



Integrating moral and social development within middle school social studies: a social cognitive domain approach

Larry Nucci, Michael W. Creane & Deborah W. Powers

To cite this article: Larry Nucci, Michael W. Creane & Deborah W. Powers (2015): Integrating moral and social development within middle school social studies: a social cognitive domain approach, Journal of Moral Education, DOI: [10.1080/03057240.2015.1087391](https://doi.org/10.1080/03057240.2015.1087391)

To link to this article: <http://dx.doi.org/10.1080/03057240.2015.1087391>



Published online: 23 Nov 2015.



Submit your article to this journal [↗](#)



Article views: 10



View related articles [↗](#)



View Crossmark data [↗](#)

Integrating moral and social development within middle school social studies: a social cognitive domain approach

Larry Nucci, Michael W. Creane and Deborah W. Powers
University of California, Berkeley, USA

Eleven teachers and 254 urban middle-school students comprised the sample of this study examining the social and moral development outcomes of the integration of social cognitive domain theory within regular classroom instruction. Participating teachers were trained to construct and implement history lessons that stimulated students' moral reasoning and conceptions of societal convention. In comparison with baselines and controls, teachers reduced didactic instruction and increased the proportion of class time devoted to small group discussions. Student engagement in transactive discourse significantly increased in participating classes with significantly greater post-test levels of moral reasoning, concepts of social convention, and cross-domain coordination. Student production of operational versus representational transacts through transactive discussion was associated with growth in moral and societal concepts. Teachers continued teaching lessons constructed in the project a year after the research ended.

Keywords: domain theory, moral education, social studies, transactive discourse

This article reports outcomes of the application of social cognitive domain theory (Smetana, 2014) for moral and social development within a school setting. Prior experimental studies pointed to the efficacy of domain theory for moral education (Nucci, 2009), and recent work on citizenship education has used a domain analysis to focus instruction on moral issues (Schuitema, van Boxtel, Veugelers, & ten Dam, 2011). However, no previous study examined whether the domain approach could be successfully adopted within mainstream public education. Domain theory draws a distinction between development of judgments about morality (issues of fairness, welfare, and rights) and concepts of societal convention (consensually determined norms of a given social system) (Smetana, 2014). Concepts within each domain follow independent courses of development, accounting for qualitatively differing

Corresponding author. Graduate School of Education, Cognition and Development, University of California, 4609 Tolman #1670, Berkeley, CA 94720-1670. Email: Nucci@berkeley.edu

aspects of social experience (Nucci & Turiel, 2009; Turiel, 1983). Contextualized social contexts may be multifaceted, including elements of both convention and morality. Decision-making in such multifaceted contexts may draw on concepts from more than one domain requiring cross-domain coordination (Smetana, 2014).

The basic principle of domain based moral education is *domain concordant instruction* in which the domain of issues under consideration is matched with the focus of discussion and written assignments. Prior experimental research using content from American History reviewed in Nucci (2009) demonstrated that development within domains of social convention and morality was enhanced when instruction was domain concordant and that growth did not occur when instruction was domain discordant (e.g., focusing discourse of an issue of convention around considerations of justice and welfare). The central goal of this study was to determine if the educational practices of the study's participating teachers would result in development of middle-school students' moral reasoning, concepts of societal convention, and spontaneous tendencies to coordinate across domains when responding to multifaceted social situations. Prior work with middle-school students indicated that their modal tendency is to subordinate a complex issue to a single domain (i.e., morality or convention) instead of incorporating and coordinating elements across domains (see Nucci, 2009). The middle school years are points of developmental transition in both domains.

Within the context of implementing a domain approach, our study had two aims. The first was to test whether engagement of students within what is referred to as *transactive discussion* (Berkowitz & Gibbs, 1983; Sionti, Ai, Rosé, & Resnick, 2011; Turner & Berkowitz, 2005) would account for observed growth in moral reasoning and understandings of convention. Transactive discussion is defined by Berkowitz, Althof, Turner, and Bloch (2008) as 'peer discussion where one discussant manifests discursive reasoning about another discussant's reasoning;' higher forms are Operational (transformation or extension of another's reasoning); lower forms are Representational (paraphrase of another's reasoning or juxtaposition of another's reasoning with one's own reasoning) (pp. 191–192). Prior research (Berkowitz & Gibbs, 1983) using dyads had demonstrated that the ratio of operational to representational transacts accounts for the effects on moral development. However, those findings have not been extensively explored within classroom settings where the pattern of conversation is more fluid. Moreover, much of the current research on transactive discourse comes from studies in the area of science education (e.g., Ai, Sionti, Wang, & Rosé, 2010; Sionti et al., 2011) rather than classroom studies of moral discourse.

In prior work using dyads, the analysis of transacts focused upon the entire conversation rather than on the production of transacts by individual members of a discussion. The assumption was that conversations rich in transacts provide for disequilibrating experiences contributing to growth for both participants (Berkowitz & Gibbs, 1983). In the present study, we assumed that transactive discussion would raise the level of moral and social conventional thinking of the classrooms as a whole. However, we also investigated the possibility that

individuals who were most active in producing ‘operational’ transactive statements would benefit more than students whose contributions were more frequently in the form of representational statements. The generation of operative transacts involves the production of statements that refute or extend the positions contained in the discourse. Findings from a recent study exploring the impact of transactive discourse upon both the actor’s and recipients growth, indicated that students benefit from both elaborating on their partner’s ideas and having their own ideas elaborated upon (Jurkowski and Hänze, 2015).

