Ethical decision making measures are widely applied as the principal dependent variable used in studies of research integrity. However, evidence bearing on the internal and external validity of these measures is not available. In this study, ethical decision making measures were administered to 102 graduate students in the biological, health, and social sciences, along with measures examining exposure to ethical breaches and the severity of punishments recommended. The ethical decision making measure was found to be related to exposure to ethical events and the severity of punishments awarded. The implications of these findings for the application of ethical decision making measures are discussed.

Keywords: integrity, ethical decision making, scientists, measure, validation

Integrity is fundamental to the success of the research enterprise (Steneck, 2004). Psychologists rely on investigators to be honest in the conduct of a study, in their analysis of the data, and in their business and professional relationships (Helton-Fauth et al., 2003). Unfortunately, recent events ranging from outright fabrication to the death of people participating in studies (Kochan & Budd, 1992; Marshall, 1996) have led to a new set of questions about the integrity of
the research enterprise (Kalichman & Friedman, 1992; Martinson, Anderson, & de Vries, 2005). Although those questions have prompted a new wave of research into the origins of scientific integrity, it has proven difficult to directly assess actual integrity related behavior.

Instead, studies of research integrity have progressed using measures intended to capture more distal manifestations of ethical behavior on the part of scientists. For example, Jasanoff (1993) examined historic conditions surrounding incidents of past misconduct. Martinson et al. (2005) asked investigators to report exposure to unethical behaviors. Although few scholars would debate the value of these studies, research on ethical behavior in the sciences, like studies of ethics in other fields (Loe, Ferrel, & Mansfield, 2000), has often employed ethical decision making scales as a primary measure of integrity. Our intent in this study is to provide some evidence bearing on the validity of ethical decision making measures as they are applied in studies examining the integrity of the research enterprise.

**ETHICAL DECISION MAKING MEASURES**

The logic underlying the application of ethical decision making measures in studies of integrity (Beu, Buckley, & Harvey, 2003), including studies of research integrity (e.g., Rest, Narváez, Bebeau, & Thoma, 1999), is quite straightforward. Essentially, it is held that ethical behavior in professions arises from the decisions people make concerning the application of relevant expertise. Although the outcomes resulting from these decisions for others, including the broader institution, may be shaped by a variety of situational variables, integrity, as it is reflected in individual-level behavior, is thought to arise from the decisions people make when they are presented with a complex situation in which ethical considerations are salient. Individuals evidencing an ethical predisposition are held to be more likely to make ethical decisions in these situations. Thus, ethical decision making measures can be viewed as a low-fidelity simulation exercise (Motowidlo, Dunnette, & Carter, 1990).

Although the application of this logic can be questioned (Miner & Petocz, 2003; Mumford & Helton, 2002; Reall, Bailey, & Stoll, 1998), ethical decision making measures are now widely applied in studies of integrity. For example, Trevino and Youngblood (1990) presented two ethical decision making scenarios to people. Responses to these scenarios were used to develop a model of individual-level characteristics, for example locus of control, contributing to ethical decisions. Sims and Keon (1999) presented ethical decision making scenarios to business graduate students and found that current work climate was a noteworthy influence on subsequent decisions. In still another study along these lines, Eastman, Eastman, and Tolson (2001) developed four ethical decision making scenarios applicable in a health maintenance organization. They found that doctors were more likely to make decisions that benefited the patient, rather than themselves.
The widespread application of ethical decision making measures in studies of integrity (Loe et al., 2000; Vitell & Ho, 1997), however, broaches a critical question. Is there reason to suspect that these measures are, in fact, valid (Messick, 1989, 1998)? Validation of measures of ethical decision making is typically based on a judge’s appraisal of content validity. More specifically, does the decision scenario call for an ethical decision? Although there is value in this kind of content validity evidence, simple judgmental appraisals of scenario relevance cannot provide a fully adequate assessment of the validity of these scenarios.

To begin, decision making scenarios vary in the number and nature of the response options presented. As a result, these response options may give rise to different decisions. And, it is possible, at least at times, that use of certain response formats may influence the ability of these measures to capture ethical decision making. For example, in some cases the responses may be so obvious that the resulting measure may reflect socially desirable responding rather than ethical decision making. Moreover, when more complex and realistic response options are presented, these responses might be scored in a number of different ways—some of which do, and some of which do not, reflect ethical decision making. These issues, bearing on the substantive meaningfulness of responses to ethical decision making measures, are further complicated by the fact that a given scenario may, or may not, activate ethical decision making processes in respondents.

VALIDATION STRATEGIES

These observations bearing on the ambiguity of appraisals of validity based on judges’ assessments of content relevance, of course, bring to the fore a new question. How might one provide evidence for the construct validity of a proposed measure of ethical decision making? Although a variety of strategies might be used to accrue evidence for the validity of ethical decision making measures (Messick, 1995), four strategies appear to be especially promising: (a) convergent and divergent validity evidence obtained through the measures’ correlations with other measures, (b) convergence and divergence with regard to different substantive frameworks for scoring response options, (c) the correlation of ethical decision making measures with expected causes of ethical decisions, and (d) the correlation of ethical decision making measures with expected outcomes of ethical decisions.

Perhaps the most common strategy used to provide evidence for the construct validity of ethical decision making measures is examination of the measure’s correlation with other individual-level constructs. For example, Mumford, Gessner, Connelly, O’Connor, and Clifton (1993) developed a measure of ethical decision making using a variation on an in-basket task, in which people were asked whether they would or would not follow a decision recommendation in which the decisions were structured such that they might result in harm to others or harm to institu-
tions. The positive correlations observed between scores on this measure and measures of other psychological constructs held to be linked to unethical behavior, such as narcissism and power motives, provided requisite construct validity evidence. Further evidence bearing on the construct validity of this measure was provided by showing that responses to the measure were not related to social desirability but were related to situational cues inducing ethical processing. Similar strategies for validating ethical decision making measures have been applied by Paolillo and Vitell (2002); Roosen, Pelsmacker, and Bostyn (2001); and Trevino and Youngblood (1990).

Another strategy that has been used to provide construct validity evidence has focused on the substantive implications of the response options. In one such study, Wittmer (2000) scored a measure of managerial ethical decision making for both the appropriateness of decisions and a ranking of the ethical issues involved in these decisions. He provided evidence for the validity of the ethical decision making measure by showing that decisions flowing from this measure were positively related to rankings of relevant ethical issues. Of course, a number of approaches along those lines might be used to provide construct validation evidence. For example, one might demonstrate that ethical decisions were negatively related to scoring systems examining relevant social psychological attributes of responses such as aggression and deception (Murphy, 1993). Alternatively, one might show that ethical decisions were positively related to the application of cognitive strategies, such as assessing downstream consequences for others, held to contribute to integrity (Mumford & Helton, 2002).

