

7-2022

Are Two Heads Better Than One? Interpreting Students' Moral Reasoning Skills

Robyn Dean

Vincent Samar
Rochester Institute of Technology

Daniel Maffia
Rochester Institute of Technology

Follow this and additional works at: <https://tigerprints.clemson.edu/ijie>

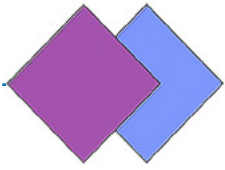


Part of the [Education Commons](#), and the [Sign Languages Commons](#)

Recommended Citation

Dean, Robyn; Samar, Vincent; and Maffia, Daniel (2022) "Are Two Heads Better Than One? Interpreting Students' Moral Reasoning Skills," *International Journal of Interpreter Education*: Vol. 14: Iss. 1, Article 4.
DOI: <https://doi.org/10.34068/ijie.14.01.04>
Available at: <https://tigerprints.clemson.edu/ijie/vol14/iss1/4>

This Research Article is brought to you for free and open access by TigerPrints. It has been accepted for inclusion in *International Journal of Interpreter Education* by an authorized editor of TigerPrints. For more information, please contact kokeefe@clemson.edu.



Are Two Heads Better Than One? Interpreting Students' Moral Reasoning Skills

Robyn K. Dean¹

Rochester Institute of Technology

Vincent Samar²

Rochester Institute of Technology

Daniel Maffia³

Rochester Institute of Technology

Abstract

The Defining Issues Test (DIT) is an internationally used instrument that measures an individual's moral reasoning skills—that is, how an individual explains right and just action. DIT scores are correlated with age and education, and they are also correlated with clinical performance when administered to professional practitioners. Practicing signed language interpreters' scores, however, were not reflective of their age and education in one study, being much lower than those of practitioners from other professions. Providing communication access for individuals who do not share the same language as their service providers is grounded in social justice and equity, yet practicing interpreters' DIT scores did not reflect higher-order justice skills. The current study investigates American Sign Language interpreting students' DIT performance. Over the course of 3 years, different classes of third-year interpreting students in an undergraduate program in the United States took the DIT. Each year, the students' average scores were significantly higher than those of working interpreters ($n = 80$). This result raises the question of whether, how, and why years of interpreting experience curtail ethical development. A follow-up study with one student class ($n = 32$) also found that taking the DIT as a collective, meaning that answers were negotiated among group members, resulted in higher scores than the group members' individual median scores. This additional finding adds credence to educational approaches that focus on ethical discussion and deliberation of interpreting practice.

Keywords: interpreting students, decision making, ethical reasoning, Defining Issues Test

1 Correspondence to: robyn.dean@rit.edu

2 Correspondence to: vincent.samar@rit.edu

3 Correspondence to: daniel.maffia@rit.edu

1. Introduction

During an interpreting faculty meeting a few years ago, the topic of students accepting professional work in the community while still in the interpreting program was raised. Interpreting faculty address this concern with some frequency: How do we instill in students the importance of discernment regarding their qualifications—when to accept or not accept an interpreting assignment—now and in the future?

Within the context of this conversation, one faculty member relayed a distressing conversation she had had with students a couple of weeks prior: “I asked the students how they should determine whether or not they accept an interpreting assignment.” She went on to contrast two responses in particular. One student suggested that they would consider the impact on the deaf and hearing people. Another student admitted that it was their *availability* that would determine whether they accepted the assignment: “I just ask myself, does this fit my schedule, and if it does, I would accept it.” This response caused the entire faculty to groan with frustration.

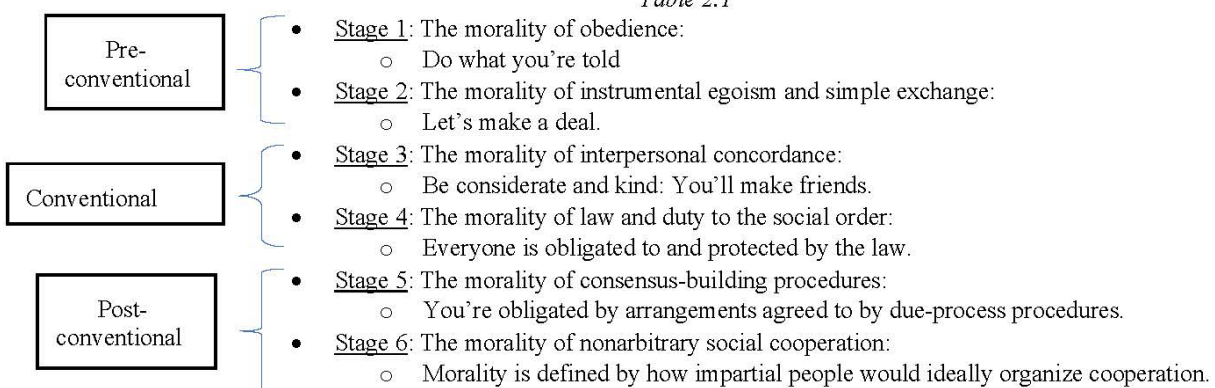
How an individual *reasons through* a decision to reach what they imagine is right and just action is of interest to moral psychologists (Gilligan, 1977; Kohlberg, 1976; Rest et al., 1999). Moral psychologists have proffered theories and empirical data for what constitutes an ethically defensible reason as well as what makes some reasoning patterns better than others (Rest et al., 1999). In more than 100 years, more than 20 instruments have been created and used to measure moral judgment, moral reasoning, and moral development (Martí-Vilar et al., 2021).

