



Volume 10 Number 1

IJHLTR

**International Journal of
Historical Learning,
Teaching and Research**

August 2011

www.history.org.uk

ISSN 1472-9466

In association with



UNIVERSITY of 
Cumbria

High School Students' Understanding Of Personal Betrayal In A Socio-Historical Context Of Ethnic Conflict: Implications For Teaching History

Michelle J. Bellino and Robert L. Selman,
Harvard Graduate School of Education

Abstract: *Research in history education frequently characterizes evaluative judgments of past events as ahistorical applications of contemporary ethical standards, citing “presentism” as a hindrance to disciplinary thinking. In accordance with this view, much of the recent educational research exploring adolescents’ capacity to understand human agency in past contexts has focused on their ability to cognitively interpret historical perspectives while minimizing the opportunity for evaluative components of those decisions. But little work has focused on how adolescents understand decision-making in historical contexts of conflict, where socio-historical understanding implies ethical reflection on the part of the historical agents. In this paper, we focus on one component of a new measure designed to assess adolescent historical understanding, where issues of intolerance and injustice are dominant (or “purposeful”) parts of the historical narrative to be understood. The authors analyzed the responses of the control sample of a national evaluation study (n=621), in which ninth and tenth grade students considered why one friend might betray another during a period of ethnic and religious conflict in the former Yugoslavia. Despite the risk of reinforcing the natural tendency to apply presentistic thinking to historical actors, our findings suggest that when asked for a descriptive interpretation of human agency, a significant number of participants in this study grappled with moral agency, not only from the perspective of historical actors, but also as moral agents in the present. We contextualize these findings within the pedagogical debate on the degree to which historical cases of social injustice can be taught in a way that helps students to integrate their natural inclination for ethical reflection with the intellectual rigor required of historical understanding.*

Keywords: *History education, Historical thinking, Ethical reflection, Moral development, Adolescence.*

Introduction

Inquiry by historians, educators, and psychologists into how to understand past human actions, agenda, and the degree to which actors have autonomy in their choices has paid much attention to the anachronistic effects of ‘presentism,’ our natural tendency to filter the past through our present sensibility. While some (Carretores & Voss, 1994) contend that asking students for moral reflection on the actions of historical agents provides too high a risk for presentism, others (Barton & Levstik, 2004) have argued that the promotion of students’ recognition of their own present context as a ubiquitous side effect of historical thinking is necessary to a deep construction of the meaning of life and times, past and present. Although we ‘cannot help’ but judge the past, a fundamental goal of historical thinking, Wineburg (2001, p. 22) suggests, is to scaffold students’ awareness of their present subjectivity so that they can take the perspective of actors within particular historical contexts, unhindered by their present beliefs and values.

Recently, some educators have gone so far as to argue that curricular material and pedagogical methods that cover histories of violence and injustice necessitate ethical discourse and discussions that go beyond the traditional tenets of evidential historical thinking (Rittner, 2004; Stern Strom, 1994) to an education ‘that teaches the common humanness of the other, that stresses the values of caring, and that emphasizes compassion and responsibility’ (Rittner, 2004, p. 3). Similarly, from a critical theory perspective, Barton & Levstik (2004) argue that historical empathy generally focuses too much on causal understanding of the perspective, decisions, and strategic inferences used by historical actors to make choices, without enough attention to the moral consequences of those actions. Nowhere is this constraint more evident than when attempts are made at a ‘value neutral’ teaching of injustice to high school students (Bellino, 2010/2011). Yet, few research tools have been available to help provide evidence for either side in this debate.

Until recently, that is. Building on past work by Ashley & Lee, more recent empirical work by Hartmann and her colleagues (Hartmann & Hasselhorn, 2008; Hartmann, Sauer, & Hasselhorn, 2009) have demonstrated how curriculum-related, psychometrically sound measurements can be used to refine the conceptualization of emergent historical perspective taking (or historical empathy) and its relation to presentistic thinking or interpretations. Specifically, in a set of cross sectional studies, the researchers assessed students’ orientations toward one of three frames of interpretation of the possible voting behavior of past citizens: present oriented perspective taking (i.e., presentism), awareness of the specific role of the historical agent (in the past), and the capacity for past/present historical contextualization. Their comparative snapshots of seventh and tenth grade students demonstrated

Acknowledgements: The authors would like to thank Facing History and Ourselves for the opportunity to work with them on their national evaluation study, as well as their permission to explore implications for historical understanding with these data. In particular, we would like to thank Dennis Barr for his close collaboration, the Facing History executive team, Angela Bermúdez, Alan Stoskopf, Melinda Fine, Ethan Lowenstein, and the research team from Abt Associates, Inc. for their contributions to the National Professional Development Evaluation Project (NPDEP), as well as the Richard and Susan Smith Family Foundation for their generous funding of the NPDEP. We would also like to thank John Willett, Andrés Molano, and Nealia Khan for their contributions to our analysis of the data. Finally, we would like to thank Harvard Graduate School of Education Dean’s Fellowship for funding the exploratory research represented in the paper.

both developmental and individual differences at both ages, with a modest decline in students' surprisingly prevalent predilection to anachronize across ages, and a relatively modest increase in the amount of past/present contextualization in their intellectual portfolio with chronological maturity, from grades 7 to 10. This research raises a new question: to what degree is the difficulty eradicating presentism a developmental and cultural issue, and to what degree is it a function of the students' experience in history classes?

Studies of the teaching and learning processes involved in history education is another rapidly growing field with important implications (Barton & Levstik, 2004; Selman & Barr, 2009; Selman & Kwok, 2010). Despite this attention, there remains a minimal amount of research on historical understanding that is immediately compatible with classroom practice and that can test alternative theories and their hypotheses about these practices. Although one can draw to some degree upon educational and psychological research on adolescent responses to moral dilemmas, these studies often situate cases in hypothetical, atemporal, or present contexts in order to evaluate civic knowledge and actions (Kohlberg, 1985). Few studies, however, explore students' responses to moral questions located in resource rich historical narratives. Furthermore, research on history education usually has been either large-scale surveys of students' factual knowledge or has relied on interpretive methods with small samples to understand learning processes (Borries, 2009). Without studies that lead to generalizable claims about the development of adolescent historical competence using methods that explore how students derive personal and shared meaning from historical evidence, we cannot easily understand how ethical awareness and historical thinking relate to one another, nor can we understand the quality of adolescents' capacity to understand themselves morally or understand the past historically.

To take a step in that direction, in this study we ask if historical understanding and moral inquiry in middle and high school education can reinforce one another at their intersection.