Still another way one might examine the validity of an ethical decision making measure is to appraise the correlation between scores on this measure and hypothesized causes of ethical decisions. One demonstration of this approach may be found in Sims and Keon (1999). In this study, examining business students employed full time, it was argued that work climate would be used to provide guidelines for responding to ethical dilemmas. The resulting positive correlation between climate perceptions and ethical decision making, in turn, provided evidence for the construct validity of this measure. Other work by Brown and Kalichman (1998) and Ulrich, Soeken, and Miller (2003) indicated that ethical training and relevant career experiences might represent noteworthy influences on ethical decisions.

In contrast to examining the causes of ethical decision making, one might, alternatively, seek to provide validation evidence by examining likely outcomes of ethical decision making. Although, at first glance, examining the relationship between ethical decision making and subsequent ethical behavior is attractive, such studies may be undermined by the low frequency of ethical violations. An alternative strategy, suggested by Sackett and Waneck (1996), is to examine less extreme consequences of ethical decision making. Thus, the penalties awarded for perceived violations or the affective intensity associated with exposure to unethical events
illustrates the kind of measures that might prove useful in validating measures of ethical decision making.

With these observations in mind, our intent in this study is to provide some construct validation evidence for a new measure of ethical decision making. In the first part of this study, correlations between the ethical decision making measure and standard individual differences measures are assessed. Next, the correlations between the ethical decision making measure and alternative systems for scoring response options in terms of social psychological variables and cognitive strategies are examined. In the final phase of this investigation, correlations and regressions are used to examine how ethical decision making was related to (a) perceived work environments—a putative cause of ethical decision making, and (b) the severity of punishments awarded—a consequence of ethical decision making.

**METHOD**

**Sample**

The sample used in this study was composed of 102 first-year doctoral students attending a large university in the Southwest. The 40 men and 60 women (2 did not designate) who agreed to participate in this study were recruited no earlier than 4 months after beginning work at the university and no later than 9 months after beginning work at the university. Sample members were recruited from programs awarding doctoral degrees in the biological (28%), health (28%), and social (44%) sciences where independent research work was required for award of a terminal degree. On average, sample members were 27 years old. They had typically completed 17 years of schooling prior to admission into their doctoral program. Although 59% of the sample was employed in nonresearch positions, 41% were employed as full-time research assistants. All sample members, however, were actively involved in research in one of the university’s laboratories.

**General Procedures**

This study was conducted as part of a larger, federally funded initiative concerned with research integrity. The university provided names, department assignments, e-mail addresses, and telephone numbers of all doctoral students attending the university in 2005. A three-stage recruitment system was used to encourage students to volunteer for this study. First, flyers announcing the study, and that $100 would be paid for participation, were placed in the university mailboxes of first-year students. Second, first-year students were called to solicit participation. Finally, e-mails were sent, up to four, recruiting students to participate in this study.
In all recruitment attempts it was noted that the study was an investigation into research integrity. The study was described as examining the effects of the students’ educational experiences on integrity and problem solving. If a student agreed to participate in the study, he or she was asked to schedule a time to complete a 4-hr battery of paper-and-pencil measures. When the students had completed the battery of measures, they were again informed as to the nature of the study and provided a debriefing document.

After students had completed the informed consent document, they were asked to start work on the battery of measures. Initially, students were asked to complete a background information form and a career experiences inventory in which they were asked to describe events that had happened to them at the university. At this time, they were also asked to review an institutional review board case and assign responsibility and sanctions for the events reported in this case. After this material had been completed, participants were asked to complete a battery of standard psychological tests examining ability and personality characteristics that might be linked to integrity. In the final component of this battery, participants were asked to complete an ethical decision making measure tailored to issues they might encounter in their day-to-day work. It is of note that these field-specific ethical decision making measures were cast as work-oriented problem-solving measures to minimize potential demand characteristics.

Ethical Decision Making Measure

The ethical decision making measure provided the primary measure to be validated in this study. This ethical decision making measure was developed to reflect ethical issues people might encounter in their day-to-day work as researchers. Development of this decision making measure was based on the taxonomy of ethical behavior developed by Helton-Fauth et al. (2003). They reviewed codes of conduct for fields in the biological, health, and social sciences. This review led to the identification of 17 dimensions of ethical behavior subsumed under four areas: data management (including data massaging and publication practices), study conduct (including institutional review board, informed consent, confidentiality protection, protection of human participants, and protection of animal subjects), professional practices (including objectivity in evaluating work, recognition of expertise, protection of intellectual property, adherence to professional commitments, protection of public welfare and the environment, and exploitation of staff and/or collaborators), and business practices (including conflicts of interest, deceptive bid and contract practices, inappropriate use of physical resources, and inappropriate management practices). Helton-Fauth et al. (2003) have provided evidence for the validity of this taxonomy in accounting for incidents of scientific misconduct.

Web sites (e.g., on-line ethics center, Office of Research Integrity, The American Psychological Association) were reviewed to identify case studies that might
be used to assess ethical decision making with respect to one or more of the dimensions included in the Helton-Fauth et al. (2003) taxonomy. Overall, 45 cases were identified, in each of the three areas (biological, health, and social sciences) that might be used to assess ethical decision making with respect to one or more of these dimensions. The criteria applied in selecting cases for development of the decision making measures were (a) relevant to day-to-day work, (b) involvement of both technical and ethical issues, and (c) potentially challenging decisions across a range of expertise. Based on these criteria, a panel of three psychologists, assisted by subject matter experts, then selected 10 to 15 cases in each area (biological, health, and social sciences) that covered a range of different ethical and unethical events. Once these cases had been selected, a short one- or two-paragraph scenario was abstracted describing the general event and the key actors in this event.

After these scenarios had been written, development of the decision making measures to be applied in biological, health, or social sciences began. A panel, composed of three psychologists and a relevant subject matter expert, was asked to review each scenario and generate a list of 8 to 12 events that might plausibly occur in this situation as action unfolded. These events were to be developed under the constraint that half of the events formulated were to have only technical implications. The remaining half of the events were to have implications for one of the 17 dimensions of ethical conduct identified by Helton-Fauth et al. (2003). Panel members were then asked to review this list of events and identify two technical events and two ethical events that might plausibly occur in the scenario at hand, and organize these events into a plausible flow of action. It is of note that both technical and ethical events were developed for each scenario to minimize demand characteristics and embed integrity-related decisions in actual work.

Following generation of these events, the panel was asked to generate a list of potential actions that might be taken in response to a given event. For each event under a scenario, panel members were asked to generate 6 to 12 potential responses. In the case of technical events, these actions reflected alternative approaches for resolving the problem broached by the event. In the case of ethical events, one third of the response options were to reflect highly ethical responses to the event given the dimension under consideration, one third were to reflect moderately ethical responses to this event given the dimension under consideration, and one third were to reflect unethical responses given the dimension under consideration. The response options and scoring of these response options were reviewed by a panel of three psychologists familiar with the literature on ethical conduct, who suggested any necessary revisions in the response options developed for each event. It is of note that the response options generated for each event, and review of these response options, were based on the ethical principles presented in relevant professional codes of conduct.