Our second aim was to engage classroom-based moral education that would be sustainable beyond our project. We attempted this by integrating educational practices for moral growth within regular academic instruction of middle-school history. History is a content area that is rich with moral issues of justice and welfare. It also reflects shifts in the conventions that regulate social systems, and conflicts between societal conventions, such as gender norms, and strivings for justice among members of society (e.g., women). Our goal was to avoid what Kohlberg (1985) referred to as the ‘psychologist’s fallacy’ of applying the findings of developmental psychology directly to classrooms. The failure of many prior efforts to sustain teacher engagement in moral education can be attributed to the disjunction between the educational goals of classroom teachers and to the commitment of time and effort associated with moral dilemma discussion directed at an educational goal defined by the researchers. More recently, there have been efforts to correct this disjunction through what is referred to as curriculum-oriented moral education (Schuitema, ten Dam, & Veugelers, 2008). The work of these European researchers also attended to the nature and quality of student discourse (other than transacts). They reported outcomes regarding student capacity for sustaining valid moral argumentation that were encouraging for the goals of this research (Schuitema et al., 2011). However, most prior efforts at moral education, even when integrated within the academic curriculum, as in the case of the European studies just cited, have provided teachers with ready-made curricula. In so doing, the researchers risk undercutting the engagement and investment of the classroom teachers to a set of lessons only partially related to their academic goals. In the present study, we addressed these shortcomings by engaging teachers in the construction of *their own* lessons within the context of their state-mandated curriculum.

We hypothesized that the research intervention would have an impact on the teachers and their approach to instruction as well as on students’ sociomoral development. As will be described in Methods, domain concordant instruction was ensured through collaboration with teachers in designing their lessons. Within the framework of these domain-concordant lessons, we hypothesized that there would be: (1) an increase in the proportion of class time students spent in discussion rather than didactic instruction; (2) an increase in the proportion of discussion that was transactive; (3) an increase in the ratio of operational to representational transacts used by students; and (4) an increase in the ratio of complex to simple forms of operational transacts employed in discussions.

With regard to students, we hypothesized that the intervention would: (1) result in the development of students' moral reasoning and their concepts of social convention; (2) increase student tendencies to employ coordination across domains when making judgments about multifaceted issues; (3) increase student engagement and class ratings for the moral development lessons relative to regular history lessons; and (4) result in students' self-reports of having learned as much or more history in the experimental than in control classrooms.

We had two specific hypotheses with regard to the relationship between the patterns of discussion and student development:

- (1) The development of individual students would be associated with their degree of usage of operational to representational transacts.
- (2) Development would be positively related to the ratio of higher to lower order operational transacts.

Finally, we hypothesized that teachers would value the inclusion of a focus on moral development within their history courses and would continue to engage in the use of moral development lessons beyond the end of the research project. We predicted that the majority of teachers would continue to use the lessons that they created a year after the project officially ended.

Methods

Participants

Participants were eleven middle-school teachers (eight participating; three control) and their 254 students from three middle schools within an urban public school district in northern California. Table 1 presents the student demographics, Academic Performance Indicator (API) scores on state tests, and distribution of participating and control teachers by grade for each school. As can be seen in Table 1 the control teachers were drawn from the two higher performing schools to ensure that any positive differences in social and moral development favoring the participating classrooms could not be attributed to a prior history of student academic achievement.

Teacher preparation and support

In October of the project year, the first author provided a series of three two-hour workshops that provided background in domain theory and development in cognition of morality and social convention in early adolescence, direction in the construction of lessons that integrated attention to morality and social convention within the history curriculum, and experiential guidance in how to engage students in transactive discussions. Following the workshops, teachers began constructing their own lesson plans to which the authors provided feedback. This feedback was designed to ensure that lessons over the course of the school year addressed issues in both the moral and conventional domains and that the design of student

Table 1. Project schools, teachers, and students, 2013

	School 1	School 2	School 3
<i>Distribution of Project Teachers</i>			
Grade 6	1 control	1 participating	2 participating
Grade 7	1 participating	1 participating, 1 control	1 participating
Grade 8	1 participating	1 control	1 participating
<i>Ethnic Breakdown of Student Population</i>			
% African American	37	36	61
% Asian	9	30	5
% Latino	19	16	12
% Multi-ethnic, other, or non-identified	9	6	8
% White	26	13	14
% Qualified for Free and Reduced Lunch;	42;	66;	56;
API Score;	789;	773;	713;
State Rank from 1 (lowest) to 10 (highest)	5	4	2

Note. Source of Academic Performance Index (API) scores and student demographics information is from the California Department of Education.

discussions and written work was domain concordant. The teachers, who were free to accept or reject suggestions, then taught the lessons. In all cases, the teachers accepted the authors' suggestions. The teachers created 44 lessons distributed across the academic year beginning in the late fall and continuing into early June. Lessons included in our analyses each lasted one 50-minute class period with the exception of four lessons that took place on shortened days with 35 minute periods. The majority of lessons (88%) were stand-alone topics. In the remaining cases two or three lessons followed one another to complete a theme. For example, a sixth grade World History teacher used three lessons to explore the moral and social conventional elements of Hammurabi's code.

Project data and assessments

Classroom observations and treatment verification. For both control and participating teachers, we conducted observations that allowed us to determine whether our intervention altered teacher behavior in line with project goals. We asked each teacher to select one class period as a focus for observations and assessments. Observations included time-series checklists, a written running narratives of classroom events, and audio recordings of classroom discourse. When classes were broken into small groups, we placed a digital audio recorder within each group. The recordings were transcribed for later coding and analysis.

