Developing and Assessing Empathy through the Study of Christian Heresies in an Introductory Christian History and Theology Course

Christopher J. Richmann  
Baylor University

Courtney Kurinec  
Washington State University

Felicia S. Osburn  
Baylor University

ABSTRACT

Empathy is a nebulous concept that is touted in many schools’ curricular goals and striven for by many instructors. Unfortunately, lack of a shared definition and reliable measures for empathy hamper efforts to determine whether this goal is realized. Defining empathy as “the ability to know the internal state of another,” this study explores the effect of a learning project focused on Christian “heresies” on students’ empathy by developing and refining tools for assessing empathy in student writing. The intervention included scripted role-playing and reflection, group discussion, and an essay in which students discussed the appeal and rejection of a particular heresy in the persona of a “heretic.” Quantitatively, we found no significant effect in time or group comparison on an empathy questionnaire; qualitatively, we found an upper-middling level of empathy in essays, and a large effect in group comparison of student responses to a simple prompt to define a “Christian heretic.”

KEYWORDS

introductory religion course, empathy, heresies, assessment, role-playing

INTRODUCTION

The study of Christianity is core to the curricula of a range of institutions. Within humanities or social studies options at public universities, Christianity forms a substantial element of introductory religion courses. Many Christian-affiliated colleges and universities require a course, or series of courses, on Christianity—its history, theology, or sacred texts. For both public and private institutions, the rationale for these courses goes beyond content. Instructors at public institutions often consider religious studies an aid to the development of ethical perspectives or helping students acquire an interest in asking their own questions and seeking answers (Walvoord 2007). At Christian institutions, the rationale for such requirements may include affirming a Christian mission,
DEVELOPING AND ASSESSING EMPATHY

Empathy is a complicated concept. Scholars disagree on the definition, the origin, and the effects of empathy. According to one expert, researchers use the term to refer to at least eight different phenomena (Batson 2011). As Batson argues, these concepts are not aspects of a single phenomenon, but “stand-alone psychological states” (2011, 3).1 For the purposes of this study, empathy is the capacity to know the internal state of another. This is not necessarily to feel as the other feels (which we believe is better captured by the term “sympathy”), nor is it the ability to know exhaustively or perfectly what another is thinking, but the ability to infer another’s thoughts with reasonable accuracy through the clues of language, context, and analogy to one’s own experiences. Thus it is connected to imagining how another is thinking and feeling and imagining how one would think and feel in another’s place—particularly in relation to the motivations for the other’s specific behavior. The desire to respond compassionately—while mentioned in many definitions of empathy—is less applicable in arenas of “observational empathy,” like the historical study of religion, where the “self” does not directly interact with the “other” (Barrett-Lennard 1981).

The English language did not have the word “empathy” until 1908, but this does not make the concept less important. David Hume suggested that the phenomenon we now call empathy is the foundation for all social perception and interaction (Batson 2011). Some today argue that the development of empathy in students is one of the central tasks of higher education (Thompson 2014). Furthermore, empathy should be seen not only as a goal of education, but as a tool for education. Empathy is necessary for serious study in any of the humanities disciplines. Critical inquiry and full understanding of human behavior, regardless of methodology, requires the ability to appreciate the perspective of the humans under scrutiny—their motivations and their intentions. At least one school of thought holds that the study of

1 For this reason, Batson urges researchers to “recognize the different phenomena, make clear the labeling scheme one is adopting, and use that scheme consistently” (2011, 8).

history demands that scholars use the human “gift of empathy” to make sense of human behavior in different times and cultures from our own (Bebbington 1990; Southgate 2005); increasingly, history teachers recognize this “perspective taking” as central to teaching history (Wineburg 2001). Religion is another helpful example. Many scholars agree that “phenomenological approaches to religious studies . . . assume the importance of empathy to engaging religious differences and developing a critical understanding of these differences” (Trothen 2016, 245). “Highly effective” religion instructors commonly include empathy in learning objectives (Walvoord 2007). A college course in religious history occurs at the intersection of multiple concerns related to empathy in students.

Empathy is not only important, it can be developed. Although some scholars despair of the ability to quantify “historical empathy” (Endacott and Brooks 2018), and we’ve heard faculty colleagues claim “you can’t teach empathy,” research indicates educational interventions can improve cognitive (as opposed to affective or behavioral) empathy. In particular, research in pre-professional education demonstrates that “role-playing and other simulations” are effective at enhancing empathy (Trothen 2016, 251; Porter 2008). As with other cognitive phenomena, many factors determine the degree to which empathy can be developed, and one major factor may be a person’s prior beliefs about the malleability of empathy, such that those who believe empathy can be developed expend greater empathetic effort in difficult situations (Schumann, Zaki, and Dweck 2014).