On average, three events were formulated for each of the 17 ethical dimensions applying in a given field (e.g., biological, health, and social sciences)—with three
to five events embedded under a given scenario. After reading through each sce-
nario, participants were asked to assume the role of the primary actor in the sce-
nario. As they read through each event, they were asked to select the two response
options that they believed would most likely resolve the problem broached by an
event. To score responses, actions selected in responding to an event were given
scores of 3, 2, or 1, based on whether the action reflected responses of high, moderate,
or low integrity. The average score of all events subsumed under a dimension
provided the basis for generating dimensional scores. Dimensional scores were
then aggregated using the Helton-Fauth et al. (2003) taxonomy to obtain overall
ethical decision making scores bearing on data, study conduct, professional prac-
tices, and business practices.

When dimensional scores were aggregated into the general dimensions of data,
study conduct, professional practices, and business practices, the resulting average
split-half reliability estimate was .74. Although these scales evidenced the ex-
pected positive relationships, the magnitude of the obtained relationships provided
some evidence for the meaningfulness of the measures. For example, strong posi-
tive relationships were observed between data and professional practices \((r = .57)\)
and between study conduct and business practices \((r = .53)\), but weaker relation-
ships were observed between data and business practices \((r = .18)\) and data and
study conduct \((r = .22)\). The Appendix provides an illustration of the nature of
these ethical decision making measures.

**Alternative Scoring**

Responses to these ethical scenarios could, of course, be scored on other bases,
aside from ethical behavior. One way scoring of these responses might occur is
through analysis of the social psychological basis of behavior. Another way scor-
ing responses might occur is through analysis of the cognitive strategies giving rise
to a response. Development of these alternative scoring systems began with a re-
view of the literature to identify dimensions that might account for the selection of
response options within these two paradigms. In the case of the social psychologi-
cal variables, these dimensions included (a) involving others, (b) retaliation, (c)
deception, (d) active involvement, (e) avoidance of responsibility, (f) selfishness,
and (g) close self to future decision/actions (Anderson, 2003; Beu et al., 2003;
Darke & Chaiken, 2005; Kahneman, 2003; Knaus, 2000; Munro, Bore, & Powis,
2005; Schweitzer, DeChurch, & Gibson, 2005; Tyler, 2006). In the case of the cog-
nitive strategies, these dimensions included (a) recognition of circumstances, (b)
seeking help, (c) questioning one’s judgment, (d) anticipating consequences, (e)
dealing with emotions, (f) analysis of personal motivations, and (g) consideration
of effects of action on others (Butterfield, Treviño, & Weaver, 2000; Street,
Operational definitions of each of social psychological and cognitive strategy were formulated. Subsequently, a panel of four psychologists was familiarized with these operational definitions and the strategies to be applied in appraising response options. More specifically, in the case of the social psychological dimensions, judges were to focus on the behavioral outcomes of a response, whereas in the case of cognitive strategies, judges were to focus on the psychological processes giving rise to a response. This training took approximately 20 hr. After training in the application of a given system, judges were asked to read through a scenario, the events, and the associated response options, and then rate each response option on each dimension, applying a 7-point rating scale.

The average interrater agreement coefficient obtained for the social psychological dimensions was .83, whereas the interrater agreement coefficient obtained for the cognitive strategies dimensions was .91. Scores on those dimensions were obtained by taking the weighted average of the responses selected by participants. As expected, both the social psychological ($\bar{r} = .42$) and cognitive strategies ($\bar{r} = .47$) dimensions, the average correlation across these dimensions, and the four ethical dimensions produced positive correlations. However, the pattern of these correlations—for example, the strong positive relationship of the active involvement dimension with data management ($r = .34$) and the weaker relationship of close self to future action with study conduct ($r = -.07$)—provides some evidence for the meaningfulness of these social psychological ratings. Similarly, for the cognitive strategies dimensions, a strong positive relationship between dealing with emotions and data management ($r = .37$) and the weaker relationship between seeking help and business practices ($r = -.05$) point to the meaningfulness of these scales.

Individual Differences Measures

In addition to examining the validity of this ethical decision making measure with respect to alternative scoring systems, evidence bearing on the meaningfulness of the ethical decision making measure was also obtained by examining its relationship with standard measures of people’s differential characteristics. Table 1 provides a summary description of these measures. The first two measures participants were asked to complete included a verbal reasoning measure of intelligence (Ruch & Ruch, 1980) and a consequences measure of divergent thinking (Merrifield, Guilford, Christensen, & Frick, 1962). In addition to these two cognitive controls, participants were asked to complete Paulhus’s (1984) measure of socially desirable responding and John, Donahue, and Kentle’s (1991) Big Five measure of agreeableness, extraversion, conscientiousness, neuroticism, and openness to experience.

In addition to these general control measures, participants also completed a series of measures that could be expected to be more directly linked to integrity.
### TABLE 1
Description of Individual Difference Measures

<table>
<thead>
<tr>
<th>Construct</th>
<th>Items</th>
<th>Subscales</th>
<th>Nature of Items (or Example Items)</th>
<th>Reliability Estimates (α)</th>
<th>Validity Cites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligence</td>
<td>30</td>
<td>No</td>
<td>Conclusions based on a set of facts in a given scenario (6 sets of facts)</td>
<td>.80</td>
<td>Ruch &amp; Ruch (1980)</td>
</tr>
<tr>
<td>Divergent thinking</td>
<td>5</td>
<td>No</td>
<td>What would be the results if the force of gravity were suddenly cut in half?</td>
<td>.93</td>
<td>Merrifield, Guilford, Christensen, &amp; Frick (1962)</td>
</tr>
<tr>
<td>Social desirability</td>
<td>40</td>
<td>No</td>
<td>I always know why I like things.</td>
<td>.62</td>
<td>Paulhus (1994)</td>
</tr>
<tr>
<td>General personality</td>
<td>44</td>
<td>No</td>
<td>I always obey laws, even if I’m unlikely to get caught.</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>(Big Five)</td>
<td></td>
<td></td>
<td>I see myself as someone who can be moody.</td>
<td>.80</td>
<td>John, Donahue, &amp; Kentle (1991)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neuroticism</td>
<td>I see myself as someone who values artistic, aesthetic experiences.</td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Openness</td>
<td>I see myself as someone who values artistic, aesthetic experiences.</td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extraversion</td>
<td>I see myself as someone who is talkative.</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>I see myself as someone who does a thorough job.</td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agreeableness</td>
<td>I see myself as someone who has a forgiving nature.</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>Narcissism</td>
<td>37</td>
<td>No</td>
<td>I can make anybody believe anything.</td>
<td>.82</td>
<td>Raskin &amp; Hall (1979)</td>
</tr>
<tr>
<td>Cynicism</td>
<td>10</td>
<td>No</td>
<td>People pretend to care more about one another than they really do.</td>
<td>.79</td>
<td>Wrightsman (1974)</td>
</tr>
<tr>
<td>Trust</td>
<td>10</td>
<td>No</td>
<td>The typical person is sincerely concerned about the problems of others.</td>
<td>.81</td>
<td>Wrightsman (1974)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>20</td>
<td>No</td>
<td>I work under a great deal of tension.</td>
<td>.73</td>
<td>Taylor (1953)</td>
</tr>
<tr>
<td>Philosophies of human</td>
<td>8</td>
<td>Variability</td>
<td>Different people react to the same situation in different ways.</td>
<td>.78</td>
<td>Wrightsman (1974)</td>
</tr>
<tr>
<td>nature</td>
<td>6</td>
<td>Complexity</td>
<td>People are too complex to ever be understood fully.</td>
<td>.78</td>
<td></td>
</tr>
</tbody>
</table>
Based on the findings of Mumford et al. (1993) indicating that narcissism is negatively related to integrity, participants were asked to complete Emmon’s (1987) measure of narcissism. In addition to narcissism, participants were also asked to complete measures of cynicism and trust (Wrightsman, 1974), based on the observation that cynicism and trust influence people’s willingness to engage in self-protective behaviors, including behaviors linked to a lack of integrity. Following Fromm’s (1973) observation that uncertainty is often linked to a lack of integrity as people take actions, often unethical actions, to reduce feelings of uncertainty, participants were also asked to complete Taylor’s (1953) manifest anxiety scale. Finally, participants were asked to complete a measure drawn from Wrightsman (1974) examining complexity and variability in people’s philosophies of human nature.