Methodology

The context for measurement construction and validation

In this study, we draw upon a sample comprising 621 students drawn from a representative national sample of urban and suburban ninth and tenth grade students who were randomly assigned (at the school level) to the control condition of an experimental evaluation of Facing History and Ourselves.¹ In this way, the possible impact of the intervention does not affect the generalizability of the findings to students in 9th and 10th grade history classes. Using a random assignment by school design, a team of researchers crafted a measure that would assess historical understanding, both as it is theoretically constructed in current research (Lee, Dickinson, & Ashby, 2001; Stearns, Seixas, & Wineburg, 2000; Wineberg, 2001) and as it is pedagogically conceptualized in alignment with Facing History's theory of educational change for students (Barr & Facing History, 2010; Selman & Barr, 2009).

The 'Assessment Measure for Students' Historical Understanding' (Stoskopf et al., 2007) is a comprehensive approach to assess the degrees to which students understand and apply three core historical understanding concepts (historical evidence, causality, and agency) to a historical situation presented to them. The overall measure is organized in case study format that takes students approximately 30 minutes to complete. It comprises fourteen document-based questions, each aligned with one of seven primary and secondary source documents embedded in the questionnaire. These documents pertain to the inter-ethnic conflict that consumed the former Yugoslavia during the early 1990s. This historical case was selected because it likely presented students with new or unfamiliar historical content focused on intergroup and ethnic conflict. The questions are of three types: closed-ended, item rating and ranking, true and false comprehension items, and short open-ended questions. (See Appendix A for examples.)

For this present report, we examined the responses to one open-ended question that follows the students' reading of the final document in the series, entitled 'Jasmina and Tanja's story' (See Appendix A.) Based on a memoir written by Jasmina Dervisevic-Cesic (1994), a Bosnian Muslim who emigrated to the United States, the document tells the story of Jasmina's teenage years as war erupted in her hometown of Visegrad. The story's synopsis, retold in the measure through a third-person narrative so as not to overtly bias students with Jasmina's perspective, centers on the lifelong friendship between Jasmina and her best friend Tanja, a Bosnian Serb. The friendship becomes a microcosm of the country's struggle with ethnic and religious pluralism. Tanja and many of the friends' Serbian neighbors begin supporting the national army, which was dominated by Serbian Orthodox Christians. As intolerance spread, Tanja began working as an informant for the Chetniks, an extremist paramilitary group that committed atrocities against non-ethnic Serbs, targeting Jasmina, her family, and other members of the Muslim community. Consequently, Jasmina and her family were forced to flee the country.

Following the written narrative that they were asked to read, students were given this prompt: 'By the end of the story, Tanja was working for the Chetniks, helping them round up her Muslim neighbors. How would you explain what she did?'

Sample specificity

This measure was administered to ninth and tenth grade students of teachers who were randomly assigned by school to either the treatment and 'as is' conditions in the experimental evaluation study. A majority of students in the NPDEP control sample are 10th graders (75%). The sample drawn from seven metropolitan areas of the United States (Los Angeles, Chicago, New England, Denver, Memphis, New York, and Cleveland), contains a roughly 50/50 split of females to males, respectively. Approximately one-third of respondents identified themselves as Hispanic (37%) or White (36%); the remaining respondents identified themselves as black, Asian, or other. The first language that respondents learned as children is predominantly English (69%). With regards to their mother's and father's highest level of educational attainment, slightly less than half of both mothers and fathers had not completed high school or had only completed high school. Demographic characteristics of the control group matched overall sample characteristics, designed to represent a balanced sample of urban and suburban schools (Boulay, et al., 2011).

Methods of analysis²

The study's full sample (both control and treatment groups) of 1371 students was used to construct the coding scheme we applied to the data. Space on the survey page allowed up to four sentences to be easily written, though responses tended to be one or two sentences in length. We developed a multi-tiered analytic coding scheme consisting of four tiers of analysis to evaluate student responses to the open-ended question about why Tanja 'did what she did.' Table 1 summarizes the four tiers, which we then describe.

TABLE 1: A summary of the four-tiered coding scheme used to assess comprehension, explanations, judgments and contextualized reflection.

Tier 1 Degree of Comprehension	
CODE	SAMPLE RESPONSE
No comprehension	No response; I don't understand this question.
Partial comprehension	She rounded them up by gathering them and hiding them.
Adequate comprehension	Tanja was probably trying to be supportive of her ethnic group.
Tier 2 Explanations	
Unclear explanation	She helped the Serbs.
Belonging and loyalty	Tanja felt loyal to her own ethnicity... maybe they gave her a higher meaning of living.
Pressure to conform	The Chetniks must have convinced her of their beliefs and why she should follow them and not bother about her Muslim friends.
Following orders	She was merely doing what she was told to do.
Dehumanizing the other	What she did was because she secretly didn't like Bosnian Muslims.
Self-preservation	She did it to protect herself.
Protecting the safety of others	She did it to protect the Muslims.
Following Tanja's own moral judgment	She did what she thought was right.
Pretending	Tanja was pretending about her unconditional loyalty to the Chetniks.
Fog of war	Perhaps she was lost in the war-whirl-pool of everything.
Choiceless choice	Tanja had no other choice.
Tier 3 Moral Judgment	
Positive judgment	She just wanted what was best for her people.
Negative judgment	BAD! Tanja isn't such a great person.
No judgment, description	I would explain it by offering evidence.
Uninterpretable judgment	It's not her fault that she had to work for the Chetniks because she had to protect herself. Either that or torturing herself. So I wouldn't blame her.
Tier 4 Contextualized reflection	
Contextualized reflection	She felt the need to be loyal to her people and probably feared getting into trouble if she did not cooperate. But as time passed I feel that she truly believed in what she was doing.
	I would say she was just doing her duty but her conscience is poor. I bet she can't sleep because she knows what she is doing is wrong. But now she can't go back and change things.

Tier one, degree of text comprehension

To account for the students' comprehension of the document's text, as well as the multiplicity of substantive responses students provided, we constructed a multi-tiered, iterative coding scheme. The first analysis addressed the quality of comprehension the students expressed in their responses. Comprehension included direct or simple (literal) reading comprehension, as well as deep (e.g. figurative, interpretive) reading comprehension and touched upon the formal realm of historical understanding, identified in literacy theory as background knowledge (Snow, Lawrence, & White, 2009).

Background knowledge directly available in the complete text of the document based assessment included: the distinction among ethnic and religious identities, the historical precedents of ethnic tension, and the social, economic, and political context of the evolving conflict. Analysis for comprehension considered how well the student understood the language and its intended meaning of these documents as used in the text, which included evidence provided across the scope of all documents.³ For instance, when a student appeared to interpret Tanja's actions as intelligent, brave, or morally upstanding, could one conclude that he/she misunderstood the evidence presented in the story, or alternatively, was the student potentially communicating an interpretation that did not agree with the presupposed ethical judgment we had assumed through the selection of Jasmina's memoir as a victim of injustice, and as aligned with the previous six documents and associated questions? Or was the student simply non-compliant with the survey intent?