The Project: Nature, Purpose, and Challenges

The primary tool for developing empathy in this study was a multifaceted project in which students in an experimental group investigated a single Christian heresy of their choosing. This project entailed, for the experimental group but not for the comparison group, role-playing exercises, a small group workday prompting students to consider their own interest in the heresy they chose to study and the commitments and assumptions undergirding the teaching, and a four-question essay (described below). At the beginning of the semester, students received a list of suggested topics, movements from the first six centuries of Christian history that were regarded as heretical by the dominant strand of the faith, as represented by the first four ecumenical councils. The chronological approach of the class allowed them to read and participate in class sessions on these subjects before selecting one. While this approach is similar to previous attempts to foster empathy through identification with a distinct religious group, the present project allowed students to choose their group rather than being assigned a historical identity (Porter 2008). Although this may seem to counteract the intent for students to encounter foreign viewpoints, since students may choose to study a group they already identify with, we believe the intrinsic motivation involved in choice outweighs the potential drawbacks. Besides, for the students in these classes, all heresies are by definition “other,” since no students self-identified as adherents of these groups; any student affinity for the comparison group, role-playing exercises, a small group workday prompting students to consider their own interest in the heresy they chose to study and the commitments and assumptions undergirding the teaching, and a four-question essay (described below). At the beginning of the semester, students received a list of suggested topics, movements from the first six centuries of Christian history that were regarded as heretical by the dominant strand of the faith, as represented by the first four ecumenical councils. The chronological approach of the class allowed them to read and participate in class sessions on these subjects before selecting one. While this approach is similar to previous attempts to foster empathy through identification with a distinct religious group, the present project allowed students to choose their group rather than being assigned a historical identity (Porter 2008). Although this may seem to counteract the intent for students to encounter foreign viewpoints, since students may choose to study a group they already identify with, we believe the intrinsic motivation involved in choice outweighs the potential drawbacks. Besides, for the students in these classes, all heresies are by definition “other,” since no students self-identified as adherents of these groups; any student affinity would be superficial and inconsequential.4

Heretics are the “other” on the margins of one’s group rather than completely beyond it; they share most theological problems with the dominant group yet differ in one or more of their solutions. We speculate that such groups are similar enough to engage feelings of solidarity yet different enough to require substantial effort to appreciate. In other words, heresies occupy the intersection of self and other that is the arena of empathy. Yet these potential benefits are also potential liabilities. On one hand, students may encounter the familiar terms used by heretical groups and fail to appreciate the different presuppositions behind the terms or the precise meaning those terms have for heretical groups. On the other hand, heretical groups pull students further from the dominant narrative, familiar concepts, and—for most students involved in this project—their own faith tradition.

3 The cognitive domain “involves the ability to know and understand that other people have diversity of perspectives.” This maps nicely to the definition used for this study. The affective domain “involves the sharing of the other’s emotional state.” The behavioral domain is “the intention to respond in a way that communicates a compassionate awareness” of the other’s emotions (Trothen 2016). The affective domain may be better captured by the term “sympathy,” and the behavioral domain may be better labeled “compassion.”

4 Although “non-orthodox” Christian groups continue to flourish, we discounted the possibility that any participants identified with a historically heretical group for two reasons. First, although some modern groups, like the Latter-Day Saints and Jehovah’s Witnesses, have a Christology similar to Arianism, there is no historical connection between the groups, and they arrive at their theology by means of different logic and framework. The same is true for the relationship between pre-Nicene modalism and Oneness Pentecostalism. Second, a recent student body survey showed that no students at this institution identified as Monophysite or Nestorian Christians, the two major early heretical groups that have survived to the present day.
This perceived difference between heretical positions and students' own faith adds an additional challenge to the empathetic study of heresies at a Christian institution. While students might willingly memorize facts about other religious viewpoints, for several reasons, they are less likely to engage with heresies empathetically, that is, to take these positions seriously as reasonable answers to the fundamental human questions religions address. Research shows that people often do not remain empathetic when they feel greatly dissimilar to the person or group to whom they are asked to relate (Schumann, Zaki, and Dweck 2014). This suggests that any instructor aiming at developing empathy must actively work to reduce the perceived difference between the self and other. Likewise, empathy is strained when people believe empathizing may be distressing or discomforting (Schumann, Zaki, and Dweck 2014). Students who view their faith as sacred space and privileged commitments might view opposing religious ideas—not to mention critical scrutiny of one's own religion—as threatening. Instructors therefore run the risk of shutting down both learning and empathy if students interpret a generous approach to heresies as an attack on their own commitments. According to the Triune Ethics Theory of decision-making, one who feels threatened is driven by the “safety ethic,” which automatically excludes the “engagement ethic” associated with empathy (Fleming and Lovat 2015).

To further counter students' tendency to dismiss heretical viewpoints, the project framed “heresy” in political and social terms rather than as incorrect theology or practice (Pagels 1979). Additionally, the texts students were required to consult for initial detailed information about the heresies they chose to study (González and González 2008; Wilhite 2015) have been noted for their fairness and generosity in interpreting heresies. It further noted the task was not to defend the orthodox position or point out the flaws of the heresies but through historical investigation to gain understanding of heretics and appreciate the role they played in prompting the “orthodox” Christians to refine and articulate their positions. That is, that heresies were not fended off by orthodox Christianity so much as they were forced out of a narrowing dominant group (see Appendix A). These ideas were reinforced in an in-class workday through guided reflection and discussion.

The first phase of the project for the experimental group involved public role-playing of scripted hypothetical discussions between a representative of a heretical group and a representative of the orthodox party. Each scripted conversation was about five minutes long and involved two students. These conversations were spread out individually over the course of nine class sessions. Previous research on empathy development through the study of religion has suggested that role-playing, particularly when coupled with listening exercises, helps facilitate students' understanding of other viewpoints (Trothen 2016). Students who were not directly role-playing watched and then reflected for about five minutes, using a series of guided questions related to their perceptions of the experience of the heretic (see Appendix B). While only sixteen students out of fifty-two students publicly role-played, all students had multiple experiences of witnessing and reflecting on the conversations. This purposeful debriefing “allows the awareness gained in the exercise to be clarified and reintegrated into [the student's] own inner life,” that is, helping to clarify the self in relation to the other (Burns 2001, 20). This phase was also meant to pique students' interest in the heresies and help them decide which group to focus on for the second phase of the project.