Causes of Ethical Decision Making

Among entry-level doctoral students, it can be argued that the laboratory practices they are exposed to in their day-to-day work will have a marked impact on their ethical decision making. To examine the impact of early career experiences on ethical decision making, a measure of laboratory work experiences was developed. Initial development of the laboratory work experiences measure began with a review of professional codes of conduct applying to fields in the biological, health, and social sciences. The review was used to identify specific work events associated with ethical practices in each professional field (e.g., selectively replace participants when gathering data or maintain anonymity of manuscript reviews). In all, some 240 ethical events were identified in the literature review.

Subsequently, three psychologists and a subject matter expert were asked to review these events. Events that would occur outside of the research context—for example, in clinical practice—were eliminated. Moreover, redundant events were collapsed. The remaining events were then assessed for clarity and relevance in the four general domains of ethical decision making under consideration: (a) data management, (b) study conduct, (c) professional practices, and (d) business practices. Based on this review, 52 ethical work events were identified—8 events in data management, 25 events in study conduct, 19 events in professional practices, and 8 events in business practices—all of which first-year doctoral students might be exposed to in the course of their day-to-day work. In addition, panel members generated a list of 50 work events applying to people’s day-to-day work that reflected purely technical decisions or actions (e.g., establish procedures for entry of scanned data or check citations of manuscripts against a reference list). These technical events were included in the final event list to minimize demand characteristics. All events included in this list, both ethical and technical events, were constructed so as to be relevant to the work conducted across a range of disciplines. Table 2 provides a description of the ethical and technical work events.
Participants were presented with this list of events where the events were presented in random order. After reading the description of an event, participants were asked to rate, on a scale from 1 (low) to 7 (high), the frequency with which they had been exposed to the event in the course of their day-to-day work. When event ratings were summed to obtain scores for exposure to data, study conduct, professional practices, and business practices, the reliability coefficients obtained for these scales were .84, .88, .87, and .66, respectively. Examination of the correlations among these evaluations of event frequency also provided evidence for the construct validity, or meaningfulness, of the frequency of exposure scales. Thus, data was found to be strongly positively related \((r = .42)\) to study conduct, but less strongly positively related \((r = .28)\) to business practices.

### Outcomes of Ethical Decision Making

The ethical decision making orientation of people has a noteworthy impact on a number of behaviors. However, one likely outcome of people’s ethical decision making orientation is the ability to make informed decisions that align with ethical standards. This section will outline the outcomes of ethical decision making, focusing on the impact on data management, study conduct, professional practices, and business practices.

#### Table 2: Examples of Work Event Questions

<table>
<thead>
<tr>
<th>Ethical Events</th>
<th>Technical Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selectively replace participants or rerun analyses when gathering data</td>
<td>1. Establish procedures for entry of scanned data</td>
</tr>
<tr>
<td>2. Publish in “least publishable units” (i.e., write up results as several</td>
<td>2. Release data or study findings prior to completion of full data analysis</td>
</tr>
<tr>
<td>small articles vs. one or two larger ones)</td>
<td></td>
</tr>
<tr>
<td>3. Discuss data that personally identifies human participants with others not</td>
<td>3. Develop a consistent procedure to assign subject codes in a data set</td>
</tr>
<tr>
<td>involved in the research</td>
<td></td>
</tr>
<tr>
<td>4. Overstate benefits of participation in a study to human participants</td>
<td>4. Provide sufficient instructions to participants</td>
</tr>
<tr>
<td>5. Conduct research without institutional review board (IRB) approval</td>
<td>5. Provide clear application forms to IRB</td>
</tr>
<tr>
<td>6. Include a participant whose health status does not quite meet standards in</td>
<td>6. Answer questions that participants may have during an experiment</td>
</tr>
<tr>
<td>a study considered to be low risk</td>
<td></td>
</tr>
<tr>
<td>7. Maintain anonymity of manuscript reviewers</td>
<td>7. Provide clear and specific comments in reviews of manuscripts of drafts or</td>
</tr>
<tr>
<td>8. Suggest what the acceptable results of a study should be to research</td>
<td>articles</td>
</tr>
<tr>
<td>assistants</td>
<td></td>
</tr>
<tr>
<td>9. Accept payment for serving as an expert witness in judicial proceedings</td>
<td>9. Inform study assistants of preliminary experimental hypotheses</td>
</tr>
<tr>
<td>10. Accurately describe program content, goals, or costs associated with a</td>
<td>10. Ask for deadline extensions on deliverables for grants and contracts</td>
</tr>
<tr>
<td>research proposal</td>
<td></td>
</tr>
</tbody>
</table>

Participants were presented with this list of events where the events were presented in random order. After reading the description of an event, participants were asked to rate, on a scale from 1 (low) to 7 (high), the frequency with which they had been exposed to the event in the course of their day-to-day work. When event ratings were summed to obtain scores for exposure to data, study conduct, professional practices, and business practices, the reliability coefficients obtained for these scales were .84, .88, .87, and .66, respectively. Examination of the correlations among these evaluations of event frequency also provided evidence for the construct validity, or meaningfulness, of the frequency of exposure scales. Thus, data was found to be strongly positively related \((r = .42)\) to study conduct, but less strongly positively related \((r = .28)\) to business practices.
making includes their perceptions of and reactions to ethical misconduct. Accordingly, participants were asked to assume the role of a member of the university’s ethical review panel. On this task, participants were presented with eight incidents of misconduct and were asked to judge the misconduct incident.