In this manner, we concluded that we only could assess evidence of clear miscomprehension or of partial and adequate comprehension. We designated responses as exhibiting either 'no comprehension expressed' ('I don't know' replies and non-responses), 'partial comprehension' (revealing some evidence of misunderstanding story evidence and/or historical details), or 'adequate comprehension' (or, to put it more cautiously, no clear evidence of miscomprehension). This tier one comprehension hierarchy, not highly differentiating, considered all points of view and moral interpretations, while recognizing that some students may have fundamentally misinterpreted the 'facts' in the historical case. Students whose responses were categorized with 'no comprehension' were only included for the comprehension analysis tier, but were removed from further analysis of the subsequent three tiers of coding operations (See Table 2).

Tier two, number and type of explanations for Tanja's actions

For the second tier of the coding scheme, 'open' codes were created to identify patterns that emerged from student words, without a prescriptive theoretical lens or ideological position (Charmaz, 2006). The first round of analysis resulted in thirteen 'explanation codes.' During the second and third rounds of analysis, we collapsed these open categorical codes to eleven interpretive themes, including one code for 'unclear explanations': a) belonging and loyalty, b) pressure to conform, c) following orders, d) dehumanizing the other, e) self-preservation, f) protecting the safety of others, g) following Tanja's own moral judgment, h) pretending (to be an aggressor, so that she can secretly help), i) fog of war, and j) 'choiceless choice' (a recasting of Langer's (1982) term for victims' lack of agency in oppressive contexts) (See Table 1). Despite the partial overlap of many of these explanations in practice, we felt it was important to recognize the nuances between each of these codes and to acknowledge the diverse, and sometimes partial, ways that students invoked them. When students listed more than one explanation for Tanja's actions, they received multiple codes. Thematic explanations and number of explanations were recorded for each student response.

Tier three, type of moral judgment expressed

The third tier of reading the open-ended responses delineates indications of moral judgment revealed in the student response. The prompt 'explain what she [Tanja] did' meant students were likely to infer either/both psychological and sociological causes, and had the liberty to make interpretations of Tanja's choice-making. Though the question asks students to offer an explanation of them, some students visibly grappled with an evaluative response, an expressed phenomenon we decided to take into an empirical account. In fact, many student responses do not directly provide a clear explanation for Tanja's actions. For example, a response that reads, 'BAD! Tanja wasn't such a great person,' does not offer an explanation or necessarily demonstrate disciplinary thinking, but it does demonstrate that the student is critiquing Tanja's actions and judging them as negative. Given the structure of our multiple reading method and tiered (semi-independent) coding scheme, it is possible for a response to be included for a simultaneous analysis of moral judgment and explanations of historical agency in context.

Tier four, 'contextualized reflection'

The fourth tier of analysis evaluates whether the student demonstrates sophisticated historical understanding as described in the research literature (Lee, Dickinson, & Ashby, 2001; Seixas, 1996; Seixas, 2000; Wineburg, 1991; Wineberg, 2001). In the case of this open-ended question, 'contextualized reflection' was operationalized as a demonstration of ethical and historical thinking, requiring an overt expression of Tanja's historical context, along with an explanation of Tanja's choice and potentially her discarded choices (Hartmann & Hasselhorn, 2008; Martin, Sokol, & Effers, 2008). This tier yielded a dichotomous classification, present or absent. It uses a holistic consideration of each student's complete response, such that participants needed to provide at least one explanation for Tanja's actions and to display sophistication in historical empathy, their understanding of historical agency, consideration of multiple perspectives, and/or use of evidence. (See Table 1 for two examples of responses coded as contextualized reflection.)

While moral reflection was not an explicit requirement of this code, many tier 4 responses used moral terms to display historical empathy, invoking the student’s own moral framework while remaining sensitive to Tanja’s perspective and historical context. Some responses that illustrate contextualized reflection indicated a recognition of the limits of empathy across cultural and temporal distances, the need to struggle with the complexity of Tanja’s action as a moral decision, exploration of the intricate relationship between individual and social forces, or the presentation of carefully contextualized cross-cultural comparisons. Several student responses so categorized offered alternative actions that Tanja could have taken instead of what she did. In short, as aligned with the theoretical foundations of recent research in history education, responses that were coded at this tier demonstrate the most analytical, contextualized, reflective historical understanding responses among the sample.

Exploring frequencies of codes, tier-by-tier

Once the coding framework was finalized, two teams of researchers applied the framework to the full dataset (n=1371), blind to experimental conditions. This framework was applied by one set of four coders and one set of three coders, both facilitated by the same lead coder. For both sets of coders, ten classes of students were consensus coded before coding 20-30 classes separately and determining inter-rater reliability. Reliability for the first set of three coders and for the second set of four coders are reported in Table 2.

TABLE 2: Inter-rater Reliability calculations for Coding Group 1 (three members) and Coding Group 2 (four members) in their application of the multi-analytic coding scheme.

Coding Group 1		
	Kappa	Standard Error
Tier 1	.97***	.04
Tier 2	.99***	.03
Tier 3	.97***	.03
Tier 4	.81***	.04
Coding Group 2		
	Kappa	Standard Error
Tier 1	.89***	.05
Tier 2	.94***	.04
Tier 3	.98***	.04
Tier 4	.5	.064
Key: $\bar{p} < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$		

The poor reliability of tier 4 with the second group of coders might be explained by the small number of student responses that were taken into account to calculate the kappa; in total, four coders only cited nine potential responses as tier 4. Because reliability was strong among both groups across nearly all tiers of analysis, the remaining classes were coded by one lead coder. Subsequently, we used the control sample student ID numbers in order to calculate tier-by-tier frequencies, to control for available background characteristics, and to analyze overall patterns in the data that most represented history classrooms nationally across seven major metropolitan areas (i.e., not classrooms that had just received an intervention that emphasized historical and ethical co-reflection).⁴

Results

Distributions among degrees of comprehension

The tier one analysis suggests that comprehension was an impediment to any potential for disciplinary understanding. Of 621 participants, 28.8% of student responses revealed no comprehension. Students either left these responses blank or explicitly stated they did not understand. The remaining responses were coded as partial comprehension (14.0%) and adequate comprehension (57.2%).

TABLE 3: Frequency of student responses by comprehension, explanation, judgment, and contextualized reflection.