The time-series checklist included categories to capture the overall structure of the classroom and the specific mode of instruction. Observers entered check marks every five minutes to capture classroom activities. The categories for classroom

Downloaded by [University of California, Berkeley], [Larry Nucci] at 12:49 04 December 2015

structure were: whole class, individual and silent work, pairs or small group work, and transition. The categories for teacher activity were: provision of procedures and directions, didactic instruction, question–answer (Q&A), organic discussion, use of media, silent observing class, and attention to individual student.

We conducted three baseline observations of participating teachers prior to the onset of project workshops. Our instructions to participating teachers were to teach lessons integrating concerns for issues of justice and human welfare. We provided these same instructions to our control teachers.

Analysis of classroom discourse. Our analyses focused on capturing the relative usage of transactive forms of discourse within discussions. Our coding scheme was adapted from Berkowitz and Gibbs (1983) and modified to include codes developed by Ai et al. (2010) and Sionti et al. (2011). We incorporated these latter codes to capture contextual elements of speech that occur in dynamic classroom environments more frequently than in the peer dyad discussions that were the focus of the original studies by Berkowitz and Gibbs (1983). We also developed two additional transactive subcodes (simple agreement–disagreement and recognition of incongruity) to capture simple interjections (e.g., ‘yeah,’ ‘nah,’ ‘that’s strange’) that appear more often in the course of small-group classroom discourse.

Transactive speech acts were subdivided into three types: Representational, Operational, and Hybrid transacts. Representational transacts are statements that refer to statements of others without acting upon them. Operational transacts are statements that act on another’s reasoning. *Hybrid* transacts contain elements of both representational and operational transacts. Our coding scheme resulted in ten sub-codes of operational transacts that describe an ascending order of complexity divided into two categories: lower level and higher level. Forms of transacts such as offering a *Comparative Critique* were considered on the higher end of the spectrum since they involve more complex reasoning upon one’s own, another’s, or multiple discussants’ reasoning. Table 2 presents the primary categories that captured student and teacher statements.

Assessments of students’ moral and social reasoning. This study employed domain-theory-based assessments of students’ moral development and concepts of societal convention. We elected to use these assessments rather than established measures of moral growth such as the DIT because those measures are predicated on Kohlberg (1985) stage-theory assumptions that moral thinking within adolescence is conflated with conceptions of convention. Our goal was to assess shifts in students’ reasoning within each domain separately as well as to obtain estimates of shifts in students’ tendencies to engage in cross-domain coordination. The levels assessed are based upon findings of major shifts in moral reasoning and concepts of convention from late childhood through early and middle adolescence (Nucci & Turiel, 2009; Turiel, 1983). These assessments were in the form of in-class short-response essays. The classroom teachers were given copies of the completed assessments for their own

Table 2. Classroom discourse codes

A. PURPOSE OF STATEMENT

1. **Off Task:** blatantly off-task contributions.
2. **Management:** management moves or announcements.
3. **On Topic:** statement related to topic being discussed.

B. NON-TRANSACTIVE SPEECH ACT CODES

1. **Elicitational:** eliciting information (e.g., What do you mean?).
2. **Externalization:** Offering an opinion, position, or stance without transacting with another individual's statement.

C. TRANSACTIVE SPEECH ACTS**Representational**

1. **Simple Agreement/Disagreement:** 'Yeah; I agree'; 'I disagree'; 'No, that's not right.'
2. **Recognition of Incongruity:** (Refers to other's reasoning.) This is an issue we haven't yet resolved (e.g., 'That's weird', 'How strange', 'How odd.')
3. **Feedback Request:** (Refers to one's own reasoning.) 'Do you understand or agree with my position?'
4. **Paraphrase:** (Refers to other's reasoning.)
5. **Justification Request:** (Refers to other's reasoning.) 'Why do you say that?'
6. **Dyad Paraphrase:** (Refers to two individuals' joint reasoning.) Here is a paraphrase of a shared position. 'I see that we both think X.'

Operational (Operates on the representation of another individual's reasoning.)

Lower Level Operations

1. **Clarification:** No, what I am trying to say is the following.
2. **Competitive Clarification:** My position is not necessarily what you take it to be.
3. **Refinement:**
 - (a) I must refine my position or point as a concession to your position or point (Subordinative mode).
 - (b) I can elaborate or qualify my position to defend against your critique (Superordinative mode). (Refers to own reasoning.)
4. **Extension:**

Here is a further thought or an elaboration offered in the spirit of your position.
5. **Contradiction:** There is a logical inconsistency in your reasoning.

Higher Level Operations

6. **Reasoning Critique:**
 - (a) Your reasoning misses an important distinction, or involves a superfluous distinction.
 - (b) Your position implicitly involves an assumption that is questionable
 - (c) Your reasoning does not necessarily lead to your conclusion.
 - (d) Your reasoning applies equally well to the opposite opinion.
7. **Competitive Extension:**
 - (a) Would you go to this implausible extreme with your reasoning?
 - (b) Your reasoning can be extended to the following extreme, with which neither of us would agree.
8. **Counter Consideration:** Here is a thought or element that cannot be incorporated into your position.
9. **Common Ground/Integration:**
 - (a) We can combine our positions into a common view.
10. **Comparative Critique:**
 - (a) Your reasoning is less adequate than mine because it is incompatible with the important consideration here.
 - (b) Your position makes a distinction which is seen as superfluous in light of my position, or misses an important distinction which my position makes.
 - (c) I can analyze your example to show that it does not pose a challenge to my position.

use. Thus, these assessments were part of the students' regular class assignments. Participating students completed these essay assessments at the beginning and toward the end of the school year. Students in the control classrooms completed the essay assessments at the end of the school year. Each assessment included four topics. Three were assessments of the students' levels of moral reasoning, concepts of societal convention, and their coordination of moral and conventional elements in a multifaceted context. The fourth essay concerned an issue that most adolescents would view as a matter of personal choice, included to reduce a response bias toward interpreting the essays as asking students to judge the actions in terms of the social expectations of their teachers and other authorities.