Development of this measure began with a review of online ethics sites to identify ethical events or cases of high, moderate, and low impact, as reflected in the consequences of the events for others. Three judges, all psychologists, rated case impact, producing an interrater agreement coefficient above .80. For the cases drawn from each field (biological, health, and social sciences), two low, three moderate, and two high impact cases were selected based on the following criteria: (a) only one ethical event could be present in a case, (b) the case was focused on a research integrity issue, (c) the case was realistic, and (d) the cases under consideration reflected different dimensions (e.g., data, study conduct) of ethical behavior.

After these cases had been identified, the judges, along with an experienced institutional review board member, summarized each case in a one-paragraph scenario. These case summaries described the action and context in which the event occurred, as well as the actual ethical event. Participants were asked to assume the role of a member of a review panel that had been asked to screen cases before they proceeded to full review. After reading through the one-paragraph summary, participants were asked to rate each scenario on a scale from 1 (low) to 7 (high) with respect to how serious the violation was, how often they thought such violations occurred, how important it was to punish such violations, and the severity of punishment they would recommend if the allegations proved true. Table 3 provides an illustration of this task.

The reliability of these ratings proved to be adequate. The internal consistency coefficient obtained for ratings of severity, frequency, importance, and punishment were .72, .77, .74, and .74, respectively. More centrally, examination of the correlations among these rating scales provided some evidence for their construct validity. The ratings of severity were found to be positively related to punishment recommendations ($r = .86$). The mean ratings of severity and punishment recommendations, however, were found to be less strongly related ($\bar{r} = −.09$) to perceptions of frequency.

Analyses

The analyses to provide evidence for the validity of the ethical decision making measure began by correlating the scores on the ethical decision making measure with individual differences variables and the scores on the social psychological and cognitive strategies dimensions. Subsequently, scores on the ethical decision making measures were treated as criteria measures and correlated with, and regressed on, the measure of day-to-day exposure to ethical events. Finally, scores on the review panel task were treated as a criterion measure and correlated with, and
Johnson is investigating the effects of intraventricular AVP infusions on parental investment in deer mice. He is expected to provide fresh food and water at least once daily and attend to the general comfort and health of the postoperative animals. During routine inspections, laboratory animal care personnel have found the deer mice without water, food, or both on two consecutive weekends. On the second weekend one animal was found dead. The director of the animal facility has handed the matter over to your review board.

A complaint has been filed against Smith, accusing her of having used the privileged unpublished work of others as the basis of a National Institutes of Aging (NIA) grant application. Her proposal to NIA was built around a novel hypothesis together with supportive data bearing on the relationship between circulating glucocorticoids and arterial elastin breakdown. The content of the proposal is claimed to be virtually identical to that of a privileged unpublished manuscript in Smith’s possession.

Deborah has been accused of misconduct arising out of research supported by National Institute of Mental Health grants. Deborah was responsible for administering and scoring neuropsychological, neurological, and cognitive tests on patients during the course of two studies. According to the complaint, Deborah failed to conduct the required tests on 3 patients in the first study and on 10 to 12 patients in the second study. Deborah is alleged to have fabricated the experimental records for those tests. The data in question were included in a publication, and will be used in widespread treatment of schizophrenic patients in 14 states in the next 6 months. If the questionable data were eliminated, no significant results would be found.

### TABLE 3
Review Panel Task

<table>
<thead>
<tr>
<th>Biological</th>
<th>Health</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johnson is investigating the effects of intraventricular AVP infusions on parental investment in deer mice. He is expected to provide fresh food and water at least once daily and attend to the general comfort and health of the postoperative animals. During routine inspections, laboratory animal care personnel have found the deer mice without water, food, or both on two consecutive weekends. On the second weekend one animal was found dead. The director of the animal facility has handed the matter over to your review board.</td>
<td>A complaint has been filed against Smith, accusing her of having used the privileged unpublished work of others as the basis of a National Institutes of Aging (NIA) grant application. Her proposal to NIA was built around a novel hypothesis together with supportive data bearing on the relationship between circulating glucocorticoids and arterial elastin breakdown. The content of the proposal is claimed to be virtually identical to that of a privileged unpublished manuscript in Smith’s possession.</td>
<td>Deborah has been accused of misconduct arising out of research supported by National Institute of Mental Health grants. Deborah was responsible for administering and scoring neuropsychological, neurological, and cognitive tests on patients during the course of two studies. According to the complaint, Deborah failed to conduct the required tests on 3 patients in the first study and on 10 to 12 patients in the second study. Deborah is alleged to have fabricated the experimental records for those tests. The data in question were included in a publication, and will be used in widespread treatment of schizophrenic patients in 14 states in the next 6 months. If the questionable data were eliminated, no significant results would be found.</td>
</tr>
</tbody>
</table>

### Rating Scales

<table>
<thead>
<tr>
<th>Severity of Violation</th>
<th>Perceived Frequency of Violation</th>
<th>Perceived Importance of Punishment</th>
<th>Severity of Punishment Recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>How serious is this alleged violation in your opinion?</td>
<td>How often do you think such violations occur?</td>
<td>How important do you think it is to punish such violations?</td>
<td>What degree of punishment would you recommend in this situation if the allegations are true?</td>
</tr>
<tr>
<td>8-point Likert-type scale ranging from 0 (not serious at all) to 7 (extremely serious)</td>
<td>8-point Likert-type scale ranging from 0 (not at all important) to 7 (extremely important)</td>
<td>8-point Likert-type scale ranging from 0 (never) to 7 (extremely frequently)</td>
<td>8-point Likert-type scale ranging from 0 (no punishment at all) to 7 (most severe punishment possible)</td>
</tr>
</tbody>
</table>
RESULTS

Individual Differences

Table 4 presents the correlations of the ethical decision making measures with the individual differences variables. As may be seen, neither of the social desirability subscales (self-deceptive enhancement and impression management) produced significant correlations with these measures of ethical decision making. Thus, scores on these measures did not appear to be influenced by social desirability. Along related lines, although intelligence and divergent thinking were found to be positively related to ethical decisions with regard to business practices ($r = .31$), intelligence and divergent thinking were not strongly related to scores on the other dimensions of ethical decision making. Thus cognitive abilities, like social desirability, are not marked influences on ethical decision making, at least within a doctoral level population.