Code	Tier 1 Comprehension		
	Number of students	Percentage of students with partial and adequate comprehension	Percentage of full sample
No comprehension	179	n/a	28.8%
Partial comprehension	87	19.7%	14.0%
Adequate comprehension	355	80.3%	57.2%
Tier 2 Explanations			
No explanation (No comp.)	179	n/a	25.9%
Unclear explanation	164	37.1%	23.7%
Belonging and loyalty	64	14.5%	9.2%
Pressure to conform	28	6.3%	4.0%
Following orders	21	4.8%	3.0%
Dehumanizing the other	7	1.6%	1.0%
Self-preservation	124	28.1%	17.9%
Protecting the safety of others	38	8.6%	5.5%
Following Tanja's own moral judgment	19	4.2%	2.7%
Pretending	6	1.4%	.87%
Fog of war	8	1.8%	1.2%
Choiceless choice	34	7.7%	4.9%
Tier 3 Moral Judgment			
Positive judgment	43	9.7%	6.9%
Negative judgment	86	19.5%	13.9%
No judgment, description	59	13.3%	9.5%
Uninterpretable judgment	254	57.5%	40.9%
Tier 4 Contextualized reflection			
Contextualized reflection	28	6.3%	4.5%

Thematic analysis of the explanations students provided for 'what Tanja did'

Each response was coded for the explanation, series of explanations, or lack of explanations that it offered. In order to measure this aspect of historical understanding, we first counted how many student responses provided one uni-causal explanation for Tanja's actions, in comparison to responses with multiple explanations. Less than half the sample (44.7%) offered one or more responses designated as explanations for Tanja's actions. Overall, 55.2% offered no explanations, 34.5% of the sample provided one explanation, while 9.2% offered two, and 1.1% offered three.

Fig. 1 illustrates the ten themes that were reliably identified, in addition to the percentage of non-responses and unclear explanations. Among the ten thematic explanations provided among student responses (that exhibited either partial or adequate comprehension), the most frequently cited reasons for Tanja's actions were self-preservation (18%) and belonging and loyalty (9%), followed by responses that Tanya was protecting the safety of others (5%), and that she was left with a 'choiceless choice' (5%). Other responses explained that she was responding to pressure to conform (4%), following orders (3%), or following her own moral judgment and doing what she thought was right (3%). The least cited explanations were that Tanja made her decision to act because of the dehumanization of the other (1%), because Tanja was only pretending (1%), and because of the 'fog of war' (1%). Although some of these themes exhibit low frequencies, we chose to include them in our analysis, because they may have been more salient in the entire student sample (n=1371), because their frequencies shift when controlling for covariates, and because their infrequent presence may reflect a developmental range of sophistication among explanation themes (Loevinger, 1976).

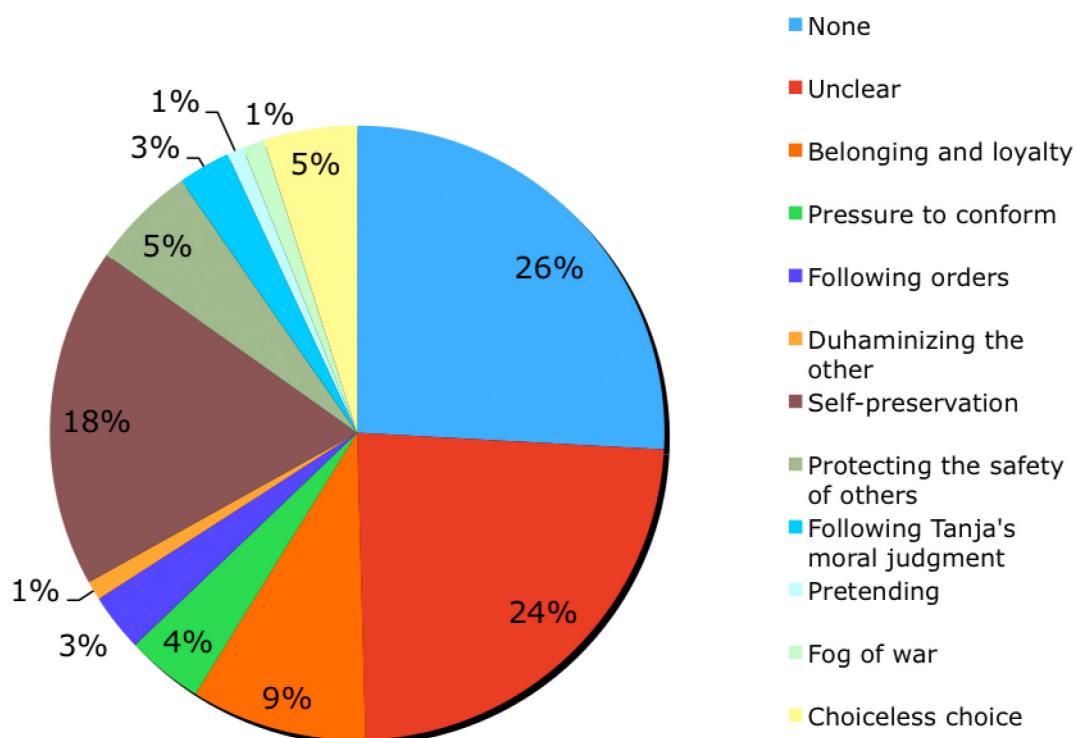


Fig. 1. Percent of student explanations offered for Tanja's Actions

Student evaluative (moral) judgments

While 13.9% of students explicitly judged Tanja's actions as negative, 6.9% explicitly considered them positive. Nine and a half percent of students articulated responses in neutral, descriptive terms, so that they seemed to abstain from judgment entirely. (See examples in Table 1.)⁵ However, 40.9% of the sample offered evaluative responses that we designated as not interpretable in terms of a valenced (neutral, positive or negative) moral judgment. Though some of these responses explicitly considered positive and negative elements of Tanja's actions, their evaluative comments, as a whole, remained unresolved. While it was not difficult to achieve reliability in coding the four categories of moral judgment (.97, $p < .0001$; .98, $p < .0001$ %), these 40.9% of student responses proved the most difficult aspect to interpret.

Contextualized reflection

Few responses in the overall sample exhibited what coders considered contextualized reflection about the potentially complex nature of Tanya's decision, while maintaining attention toward the particular historical context. Of 621 adolescents, only 28 respondents (4.5%) achieved contextualized reflection.

Correlations among the codes

Comprehension levels and their connection to moral judgment, number of explanations, and contextualized reflection

To understand the historical understanding correlates of quality of comprehension, we examined the degree to which responses that provided an 'adequate' comprehension were likely to offer one or more explanations, compared to those responses that exhibited a 'partial' degree of comprehension. Although degree of comprehension and number of explanations are highly and significantly correlated (.52, $p < .0001$), when we analyzed the number of explanations student responses provided by degree of comprehension, we found that nearly 34% of students with adequate comprehension did not offer any explanations in their response. Therefore, degree of comprehension alone did not explain the lack of expressed historical reasoning among the sample. See Fig. 2 for the number of explanations offered among student responses, sorted by partial and adequate degrees of comprehension.