Assessment of moral development. The assessment of moral development was adapted from the interview protocol and coding scheme employed to generate levels of moral reasoning reported in previous studies (Nucci & Turiel, 2009). A detailed discussion of this assessment is beyond the scope of a journal article. Because this is a new assessment, we provide some illustrative examples along with brief descriptions of the basic levels assessed in this study. The scenario in the examples depicts a situation in which the protagonist must decide whether to return money dropped by a stranger. Students responded to two essay questions: 'Would it be wrong or all right for the protagonist to keep the money,' and 'Would the protagonist have a right to keep the money if he or she so chose?' The second question was an indicator of whether the students felt that the protagonist is bound by a judgment that keeping the money is wrong.

Participant responses were coded into three levels that reflected the students' coordination of potentially ambiguous moral elements ('grey areas') and the students' perceptions of the context as a matter choice (a personal domain matter). Responses at Level 1, classified as *Simple/Straightforward* reasoning, were dominated by straight-forward considerations of harm and welfare. For these students, the moral situation posed little ambiguity. This was not because the students were inattentive since several of the story elements might be identified within a given student essay. However, there was no evidence in the responses at this level that the students attempted to integrate those factors into their judgments. This form of reasoning is illustrated in the following example. (These were free-response essays; students were not given time to edit for grammar or spelling.)

- Question: Would it be wrong or all right for Brandon to keep the money?
 Student Response: It would be wrong. If he pick it up and keeping the money that means that he is taking stuff from other people so it's like stealing.
- Question: Suppose Brandon decides to keep the money. Does he have the right to keep the money if he wants to?
 Student Response: No, because it's someone else's \$10 bill.

Responses coded as Level 2, *Uncoordinated*, were characterized by attention to elements that added complexity and lent a 'grey area' to their moral judgments.

These sources of complexity included the fact that the protagonist did not set out to take the other person's money and did not actively cause the money to fall to the ground. The arguments offered by students at this level often disputed continuous ownership of the money once it hit the ground in a public location. In addition, the students at this level were more likely to link these ambiguities to arguments that included considerations of personal choice and free will, often connected to statements about the US as free country. A primary characteristic of reasoning at this level was lack of integration of disparate moral arguments to support judgments, illustrated by the following excerpt from a Level-2 child quoted in (Nucci & Turiel, 2009): 'He's not necessarily doing something wrong, but the right thing to do would be to give it back, but he's not necessarily, he doesn't necessarily have any wrongdoing.' The following excerpt illustrates reasoning at Level 2.

- Question: Suppose Brandon decides to keep the money. Does he have the right to keep the money if he wants to?
- Answer: It is the same thing as finding \$10 on the bus with nobody else around. Lost money is up for grabs. He has the right to keep it because that person dropped and it's a free country. He has a right because it didn't said it was the property of the lady, it just said the property of the United States, and he is an American.

Student responses coded as Level 3 also addressed the potential complexity and ambiguity posed in a moral situation. They acknowledged the issues raised as sources of ambiguity by students' reasoning coded as Level 2 as in this example: 'Now, if he just found money on the bus and didn't see the man drop it, it's OK to keep it because he couldn't have returned it to anyone.' Responses coded as Level 3, however, did not end with this 'finders keepers' argument. In the following excerpt, the student explains how this situation is not simply a matter of finding 'lost' money in response to the question, 'Would Brandon have a right to keep the money if that is what he decides to do?':

Brandon has not earned the money and even if the man dropped the money it does not mean that it has stopped being the man's because there is a possibility he can get it back. Brandon does not deserve that money and it is not his right to keep because he could return the man's money, which is his, the man's.

The distribution of action decisions were similar for students employing Level 3 moral reasoning to Level 1. For students at Level 3, however, this decision entailed accounting for potentially mitigating situational factors that were not addressed as part of Level-1 moral reasoning. These shifts in moral thinking reflect changes in the ways in which the elements of the situation are evaluated in *moral* terms and the salience of these elements for rendering a moral decision. Similar age-related patterns were found in judgments about helping (Nucci & Turiel, 2009). A complete analysis of the elements that comprise moral judgments at each level for this scenario, as well as others entailing issues of harm and helping, are available from the first author on request.

Assessment of concepts of societal convention. The assessment of students' levels of understanding societal convention employed a measure developed in 2004 by Nucci, Becker, and Horn (2004) based upon the levels of conceptual development of social convention described by Turiel (1983). The measure we employed was developed for use with American history classes. As with morality, there were three levels of development across the age groups.

Responses coded as Level 1 generally affirmed the convention under consideration based on the connection between convention and a concrete understanding of social hierarchy. Beyond this basic notion of hierarchy, there was no evidence that a given convention was connected to the structuring or organization of society as a social system.