For the second phase of the project, students in the experimental group researched, developed, and wrote a response to a series of questions involving their Christian heresy. In order to help “loosen the boundaries of the self, encourage flexibility, and increase the permeability of the self’s enveloping membrane” (Burns 2001, 20), the prompt required students to craft responses from the point of view of an adherent of the heresy they were studying (see Appendix A). Previous work has found that, when asked to imagine oneself or another particular person in a specific historical circumstance, students display enhanced empathy, and their writing includes more details and emotions (Brooks 2011; Leur, Boxtel, and Wilschut 2017). Additionally, because peer discussion has shown promise in developing students' empathy (Ashby and Lee 1987), students spent a whole class session (one hour and fifteen minutes) in small groups of five-six discussing questions designed to prompt them for the essay:

---

5 Leur et al. (2017) also found, however, that a first-person imagination yielded greater presentism and moral judgments.
1. Why did you choose to study this heresy?
2. Identify the theological, philosophical and/or religious commitments of this group. For instance, What is authoritative? What makes creation different from God? How does Christianity relate to Judaism? (or, How does the ‘Christian God’ compare with the ‘Jewish God’?) What does the work of Jesus accomplish? How do humans get the benefit of Jesus’ work?
3. In one or two sentences, describe the central teaching of this heresy in the most charitable (yet accurate) way possible (that is, explain it in a way an adherent of this teaching would approve of).

The instructor and a graduate teaching assistant engaged the groups during the session, and each group wrote their thoughts on poster board and presented them to the instructor or teaching assistant.

For the essay, students responded to these questions in 1,000-1,500 words:

1. What theological problem are you trying to address?
2. What is your solution?
3. Why is this solution attractive to you and other people you know?
4. Why do you think your position was rejected by others?

Before final submission of their essays, students were required to submit an outline of their answers to these questions. Each outline was read and commented on by the instructor or teaching assistant.

In addition to engaging the challenge of developing academic virtues, this assignment aimed to meet two additional challenges. First, through guided questions rather than an open-ended argument, it provided a focus for writing that, we hope, leads students through lower-order to higher-order thinking, in which students not only remember, but understand, apply, and analyze information (Bloom 1956). For our purposes, these higher-order skills are synonymous with “critical thinking,” in which students do more than identify or repeat information but actually manipulate information in some way—for example, through comparisons, application to novel situations, or putting knowledge in their own words. This guided-questions approach is framed, in part, by our belief that writing in lower-level courses should not be guided by hopes that students will learn to write according to disciplinary conventions (e.g., write like a historian of religion, crafting fine-tuned thesis statements, charting their own original historical arguments, etc.) but acquire critical thinking skills that can serve them well in their own fields. Importantly, the goal of achieving empathy works in tandem with the goal of leading students to higher-order thinking, since research shows “an inverse relationship between empathy and dogmatic thinking or intolerance of difference” (Burns 2001, 16).

Second, the empathetic study of heresy can give instructors a fresh approach that incorporates the critical tools of religious studies without sacrificing personal or institutional convictions. Since critical study is greatly enhanced by attention to difference, a Christianity course which has no benefit of a comparative religions approach may particularly profit from deeper investigation of heresies. It is also important for teaching at a Christian institution, where there is a likelihood that students and instructors alike may perceive the task to be understanding or confirming one’s own religious tradition, which can result in little attention to marginal voices or—when they are treated—focusing on their theological errors or faults. More intentional engagement with Christian heresies may counteract this tendency.

Methods

Participants

Participants were students enrolled in two sections of The Christian Heritage. One section (n = 60) served as the comparison group and was taught by an instructor otherwise not involved in this study. This course was not a control group, strictly speaking, as the content and pedagogy were neither monitored nor controlled. Rather, this course serves as a loose proxy for the results of a typical instructor at our institution not specifically concerned with teaching for empathy. The other
section (n = 52), taught by one of this study’s authors (Richmann), served as the experimental group. The experimental group was distinguished by the role-playing exercises, small group discussion workday on students’ chosen heresy, and the four-question essay described above. This study received our institution’s IRB exemption, part of which entailed students consenting for the results to be shared for the potential benefit of future instructors.

**Tools and Procedures**

This study used mixed methods, gathering qualitative and quantitative data to determine if the project enhanced students’ empathy. Qualitative data came from two student artifacts: the four-question essays (experimental group only) and a post-project questionnaire with an open-ended question (comparison and experimental groups). Because written work often prioritizes the finished product or “correct” answers, it may give less insight into students’ progression of thought than, say, interview or group work observation data. Yet written work can be carefully crafted to prompt students to articulate their thought processes (Ashby and Lee 1987, 65).