One of the top two measures used in the last 30 years is the Defining Issues Test (DIT; Martí-Vilar et al., 2021). The DIT measures an individual’s pattern of moral reasoning and has been used with professional practitioners in the fields of teaching, medicine, business, and law (Christensen et al., 2016; Rest, 1994; Roche & Thoma, 2017). The DIT more recently has been used to measure the moral reasoning abilities of American Sign Language–English (ASL-ENG) interpreters in the United States (Dean, 2015), the results of which reveal moral reasoning patterns commensurate with individuals in their teenage years, more than half of the average age of the study participants. Teenagers and young adults (~ age 22) on average demonstrate a moral reasoning pattern that is concerned with following rules or social conventions. The results of this initial study beg the question of why these interpreters have not developed beyond this stage. This current study investigates whether student interpreters, representative of the teenage to young adult category, show similar reasoning patterns to their peers and to practicing interpreters.

2. Stages and Schemas of Moral Reasoning

At the interpreting faculty meeting described above, no one felt the need to question what was wrong with the second student’s response or why the first student’s response was better. The faculty all just seemed to tacitly agree that the first was acceptable and the second wasn’t. Kohlberg’s (1976) six stages of moral development readily explain why the faculty assessed each student’s response differently. Kohlberg’s six stages include a notable hierarchical nature to ethical reasoning, from preconventional to postconventional. Table 2.1 defines Kohlberg’s moral stages (Rest 1994, p. 5).

Table 2.1



According to Kohlberg, the second student's response of "Does this fit my schedule?" would be indicative of the *preconventional stages*, where the focus is on the self and limited to only those who are known (i.e., kinship). In other words, an individual in the preconventional stage might ask, "What's in it for me and my ingroup?" By being concerned about the impact the decision could have on others, the first student demonstrated a type of *postconventional reasoning*. Postconventional reasoning, or principled reasoning, as it is also called, is concerned with shared values; these individuals might ask, "How can I cooperate?"

There was not a third student in this anecdote, but as a means of completing the illustration, imagine a student reasoning at the conventional stages. They might have offered, "I should only accept work based on those in authority; if my boss says I should take it, then I should." Those in the conventional stages tend to ask, "What is the rule, and what is my duty?"

What Kohlberg's theory does not account for is why one 20-something-year-old college student was reasoning at an age typical of preadolescence while the other 20-something was reasoning at a level *theorized to be* beyond their years (Rest et al., 1999). According to Kohlberg (1976), these similarly aged individuals should have been drawing on a shared moral stage and, as a result, should have reflected similar reasoning abilities. Indeed, 20-somethings tend to fall into the conventional stages of moral development, but neither of these students gave a conventional response—the conventional response was created and added for the purposes of illustration.

The moral psychologist James Rest was a former doctoral student of Lawrence Kohlberg. Along with scores of other graduate students, Rest helped advance the empirical value of Kohlberg's instrument, the Moral Judgment Interview (MJI). In the MJI, participants are offered a hypothetical scenario of an overarching social nature, asked what the character in the scenario should do, and then queried for their reasoning behind the proposed ethical action. The participants' moral reasoning statements are then coded based on Kohlberg's six stages.

Years later, Rest's own research would address some of the weaknesses apparent in Kohlberg's research design and protocol (Rest et al., 1999). The first of those weaknesses was the concept of stages. Kohlberg's notion that individuals follow along a steplike, age-based series of stages as they mature was not compelling and not sufficiently evidenced in the data. Rest instead proposed *moral schemas*. Although schemas are not completely irrespective of age, he found that all individuals (old enough to be verbal) have the capacity to think and reason in preconventional, conventional, and postconventional ways. Instead of Kohlberg's six stages, Rest proposed three schemas that still include the preconventional, conventional, and postconventional categories. Table 2.2 lists Rest's moral schemas, descriptors for each, and how they correspond to Kohlberg's stages (Narvaez & Bock, 2002).

Table 2.2: Rest's Moral Schemas

Rest's Moral Schemas	Schema Features	Compared to Kohlberg
Personal Interest	Arbitrary or negotiated cooperation Self-focused Advantage to self is primary Survival orientation Scope includes known others In-group reciprocity	Stages 2 & 3
Maintaining Norms	Need for norms Self-focused Uniform categorical application Partial society-wide reciprocity Duty orientation	Stage 4
Postconventional	Appeal to an ideal Sharable ideals Primacy of moral ideal Full reciprocity Rights orientation	Stages 5 & 6

Although all individuals have the *capacity* to draw on any of the three schemas while thinking and reasoning ethically, as they move through life stages (e.g., age, education, life experience), individuals develop a dominant schema—one that is more consistently relied upon (Rest, 1994). Deviations from that dominant schema would be expected when particular situations are encountered. For example, a person could be dominant in postconventional reasoning, but when faced with a particular topic and the decisions associated with that topic (e.g., military service or diet and exercise), they might reason at a level more reflective of the maintaining-norms schema. Emotional states, such as anxiety and fear, can also shift an individual's dominant schema. Rest's research showed that dominant schemas do tend to cluster around age and education, but there is also potential for deviation from those trends.

3. Tacit Processes

Rest (1994) also proposed that moral schemas function tacitly, or *beyond one's awareness*. As a result, when an individual is asked directly to reason or to justify their proposed ethical action (like with the design of Kohlberg's MJI), they may struggle to express their reasoning. For example, in the scenario above, the faculty collectively expressed their frustration at the reply of "If it fits my schedule. . ." They all appeared to be agreeing that the first response (concern for the impact of the decision) was more ethically compelling than the second (a concern for personal circumstances and benefits). No explanations were needed; tacitly, the faculty agreed with the morally superior response.

However, if someone had asked what was wrong with the second student's answer, the tacit nature of moral schemas would have become more evident. Some might have only replied, "Isn't it obvious?"; others could have explained how *concerns for others* morally supersede *concerns for the self*. This variation in response is not necessarily because of variations in moral reasoning abilities but in verbal abilities (Narvaez & Bock, 2002). In other words, the ability to articulate tacit reasoning is linked more to verbal skills than to reasoning skills (Rest et al., 1999).