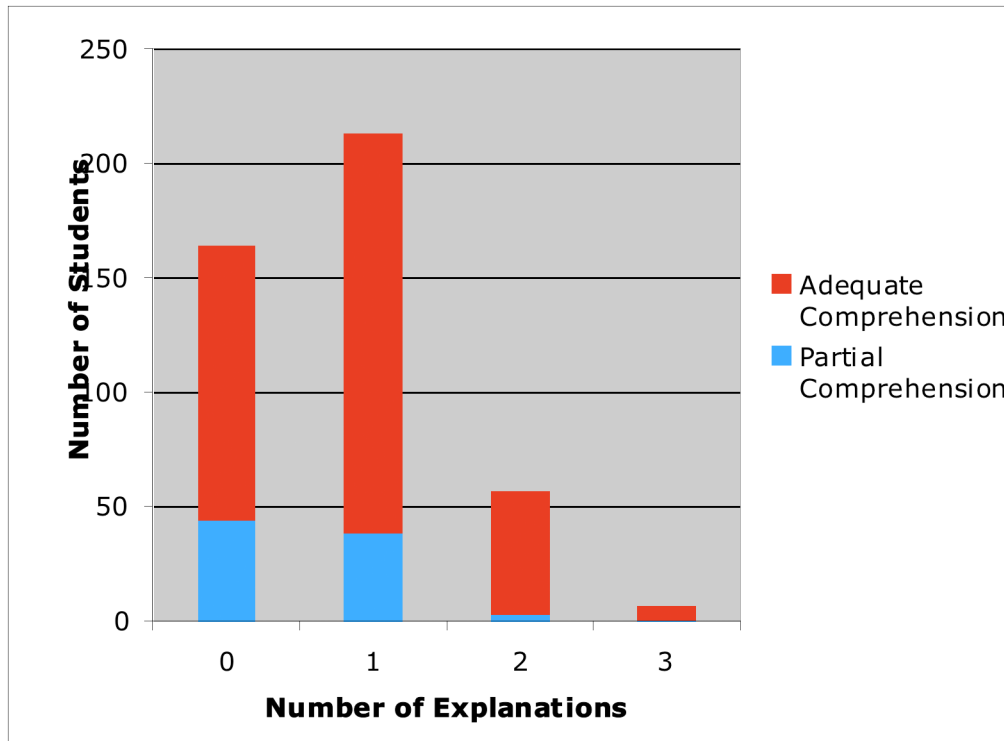


Fig. 2. Number of explanations selected by students with partial and adequate degrees of comprehension.

Among student responses rated as displaying an adequate degree of comprehension, the most frequently offered explanations remained the same as the overall thematic frequencies. Self-preservation persisted as the most frequent code (33.8%), followed by belonging and loyalty (17.7%), and choiceless choice (9.3%). Among students who exhibited partial comprehension, 36.8% of student responses explained Tanja's actions as protecting the safety of others, a code highly correlated with partial comprehension of the story (.516, $p < .0001$). Pretending (3.4%) was somewhat correlated with partial comprehension (.102, $p < .01$). The remaining explanations were selected by less than 2.3% of students with partial comprehension (1-2 students).

Among the 28 responses that exhibited contextualized reflection, the thematic explanations cited match the two most frequent overall codes: self-preservation and belonging and loyalty. However, no significant relationship emerged between a response's contextualized reflection and the thematic explanations offered, or the number of explanations offered.

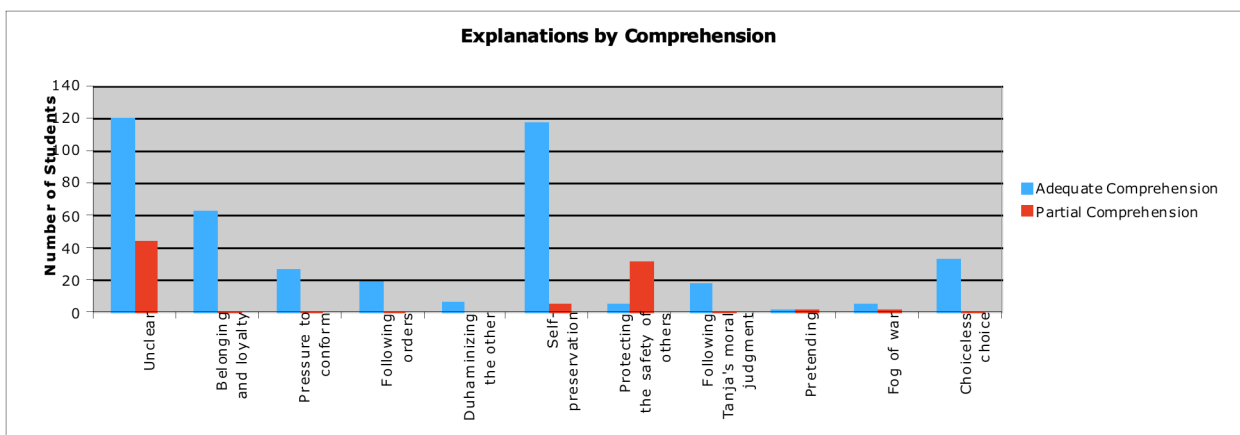


Fig. 3. Explanations by Number of Students Demonstrating Partial and Adequate Degrees of Comprehension.

Partial comprehension was significantly correlated with the presence of a positive moral judgment of Tanja's actions (.274, $p < .0001$), possibly related to miscomprehension of Tanja's role as a helper, 'rounding up' Muslims to save them, rather than to turn them in to the Chetniks. On the other hand, contextualized reflection fully aligned with adequate comprehension.

The Presence or absence of moral judgment in correspondence with explanation and contextualized reflection

Nearly 21% of students responded to the prompt, 'How would you explain what she did?' with strongly positive or negative judgment regarding Tanja's actions, and nearly half of these students (11.6% of the sample) expressed valenced moral consideration in the absence of an explanation for why Tanja made the choice to deceive and betray Jasmina. This means that students did not necessarily respond with judgment and explanation hand in hand, a critical distinction for both disciplinary and civic learning goals. Meanwhile, 35.4% of student responses offered clear thematic explanations associated with the classification of un-interpretable moral judgment. The relationship between un-interpretable judgment and number of thematic explanations was significant ($\chi^2=178.62$ $p < .0001$).

To test the relationship between responses that displayed valenced moral judgment and an outcome variable that would demonstrate some form of disciplinary understanding, we collapsed the number of explanations into a dichotomous variable, asking whether the student provided at least one or more explanations compared to none. Similarly, we collapsed the judgment tier into a dichotomous variable, asking whether the evaluation was valenced (i.e. explicitly positive or negative judgment) or unresolved (i.e. un-interpretable judgment). Contingency table analyses yielded a series of Pearson χ^2 statistics, determining that there is a significant relationship between the kind of judgment expressed and whether or not a student offers an explanation ($\chi^2=76.94$ $p < .0001$). On average, the percentage of students who both express an unresolved judgment and offer one or more explanations is 1.959 times the percentage of students who express explicitly valenced judgment and offer one or more explanations.