**Table 1. Modified Five Levels of Empathy (substantive changes in italics)**

<table>
<thead>
<tr>
<th>Level</th>
<th>Ashby and Lee</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1: “Divi” Past (“divi” is British slang for “thick”)</td>
<td>regards the behavior of those in the past as stupid and unintelligible</td>
<td>regards the behavior of those in the past as stupid and unintelligible</td>
</tr>
<tr>
<td>Level 2: Generalized Stereotypes</td>
<td>uses conventional accounts of people’s intentions, situations, values, and goals, without reference to context</td>
<td>misrepresents context or uses conventional accounts of people’s intentions, situations, values, and goals, without reference to context</td>
</tr>
<tr>
<td>Level 3: Everyday Empathy</td>
<td>uses evidence of the historical situation but treats it largely in modern terms and often involves some form of personal projection (“What would it have been like for me if I had been there?”)</td>
<td>uses evidence of the historical situation but leaves context implicit or incomplete; often involves some form of personal projection (“What would it have been like for me if I had been there?”)</td>
</tr>
<tr>
<td>Level 4: Restricted Historical Empathy</td>
<td>acknowledging that standards for historical behavior differ from ours, uses evidence of the specific situation to understand historical behavior, giving individual actors the benefit of the doubt and explaining institutions somewhat narrowly in terms of their explicit functions</td>
<td>uses evidence of the historical situation, making the context explicit but not well-connected to historical actors’ beliefs or behaviors</td>
</tr>
<tr>
<td>Level 5: Contextual Historical Empathy</td>
<td>considers a wider picture beyond the overt goals or intentions of people and institutions and explores the ways in which standards for historical behavior differ from ours</td>
<td>considers a wider picture beyond the overt goals or intentions of people and institutions, attempts to integrate the concerns of opposing stances, or acknowledges and explicitly connects multiple relevant contexts to historical actors’ beliefs or behaviors</td>
</tr>
</tbody>
</table>
Essays. In assessing the four-question essays, raters used a modified set of five empathy levels (Table 1) originally developed by observing school-age children working in discussion groups as they explain and discuss historical topics (Ashby and Lee 1987).6

To establish some context and baseline for rating, two raters assessed six essays independently and compared the results. Ratings on five of the six essays were identical; one essay received different ratings (3 from one rater; 4 from the other). The rest of the essays were rated by a single rater. While individual student writing often displayed multiple levels, each essay was given an overall level, generally associated with the highest level unambiguously achieved.

Open-Ended Response. At the end of the semester, all participants (comparison and experimental) were asked to respond to this question:

*How would you describe what a “Christian heretic” is to someone who has not taken this class?*

Through inductive analysis, student responses to this question were coded in three groups: Descriptive, Non-understanding, or Judgmental. Descriptive responses generally explained heresy as a social construct: deviance from dominant or majority belief with no inherent positive or negative value. While recognizing that this principle of contextualization does not guarantee empathy in any given situation, it is evident in all empathy. We suspect that recognition of this principle can be a useful foundation for empathy in more specific situations. Non-understanding responses were those that were either left blank, contained blatant inaccuracies, were nonsensical, irredeemably vague, or admitted not understanding the term “Christian heretic.” Judgmental responses displayed a negative view of heretics and/or made no attempt to contextualize heresy, often referring implicitly to deviance from static or objective theological truth. While it is possible to empathize with an idea or group and still have a negative view of it, such a position requires more nuance than was evidenced in these responses, which were completely negative and/or totally lacked any contextualization.

Empathy Questionnaire. In addition to quantifying the empathy level of essays and categories for open-ended responses, quantitative data came in the form of a validated empathy questionnaire, the Interpersonal Reactivity Index (IRI) (Davis 1980, 1983; See Appendix C) administered at the beginning of the semester and again at the end of the semester to all participants. The IRI is composed of four independent subscales each measuring a facet of empathy.7 Each subscale consists of seven items, each of which participants respond to using a 5-point Likert-type scale (0 = Does not describe me well to 4 = Describes me very well). Although this tool was designed as a general empathy scale rather than for measuring historical empathy, its use is an advance over previous methods which have been tied generally to specific historical topics rather than a generalized empathetic capacity (Endacott and Brooks 2018; Hartmann and Hasselhorn 2008; Huijgen et al. 2014). Furthermore, historical empathy has been linked to pro-social behaviors like conflict resolution and self-awareness, suggesting that empathy is to some degree transferrable from historical subject to interpersonal subject (Gehlbach 2004). As Ashby and Lee (1987, 64) put it, “pupils who . . . have learnt the kind of moves that help them achieve empathy, have not just taken a step forward in their history [education], but are more likely to be able to cope with the present world.” Since this study defines empathy as *the capacity to know the internal state of another*, the “perspective taking” subscale of the IRI is particularly applicable, although a previous study on historical empathy found no change on IRI scores or the “perspective taking” subscale (Dillenburg 2017).

To explore correlations between empathy levels and beliefs in the malleability of empathy, we administered the Theories of Empathy (ToE) scale (Schumann et al., 2014, see Appendix D) to all participants at the beginning of the semester. Students

---

6 This tool has limitations for the current project, since it emerged inductively from discussion groups (not written work) with school-aged children (not college students). Yet it is most applicable framework available in the literature. The fact that students in the present study scored on a range of the levels (see Results section) suggests that developmental differences between school-aged children and college students are not a significant factor. On the other hand, the need to modify Ashby and Lee (1987) arose from the difference in student task. The original levels developed from observation of students, in group discussion prompted by open-ended questions, “making sense” of historical behavior researchers expected to appear strange to students. As such, the original levels failed to capture some of the nuances of student responses in this study, such as nominal empathy without rationale and application of historical evidence without referencing potential differences in meaning across time.

7 Davis’s IRI includes subscales Fantasy (Imagination), Perspective Taking (Cognitive), Empathetic Concern, and Personal Distress (Affective).
indicated their level of agreement with six items on the ToE using a 7-point Likert-type scale (1 = Strongly disagree to 7 = Strongly agree). Negative items were recorded, and a mean ToE score was calculated such that higher scores indicate greater belief in the malleability of empathy.

As explored in the “Limitations and Further Research” section below, this type of quantitative data, while potentially useful in offering generalizations and comparisons, risks obscuring the nuances of students’ experiences and performance. As will be seen below, qualitative data in the form of student writing provides greater insight into the unanticipated themes of student learning and depth of student understanding.