In Kohlberg's MJI research, the protocol was to collect spontaneous responses and measure them. Kohlberg's MJI instrument uses hypothetical scenarios and poses the questions of "What should be done?" and "Why is that the right thing to do?" However, decades of data collection did not produce sufficient evidence of participants' ability to reason at the most sophisticated Stages 5 and 6, or *postconventional reasoning*. In what would eventually be called Kohlberg's *fatal flaw*, Kohlberg's data were thought to be representative of verbal skills and not reasoning skills (Rest, 1994).

Kahneman (2011) sums up this phenomenon by explaining that moral reasoning "requires a richer vocabulary than is available in everyday language" (p. 4). Without access to or use of such vocabulary, it can be argued that what a person is drawing on is just what is *cognitively available* (Kahneman, 2011) and not necessarily reflective of their actual abilities. What becomes cognitively available to an individual depends on the kinds of normative messages that they have directly or indirectly acquired (Dean, 2014, 2015).

Normative messages—being told directly or indirectly what is right and wrong—are ubiquitous. For professionals, normative messages about effective and ineffective practices are found in formal and informal sources—that is, they can be found in sanctioned documents and expressed through common tropes (e.g., "You're just the interpreter"). They come from a profession's ethical codes and standards of practice documents and are equally acquired from the "ethical parlance adopted by educators and the discursive norms and heuristics used by practitioners" (Dean, 2015, p. 96).

In one study, I (Dean) identify a series of normative messages prevalent in signed language interpreting textbooks and in interpreting discourse. Normative messages can sometimes morph into heuristics or mental shortcuts. Heuristics help the brain simplify ethical issues into *rules of thumb* and, by doing so, intimate solutions (Kahneman, 2011):

[I]n order to function, the brain creates heuristics allowing for cognitive ease, to think and respond quickly. Heuristics [are] "simple procedures that help find adequate, though often imperfect, answers to difficult questions" (Kahneman, 2011, p. 98). Therefore, it is likely that . . . normative messages would function as an availability heuristic for [sign language interpreting] students. (Dean, 2014, p. 65).

Consider the heuristic of “better safe than sorry”—something we might say to ourselves or to each other when faced with a decision. This heuristic distills the complexity of the decision and proposes two possible outcomes—being safe or not being safe, and in not being safe, experiencing regret. The solution further intimated by the heuristic is “choose whichever leads to safety.” Consider the statement “You’re just the interpreter,” which is sometimes used in the interpreting field and similarly can function as a heuristic. “You’re just the interpreter” implies that whatever the expressed concern is, in actuality, *should be of no concern to you*. The solution is further implied from this simplification—if it should be of no concern to you, then you should *do nothing, not address it, or take no action*.

Doing nothing, taking no action, and the like are common themes found in the ethical content material used in interpreter education (Dean, 2014). Interpreting texts and interpreting discourse reveal a series of normative messages that can be summarized as: When faced with a decision to take action or not take action, interpreters should almost always not take action. Exceptions include those associated with effective message transfer (e.g., not hearing the speaker), the preferences and directives of the deaf person (e.g., “Please don’t interrupt the speaker”), or those resulting in untoward consequences, as in life-or-death decisions (see Dean, 2014, p. 65, for a more detailed description).

If you ask a person why they consider a particular decision to be ethically correct, you are more likely to access what they are *verbally* capable of than what they are *morally* capable of—that is, they might have the tacit ability to think and function in a more sophisticated way than what they can express. Moreover, what they are verbally capable of may well be informed by those normative messages that are elicited from stimuli, such as familiarly themed ethical scenarios or ethical dilemmas.

4. Ethical Scenarios in Interpreting Education and Research

Posing ethical scenarios, dilemmas, or cases (real or made-up) as a means of eliciting dialogue on decision making and reasoning is a common educational technique in professional ethics (Hill, 2004). Although it is also common in interpreter education (Cartwright, 2009; Dean, 2014; Drugan & Megone, 2011; Humphrey, 1999; Seal, 1998), some have highlighted the kinds of problems that occur with overreliance on this teaching technique (Dean, 2014; Marin, 2020; Wilbeck, 2017).

Ethical scenarios typically focus on the boundaries between right and wrong (Hill, 2004), requiring the decision maker to consider only two obvious choices: Do *this* and don’t do *that*. This type of ethics education only advances *mandatory minimal standards* of a profession’s ethics (Hill, 2004). As a result, ethical dilemmas “do very little to advance ethical awareness . . . [or] conversations [that] foster and ensure *effective practice*” (Dean, 2014, p. 62). The use of ethical dilemmas might also lead to a type of defensive practice (Hill, 2004), causing practitioners to await and expect some ethical breach (Dean, 2015). As a result, practitioners “might limit their behaviour even more severely than necessary and thus fail to serve their client’s best interests” (Hill, 2004, p. 140; see also Turner & Best, 2017). Posing ethical scenarios is not only an educational technique; it is used as a means of assessing interpreters’ professional ethical knowledge and skills (as in education and professional accreditation), and it is used as a research tool.

Collecting data through discussion of ethical scenarios or ethical dilemmas is a popular technique in interpreting research (Bergson & Sperlinger, 2003; Dean & Pollard, 2009; Stewart & Lindsey, 1990; Tate & Turner, 2002; Thomas, 2012). Often, the respondents’ spontaneous speech or self-explanation is prompted by a scenario, and then that speech is recorded and analyzed for ethical thought and reasoning. In other instances, *in vivo* interpreted situations are observed and followed up by interviews about interpreters’ actions to collect and analyze the interpreters’ ethical thoughts and reasonings (Hsieh, 2007, 2009).