This relationship persisted across nearly all background variables. Valenced vs. unresolved judgment and presence of explanation remain related at partial and adequate levels of comprehension, for males and females, for all levels of mother's education, for native and non-native English speakers, and for students of all ethnic/racial backgrounds with the exception of African American students. This relationship was found to be significant for all regions included in the sample, with the exception of Cleveland (the smallest regional sample) and Memphis. See Fig. 6 for Pearson χ^2 statistics.

Pearson χ^2 statistics demonstrating the relationship between explicitly valenced or unresolved forms of moral judgment and number of explanations for historical agency.

	Pearson χ^2 statistics
The relationship between valenced/ unresolved judgment and 0/ 1 or more explanations	76.94***
Controlling for covariates	
Partial comprehension	8.54**
Adequate comprehension	83.46***
Male	17.42***
Female	67.84***
African American	1.416
Hispanic	33.045***
Asian	11.72***
Native	5*
White	19.22***
Other	11.4286***
Minority	54.89***
9 th Grade	11.43***
10 th Grade	66.98***
Non-native English speakers	25.69***
Native English speakers	49.83***
Chicago	4.03*
Cleveland	3.43~
Denver	7.166**
Los Angeles	18.87***
Memphis	1.028
New England	6.857**
New York, New Jersey	23.299***
San Francisco	20.969***
Key: ~ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$	

Discussion

Sophisticated historical understanding hinges on sophisticated comprehension processes. The relationship between disciplinary thinking and what literacy experts consider 'deep reading' of text (Snow, Lawrence, & White, 2009), analytical and reflective processes that motivate comprehension, is critical for historical understanding, both in the way that students read and interpret texts, and in the way they use historical accounts to construct their own meaning of past and present. That 57% of students exhibited an adequate degree of comprehension in this evaluation is not unexpected; perhaps more surprising is that so few students who exhibited either partial and adequate comprehension attempted to explain Tanja's actions.

It is important to recall that all of the students in the study were in a position to consider a good deal of background knowledge and evidence presented throughout the measure—not only the information from Document 7, 'Tanja and Jasmina's story.' With this cumulative evidence of historical agents' perspective, agency, and historical context to 'explain what she did,' why did students' historical thinking stop short? We cannot be sure why so few students who comprehended the text and historical case offered explanations for Tanja's actions, and why even fewer students offered multiple explanations, or explanations that revealed contextualized reflection.

One possibility is that students are not sufficiently or effectively exposed to history's 'disciplinary habits of mind' (Wineburg, 2001), or that they are not often held to these standards in history classrooms. Despite some push in the field toward history-as-interpretation, history-as-fact still guides much classroom practice; accordingly, students are acculturated into an 'epistemology of text' in which historical narratives offer information to be gathered, memorized, and reported, rather than interpreted or critiqued (Wineburg, 2001, p. 76). Another possibility is that students are not sufficiently or effectively held to rigorous disciplinary standards when exposed to content that is morally complex and emotionally troubling, such as the conflict in the former Yugoslavia. Does students' propensity to evaluate, rather than interpret, Tanja's actions legitimate the fear that strict historical understanding constructionists have when epistemological domains are traversed? Or perhaps the way we teach contentious and unsettling historical events requires attention to the emotional and ethical responses that students have, as well as insisting on deep historical inquiry. In particular, more work needs to be done to determine whether there is a potentially developmental need for adolescents to simultaneously consider historical thinking and ethical reflection, or at least some synergy.

Student explanations and evaluations of historical agency

Among the explanations that students selected, self-preservation is by far the most frequent, remaining the preferred response even when controlling for comprehension levels. This student interpretation occurs despite the lack of documented or 'hard' evidence presented in the measure that illustrates that Tanja actually would have been in danger had she not participated. Is it, then, problematic that many students seemed to believe, in the absence of historical evidence, that Tanja not only was not in a position to *resist* the injustice, but that she had to actively *participate* as well in its commission? If students are simply taught that actors 'on the dominant side' during periods of genocide must participate in atrocities for their own safety—or, perhaps more detrimental, agency in times of human rights violations is not explicitly addressed at all in history classes— it obscures the scope of agency among those perpetrators who mobilize and participate in mass violence, as well as the critical role of resisters, rescuers, and 'upstanders' during instances of injustice (Stern Strom, 1994). Further, this explanation may indicate students' assumption that conflict takes place strictly between perpetrators and victims, denying the role of bystanders, who themselves have agency to be explored and choices to understand.

Adolescent understanding of historical agency is often theorized as a link between recognizing the decisions of past actors in past contexts and one's own potential civic agency in the present (Boix Mansilla, 2001; Selman & Barr, 2009; Stern Strom, 1994). If students see Tanja's decision as a 'choiceless choice,' are they likely to translate this awareness to their own decisions in the present? On the other hand, might some students see their own limited capacity to act in the present as also operating within Tanja's historical case (Selman & Kwok, 2010)?

These questions illuminate an under-theorized tension between historical actors—as well as students engaging in historical thinking—as both autonomous *and* situated beings. Individuals may have the desire to act but cannot, or will not; alternatively, they may not be conscious of the agency that they do have (Nussbaum, 1999). Contending with this tension has implications for human rights awareness and 'informed civic engagement' (Selman & Barr, 2009) where agency, empowerment, and awareness of agency are equally critical. Though it is not our aim to recommend that students uncritically upstand without considering physical safety and the consequences of their actions, we do suggest that students learn to practice the moral agency they do have when situations seem to close in on them.

Moral judgments in historical context

While strong, valenced points of view were found, the majority of student responses considered the moral complexity of Tanja's decision in a way that we coded as un-interpretable, or unresolved, judgment. It may seem that this tier of the coding scheme fell short of scientific expectations for an agreed upon and interpretable set of responses. For example, we cannot be sure that students who openly praised Tanja's actions did not fully comprehend them in context. An alternative possibility is that moral praise was a clumsy attempt at authentic historical thinking, i.e. that Tanja did the right thing considering the limitations of her situation. Similarly, we cannot be sure that students who openly denounced Tanja's actions were not overlooking contextual elements and making presentistic decisions based on their own social contexts, or in the context of their own belief systems, or based on universalist claims of what is morally correct.

We decided, however, there was merit to report the analysis of moral judgment in tier three for two reasons: first, a significant portion of students provided responses that were clearly judgmental, even if the valence or structure of the judgment could not always be identified, and second, a significant portion of students expressed some form of judgment at the expense of offering an explanation for why Tanja did what she did. Ignoring the student's willingness to engage with the moral agency at stake may overlook important elements of historical understanding among adolescents. Furthermore, given the structure of our method and coding scheme, a response can be included in an analysis of distinct moral judgments and clear interpretations of historical agency in context, positing ethical reflection as a potentially meaningful entry point for disciplinary thinking.