Ethical decision making has often been linked to personality or dispositional variables (Mumford & Helton, 2002; Murphy, 1993). The results obtained in this study, however, underscore a noteworthy caveat in this regard. More specifically, it

<table>
<thead>
<tr>
<th>Measure</th>
<th>Data</th>
<th>Study Conduct</th>
<th>Professional Practices</th>
<th>Business Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligence</td>
<td>.17</td>
<td>.17</td>
<td>.16</td>
<td>.25*</td>
</tr>
<tr>
<td>Divergent thinking</td>
<td>.04</td>
<td>.15</td>
<td>.11</td>
<td>.37*</td>
</tr>
<tr>
<td>Self-deceptive enhancement</td>
<td>.11</td>
<td>-.04</td>
<td>.05</td>
<td>-.14</td>
</tr>
<tr>
<td>Impression management</td>
<td>.03</td>
<td>-.04</td>
<td>.05</td>
<td>-.11</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>.08</td>
<td>.10</td>
<td>.19*</td>
<td>.02</td>
</tr>
<tr>
<td>Extraversion</td>
<td>.02</td>
<td>-.10</td>
<td>.05</td>
<td>-.18</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.03</td>
<td>.06</td>
<td>.13</td>
<td>.16</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>-.07</td>
<td>.07</td>
<td>-.10</td>
<td>.14</td>
</tr>
<tr>
<td>Openness</td>
<td>.13</td>
<td>.10</td>
<td>.17</td>
<td>.05</td>
</tr>
<tr>
<td>Narcissism</td>
<td>-.07</td>
<td>-.01</td>
<td>-.22*</td>
<td>-.08</td>
</tr>
<tr>
<td>Cynicism</td>
<td>-.26*</td>
<td>-.21*</td>
<td>-.29*</td>
<td>-.26*</td>
</tr>
<tr>
<td>Trust</td>
<td>.14</td>
<td>.03</td>
<td>.25*</td>
<td>.26*</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-.10</td>
<td>-.08</td>
<td>-.12</td>
<td>.01</td>
</tr>
<tr>
<td>Complexity of philosophies of human nature</td>
<td>.06</td>
<td>.01</td>
<td>.03</td>
<td>.13</td>
</tr>
<tr>
<td>Variability of philosophies of human nature</td>
<td>-.06</td>
<td>.01</td>
<td>.07</td>
<td>.07</td>
</tr>
</tbody>
</table>

* $p < .05$. 

Downloaded By: [Antes, Alison] At: 16:54 4 August 2009
appears that ethical decision making is not strongly linked to general dispositional characteristics such as agreeableness, extraversion, conscientiousness, neuroticism, openness, anxiety, variability, and complexity. Instead, ethical decision making was found to be more strongly (although still modestly) related to personality variables involving the assumptions people make about themselves and others. Thus, cynicism was found to be negatively related ($r = -0.25$) to all four types of ethical decisions, and trust was found to be positively related ($r = 0.26$) to ethical decisions, with regard to professional practices and business practices. Apparently, cynical individuals who do not trust others are especially prone to make unethical decisions, perhaps as a self-protection strategy. Moreover, in keeping with the observations of Mumford and Helton (2002), narcissism, an overinflated sense of self-worth, was found to be negatively related ($r = -0.22$) to ethical decisions with regard to professional practices.

**Alternative Scoring**

As noted earlier, ethical decision making measures, by virtue of their complexity, can often be scored in a number of ways. The correlations observed among these alternative scoring systems, of course, also provide some evidence bearing on the construct validity of these measures. Table 5 presents the correlations of the ethical decision making measures with alternative scoring systems.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Study Conduct</th>
<th>Professional Practices</th>
<th>Business Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social psychological</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Involving others</td>
<td>0.20*</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>2. Retaliation</td>
<td>-0.22*</td>
<td>-0.29*</td>
<td>-0.25*</td>
</tr>
<tr>
<td>3. Deception</td>
<td>-0.47*</td>
<td>-0.45*</td>
<td>-0.47*</td>
</tr>
<tr>
<td>4. Active involvement</td>
<td>0.34*</td>
<td>0.16</td>
<td>0.24*</td>
</tr>
<tr>
<td>5. Avoidance of responsibility</td>
<td>-0.49*</td>
<td>-0.49*</td>
<td>-0.51*</td>
</tr>
<tr>
<td>6. Selfishness</td>
<td>-0.50*</td>
<td>-0.51*</td>
<td>-0.51*</td>
</tr>
<tr>
<td>7. Close self to future action</td>
<td>-0.24*</td>
<td>-0.07</td>
<td>-0.21*</td>
</tr>
<tr>
<td>Cognitive strategies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Recognition of circumstances</td>
<td>0.46*</td>
<td>0.52*</td>
<td>0.52*</td>
</tr>
<tr>
<td>2. Seeking help</td>
<td>0.25*</td>
<td>0.12</td>
<td>0.18</td>
</tr>
<tr>
<td>3. Questioning one’s judgment</td>
<td>0.45*</td>
<td>0.39*</td>
<td>0.50*</td>
</tr>
<tr>
<td>4. Anticipating consequences</td>
<td>0.41*</td>
<td>0.45*</td>
<td>0.44*</td>
</tr>
<tr>
<td>5. Dealing with emotions</td>
<td>0.37*</td>
<td>0.34*</td>
<td>0.41*</td>
</tr>
<tr>
<td>6. Analysis of personal motivations</td>
<td>0.47*</td>
<td>0.40*</td>
<td>0.51*</td>
</tr>
<tr>
<td>7. Consideration of others</td>
<td>0.50*</td>
<td>0.44*</td>
<td>0.52*</td>
</tr>
</tbody>
</table>

*p < .05.
decision making measures with two alternative scoring systems: one based on dimensions of social psychological behavior as evident in response options, and one based on the cognitive strategies giving rise to a response option.

In examining the correlations of the social psychological dimensions with ethical decision making, involving others produced only one significant positive relationship \( (r = .20) \) with ethical decisions bearing on data—a finding that may reflect the fact that unethical manipulations of data often occur in private. However, involving others was not found to be strongly related to other ethical decisions. Most forms of ethical decision making were found to produce significant relationships with the remaining social psychological dimensions. Of these dimensions, selfishness \( (\bar{r} = -.50) \), avoidance of responsibility \( (\bar{r} = -.50) \), and deception \( (\bar{r} = -.46) \) produced particularly strong relationships with ethical decisions, suggesting that individuals who are not accountable, are concerned with their own outcomes, and are unwilling to trust others are especially unlikely to make ethical decisions.

In examining the correlations of the cognitive strategies scoring with the ethical decision making measures, some further evidence pointing to the measures’ construct validity was obtained. Seeking help, like involving others, produced a positive relationship \( (r = .20) \) with ethical decisions involving data but otherwise produced weak relationships. The remaining cognitive strategies dimensions, however, produced stronger relationships with ethical decision making. In particular, strong relationships were obtained for recognition of circumstances \( (\bar{r} = .50) \), consideration of the effects of actions on others \( (\bar{r} = .46) \), anticipating consequences \( (\bar{r} = .45) \), and analysis of personal motivations \( (\bar{r} = .40) \). Apparently, high scores on ethical decision making measures are related to forecasting the effects of actions in relation to the circumstances and likely motivation of the people involved in the event.