Results

Essays

The median empathy rating for essays (n = 50) was 3. Although writing displayed a range of empathetic levels from 2 to 5, most students’ essays were rated level 3 (see Table 2). Given the ordinal nature of our ratings, we used Spearman’s correlation to explore the relationship between empathy ratings and student essay grades, which were given independently. We found a moderate correlation between empathy ratings and student grades on the essay, $r = .49986, p = .0002$; significant at $p < .05$.

Table 2. Essays Rated on Empathy Levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>0</td>
</tr>
<tr>
<td>Level 2</td>
<td>9</td>
</tr>
<tr>
<td>Level 3</td>
<td>30</td>
</tr>
<tr>
<td>Level 4</td>
<td>9</td>
</tr>
<tr>
<td>Level 5</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
</tr>
</tbody>
</table>

**Level 1.** No student writing was assessed at this level.

**Level 2.** Writing at this level overgeneralized to the point of distortion. One student, in ostensibly giving reasons for Docetism’s appeal, claimed, “Most early Christians were not very intelligent due to lack of education[,] so Docetism was very convenient for many.” Another student also pushed the appeal so far as to misrepresent the heresy (Arianism) she was trying to explain. “How can one be the same as the other? . . . Jesus came from human flesh, God did not, so what exactly makes them the same?” Not only did this student omit the importance of monotheism as theological context for Arianism, she also implied that Arians did not teach Christ’s incarnation. It was not only the heretics or Christians in general who were stereotyped. One student, in explaining the orthodox reaction to Docetism, stated, “This concept was impossible to grasp by the traditional and widespread teachings of the mainstream church.”

Students also rated level 2 if they misrepresented the context. One student, writing about resistance to Pelagianism, argued, “Most [people] enjoyed the thought of being able to live how they pleased and just paying for their faults to reach salvation.” This student appears to be importing notions about medieval indulgences into her discussion of Pelagianism. Furthermore, since the prompt asked students to write in the persona of an adherent of a heresy, some students took this as license to caricature the orthodox. According to one student, “Pelagius’s theology was rejected solely because of Augustine and the sheer amount of power that he had over the public and the mainstream church.” This “control,” claimed the student, resulted in “closed mindedness” among the orthodox. Another student referred to the orthodox opponents...
of Pelagianism as “those who wish to deny the truth of God’s intention for human nature.” This hints at a challenging subtlety of empathy; it is not the same as identification. Said another way, the distinction between empathy and sympathy is crucial, since empathy requires perspective-taking while maintaining critical perspective.

**Level 3.** Student writing in level 3 often referred to the logic or simplicity of the heresy when seeking to explain its appeal, while making only minimal attempt to contextualize the logic or failing to demonstrate full understanding of the context. Such students portrayed logic and simplicity as self-evident and timeless. As one student wrote in persona of a Pelagian, “By believing in our ability to keep the commandments, we ascribe to the command[']s fairness.” This student does not acknowledge that concepts such as “fairness” are historically determined and socially constructed. In discussing Monophysitism, one student claimed, “Logically by the fact that one is infinite and the other is finite, Jesus' human portion was overwhelmed [by divinity] to the point that while it was still there, it was likely undetectable.” “Many Christians chose Modalism,” wrote another student, “because it was closely related to something that was already being praised by high members of the church [i.e., the Trinity] . . . But I think that the handful of us Christians that didn't fully understand it, looked for a similar concept, but with a more simplistic and coherent reasoning.” If this student had demonstrated an appreciation for the fact that “the Trinity” was not yet a fully formed concept in the third century and explained why Modalism was simpler and more coherent than orthodox views of the godhead, these statements could have reached level 4.

This appeal to logic was sometimes implicit and could even factor into a negative assessment of the heresy, even though they were prompted to write in persona of a heretic. One student wrote that Arius’ “literal interpretation of the New Testament is ultimately flawed in that it tries to describe the unknowable God and the Trinity in limited, human terms.” Another student, writing on Pelagius, wrote, “Where Pelagianism failed and the reason it ended up being rejected by so many . . . is because of the many fallacies it contains . . . These clear ideas that are so contradictory to scripture is arguably the greatest reason that Pelagianism was declared a heresy.” These types of comments frequently appeared in the “why do you think your position was rejected by others?” section of the essay, which suggests that some students had difficulty fairly presenting the opposition through a lens that still maintained a basic posture of openness to the heresy—a challenging sort of double-perspective-taking exercise.

While mentioning historical context, student writing in level 3 often left the origin or impact of context implicit or incomplete. For instance, one student rightly noted the moral and social chaos of Rome in the fourth century (“Rome was a mess”). But rather than concretely connect the social conditions with Pelagius’ ascetic alternative, the student simply presented Pelagius as a sauntering hero, who offered “an alternative to the current way of living in Rome.” More detail on Rome’s “mess” and Pelagius’ moral “alternative” would likely have moved these reflections to level 4. Many students took for granted rather than explained why early Christians regarded Hebrew scripture or apostolic tradition as authoritative. As one student put it, “If you look to scripture there is evidence or proof that what Pelagius believed could very well be true.” Another student commenting on Pelagius overlooked the biblical/theological context even while emphasizing social context: “Many of these [new converts] continued to live in their old, sinful ways, yet carried the Christian title . . . Pelagius soon spoke out and pressed for better behavior from Christians.” This statement may have rated level 4 if the student had explained that Pelagius judged Roman Christian behavior by certain inherited biblical standards and admonitions.