At Level 2, the function and meaning of social convention is reevaluated in terms of the consensual and arbitrary basis of conventions. Level-2 responses negated conventions as merely the 'dictates of authority' or the views of a particular group. The conventions themselves were seen as having no prescriptive force and were thus judged not especially important. There was no connection in these responses between individual conventions and societies as normative systems.

At Level 3, conventions are affirmed as constituent components of social systems. Conventions are understood to be arbitrary in and of themselves, but collectively form the 'rules of the game' for members of a social system. There is at this level a clear connection between social conventions and societies as normative systems.

Assessment of cross-domain coordination. In addition to estimates of students' levels of development within each domain, we assessed their tendencies to coordinate moral and conventional elements in contexts that pitted moral considerations for fairness and welfare against prevailing societal conventions. For example, one scenario depicted a situation in which the government of India offered money to build a school if it served both genders. However, the male leaders of the community turned down the offer because it contradicted their cultural position of girls and women. Students were asked to respond to two questions: (1) Was the decision made by the men right or wrong; (2) Is there a resolution to the given situation?

Following procedures developed by Nucci and Weber (see Nucci, 2009), we coded responses into two types. Type-1 responses subordinated these overlapping issues to a single domain. Type-2 responses reflected an effort to coordinate moral and conventional elements. We further divided these responses into two levels. Responses coded as 2A reflected efforts by the student to address concerns from both domains without, however, fully resolving at a societal level the source of the underlying conflict. Responses coded as 2B entailed domain coordinations where students identified elements from both domains using these elements to develop a system-level solution.

Students' evaluations of teachers, lessons and amount learned

Toward the end of the school year, we distributed a Likert-format survey to all of the participating and control students to get anonymous evaluations of their teachers, their history classes, and their sense of how much they felt they had learned over the course of the year. Students in participating classes also provided an evaluation of the moral-development lessons.

Teachers' evaluations of lessons, the project and second year follow-up

At the close of the school year, we obtained feedback from participating teachers regarding the project. Teachers completed a survey asking on a five-point Likert scale from 'strongly agree' to 'strongly disagree' how much they concurred with 11 positively and negatively worded statements about the following: the professional development sessions associated with the project, how much they learned about moral and social development of their students, the degree of new insights they learned about teaching history, how much students benefitted from the lessons, whether they would recommend the project to colleagues, and whether they would continue teaching the lessons in the future. In addition to the survey data, we engaged the teachers in a year-end open meeting in which they were asked to share reactions to their experiences in the project and their overall sense of the value to students of including moral and social conventional issues within the teaching of history. This discussion was audio-recorded and transcribed for later analysis. The transcript was shared with the district coordinator for social studies.

In the year following the end of project activities, we contacted the participating teachers via email to fill out a survey indicating whether or not they were using any of the project lessons (if so, how many) and whether they had created new lessons.

Results

Impact on classroom structure and teaching practices

Comparison of control and participating classrooms during baseline lessons. An initial analysis of the patterns of teaching between the baseline lessons of participating teachers and control classrooms indicated that the overall approach to instruction was quite similar for both groups. Whole-class instruction and individual seat-work took up the majority of class time with small-group and pairs discussions, accounting for less than 25% of class time for the control classrooms and participating classroom baseline lessons. The modes of instruction were also similar. The primary modes of instruction for both were didactic instruction (25–31% class time), individual seatwork (10–24% class time), and Q&A sessions (14–17% of class time). The average engagement of students in whole-class organic discussions was actually higher for our control teachers (14% of class time) than during the baseline lessons of our participating teachers (5% of class time), although in neither case was engagement of students in free-flowing discussion a predominant aspect

of their classroom teaching. Finally, the range of scores on all of the observed measures for the control teachers fell within the range of scores for the participating teachers. Thus, we do not attribute our study outcomes to a systematic pre-selection of control or participating teachers.

Comparison of baseline and project lessons of participating classrooms

There were marked shifts both in classroom format and teaching practices by the participating teachers between the baseline and project lessons. These changes were in the direction consistent with the educational practices promoted in project workshops. Seven of the eight participating teachers reduced the amount of class time spent in whole-class formats and increased the amount of time spent in pairs and small-group student discussions (binomial $p < .05$). Summing across teachers, the ratio of whole-classroom to small-group instructional formats shifted from a 2:1 ratio during baseline lessons to nearly 1:1 during the project lessons. This shift toward small-group formats was designed to promote an increase in student participation in discussion and to a corollary increase in the engagement of individual students as active contributors to transactive discussions of issues of morality and societal convention.

In line with this goal, we also observed a significant shift in teachers' modes of instruction when engaging with students during whole-classroom instruction. All eight teachers reduced usage of didactic instruction and increased engagement with students in organic discussions (binomial $p < .001$). On average, the teachers dropped the amount of didactic instruction by two-thirds and increased engagement of students in organic discussion by a similar amount.

Impact on usage of transactive forms of discourse. Transcripts of discourses were coded by primary and reliability coders. We calculated interrater reliability using Cohen's Kappa among the categories of speech acts and transactivity with scores ranging from .72 to .82. In line with changes observed in teachers' modes of instruction, we saw significant shifts in proportions of transactive rather than nontransactive statements or questions in classroom conversations. This analysis included all discourse that took place within the class period. Within project classrooms, rates of nontransactive student-generated speech acts shifted from 80% during baseline to 51% during project lessons. Moreover, the majority of students' transactive statements (51%) during project lessons were higher-order transacts.