Shortcomings of the study

This research draws on one of the few large-scale assessments of adolescent historical understanding, using both qualitative (categorical, thematic analysis of open-ended but brief written responses) and quantitative analyses (codification and quantification of degree of comprehension, number and type of explanation, evidence of contextualized reflection). However, even with a sample representative enough to make inferences about the distribution of variations in comprehension, explanations, and moral evaluations in ninth and tenth grade students' understanding of personal betrayal in a historical context, these data lack the depth of probed written or oral responses necessary to shed light on the potentially collaborative processes of historical thinking and ethical reflection—for example, the meaning of unresolved judgment.

Furthermore, students may have responded critically to Tanja's actions because of the social desirability of altruistic, courageous claims in the context of their classroom based assessment atmosphere, but while that might have led to a particular distribution of moral judgment, it should not have had much influence on performance in the other three tiers of analysis. Nor are we sure what bias exists when the same data are read for four different, we would argue distinct, phenomena, even when blind to coding at the other levels.

Finally, resources limited opportunity for students to express themselves, with both time and space constraints; this needs to be weighed against the opportunities large scale samples afford to study the questions that provide direction for this research. This is particularly important when research needs to explore the degree to which students at different ages and with different educational experiences are able to generate and transfer contextual reflection abilities in reading history. However, this study begins a programmatic attempt to develop and validate measures that will allow questions such as these to be more adequately answered.

Conclusion and educational implications of this study

Though genocide and human rights education philosophies are frequently inserted into history curriculum, there is relatively little empirical research on what constitutes sophisticated historical understanding of lived experience of past and distant ethnic conflict and grave human rights violations, especially when students do not identify these experiences as close to their own. The key tenets of historical understanding undoubtedly shift when situated in in-depth case studies of morally complex situations where questions of loss, injustice, power, and responsibility—personal and societal—play critical roles in the decisions and outcomes of historical events.

Perhaps this is a question not limited to cases of injustice, but deeply rooted in the ideological goals of purposefully teaching the past in the context of present events. When should history be taught with disciplinary standards in mind, emphasizing interpretation and skepticism, and without misconceptions that might arise as a consequence of promoting any secondary uses (Boix-Mansilla, 2000); and when should history be put to use to teach civic values in pluralistic democracies (Barton & Levstik, 2004)? Those who take the most stringent-boundary position insist that any mixing will relegate one to a secondary, instrumental use for the intrinsic purpose of the other. But we rarely question whether disciplinary rigor and ethical reflection might actually require or co-construct one another. What if emotional engagement, ethical reflection, and intellectual rigor are actually best understood by students when studied as collaborative processes; what if the 'unnatural act' (Wineburg, 2001) is not historical thinking on its own but a separation of that which cannot be easily studied by being pulled apart?

Notes

¹ Facing History and Ourselves is an international educational organization that provides professional development to educators and an educational approach for adolescents that is positioned at the intersection of historical understanding, civic engagement, and ethical reflection. One educational aim of Facing History is to connect students' own contemporary 'faced choices' in civil society to their study of historical cases. For more information on the context and findings of the national evaluations, see Barr, D. & Facing History and Ourselves. (2010). *Continuing a Tradition of Research on The Foundations of Democratic Education: Facing History and Ourselves National Professional Development and Evaluation Project (NPDEP)*.

² The authors have remained blind to the treatment and control groups throughout the analytical process. For experimental findings following the first year of data collection, see Boulay, et al., 2011.

³ As part of the early measurement construction debate as to where to put this component of the questionnaire (front, back, middle), the access to background knowledge was an important consideration in its location as the final set of questions. One alternative argument was that putting this personal drama as the introduction to the questionnaire would make it more engaging for the students, and that by the end of the measure, the students would be less energetic. Field testing did not show demonstrable fatigue.

⁴ The Facing History research department manages the original dataset and coordinated our use of the data.

⁵ These percentages do not add up to 100%, as students coded as 'No comprehension' were removed from subsequent tiers of analysis. For example, we did not code 'no response' as uninterpretable judgment. Rather than suggest that 28.8% of the students in the sample did not consider judgment, explanation, or reflection, we prefer to conclude that they revealed no comprehension and therefore could not be further analyzed.

Correspondence

Michelle J. Bellino, Harvard Graduate School of Education, Cambridge, MA, mjb033@mail.harvard.edu

References

- Arendt, H. (1978) *The Life of the Mind: The Groundbreaking Investigation on how we Think, Volume 1 & 2* Orlando, FL, Harcourt.
- Barr, D. J. and Facing History and Ourselves (2010) *The Facing History and Ourselves National Professional Development and Evaluation Project: Continuing a Tradition of Research on the Foundations of Democratic Education* Brookline, MA, Facing History and Ourselves National Foundation, Inc.
- Barton, K.C., Levstik, L.S. (2004) *Teaching History for the Common Good* Mahwah, NJ, Lawrence Erlbaum Associates.
- Bellino, M. (2010/2011) 'A History of Violence Argues against a Culture of Violence' *ReVista*, 10, (1), pp 13-16.
- Bermúdez, A. & Jaramillo, R. (2001) 'Development of Historical Explanation in Children, Adolescents and Adults', in Dickinson, A., Gordon, P. and Lee, P. (eds) *Raising Standards in History Education* London and Portland, OR, Woburn Press.
- Boix-Mansilla, V. (2000) 'Historical Understanding: Beyond the Past and into the Present', in Stearns, P.N., Seixas, P., & Wineburg, S. (eds) *Knowing, Teaching, and Learning History: National and International Perspectives* New York and London, New York University Press.
- Borries, B. von (2009) 'Historical understanding of students: An interpretation of a single case', in Martens, M., Hartmann, U., Sauer, M., & Hasselhorn, M. (eds) *Interpersonal understanding in Historical Context* Rotterdam, The Netherlands, Sense Publishers.
- Boulay, B., McCormick, R., Gamse, B., Barr, D., Selman, R.L., Lowenstein, E., & Fine, M. (2011) *Effects of Facing History and Ourselves on Teachers and Students: Findings from the National Professional Development and Evaluation Project*, Manuscript submitted for publication.
- Carr, D., Flynn, T.R. & Makkreel, R.A. (eds) (2004) *The Ethics of History* Evanston, IL, Northwestern University Press.
- Carretores, M. & Voss, J.V. (eds) (1994) *Cognitive and Instructional Processes in History and the Social Sciences* Hillsdale, NJ, Lawrence Erlbaum Associates.
- Charmaz, K. (2006) *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis* Los Angeles and London, Sage Publications.
- Dervisevic-Cesic, J. (1994) *The River Runs Salt, Runs Sweet: A Memoir of Visegrad, Bosnia* Eugene, OR, Panisphere.
- Dulberg, N. (2002, April) 'Engaging in History: Empathy and Perspective-taking in Children's Historical Thinking', Paper presented at the annual meeting of the *American Educational Research Association* New Orleans, LA.