**Level 4.** Level 4 writing made historical and theological context explicit and loosely connected this context to heretical teachings and orthodox responses. After mentioning the Jewish Revolt and Roman hostility to Judaism, one student wrote, “Marcion's teachings provided a reason for Gentile Christians to break from the Jewish roots of Christianity.” In establishing the “problem” that Modalistic Monarchianism sought to solve, another student wrote simply, “Wanting to uphold the belief in there only being one God could be traced back to Judaism.” Such responses indicated students appreciated the origins and impact of the context for heresies but often failed to connect all the historical dots (e.g., why was Christianity largely beholden to Jewish monotheism?).

**Level 5.** The distinguishing feature of this level is an attempt at comprehensiveness: students acknowledge multiple issues in context as relevant for explaining historical beliefs and behaviors. As one student explained:
Ebionism is a broad stroke for mitigating conflicting theologies. There are a lot of differences to be found between Jesus’ teachings and what is written in the Old Testament that are resolved in this way of thinking, as well as the issue of monotheism is addressed directly in this wonderful sect. The greatest strength of this religion is that it solves multiple theological problems that have been the topic of debate for some time now.

This student’s articulation of the multiple problems of Christianity’s relationship to Judaism set the stage for empathetic treatment of both Ebionism and orthodoxy. This implies an appreciation for how opposing historical stances can be integrated by explaining their common concerns. In this student’s words, Ebionism provides “a way to follow Jesus while still remaining faithful to the Mosaic teachings,” something all early Christians favored.

Open-Ended Response

Overwhelmingly, a greater percentage of students in the experimental group gave Descriptive responses (see Table 3). They were much more likely to explain heresy as a concept relating to social or historical realities, rather than religious truth as such. In a particularly articulate response, one student asserted that a Christian heretic is “Someone w/ unconventional theological beliefs for a specific time in history. It really depends on the historical time period because Christianity is not a static set of beliefs.” Correspondingly, students in the experimental group were far less likely to give Judgmental responses. It was much more typical of the comparison group, for instance, for a student to respond that a heretic is “someone who takes Christianity and twists it for their own purposes,” to call heresy “untrue,” “incorrect and/or harmful,” or to refer to a heretic as “one who distorts or perverts the Gospel.” Also worth noting, students in the experimental group were much less likely to give a Non-understanding response.

Table 3. Comparison and Experiment Group Open-Ended Responses

<table>
<thead>
<tr>
<th></th>
<th>Comparison Group (n = 51)</th>
<th>Experimental Group (n = 44)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>9.8% (5)</td>
<td>81.8% (36)</td>
</tr>
<tr>
<td>Non-understanding</td>
<td>43.3% (22)</td>
<td>4.5% (2)</td>
</tr>
<tr>
<td>Judgmental</td>
<td>47.1% (24)</td>
<td>13.6% (6)</td>
</tr>
</tbody>
</table>

Figure 1. Comparison and Experimental Group Open-Ended Responses
In the comparison group, many students seemed to have confused “heretic” with “hypocrite,” leading to some of the judgmental responses. As one student remarked, a Christian heretic “is someone who goes against his or her own beliefs.” Such responses may more properly belong in the Non-understanding category, but there is insufficient evidence to make this claim.

**Empathy Questionnaire**

All statistical analyses were done using IBM SPSS Statistics V. 25. Pre- and post-test data was available for 43 students in the comparison group and 38 students in the experimental group. Students who were missing either the entire pre- or post-test were removed from analyses, and missing values on individual items were replaced with the median. We ran Pearson correlations between students’ scores on the Perspective Taking (PT) subscale of the IRI and their ToE scores. Students’ pre-test PT scores were not significantly correlated with their ToE scores, $r = .12, p = .272$, but their post-test PT scores were marginally related to their ToE scores, $r = .22, p = .051$. Greater perspective-taking showed a small- to medium-sized association with greater belief in the malleability of empathetic ability. This trending effect disappeared, however, when students were split into groups.

To explore how the experimental manipulation affected students’ self-reported empathy over the course of the semester, we conducted a 2 (time: pre- or post-test) x 2 (group: comparison or experimental) repeated-measures analysis of variance (ANOVA) on students’ scores on the PT subscale. Time was a within-subjects variable and group was a between-subjects variable. There was no effect of time, $F(1, 79) = 0.35, p = .554$, group, $F(1, 79) = 2.40, p = .125$, or their interaction, $F(1, 79) = 10.36, p = .119$. We next ran a 2 x 2 repeated-measures analysis of covariance ANCOVA with ToE included as a covariate. This allowed us to control for any effects of students’ beliefs about the malleability in empathy on their PT scores over time. In line with expectations, ToE scores approached significance in accounting for variance in students’ PT scores, $F(1, 78) = 3.33, p = .072$. The size of the effect of ToE scores on PT scores was small to medium, $d = 0.41$. With ToE scores included as a covariate, time remained non-significant, $F(1, 78) = 0.99, p = .322$, as did the time by group interaction, $F(1, 78) = 2.76, p = .101$. However, group became marginally significant, $F(1, 78) = 2.88, p = .093$, and had a relatively medium-sized effect on PT score change over time, $d = 0.41$. After controlling for ToE scores, those in the comparison group had slightly higher PT scores than those in the experimental group ($M_{\text{Comparison}} = 19.61, SE = 0.58; M_{\text{Experimental}} = 18.18, SE = 0.61$).

**Discussion**

The results from the project are mixed. The essays give consistent evidence that, following this intervention (although not necessarily as a result of it), students display “Everyday Historical Empathy,” meaning that in explaining a heresy, they use evidence of the historical situation but often leave context implicit or incomplete, frequently with some form of implicit personal projection, such as appeal to what seems “logical” or “simple.”