In sum, teachers changed their classroom structures and modes of instruction in line with the goals of teacher inservices. These shifts in the modes of instruction were associated with parallel changes in student classroom discourse aligned with our project goals. Thus, we felt confident that we achieved the intended experimental treatment effects with participating teachers and students.

Impact on students' moral reasoning, concepts of social convention, and tendencies to engage in domain coordination

Issues of scoring reliability. Primary and reliability coders also scored students' responses to essay assessments of moral reasoning, concepts of convention, and forms of cross-domain coordination. All of the student essays were stripped of identifiers and randomized by grade level and treatment condition. The primary coder scored the essays for all 254 students. The reliability coder scored the first 50 sets of essays of the first 100 assessments, and then 50 of the next 100, and 25 of the final 54.

We ran ongoing reliability estimates using Cohen's Kappa in order to control for potential rater drift in the scores provided by the primary coder. Interrater reliabilities ranged from .82 to .88. For the 125 transcripts scored by both raters, all scoring discrepancies were resolved and a single score assigned.

Impact on students' moral reasoning. Figure 1 presents the levels of moral reasoning for students in control classrooms and for participating students at pre-test and post-test in grades six through eight.

As we had expected, participating students' moral-reasoning post-test scores were significantly higher than their pre-test scores (Grade 6, $t = 3.83$, $p < .001$; Grade 7, $t = 3.61$, $p < .001$; Grade 8, $t = 1.99$, $p < .03$) and scores for control students (Grade 6, $t = 1.85$, $p = .03$; Grade 7, $t = 2.59$, $p = .005$; Grade 8, $t = 2.55$, $p = .006$) at each grade level. Moreover, control students' scores were not significantly different from participating students' pre-test scores at each grade level.

Impact on students' concepts of social convention. Figure 2 presents students' levels of understanding of societal convention.

Participating students' post-test scores were significantly higher than pre-test scores for grades seven ($t = 1.84$, $p < .04$) and grade eight ($t = 4.68$, $p < .001$), but

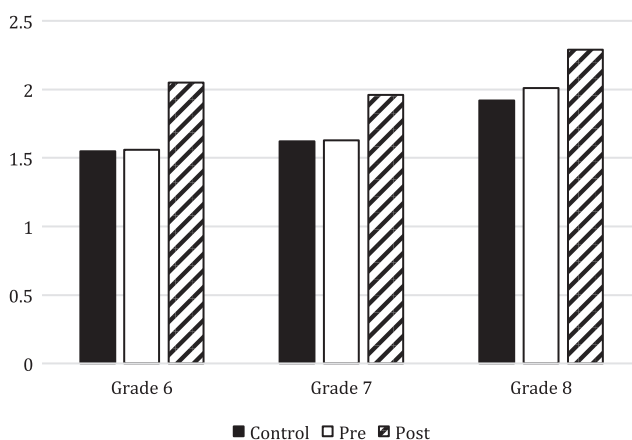


Figure 1. Impact on students' moral reasoning

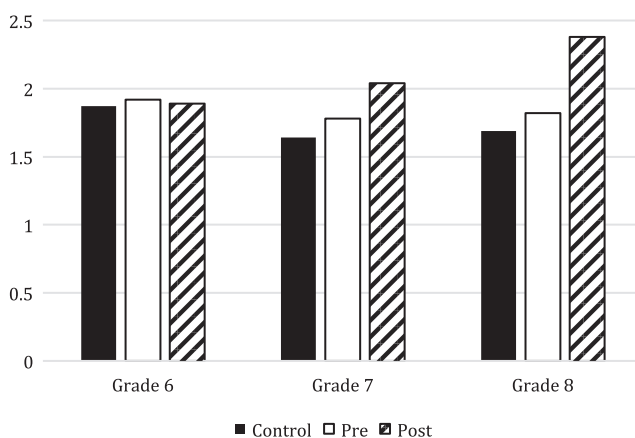


Figure 2. Impact on students' concepts of social convention

not for grade six ($t = .26, p = .60$). Similarly, participating students' post-test scores were significantly higher than those of control students for grades seven ($t = 1.95, p < .03$) and grade eight ($t = 2.7, p < .005$) but not for grade six ($t = .2, p < .45$). Thus, as we had expected, the educational intervention did raise students' levels of understanding about convention. However, this effect did not occur with our youngest participants.

Impact on students' cross-domain coordination. There were too few students who provided the highest Type-2B form of coordination to include them separately within our final analysis. Thus, we collapsed Types 2A and 2B as coordinated reasoning. When we further examined the data from the 20 students who provided Type-2B coordinated reasoning, we found that 19 of them also were at Level 3 in their understandings of societal convention, indicating a possible developmental prerequisite for Type-2B coordination of moral and conventional elements. Chi-square tests indicated that there was a significantly greater proportion of participating students who provided coordinated reasoning at the post-test than pre-test at each grade level: (Grade 6, $X^2 1, 104 = 14.5, p = .0001$; Grade 7, $X^2 1, 124 = 10.3, p = .0014$; Grade 8, $X^2 1, 91 = 2.6, p = .05$). Also, as expected, a higher proportion of participating students provided coordinated responses to multifaceted scenarios than did control students at grade levels six and eight (Grade 6, $X^2 1, 83 = 144.0, p < .0001$; Grade 7, $X^2 1, 93 = .16, p = .7$; Grade 8, $X^2 1, 75 = 16.0, p < .0001$).