- Hartmann, U., Sauer, M. & Hasselhorn, M. (2009) 'Perspektivenübernahme als Kompetenz für den Geschichtsunterricht' ['Perspective Taking as a Competence for History Education: Theoretical and Empirical Relations between Subject-specific and Socio-cognitive Characteristics in Students'] *Zeitschrift für Erziehungswissenschaft*, 12, pp 321-342.
- Hartmann, U. & Hasselhorn, M. (2008) 'Historical Perspective Taking: A Standardized Measure for an Aspect of Students' Historical Thinking', *Learning and Individual Differences*, 18, (2), pp 264-270.
- Kohlberg, L. (1985) *The Philosophy of Moral Development, Volume 1 & 2* New York, Harper & Row.
- Langer, L. (1982) *Versions of Survival: The Holocaust and Human Spirit* Albany, NY, State University of New York Press.
- Lee, P., Dickinson, A., & Ashby, R. (2001) 'Children's Ideas about Historical Explanation,' in Dickinson, A., Gordon P., & Lee, P. (eds) *Raising Standards in History Education* London, Woburn Press.
- Loevinger, J. (1976). *Ego development*. San Francisco: Jossey-Bass.
- Nussbaum, M.C. (1999) *Sex and Social Justice* Oxford and New York, Oxford University Press.
- Reisman, A. & Wineburg, S. (2010, October) 'Reading like a Historian': *Bringing Disciplinary Thinking into High School History Classrooms*, Paper presented at the meeting of Universidad Autónoma, *Understanding History and the Construction of Identities in a Global World: De-nationalizing History Teaching?* Las Navas, Spain.
- Rittner, C. (2004) 'Educating about Genocide, Yes: But what Kind of Education?', in Totten, S. (ed) *Teaching about Genocide: Issues, Approaches, and Resources* Greenwich, CT, Information Age Publishing.
- Seixas, P. (1996) 'Conceptualizing the Growth of Historical Understanding,' in Olson, D.R. & Torrance, N. (eds) *Handbook of Education and Human Development: New Models of Learning, Teaching, and Schooling* Malden, MA, Blackwell.
- Seixas, P. (2000) 'Schweigen! die kinder! or, Does Postmodern History have a Place in the Schools?', in Stearns, P.N., Seixas, P., & Wineburg, S. (eds) *Knowing, Teaching, and Learning History: National and International Perspectives* New York and London, New York University Press.
- Selman, R.L. & Kwok, J. (2010) 'Informed Social Reflection: Its Development and Importance for Adolescents' Civic Engagement', in Sherrod, L. R., Torney-Purta, J., & Flanagan, C. A. (eds) *Handbook of Research on Civic Engagement in Youth* Hoboken, N.J., Wiley.
- Selman, R.L. & Barr, D. (2009) Can Adolescents Learn to Create Ethical Relationships for themselves in the Future by Reflecting on Ethical Violations faced by others in the Past? in Martens, M., Hartmann, U., Sauer, M., & Hasselhorn, M. (eds) *Interpersonal understanding in Historical Context* Rotterdam, The Netherlands, Sense Publishers.
- Snow, C.E., Lawrence, L.F., & White, C. (2009) 'Generating Knowledge of Academic Language among Urban Middle School Students', *Journal of Research on Educational Effectiveness*, 2, (4), pp 325-344.
- Stearns, P., Seixas, P., & Wineburg, S. (2000) 'Introduction', in Stearns, P.N., Seixas, P., & Wineburg, S. (eds) *Knowing, Teaching, and Learning History: National and International Perspectives* New York and London, New York University Press.
- Stern Strom, M., (ed) (1994) *Facing History and Ourselves: Holocaust and Human Behavior Resource Book* Brookline, MA, Facing History and Ourselves National Foundation, Inc.
- Stoskopf, A., Bermudez, A. Hartmann, U., Selman, R. L., Barr, D. J. & Facing History and Ourselves. (2007) *An Assessment Measure for Students' Historical Understanding* Brookline, MA, Facing History and Ourselves National Foundation, Inc.
- Wineburg, S. (1991) 'Historical Problem Solving: A Study of the Cognitive Processes used in the Evaluation of Documentary and Pictorial Evidence', *Journal of Educational Psychology*, 83, (1), pp 73-87.
- Wineburg, S. (2001) *Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past*, Philadelphia, Temple University Press.

Appendix A

Document 7: Jasmina and Tanja's Story

Excerpted from 'Assessment Measure for Students' Historical Understanding' (Stoskopf et al., 2007).

In June 1991, Jasmina and Tanja were 18 years old and best friends. They had grown up together in Visegrad, a town in Bosnia-Herzegovina. Jasmina was a Bosnian Muslim and Tanja was a Bosnian Serb. Ethnic and religious differences, however, were not important to the girls.

(Bosnian Serbs were Serbian Orthodox Christians who either were born in Bosnia or had lived in the nation for most of their lives).

Such differences were common in Yugoslavia, where the population was very diverse and where people of different ethnicities and religions frequently had close relationships with one another.

Tanja had become an honorary member of Jasmina's extended family. In Tanja's own family, money was hard to come by, and her father was often drunk and violent. She frequently went to members of Jasmina's family for comfort and support. But in the summer of 1991, things began to change.

That summer, a war broke out when two of the six nations of Yugoslavia (Croatia and Slovenia) declared their independence. Serbia, where the central government of Yugoslavia was located, objected and sent the Yugoslav army into both countries, claiming they were protecting the rights of local Serbs. The Yugoslav army was dominated by Serbian officers and soldiers who were mostly Orthodox Christian.

In July and August of 1991, the army frequently marched through Visegrad, where Jasmina and Tanya lived. Many Serbian residents of the town, including Tanja, cheered them on. Chetniks, a paramilitary group known for their extreme nationalistic views and their intolerance of any Yugoslavs who were not ethnic Serbians, also marched through Visegrad. Tanja cheered for them too and was seen giving the special Chetnik three-fingered sign of support.

(A Paramilitary Group is a fighting unit used during conflict situations. They are not part of an official army but are often used by governments and private organizations. Paramilitary groups such as the Chetniks committed numerous atrocities during World War II and in the 1990s in the former Yugoslavia.)

When Bosnia declared its independence from Yugoslavia in March 1992, the Yugoslav army and Chetnik troops attacked Visegrad. Most Muslims had to flee the city, including Jasmina and her family - leaving all of their possessions behind.

Tanja stayed in Visegrad. She helped the occupying army and Chetnik forces round up Bosnian Muslims, many of whom she knew. The Serbian and Chetnik authorities asked her to spy on the homes of Bosnian Muslims and report any who tried to return. Tanja even watched over the apartment of Jasmina's cousins. Many Muslims who tried to return were captured and then tortured, raped or murdered.

Today, Tanja still lives in Visegrad. After the war, she moved to a nicer apartment once owned by Bosnian Muslims. Few Muslims ever returned to Visegrad. Jasmina now lives in the United States.