Mendoza’s (2012) study used surveys and interviews to investigate the ways in which novice and expert ASL-ENG certified interpreters discuss how they make ethical decisions. In the second part of the study, Mendoza interviewed six of the survey participants and asked them to recall and report on an actual interpreting situation from their own practice. They were asked to consider those situations that highlighted one of the tenets of the professional ethical code, such as confidentiality, impartiality, professionalism, or business practices. Mendoza then compared their responses, three from the novice group (3 years and under of interpreting experience) to three of the expert group

(having more than 10 years of experience). Mendoza found that the novice group was limited in their assessment of the ethical dilemma (i.e., lacking complexity), while the expert group explored greater nuance. The expert group was able to discuss how consequences of decisions would affect all individuals in the situation, whereas the novice group appeared to be more self-focused.

Posing ethical scenarios to elicit ethical responses is a common educational, assessment, and research tool. Research data that rely on the verbal data of participants need to be interpreted carefully. If ethical reasoning skills are indeed being measured through these data, then there may be concerns for research validity. The data need to be critiqued in light of what might be a function of verbal skills as well as how those verbal skills might be influenced by what is made cognitively available (Dean, 2014).

5. Rest's DIT

The DIT, now in its second iteration (DIT-2), was developed by Rest (1986) as an alternate measure of individuals' reasoning skills. Similar to the MJI, the DIT proposes ethical scenarios and follows each scenario with a series of questions. Also similar to the MJI, the research participant is queried for the optimal ethical action, but instead of having to rely on their verbal skills to explain their reasoning, they are given verbal assistance (Narvaez & Bock, 2002) in the form of a series of argumentation statements for each of the five scenarios. As such, the DIT's design is based on the participant's ability to recognize and affirm compelling rationales for their chosen ethical action. This is done by the participant's rating (a Likert scale of importance from 1 to 5) and ranking each of the 12 argumentation statements and choosing the top four most compelling.

The Center for the Study of Ethical Development (CSED; n.d.) explains the DIT's design rationale:

A common assumption in the field of morality, and one with which we disagree, is that reliable information about the inner processes that underlie moral behavior is obtained only by interviewing subjects. Contrary to assuming that interviewing presents a clear window into the moral mind, researchers in cognitive science and social cognition contend that self-reported explanations of one's own cognitive process have severe limitations. There is now a greater appreciation for the importance of implicit processes and tacit knowledge on human decision making, outside the awareness of the subject and beyond his or her ability to verbally articulate them.

For example, consider this frequently cited example, the Heinz Scenario. In summary, Heinz's wife is dying from a disease. The druggist in town has invented a medicine that will save the wife, but his price is too high. Heinz cannot raise the money. He is faced with the moral choice of whether or not to steal the drug. Sample argumentation statements offered to respondents to rate and rank include the following:

- It really depends on how much Heinz likes his wife and how much risk there is in taking the drug. If he can't get the drug in another way, and if he really likes his wife, he'll have to steal it.
- Regardless of his personal feelings, Heinz has to realize that the druggist is protected by the law. Because no one is above the law, Heinz shouldn't steal it. If we allowed Heinz to steal, then all of society would be in danger of anarchy.
- I think that a husband would care so much for his wife that he couldn't just sit around and let her die. He wouldn't be stealing for his own profit; he'd be doing it to help someone he loves.

Listed here are only three sample argumentation statements. In the DIT-2, each scenario has 12 to be rated and ranked. The argumentation statements are devised to reflect features of Rest's moral schemas: personal interest (PIS), maintaining norms (MNS), and postconventional (PCS).

The DIT is administered and scored by the CSED in the United States. Since the 1970s, the CSED has collected thousands of DIT scores on an international level. Multiple indices are offered for validity (see also Thoma & Dong,

2014). The DIT is correlated with age, education, and clinical performance; it is not correlated with personality trait measures, IQ, or socioeconomic status (CSED, n.d.). In a recent study, the DIT was found to be one of the top two most used moral development measures (Martí-Vilar et al., 2021).

Three types of scores are returned to the researcher: the *P-score*, the *Type Indicator*, and the *Utilizer Score*. The P-score, or the *principled reasoning score*, is the score that most researchers report. It accounts for the degree to which a respondent endorses statements indicative of PCS, the most sophisticated schema. The P-score is reported in percentages, as are the reports of the other two schemas, PIS and MNS. Whichever percentage is the highest of the three (PIS, MNS, PCS) is the respondent's current dominant schema. As mentioned earlier, these do tend to cluster around age and education. Table 5.1 shows the average scores on the normed data based on education.

Table 5.1: Average P-scores, based on age and education

Grade level	Average PIS	Average MNS	Average PCS
Grades 10–12	27.7%	35.3%	31.6%
Undergraduate	25.0	35.1	35.1
Graduate	20.6	34.1	41.1

The numbers for PIS and MNS decrease based on age and education, while the PCS score increases. This is an expected response to increasing age and educational attainment.

Table 5.2 lists the normed data for students in professional education and practitioners, including two anchors for comparison: seniors in high school (~ age 18) and adults in general.

Table 5.2: Normative data for students in professional education and practitioners

P-score	Group
65.2	Moral philosophy and political science graduate students
52.2	Law students
50.2	Medical students
49.2	Practicing physicians
46.3	Staff nurses
42.3	College students in general
40.0	Adults in general
31.8	Senior high school students

In addition to collecting normed data on students and professionals, the DIT has been used to measure the efficacy of ethics education in a pre-test/post-test measured design (Bebeau, 2002; Mayhew et al., 2014). McGeorge (1975) used the DIT to investigate the hierarchical nature of ethics reasoning, examining whether respondents are able to *will* themselves to do worse (they can) or better (they can't) than their own score. In other words, respondents can fake low, but they can't fake high.