Causes of Ethical Decision Making

Although these relationships provide some evidence for the construct validity of the ethical decision making measure, more compelling evidence in this regard is provided by the relationships of these measures with causes and outcomes of ethical decision making. Table 6 presents the results obtained when scores on the ethical decision making scales were correlated with and regressed on the frequency of exposure to unethical practices in day-to-day work. Exposure to unethical practices in the day-to-day work of entry-level doctoral students was negatively related to ethical decisions with respect to data \( (\bar{r} = -.33) \), study conduct \( (\bar{r} = -.24) \), professional practices \( (\bar{r} = -.38) \), and business practices \( (\bar{r} = -.31) \). This early exposure to poor ethical practices, apparently, undermines subsequent ethical decision making.
In fact, the regression analyses indicate that exposure to poor day-to-day practices may have a sizeable effect on ethical decision making by first-year doctoral students. Across the data, study conduct, professional practices, and business practices criteria for ethical decision making, significant prediction ($p \leq .05$) was obtained, producing an average multiple correlation of .45. Notably, however, exposure to a certain class of negative career events, for example events related to professional practices, did not ensure that poor decision making could be accounted for by this dimension. Instead, in the case of professional practices, significant regression weights were produced by exposure to poor study conduct ($\beta = –.31$). Exposure to poor professional practice events was the strongest predictor of ethical decisions with regard to business practices ($\beta = –.41$), whereas exposure to poor professional practice events was found to be the strongest predictor of ethical decisions with regard to study conduct ($\beta = –.27$). Similarly, decisions with regard to data was related to exposure to study conduct events ($\beta = –.26$), as well as exposure to data events ($\beta = –.29$). Although these relationships may reflect the positive correlations observed among the events scales, it is also possible that career events exert a rather complex set of effects on ethical decision making.

Outcomes of Ethical Decision Making

Table 7 presents the correlations and regression analysis results obtained when ethical decision making was used to account for the punishments awarded on the re-
view board task. As may be seen, the correlational analysis indicated that scores on the ethical decision making measure were positively related to the perceived severity of violations ($r = .40$), the perceived importance of punishment ($r = .40$), and the severity of the punishment recommended ($r = .32$). Apparently, ethical decision making, specifically people’s integrity, leads to negative reactions to unethical behavior. The notable exception to this general trend may be found in the perceived frequency of violations ($r = –.04$). These weak relationships, however, may simply reflect the known inaccuracy in people’s forecasts of the occurrence of events (Josephs & Hahn, 1995), or, alternatively, the lack of adequate baselines in a population lacking substantial expertise.

The results obtained in the regression analysis confirmed these conclusions. For perceived severity, and the perceived importance of punishment, significant ($p \leq .05$) multiple correlations were obtained when the ethical decision making measures were used as predictors. The average multiple correlations obtained when scores on the ethical decision making measures were used as predictors of these three outcomes was .55. Again, however, ethical decision making scores were not related to the perceived frequency of violations.

Examination of the regression weights obtained in this analysis, however, indicated that perceived severity, perceived importance, and the severity of the punishment recommended were most strongly related to scores on the professional practices and business practices measures. Professional practices and business practices yielded significant ($p \leq .05$) regression weights in all of these analyses. The average regression weight produced by professional practices was $\bar{r} = .30$; the average regression weight produced by business practices was $\bar{r} =$

<table>
<thead>
<tr>
<th></th>
<th>Severity of Violation</th>
<th>Perceived Frequency of Violation</th>
<th>Perceived Importance of Punishment</th>
<th>Severity of Punishment Recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$\beta$</td>
<td>$r$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Data</td>
<td>.38*</td>
<td>.13</td>
<td>–.04</td>
<td>–.04</td>
</tr>
<tr>
<td>Study conduct</td>
<td>.29*</td>
<td>–.02</td>
<td>–.02</td>
<td>.04</td>
</tr>
<tr>
<td>Professional practices</td>
<td>.48*</td>
<td>.35*</td>
<td>–.01</td>
<td>.01</td>
</tr>
<tr>
<td>Business practices</td>
<td>.42*</td>
<td>.34*</td>
<td>–.09</td>
<td>–.11</td>
</tr>
<tr>
<td>Multiple $R^2$</td>
<td>.60</td>
<td>.10</td>
<td>.54</td>
<td>.47</td>
</tr>
<tr>
<td>Significance of $R^2$</td>
<td>.001</td>
<td>.893</td>
<td>.001</td>
<td>.001</td>
</tr>
</tbody>
</table>

*p < .05.*
These findings, of course, suggest that broader, less work focused, ethical decision making orientations may have a stronger impact on people’s reactions to others’ ethical violations.

DISCUSSION

Before turning to the broader conclusions flowing from this study, certain limitations should be noted. To begin, the evidence accrued in this study for the validity of this measure of ethical decision making was obtained in a unique sample. More specifically, in this study we examined ethical decision making in first-year doctoral students. Although entry-level doctoral students represent a population of particular interest in studies of integrity, caution is called for in generalizing our findings to other, more experienced, populations.

Along related lines, although in this study we examined ethical decision making across three fields—biological, health, and social sciences—it is also true that all doctoral students participating in this study were drawn from a single institution. The relatively constant nature of doctoral training programs across institutions suggests this single institutional strategy may well be appropriate. Nonetheless, some action is called for in extending the findings obtained in this to other institutions.

In this study, in keeping with the recommendations of Messick (1998), we have sought evidence for the validity of an ethical decision making measure by examining the relationship of scores on this measure to a variety of other measures. However, as is the case in any study, it should be clear that we have not sought to examine every potential relationship. Moreover, it should be recognized that the relationships examined in this study were based on a cross-sectional data collection strategy. Thus, the question remains as to whether similar findings would be obtained in studies applying alternative measures and alternative data collection strategies.

Even bearing these limitations in mind, we believe that the results obtained in this study do point to the validity of the measure of ethical decision making developed in this study. To begin, it was found that this measure was correlated with individual differences bearing on the appraisal of self and others—particularly cynicism, trust, and narcissism. Moreover, scores on this measure were not strongly related either to general cognitive ability or impression management, suggesting that the measure of ethical decision making was not unduly biased. More centrally, however, general personality characteristics such as openness, conscientiousness, and anxiety were not found to be strongly related to ethical decision making, indicating that these decisions may arise more from the people’s view of themselves and others, rather than from general personality.
In fact, the results obtained in examining alternative scoring systems provide some support for this argument. The social psychological dimensions suggested that scores on measure of ethical decision making are related to a selfish concern with personal outcomes. In fact, this self-centered orientation has been linked to integrity in studies by Fromm (1973) and O’Connor, Mumford, Clifton, Connelly, and Gessner (1995). Moreover, the results obtained for the cognitive strategies scoring suggest that the effect of this self-centered focus may be that it diminishes effective forecasting of the effects of one’s action in a given social situation (Gollwitzer, 1999). In fact, the social psychological and cognitive strategies dimensions produced far stronger relationships with scores on the ethical decision making measure than more traditional personality measures.

Of course, it is possible that the stronger relationships obtained for the social psychological and cognitive strategies dimensions might, in part, arise from common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). By the same token, however, it should be recognized that the cynicism, trust, and narcissism relationships reflect a similar general trend. Moreover, the picture both these analyses paint suggests that ethical decision making arises from one’s view of self and others. The view of self-operating, however, may lead to losses in the anticipation of consequences of one’s actions for self and others, thereby promoting decisions that evidence a lack of integrity.