The open-ended responses showed clearly that the experimental group was more able to describe “Christian heretic” as a socially constructed category rather than judge it as theologically, morally, or intellectually deficient. The intervention, with multiple learning activities centered on contextualizing heresies, seems to have enabled students to better understand the nature of heresy as a historical phenomenon.

In contrast, the intervention did not have significant effects on participants’ PT scores on the IRI. In other words, this exercise in historical empathy has not resulted in measurable gains in interpersonal empathy, results consistent with Dillenburg (2017). Although the comparison group had slightly higher PT scores, this was not observed as a time by group interaction, pointing to a pre-existing characteristic of the comparison sample rather than a result of the intervention.
Limitations and Further Research

While it may be good that no essays were rated level 1, there are many ways to explain this. As noted above, the prompt itself led students away from any hostile or dismissive approach to the heretics they studied. If given an open forum to “make sense” of heresies, these students may have performed differently. Evidence from the essays does not illuminate the crucial difference between student performance on a graded assignment and genuine student attitudes and dispositions. For a fuller picture, additional research could compare graded written work with observation of open student discussion. This may help us to see how much effect the wording of a prompt and the nature of the task (graded or non-graded) have on empathy levels.

The issue of transferability remains opaque. We do not know to what degree the experimental group’s penchant for non-judgmental description of “Christian heretic” may transfer to how students conceive of other historical actors and institutions. Further studies could, for instance, focus an intervention on one often-disparaged historical subject (e.g., heresies) and measure students’ attitudes about another disparaged or stereotyped subject (e.g., liberal theology in a more conservative institutional context or fundamentalism in a more liberal environment).

The IRI purports to measure empathy in personal, everyday interactions. To register effects of the intervention on the IRI, participants ostensibly need to develop, transfer, enact, and report interpersonal empathy. Each of these steps provides an explanation of our findings that this intervention did not significantly affect students’ PT score on the IRI. Participants may not have become more empathetic in a general sense, in which case this intervention was not pedagogically effective for teaching empathy. They may have developed historical empathy (as suggested by the open-ended responses and the essays) without transferring this to everyday interpersonal empathy. Cognitive research shows that transfer is among the most sophisticated learning tasks and is especially challenging when knowledge is overly contextualized (National Research Council 2000). If historical empathy is regarded as an important outcome on its own (without regard to interpersonal empathy) it may require its own psychometric tools to assess—a task for further research. Still, we cannot conclude from this limited study that historical empathy does not transfer to interpersonal empathy; perhaps this process takes longer than a semester, and studies of greater longitude could provide more insight. Items on the PT subscale also focus on enacting empathetic behaviors and attitudes (e.g., “I try to look at everybody’s side of a disagreement before I make a decision.”). Again, time is a factor; perhaps a semester is not long enough to give participants enough opportunities to enact such empathetic dispositions in daily life. Finally, as with all questionnaires, the IRI relies on self-reporting, measuring self-perceptions of attitudes abstracted from interactions, not actual attitudes in the moment of interaction. Participants could be behaving in more empathetic ways without realizing it.

The weak correlation between IRI and ToE might challenge previous findings (Schumann, Zaki, and Dweck 2014). It should be noted, however, that these previous studies explored the relationship between a malleable mindset regarding empathy and expending empathetic effort, not changes over time of self-reporting of interpersonal reactivity. Additional studies could confirm or refute our finding, but instructors attempting to “teach empathy” may be heartened by the suggestion that the effect is not necessarily predetermined by students’ existing malleability mindset.

A major limitation of this study is its small scope. In particular, a larger comparison group and/or a more rigidly controlled comparison group would give more reliable data on the open-ended response. If, for instance, the comparison group class simply did not spend much time learning about heresies, their inability to describe “Christian heretic” as a social construct may be more related to lack of familiarity than lack of empathy as such.

Finally, the lack of comparison group for essays limits the inferences we can draw. The essays provided key data but were also part of the intervention itself. In this model, a comparison group writing the same essay is not possible. Additional studies could devise writing assignments common to both experimental and comparison groups that are not part of the intervention.
Implications for Teaching

A benefit of assessing student writing with the modified empathy levels is that it can help instructors distinguish between lower- and higher-order thinking. The moderate correlation between essay levels and grades suggests that empathy is an element of critical thinking, particularly as a major category for grading was development of argument, including “explanation of points of comparison/contrast . . . main points are . . . legitimate, based on sources, and establish relevance.” Many students included a great deal of accurate and pertinent information (lower-order thinking) without displaying much empathy. Essays on Modalism, for example, often mentioned the importance of monotheism in their appeal but also failed to explain why monotheism mattered to early Christians. In discussing Monophysitism, one student explained at length the challenge of reconciling divinity and humanity in the single person of Jesus, but he never explained the Greek thought-world that gave Christians their notions of the attributes associated with divine and human. An essay on Marcion presented the “problem” of “the incompatibility between Jesus of Nazareth’s teachings about the love of God and the violent acts accredited to, or ordered by, YHWH, the god of Israel as reported in the Old Testament.” While this student named the central matter of Marcionism, she did not make explicit why this incompatibility posed a problem—that is, the dominant Christian belief that Jesus represented the God of the Old Testament. Insofar as empathy or perspective-taking coincides with historical understanding of cause-and-effect relationships, the concept of empathy can provide instructors a helpful view of the ways to encourage, support, and assess higher-order thinking.