The DIT has also been used to determine whether working together as a collective produces better scores (i.e., higher P-scores) than the scores achieved when those individuals take the DIT alone (Dukerich et al., 1990; Nichols & Day, 1982). This latter design of individual versus collective scores is of particular interest to interpreter education for a couple of reasons. First, ethics education is often conducted by using real or constructed ethical scenarios to prompt ethical discussions among a group of students, typically in an ethics or practicum seminar class, or among interpreter practitioners during a professional development event. Both of these designs are based on the idea that two heads are better than one when it comes to ethical reasoning. Yet if such a hierarchy in moral

reasoning exists, do those who reason at higher levels pull up those who reason at lower levels, resulting in an influential modeling of advanced abilities? Or does the opposite occur, with those who are reasoning at an advanced level being pulled down? In other words, maybe two heads are not better than one.

6. The DIT and Interpreters

In 2013, 25 American Sign Language (ASL) interpreters agreed to take the DIT-2. This cohort was a subset of the approximately 65 participants involved in a related qualitative study (Dean, 2014 & 2015). Although the qualitative data collected from this study revealed a very rule-bound and prescriptive understanding of interpreting ethics (Dean, 2014 & 2015), those data had to be reconsidered in light of the potential concerns flagged by moral psychology researchers—that is, relying solely on interview data as a “window into the moral mind” (CSED, n.d.) is flawed.

Much of what interpreters use to justify decisions can be found in what is cognitively available from traditional normative messages (Dean, 2014). Posing the same types of questions to interpreters results in formulaic and predictable answers. Those answers may not accurately reflect a person’s reasoning capabilities because the formulaic answers may well be governed by tacit processes—that is, these predictable questions may summon a moral schema that might not be a true reflection of the participant’s dominant moral schema. However, if interpreters are provided with a different type of question, outside the predictable ones, then they might not follow a well-worn discursive pathway (Dean, 2014 & 2015). Instead, they might reveal their greater ethical reasoning capacity.

The cohort of interpreters (n = 25) who agreed to the follow-up study completed an online version of the DIT-2. As part of the DIT-2 data collection, respondents were asked to input demographic data, such as gender, age, and highest level of education. The average age of the cohort was 49 years, more than half had an undergraduate degree or above, and all but one held national- or state-level qualification.

The average PCS or P-score for the interpreting practitioners was 33.6. As shown in Table 5.1, when compared to age and educational attainment, their PCS or P-scores and PIS were similar to those in their early 20s, yet their average age was more than double that. Notably, their MNS scores were the highest, at 35.0, making this the dominant moral schema of the group. Their average MNS score was commensurate with that of teenagers, making them appear even younger than their PCS and PIS scores indicated. A dominant MNS schema means that as a collective, they were most frequently compelled by arguments that focused on maintaining the rules and following one’s duties—that is, “[t]hey appear to think that following the rule will yield a moral result—that the rule contains the moral ideal” (Dean, 2015, p. 207).

Comparing the interpreters’ P-scores to those of other professionals (and their students) yielded similarly unexpected results (see Table 5.2). Not only did the interpreters not even *make it on the list* of the other professionals; they scored lower than adults in general (a P-score of 40.0) and closer to seniors in high school (a P-score of 31.8). These unexpected results prompted the following questions: Does the prescriptive nature of interpreting ethics come to influence interpreters’ broader perception of ethical behavior? Or does the interpreting profession, because of its frequently rule-bound normative messages, attract those who already have a propensity for prescription? As students, do interpreters arrive with a prescriptive, duty, and rule-bound perspective on ethical action? Accessing the DIT scores of interpreting students was the most direct way to get help in answering these questions.

7. Methodology and Results

7.1. Individual DITs: Study 1

Third-year interpreting majors enrolled in a 4-year program in the northeastern United States were asked to participate in this study. Approximately 30 students across two sections of an interpreting ethics course took a paper-and-pencil version of the DIT-2. This study was repeated in their ethics class, using about 1 hour of class time

for the fall semesters of 2016, 2017, and 2018. Institutional Review Board (IRB) approval was granted for the three iterations of this DIT study.

The same protocol was used for each iteration of the study: Test booklets, answer sheets, and number two pencils were distributed to all students. Each student created a code name, one that would be unique to them and not easily guessed by their classmates or their instructors (e.g., the name of a pet or a special date). They were instructed to write that code name on the answer sheet and to remember or write it down for safekeeping. This code name was then linked with a random number code used by the researchers and written in the text field of the answer sheets. A key was created that linked the students' chosen code names and the researchers' assigned numbers. The answer sheets were sent to CSED, where they were scored. The scores were returned by email approximately 1 week later.

The results of the DIT included individual scores and the group's aggregate score. The aggregate scores were presented to the whole class; after that, students were asked to claim their individual scores, which were laid out on folded sheets of paper, identified only by the code name. Students were encouraged not to share or ask each other about their individual scores because a higher P-score is considered "better." Time was allowed for a discussion as well as an explanation of what the DIT can say about moral reasoning and the DIT's limitations (Bebeau & Faber-Langendoen, 2014).

7.2. Results: Study 1

Table 7.2.1 lists the cohorts from all 3 years and their P-scores, which represent PCS.

Table 7.2.1: Mean score of each P-score for third-year students in 2016, 2017, and 2018

Interpreting students 2016	n = 31	Mean P-score (PCS)	45.1
Interpreting students 2017	n = 32	Mean P-score (PCS)	47.3
Interpreting students 2018	n = 27	Mean P-score (PCS)	41.9

The P-scores for all three classes are higher than the average P-score of 33.6 for the practicing interpreter cohort in 2013 (Dean, 2015). Table 7.2.2 adds the students' PIS, MNS, and PCS scores within the table of normed DIT data for undergraduate students.