Relationship Between Engagement in Transactive Discourse and Students' Development

To determine whether the observed student growth was associated with the degree to which students' generated operational rather than representational transacts during class discussions, we identified students who had provided at least four lines of conversation in our transcripts and who had completed all of the pre and post-assessments. We combined the scores obtained by these participating students

on our developmental assessments to generate an overall sociomoral cognition score for each student at pre-test and post-test. We then subtracted the pre-test from post-test scores to generate a growth score and divided the students at the midpoint into two groups: high-change and low-change.

Following procedures established by Berkowitz and Gibbs (1983), we also calculated the ratio of operational to representational transacts for each student. For the purpose of analyzing the ratio of operational to representational transacts, the hybrid transacts were merged with the lower-order operational statements. These procedures allowed us to account for students' usage of hybrid transacts in the most conservative manner with respect to our study hypotheses. We then compared this transactive discourse ratio for the low- and high-change groups. This analysis revealed that the transactive discourse ratio for high-change students ($M = 3.67$, $SD = 2.7$) was significantly greater than the ratio for low-change group ($M = .57$, $SD = .57$) ($t = 4.8$, $p < .0001$). We also analyzed whether amount of growth was associated with the degree to which students generated higher-order rather than lower-order operational transacts. Using the same procedure as in the previous analysis, we calculated the ratio of higher-order to lower-order operational transacts and compared the resulting ratios for the high-change and low-change students. This analysis revealed that the operational transact ratio for the high-change students ($M = .58$, $SD = .78$) was significantly greater than the ratio for the low-change group ($M = 0.11$, $SD = .27$) ($t = 2.46$, $p < .02$).

Student and teacher evaluations of the educational intervention

Students' ratings of teachers, courses and project lessons. There were no significant differences in the ratings of teachers provided by control and participating students. Using a four-point scale with 4 = excellent being the highest rating, both control and participating students rated their teachers equally (control $M = 3.62$, $SD = .64$; participating $M = 3.58$, $SD = .61$; $t = .49$). However, employing a similar four-point scale, participating students rated their history classes significantly higher ($M = 3.34$, $SD = .76$) than did students in control classes ($M = 3.18$, $SD = .75$, $t = 1.61$, $p = .05$). In addition, in response to a question asking students about the amount learned in history classes (from 1 = Nothing at all to 4 = A great deal), participating students provided a significantly higher mean rating ($M = 3.53$, $SD = .55$) than the control students ($M = 3.38$, $SD = .58$, $t = 1.94$, $p = .03$). Thus, although the students in both the participating and control classes had equally favorable views of their teachers, the students in the participating classes indicated that they had learned more history and rated their history classes higher than did control students. Finally, participating students evaluated project lessons that included a focus upon issues of morality and convention as 'more enjoyable than most of the other lessons' (52%), or 'among my favorites' (23%). Only 2% of participating students indicated that they did not like the project lessons 'at all,' $X^2(3, N = 169) = 50.64$, $p < .0001$.

Teachers project evaluations and continued engagement in moral development instruction. Teachers provided ratings ranging from 4.25 to 4.85 on a five-point scale (with five being the highest value) in response to questions on an assessment having to do with their perceptions of the value of the lessons for student moral development and whether they would recommend the project to their colleagues. These ratings indicating general satisfaction with their experience in the project were sustained by statements the teachers made during a focus group held at the end of the year. More compelling, however, were two other outcomes. First, our follow-up surveys conducted with project teachers during the year following the end of the project and 18-months later during the following year revealed that seven of the eight teachers were continuing to employ lessons integrating morality and social convention. Finally, in response to the positive outcomes of the project and the positive reactions from the participating teachers, the district requested that we extend our work to provide training to *teacher leaders* at all 21 of the district's middle schools.

Discussion

This study examined the impact of an educational application of social cognitive domain theory in which participating middle-school teachers constructed and implemented history lessons that incorporated attention to the issues of morality and societal convention embedded within the academic history content. This instructional strategy was combined with a shift in the ways in which teachers structured their classroom time and in the patterns of discourse. Compared with control teachers and their own teaching during baseline lessons, participating teachers reduced the amount of class time spent in whole-class formats and increased the amount of time their students spent in small-group discussions. These shifts in classroom structure were accompanied by an overall shift away from didactic forms of instruction toward an increase in the amount of organic discussion in whole-class formats and in increased engagement in transactive forms of discourse among participating students in both whole-class and small-group discussions. These teaching practices were in turn associated with increases in students' levels of moral reasoning, their concepts of societal convention, and their spontaneous tendencies to employ cross-domain coordination when reasoning about multifaceted issues.

These findings were in line with the outcomes of prior experimental work (see Nucci, 2009) illustrating the importance of attention to social-cognitive domains for the development of social and moral cognition. We saw growth in the moral domain across grade levels exceeding that of control groups. Thus, the growth cannot be simply attributed to normal developmental changes over the course of the school year. We saw similar growth in the conventional domain among students in the seventh and eighth grades but not in grade six. This may be because the sixth-grade students were still firmly within their level of development and

refractory to the educational experiences afforded in this study. Whether this was the case, or if we would have seen growth at this grade level with more lessons attending to matters of societal convention, will await future studies. The overall findings for growth across grades in the moral and conventional domains as a result of the educational intervention provide additional confirmatory evidence for these descriptions of development within each domain (Nucci & Turiel, 2009; Turiel, 1983). Finally, as expected, we observed an increase in students' spontaneous tendencies to engage in cross-domain coordination between pre-test and post-test across grade levels. These post-test levels were also higher than those of the control group for students in the sixth and eighth grades but not seventh grade. This lack of difference at the seventh grade was due to the relatively high rate of cross-domain coordination exhibited by the seventh-grade controls in comparison the other control groups. It is not clear why that occurred.