These findings, of course, provide some evidence for the validity of the ethical decision making measures developed in the course of this study. Some further evidence bearing on the validity of this measure was provided in our examination of the ability of these decision making measures to account for people’s reactions to ethical violations by others (Sackett & Waneck, 1996). Here it was found that perceptions of the severity of the violation, the need for punishment, and the strength of the punishment recommended were all strongly related to ethical decision making, particularly ethical decision making with regard to professional and business practices. Although the unique effects exerted by professional and business practices on decisions about ethical violations suggest that broader, nonstudy specific manifestations may guide general evaluations of ethical violations (Estes & Ward, 2002), it should be recognized that all of the ethical decision making dimensions under consideration produced sizeable correlations with the review panel task examining reactions to ethical violations.

The other source of evidence bearing on the validity of this ethical decision making measure, of course, is the examination of day-to-day work experiences as a potential cause of ethical decisions. Prior studies by Brown and Kalichman (1998); Sims and Keon (1999); and Ulrich, Soeken, and Miller (2003) indicate that work experiences may be noteworthy influences on ethical decision making. Indeed, one might expect that work experiences would have a particularly powerful impact on ethical decision making among entry-level doctoral students. And, in fact, the results obtained in this study indicate that the frequency with which doctoral stu-
dents are exposed to unethical practices in the course of their day-to-day work is strongly negatively related to ethical decision making, at least in this early stage of their career. However, the exact mechanisms by which exposure to unethical practices influences subsequent decision making appear quite complex. Thus, exposure to poor study conduct practices did not, in the regression analysis, ensure that unethical decisions in the domain of study conduct would be made. In fact, exposure to poor professional practices appeared to have a larger association with unethical decisions in the domain of study conduct. Further research is needed, examining how exposure to different career events is synthesized and acted on by individuals in their ethical decision making.

Even bearing this caveat in mind, however, it is clear that experiences in one’s doctoral program provide a foundation for subsequent career development (Mumford et al., 2005). Accordingly, it seems plausible to argue that these early experiences, in conjunction with other later experiences, will influence ethical decision making for a substantial period of time. One implication of this observation is that one way society might improve ethical decision making is by training senior investigators in the responsible conduct of research—focusing in such interventions on the practices they follow in managing their laboratories.

In addition, however, the results obtained in this study suggest that educators might want to consider delivering responsible conduct of research training in a different way. Traditionally, responsible conduct of research training has focused on legal obligations and knowledge about expectations for appropriate professional behavior. Unfortunately, this kind of training, however useful and necessary it may be, has not proven highly effective in enhancing ethical decision making (Kalichman & Paik, 2004). Given the findings obtained in this study, however, it is not surprising that traditional training has limited effects on ethical decision making. Instead, training that focuses on centering people from their own personal concerns and encourages the analysis of the consequences of one’s actions for others may prove more beneficial given the findings obtained for the alternative scoring systems.

Aside from these implications for training, the results obtained in this study open up some potential new directions for studies of integrity in the sciences. Ethical decision making has long been considered a key criterion of interest in studies of integrity (Loe et al., 2000). The availability of reliable and valid measures of ethical decision making provides a basis for examining valid influences on this key component of integrity. For example, measures of the sort developed in this study might be used to assess the effects of climate or environmental influences on ethical decision making. Alternatively, the cognitive strategies influencing ethical decision making on these measures might be used to give us a better understanding of how scientists reason through ethical problems. These observations, of course, point to a broader conclusion. Progress in studies of integrity will ultimately depend on the availability of reliable and valid measures of ethical decision making.
In fact, the prior lack of these measures may be one reason why progress has been slow in studies of scientific integrity.

These observations about the practical and theoretical implications of our findings are noteworthy for a broader reason. Certainly the findings obtained in this study argue for the validity of our measure of ethical decision making. Not only was a substantively meaningful pattern of relationships with relevant individual differences variables obtained, scores on this measure were found to be associated with certain social psychological dimensions and cognitive strategies, and also related to plausible causes and outcomes of ethical decision making. The conclusions drawn about ethical decision making as a result of these analyses, moreover, point to new and different interventions that might be used to improve ethical decision making in the sciences. We hope this study will provide an impetus for future research along these lines.

ACKNOWLEDGMENTS

We thank Whitney Helton-Fauth, Blaine Gaddis, and Gina-Marie Scott Ligon for their contributions to this effort. The data collection was supported in part by the National Institutes of Health, National Center for Research Resources, General Clinical Research Center Grant M01 RR–14467. Parts of this work were sponsored by a Grant No. 5R01–NS049535–02 from the National Institutes of Health and the Office of Research Integrity, Michael D. Mumford, principal investigator.

REFERENCES


Moss is a researcher in the laboratory of Dr. Abrams, a well-known researcher in the field of economics. Moss is trying to develop a model to predict performance of stocks in the technology sector, but she is having difficulty analyzing and selecting trends to include in the model. She enlists the help of Reynolds, another experienced researcher working on a similar topic. With Reynold’s help, Moss eventually analyzes and identifies some key trends, working them into a testable model. She also discusses some of her other research ideas with Reynolds. Two weeks later, Moss comes across a grant proposal developed by Reynolds and Abrams. She sees that it includes ideas very similar to those she discussed with Reynolds. She takes the matter to Abrams, who declines to get involved, saying that the two researchers should work it out on their own.

1. Reynolds admits to Abrams that he used slightly modified versions of Moss’s ideas. Abrams is upset with this, but Reynolds is a key person on the proposal team and the grant application deadline is soon. What should Abrams do? Choose two of the following:
   a. Fire Reynolds from the lab on the grounds of academic misconduct.
   b. Leave Reynolds as first author on the proposal because he wrote up the ideas.
   c. Remove Reynolds from the proposal team, and offer Moss the position if she allows her ideas to be used.
   d. Ask Moss to join the grant team, placing her as third author on the proposal if she allows her ideas to be used.
   e. Acknowledge Moss in the grant proposal because the ideas were hers originally.
   f. Apologize to Moss and indicate that the proposal must go out as is to meet the deadline.
   g. Remove Moss’s ideas from the proposal and try to rework it before the deadline.

2. Moss is upset about Reynolds using her ideas and she decides to do something about it. Given that Moss works very closely with Reynolds and their boss Abrams, evaluate the likely success of the following plans of actions Moss can take. Choose two of the following:
   a. Moss asks Reynolds to give her credit by putting her name on the grant proposal as well.
   b. Moss asks Reynolds about the incident and tape records his reaction to later show Abrams.
   c. Moss searches for annotated notes about her ideas that are dated prior to her conversation with Reynolds.
d. Moss appeals for a “mock trial” for Reynolds to testify under oath to his superiors that the information was his.
e. Moss searches for Reynold’s lack of understanding of the concepts he claims were his own by questioning him in front of other students.
f. Moss attempts to sway other researchers to support her to Abrams.
g. Moss visits Reynolds’ office in hopes of finding evidence that she contributed to the proposal.
h. Moss asks Reynolds to write an account of their conversation on the day in question and shows her comparison account to him as evidence that he is using her ideas.