Table 7.2.2: Newer normed DIT data for each university year for college students (CSED, n.d.)

<u>University year</u>	<u>Average PIS</u>	<u>Average MNS</u>	<u>Average PCS</u>
First year: freshman	26.5	34.3	34.1
Second year: sophomore	25.7	34.3	35.2
Third year: junior	24.9	35.5	34.9
Fourth year: senior	23.7	35.7	35.9
2016 students	24.8	24.4	45.1
2017 students	25.5	22.3	47.3
2018 students	27.8	23.4	41.9

Comparatively, all classes of interpreting students did better than their peers in PCS (the P-score). They also had much lower scores for MNS compared to their peers. They also were at least 10 points lower (i.e., better) than the mean for the practicing interpreters, who had an MNS score of 35.0. Their PIS scores were more comparable to those of their peers and in 2018 were slightly higher than those of their peers.

Table 7.2.3 combines all P-scores from students and practitioners and compares them to other students and practitioners from other professions.

Table 7.2.3: Comparing interpreting student cohorts, interpreter practitioners, and other professionals

P-score	Group
65.2	Moral philosophy and political science graduate students
52.2	Law students
50.2	Medical students
49.2	Practicing physicians
47.3	Interpreting students (2017)
46.3	Staff nurses
45.1	Interpreting students (2016)
42.3	College students in general
41.9	Interpreting students (2018)
40.0	Adults in general
33.6	Interpreter practitioners (2013)
31.8	Senior high school students

In every year that the DIT was administered, the interpreting students did better than their practitioner counterparts, scoring within the realm of other professionals and graduate students. Therefore, the answer to the question that arose out of the lower-than-expected scores of the practitioners—are interpreters conventional thinkers because they arrive to educational programs already ensconced in a rule-bound perspective on ethical action?—appears to be *no*.

7.3. Collective DITs: Study 2

In 2017, we added a follow-up component to the individually administered DIT. In addition to each student's taking the DIT-2 individually, they also took it as a group, with a negotiated answer to each ethical scenario recorded on the DIT answer sheet. This collective version of the DIT was conducted before the students were informed of their class's aggregate score or their individual scores. This study was also approved by the university's IRB.

A total of seven groups, each with four or five members, was created based on the members' individual scores. Every attempt was made to balance the groups as much as possible—that is, a balanced number of P scores in the 30s, 40s, 50s, and so forth as well as a balanced number of students. This quasi-matched design was relatively effective; however, in at least two cases, students were absent on the day that the collective DIT was administered. In those cases, groups were missing members and representative scores. This quasi-matched design was not known to the students because at this point, none of the students was aware of the scores. Code names were used to divide the students, which to them likely appeared to be random.

Approximately seven students had scored in the high range of the DIT's principled reasoning or P-score (60 or above). Each group included one of the high scorers, who was assigned to be the group's leader. The leaders were given unique instructions, which are explained later. The scenario and the three questions were read aloud by each leader to each group of students. The group was asked to deliberate and come to consensus for each of the three questions across the five scenarios of the DIT-2. The DIT-2 ends with a series of demographic questions; for the collective responses, these questions were skipped. Throughout the 45- to 60-minute process, we kept tabs on each group's progress. If a group was running behind, they were asked to be cognizant of the time they were spending in deliberating over and negotiating answers.

Certainly, not all students participated equally in the deliberative process. This is not dissimilar from what might happen in an ethics course where students are asked to discuss some ethical issue pertaining to interpreting. In that way, it could be argued that the processes are similar. No special instructions were given to the students to make this process any different from what might naturally occur in class.

One question that was intriguing was whether the high scorers would have a stronger influence on the deliberative group process if they *knew* that they were high scorers. To answer this question, the leaders (i.e., high scorers) were placed in two groups with the conditions of *tell* and *no tell*.

The first section of the ethics class had five groups and, therefore, five leaders. Two of the five leaders were asked to get one test booklet, one answer sheet, and a pencil for their group. Those two leaders were given those items and sent back to their groups. This was the no-tell condition. Next, the other three leaders were called up and given the test items and then told quietly that they were high scorers (without being told their exact scores) and that, as such, they should make their opinions heard in the group. They were told not to tell anyone about their high-scorer status and encouraged not to dominate the discussion but to just *make sure they expressed their opinions*. They agreed and were sent back to their groups. This was the tell condition. The very same design occurred in the second section except, with only two groups in the smaller class, one was the tell condition and one was the no-tell condition. With a total of seven groups from two sections of the ethics course, four groups were the tell condition and three groups were the no-tell condition.

The collective DITs with negotiated scores were coded to designate group numbers, and the answer sheets were sent by mail to the CSED. The CSED scored them and returned the results in PDF through email about a week later. A presentation of the individual and collective scores was delivered toward the end of the semester; all students were informed of the study's outcome.

7.4. Results: Study 2

The 2017 third-year class, a total of 32 students, was unique in having the highest P-scores but also the largest range of P-scores. The high scorer got a 70, while the low scorer got a 20. Every effort was made to create quasi-balanced groups (see column 2 in Table 7.4.1). Column 2 in Table 7.4.1 lists each group's individual scores, with at least a 30-point difference between the lowest scorer and the highest scorer. Column 3 contains each group's median P-score. The median is preferred over the mean because of the significant spread in P-scores in the 2017 cohort. Column 4 shows the results of the negotiated P-scores, and column 5 shows each group's gain from the median P-score to the higher negotiated score.