By examining the relationship between developmental outcomes and the degree to which students employed transactive speech acts in their discussions, we were able to demonstrate that the shifts in sociomoral cognition were the result of the students' engagement in transactive discourse. These findings reinforce and extend outcomes from prior experimental work conducted with adults (Berkowitz & Gibbs, 1983) and high-school-aged dyads (Sionti et al., 2011), indicating that the generation of operational transactive speech acts is particularly effective in stimulating cognitive growth. Findings regarding the association of amount of engagement by individual students in operational transacts for their growth of moral and conventional concepts adds to research (Jurkowski & Hänze, 2015) indicating that the production of statements that build from those of others is associated with development. The engagement in these discussions occurred organically among groups and, in some cases, within entire classrooms. These findings accord with recent reports indicating that eighth-grade students are quite capable of maintaining value-laden arguments in the context of discussion around historical events (Schuitema et al., 2011). Thus, this study adds to an emerging literature that these younger students are capable of generating sophisticated forms of discourse and that such discussion can take place effectively within normal classroom settings.

Finally, our survey outcomes revealed that both students and teachers had positive views of the study's approach and that both constituencies viewed the inclusion of moral education as adding to the educational value of their classes. Our most powerful piece of evidence for this positive reaction is that the teachers have continued to engage in the use of lessons developed in the project more than a year after the end of our study. Based upon teachers' statements, we attribute this to the fact that the teachers saw no discontinuity between the goals of moral education as enacted through this integration within their history classes and their academic goals as history teachers. Taken together with outcomes from other recent work (Schuitema et al., 2008), our findings indicate that possibly the most effective way to sustain moral education within school settings is to integrate it within the existing academic curriculum.

Notes for Contributor

“This work was supported by a research gift from the S. D. Bechtel Jr. Foundation.”

Larry Nucci, Graduate School of Education, Cognition and Development, University of California, 4609 Tolman Hall #1670, Berkeley, CA 94720-1670, USA. Email: nucci@berkeley.edu

References

- Ai, H., Sionti, M., Wang, Y. C., & Rosé, C. P. (2010). Finding transactive contributions in whole group classroom discussions. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Proceedings of the International Conference of the Learning Sciences* (pp. 976–983). Chicago, IL: International Society of the Learning Sciences.
- Berkowitz, M. W., Althof, W., Turner, V. D., & Bloch, D. (2008). Discourse, development, and education. In F. Oser & W. Veugelers (Eds.), *Getting involved: Global citizenship development and sources of moral values* (pp. 89–201). Amsterdam, Netherlands: Sense Publishers.
- Berkowitz, M. W. & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly*, *29*, 399–410.
- Jurkowski, S., & Hänze, M. (2015). How to increase the benefits of cooperation: Effects of training in transactive communication on cooperative learning. *British Journal of Educational Psychology*, *85* (3), 357–351. doi:[10.1111/bjep.12077](https://doi.org/10.1111/bjep.12077)
- Kohlberg, L. (1985). The just community approach to moral education in theory and practice. In M. W. Berkowitz & F. Oser (Eds.), *Moral education: Theory and application* (pp. 27–87). Hillsdale, NJ: Erlbaum.
- Nucci, L. (2009). *Nice is not enough: Facilitating moral development*. Saddle River, NJ: Merrill.
- Nucci, L., Becker, K., & Horn, S. (2004). Assessing the development of adolescent concepts of social convention. Paper presented at the annual meeting of the Jean Piaget Society. Toronto, CA.
- Nucci, L. & Turiel, E. (2009). Capturing the complexity of moral development and education. *Mind, Brain, and Education*, *3*, 151–159. doi:[10.1111/j.1751-228X.2009.01065.x](https://doi.org/10.1111/j.1751-228X.2009.01065.x)
- Schuitema, J., ten Dam, G. T., & Veugelers, W. (2008). Teaching strategies for moral education: A review. *Journal of Curriculum Studies*, *40*, 69–89. doi:[10.1080/00220270701294210](https://doi.org/10.1080/00220270701294210)
- Schuitema, J., van Boxtel, C., Veugelers, W., & ten Dam, G. (2011). The quality of student dialogue in citizenship education. *European Journal of Psychology of Education*, *26*, 85–107. doi:[10.1007/s10212-010-0038-1](https://doi.org/10.1007/s10212-010-0038-1)
- Sionti, M., Ai, H., Rosé, C. P., & Resnick, L. (2011). A framework for analyzing development of argumentation through classroom discussions. In N. Pinkwart & B. McClaren (Eds.), *Educational technologies for teaching argumentation skills* (pp. 28–55). Oak Park, IL: Bentham Science.
- Smetana, J. G. Jambon, M., & Ball, C. (2014). The social domain approach to children’s moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 23–45). New York, NY: Taylor and Francis. doi:[10.4324/9780203581957.ch2](https://doi.org/10.4324/9780203581957.ch2)
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.
- Turner, V. D. & Berkowitz, M. W. (2005). Scaffolding morality: Positioning a socio-cultural construct. *New Ideas in Psychology*, *23*, 174–184. doi:[10.1016/j.newideapsych.2006.04.002](https://doi.org/10.1016/j.newideapsych.2006.04.002)