*(5), 893–906.
- Kitchener, K. S., Wood, P. K., & Jensen, L. (2003). *Individual differences in gains in reflective judgment and their relationship to college experiences*. University of Denver, CO.
- Kitchener, R. (2002). Folk epistemology: An introduction. *New Ideas in Psychology, 20*(2/3), 89–105.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–450). Chicago: Rand McNally.
- Kohlberg, L. (1984). *The psychology of moral development: The nature and validation of moral stages*. San Francisco: Harper & Row.
- Kohlberg, L. (1991). My personal search for universal morality. In L. Kuhmerker (Ed.), *The Kohlberg legacy for the helping professions*. Birmingham, AL: R.E.P. Books.
- Kroll, B. (1992). *Teaching hearts and minds: College students reflect on the Vietnam War through literature*. Carbondale, IL: Southern Illinois University Press.
- Lamborn, S. D., & Fischer, K. W. (1988). Optimal and functional levels in cognitive development: The individual's developmental range. *Newsletter of the International Society for the Study of Behavioral Development, 2*(Serial No. 14), 1–4.
- Loevinger, J. (1976). *Ego development: Conceptions and theories*. San Francisco: Jossey-Bass.
- Pace, R. (1990). *College Student Experience Questionnaire: Norms for the third edition*. Bloomington, IN: Center for Postsecondary Research and Planning, Indiana University.
- Perry, W. (1968). *Patterns of development in thought and values of students in a liberal arts college: A validation of a scheme* (Final Report, Project No. 5-0825, Contract No. SAE-8973). Washington, DC: Department of Health, Education, and Welfare.
- Perry, W. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. Troy, MO: Holt, Rinehart & Winston.
- Piaget, J. (1965). *The moral judgment of the child*. New York: Free Press.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Polkosnik, M. C., & Winston, R. B. (1989). Relationships between students' intellectual and psychological development: An exploratory investigation. *Journal of College Student Development, 30*(1), 10–19.
- Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis: University of Minnesota Press.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.
- Rest, J. R., Narvaez, D., Bebeau, M., & Thoma, S. (1999). *Postconventional moral thinking: A neo-Kohlbergian approach*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Sakalys, J. (1984). Effects of an undergraduate research course on cognitive development. *Nursing Research, 33*, 290–295.
- Schmidt, J. A. (1985). Older and wise? A longitudinal study of the impact of college on intellectual development. *Journal of College Student Personnel, 26*, 388–394.
- Van Tine, N. B. (1990). The development of reflective judgment in adolescents. *Dissertation Abstracts International, 51*, 2659.
- Welfel, E. R., & Davison, M. L. (1986). How students make judgments: Do educational level and academic major make a difference? *Journal of College Student Personnel, 23*, 490–497.
- Wood, P. K. (1983). Inquiring systems and problem structure: Implications for cognitive development. *Human Development, 26*, 249–265.
- Wood, P. K. (1993). *Context and development of reflective thinking: A secondary analysis of the structure of individual differences*. Unpublished manuscript, University of Missouri at Columbia.
- Wood, P. K. (1997). A secondary analysis of claims regarding the Reflective Judgment Interview: Internal consistency, sequentiality and intra-individual differences in ill-structured problem solving. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 12, pp. 243–312). Edison, NY: Agathon Press.
- Wood, P. K., Kitchener, K. S., & Jensen, L. (2002). Considerations in the design and evaluation of a paper-and-pencil measure of epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 277–294). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Wood, P. K., Kitchener, K. S., & Jensen, L. (2003). College students' concepts of belief justification regarding open-ended controversies. Manuscript submitted for publication.
- Wood, P. K., & Lynch, C. L. (1998). Using guided essays to assess and encourage reflective thinking. *Assessment Update, 10*(2), 14–15.