Table 7.4.1: Comparison and analysis of the individual DIT and collective DIT

Group 1	1. 62	44	66	+22
Groups	Individual P-scores	Median P-score	Negotiated P-score	Gains in P-score
	3. 44			
	4. 34			
	5. 30			
Group 2	1. 68	46	58	+12
	2. 52			
	3. 46			
	4. 36			
	5. 34			
Group 3	1. 72	48	60	+12
	2. 48			
	3. 48			
	4. 32			

Group 4	1. 64 2. 42 3. 38 4. 34	40	54	+14
Group 5	1. 68 2. 56 3. 36 4. 32 5. 20	36	48	+12
Group 6	1. 66 2. 48 3. 44 4. 36 5. 26	44	48	+04
Group 7	1. 68 2. 56 3. 46 4. 40	51	54	+03

In every instance, the negotiated P-score was better than the median P-score, with gains ranging from 3 points to 22 points. The average gain was 11.3 points (with a median and mode of 12 points), a significant difference by the sign test ($p = .008$). In other words, collective DIT takers scored reliably higher than individual DIT takers.

As mentioned above, four groups across the two sections of ethics class were led by those who knew they were the high scorers (the tell condition). Three of the groups were led by those who did not have any knowledge as to their high-score status or why they were chosen as leaders. This was to test whether or not the leader's ethical strength (as defined by the DIT) and influence would explain any potential gains or whether it was more likely the result of the deliberative process. Table 7.4.2 reveals which groups were in the tell groups versus the no-tell groups and the gains in their P-scores:

Table 7.4.2: Tell and no-tell conditions and their effects

	Tell	No tell	Differential in scores
Group 1	√		22
Group 2		√	12
Group 3	√		12
Group 4		√	14
Group 5	√		12
Group 6		√	04
Group 7	√		03

The average increase in the no-tell condition (groups 2, 4, and 6) was 10, while the average increase in the tell condition (groups 1, 3, 5, and 7) was 12.3. This is not a statistically significant difference. Although the sample size is too small to rule out an effect of high-score knowledge, an actual difference of between 10 and 12.3 is arguably minimal, suggesting that the leaders did not overly influence the P-score outcomes of their groups. The gains can therefore be assumed to have come from the deliberative process—collectively discussing, conferring, and settling on the best moral answer.

8. Conclusions, Limitations, and Further Research

Posing ethical scenarios as a means of eliciting ethical thought can be indicative of *something*, but it would be wrong to assume that it is indicative of moral reasoning abilities. As such, it cannot be assumed that the two different student responses in the faculty member's story meant much of anything, at least not in everyday application. Individuals need some verbal assistance to express what they are capable of cognitively. The DIT was designed to specifically offer that assistance.

When working interpreters were given that verbal assistance, it did not appear to elevate their expressions of principled reasoning. If the normative messages of the interpreting profession consistently reflect a rigid, maintaining-norms approach to ethics, then perhaps their conventional reasoning at work bled over into their broader moral reasoning in life. This was one of the considered conclusions of Dean's (2015) study, now reconsidered with greater credence, given the results of this study. In other words, because the students consistently scored higher than those in their age group, and the working interpreters scored much lower, is it possible to consider that exposure to professional discourse over their years of practice made them this way? The data showed that interpreting students consistently scored better than practitioners, even without their age and educational advantages. Therefore, the hypothesis of the profession's attracting conventional thinkers is not born out by these data.

Interpreting students also did better on average when they worked together to negotiate responses than when they worked alone. The results of this second study on collective DITs are encouraging because they appear to add some evidence to the traditional design of group discussion and deliberation employed in ethics courses and in professional development techniques, such as case report and case analysis, often used in the reflective practice of supervision (Curtis, 2017; Dean & Pollard, 2011, 2013; Hetherington, 2011; Interpreting Institute for Reflection-in-Action and Supervision, n.d.).

The generalizability of these data has limitations. The program in which all students were enrolled espouses a type of ethics across the curriculum design. Demand control schema (DC-S), a work analysis and decision-making tool, is taught from the beginning of the program and either reinforced throughout or used more extensively in subsequent courses. DC-S directly teaches principled reasoning—it requires an assessment of the context, including the values of the setting and a consideration of each stakeholder's communication aims and preferences (Dean & Pollard, 2011, 2013). It also encourages behavioral flexibility as a means of working cooperatively toward shared values (Dean & Pollard, 2011, 2013). This study was not designed to investigate the impact of DC-S on DIT scores, although that could be a topic for future research. A follow-up study with these students-now-practitioners to measure to what degree their reasoning patterns might have shifted, as is evidenced in other professions, has been discussed. Future research should also seek to increase the reliability of the data by replicating the study with greater diversity in types of interpreters (e.g., signed language interpreters in other countries, interpreters who work as deaf interpreters, and spoken language interpreters) in addition to greater sample sizes. Lastly, it would be useful to use the DIT to account for the reasoning abilities of teachers and mentors who influence students' ethical-reasoning skills.

9. Acknowledgements

We wish to thank all the students who were willing to participate in this research over 3 years.

10. References

- Bebeau, M. J. (2002). The defining issues test and the four component model: Contributions to professional education. *Journal of Moral Education*, 31, 271–295.
- Bebeau, M. J., & Faber-Langendoen, K. (2014). *Remediating lapses in professionalism, remediation in medical education*. Springer.

- Bergson, M., & Sperlinger, D. (2003). "I still don't know what I should have done": Reflections on personal/professional dilemmas in sign language interpreting. *Deaf Worlds*, 19, 6–23.
- Cartwright, B. E. (2009). *Encounters with reality: 1,001 interpreter scenarios*. RID Press.
- Center for the Study of Ethical Development. (n.d.). *About the DIT*. <http://ethicaldevelopment.ua.edu>
- Christensen, A. L., Cote, J., & Latham, C. K. (2016). Insights regarding the applicability of the defining issues test to advance ethics research with accounting students: A meta-analytic review. *Journal of Business Ethics*, 133(1), 141–163.
- Curtis, J. (2017). *Supervision in signed language interpreting: Benefits for the field and practitioners* [unpublished master's thesis, Western Oregon University].
- Dean, R. K. (2014). Condemned to repetition? An analysis of problem-setting and problem-solving in sign language interpreting ethics. *Translation and Interpreting*, 6, 60–75. <https://doi.org/10.12807/ti.106201.2014.a04>
- Dean, R. K. (2015). Sign language interpreters' ethical discourse and moral reasoning patterns [PhD thesis]. Heriot-Watt University.
- Dean, R. K., & Pollard, R. Q. (2009). Effectiveness of observation-supervision training in community mental health interpreting settings. *REDIT E-journal on the Didactics of Translation and Interpreting*, 3, 1–17.
- Dean, R. K., & Pollard, R. Q. (2011). Context-based ethical reasoning in interpreting: A demand control schema perspective. *Interpreter and Translator Trainer*, 5, 155–182.
- Dean, R. K., & Pollard, R. Q. (2013). *The demand control schema: Interpreting as a practice profession*. CreateSpace Independent Publishing Platform.
- Drugan, J., & Megone, C. (2011). Bringing ethics into translator training: An integrated, inter-disciplinary approach. *Interpreter and Translator Trainer*, 5(1), 183–211.
- Dukerich, J. M., Nichols, M. L., Elm, D. R., & Vollrath, D. A. (1990). Moral reasoning in groups: Leaders make a difference. *Human Relations*, 43(5), 473–493.
- Gilligan, C. (1977). In a different voice: Women's conception of the self and morality. *Harvard Educational Review*, 49, 431–446.
- Hetherington, A. (2011). Supervision and the interpreting profession: Support and accountability through reflective practice. *International Journal of Interpreter Education*, 4, 46–57.
- Hill, A. L. (2004). Ethical analysis in counseling: A case for narrative ethics, moral visions, and virtue ethics. *Counseling and Values*, 48, 131–148.
- Hsieh, E. (2007). Interpreters as co-diagnosticians: Overlapping roles and services between providers and interpreters. *Social Science and Medicine*, 64, 924–937.
- Hsieh, E. (2009). "I am not a robot!" Interpreters' views of their roles in health care settings. *Qualitative Health Research*, 18, 1367–1383.
- Humphrey, J. (1999). *Decisions? Decisions! A practical guide for sign language professionals*. H and H Publishers.
- Interpreting Institute for Reflection-in-Action and Supervision. (n.d.). <https://www.iirasinternational.com>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus, and Giroux.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behaviour: Theory, research and social issues* (pp. 31–53). Holt, Reinhart and Wilson.
- McGeorge, C. (1975). Susceptibility to faking of the Defining Issues Test of moral development. *Developmental Psychology*, 11(1), 108.
- Marin, J. (2020). A medical interpreter training program and signed language interpreters' decision latitude: Exploring the impact of specialized training. In I. Souza, & E. E. Fragkou (Eds.), *Handbook of research on medical interpreting* (pp. 421–455). IGI Global.
- Mayhew, M. J., Pascarella, E. T., Trolan, T., & Selznick, B. (2014). Measurements matter: Taking the DIT-2 multiple times and college students' moral reasoning development. *Research in Higher Education*, 56(4), 378–396.
- Mendoza, E. (2012). Thinking through ethics: The processes of ethical decision making by novice and expert American Sign Language interpreters. *International Journal of Interpreter Education*, 4, 58–72.
- Martí-Vilar, M., Escrig-Espuig, J. M., & Merino-Soto, C. (2021). A systematic review of moral reasoning measures. *Current Psychology*, 1–15.
- Narvaez, D., & Bock, T. (2002). Moral schemas and tacit judgement or how the Defining Issues Test is supported by cognitive science. *Journal of Moral Education*, 31, 297–314.
- Nichols, M. L., & Day, V. E. (1982). A comparison of moral reasoning of groups and individuals on the "Defining Issues Test." *Academy of Management Journal*, 25(1), 201–208.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. Praeger.

- Rest, J. R. (1994). Background: Theory and research. In J. R. Rest & D. Narvaez (Eds.), *Moral development in the professions: Psychology and applied ethics* (pp. 1–26). Lawrence Erlbaum Associates.
- Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). A neo-Kohlbergian approach: The DIT and schema theory. *Educational Psychology Review*, *11*, 291–324.
- Roche, C., & Thoma, S. (2017). Insights from the Defining Issues Test on moral reasoning competencies development in community pharmacists. *American Journal of Pharmaceutical Education*, *81*(8), 5913.
- Seal, B. (1998). *Best practices in educational interpreting*. Allyn and Bacon.
- Stewart, D. A., & Lindsey, J. D. (1990). Code of ethics: Implications for interpreters for the deaf. *European Journal of Special Needs Education*, *5*, 211–220.
- Tate, G. T., & Turner, G. H. (2002). The code and the culture: *Sign language interpreting—in search of the new breed's ethics*. Douglas McLean.
- Thoma, S. J., & Dong, Y. (2014). The Defining Issues Test of moral judgment development. *Behavioral Development Bulletin*, *19*(3), 55–61. <http://dx.doi.org/10.1037/h0100590>
- Thomas, E. M. (2012). *In search of the newer breed's ethics—investigating the ethical decision making of British sign language interpreters with NVQ and postgraduate qualifications who qualified between 1999 and 2012* [unpublished master's thesis, University of Leeds].
- Turner, G. H., & Best, B. (2017). From defensive interpreting to effective professional practices. In M. Biagini, M. S. Boyd, & C. Monacelli (Eds.), *The changing role of the interpreter: Contextualising norms, ethics and quality standards*. Routledge Advances in Translation and Interpreting Studies. Routledge.
- Wilbeck, D. K. (2017). *An investigation of student perception how to better prepare Signed Language/English interpreters for the real world* [unpublished master's thesis, Western Oregon University].