The results of this study suggest that role-playing exercises and written explanation of a Christian heresy, in persona of a heretic and refined in discussion with peers, can help students develop empathy for Christian heresies. But there is room for improvement. In mid-semester feedback forms, several students in the experimental group, without specific prompting, mentioned the helpfulness of the role-playing exercises, suggesting more and varied role-playing might have a greater effect on empathy. As empathy has been shown to develop fluidly in conversation with peers (Ashby and Lee 1987, 86), we speculate that more group work and small group discussion related to Christian heresies may increase the effect.

Furthermore, instructors who wish to increase the likelihood of empathy transfer to other areas of historical study or to interpersonal interactions should consider learning activities associated with transfer, such as balancing instruction on the specifics of Christian heresies with more general discussion about the nature of heresy, solving multiple distinct cases that help students abstract key principles themselves, prompting students with analogous experiences when addressing new situations (without telling them what or how to apply this information), or requiring students to reflect on their own thinking (National Research Council 2000). For example, the project outlined above could be extended to focus on a second heresy. Rather than research in books and articles, however, students would be asked to learn about the heresy by conducting an interview with the instructor playing the role of an adherent of the heresy. The instructor would answer questions politely but tersely, meaning students who wanted fuller information would need to formulate probing, gracious, and empathetic questions. They would then craft questions for a hypothetical interview attempting to understand someone who has very different views from them on some religious, political, or social issue. Finally, they would write a brief reflection describing how they connected these three major assignments. Such an exercise would engage students in the abstraction and particularizing, application to new situations, and metacognition associated with transfer.

BIBLIOGRAPHY


DEVELOPING AND ASSESSING EMPATHY


**About the Authors**

Christopher Richmann is Assistant Director for the Academy for Teaching and Learning and affiliate faculty in Religion at Baylor University.

Courtney Kurinec is a postdoctoral fellow in the Department of Psychology at Washington State University. She specializes in the study of memory, decision-making, and metacognition and learning.

Felicia Osburn is a Ph.D. candidate in the Department of Biology and graduate fellow with the Academy for Teaching and Learning at Baylor University.
APPENDIX A: Essay Prompt

Heresies: The Forgotten Heroes of the Christian Tradition

A major theme in the study of Christianity is “heresy vs. orthodoxy.” Heresies are schools of thought that came to be considered incorrect or even dangerous by most other Christians. This majority, then, defined its core theological stance as orthodoxy, or “correct belief.” Although there may be good theological reasons for rejecting particular heresies, it is important to understand that heresies played an important role in shaping Christianity. Oftentimes, the orthodox stance on an issue was only defined in the course of reacting against theological ideas that didn’t “sit right” with the majority. This majority then worked hard to explain its position in positive terms that clearly excluded these heresies. In other words, orthodox Christians have heretics to thank for helping them articulate their positions!

In this essay, you will focus on one of the major heresies of Christianity. Your task is NOT to defend the orthodox position or point out the flaws of the heresies. In fact, it is almost the opposite: I want you to gain sympathy and understanding for the “heretics” of the Christian tradition by attempting to explain and analyze its historical context and its theological content. You will do this by answering four questions, in the persona of an adherent of this school of thought:

1. What theological problem are you trying to address? (Give full historical context and explain the logic that led to this way of thinking.)

1. What is your solution? (Be detailed and precise.)

1. Why is this solution attractive to you and other people you know? (Give attention to both historical and theological issues. Historical arguments focus on the natural human forces that were at play, such as power structures, wealth, or personalities. Theological arguments focus on the internal logic, consistency with sacred texts and traditions, compatibility with cultural assumptions, or rhetorical power of the position.)

1. Why do you think your position was rejected by others? (Also consider historical and theological issues.)

Heresies/Heretics to consider:

Docetism
Marcionism
Modalism/Modalistic Monarchianism
Arianism
Montanism
Ebionism
Donatism
Pelagianism
Monophysitism
APPENDIX B: Role-Playing Response Form

**Reflection Questions (for observers of role-playing)**

1. What do you think was the most important issue for Joseph in this conversation? What makes you think this?
2. Do you think the Joseph got points across? What makes you think this?
3. Imagine yourself in Joseph’s shoes. How do you feel about this conversation when it is over? Include in your response a specific exchange from the conversation.

APPENDIX C: Interpersonal Reactivity Index

**NAME__________________________**

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D, or E. When you have decided on your answer, circle the letter on the answer sheet below the item number. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

**ANSWER SCALE:**

A   B   C   D   E

DOES NOT DESCRIBES ME
DESCRIBE ME WELL   VERY WELL

1. I daydream and fantasize, with some regularity, about things that might happen to me.
   A   B   C   D   E

2. I often have tender, concerned feelings for people less fortunate than me.
   A   B   C   D   E

3. I sometimes find it difficult to see things from the “other guy’s” point of view.
   A   B   C   D   E

4. Sometimes I don’t feel very sorry for other people when they are having problems.
   A   B   C   D   E

5. I really get involved with the feelings of the characters in a novel.
   A   B   C   D   E
6. In emergency situations, I feel apprehensive and ill-at-ease.
   A B C D E

7. I am usually objective when I watch a movie or play, and I don’t often get completely caught up in it.
   A B C D E

8. I try to look at everybody’s side of a disagreement before I make a decision.
   A B C D E

9. When I see someone being taken advantage of, I feel kind of protective towards them.
   A B C D E

10. I sometimes feel helpless when I am in the middle of a very emotional situation.
    A B C D E

11. I sometimes try to understand my friends better by imagining how things look from their perspective.
    A B C D E

12. Becoming extremely involved in a good book or movie is somewhat rare for me.
    A B C D E

13. When I see someone get hurt, I tend to remain calm.
    A B C D E

14. Other people’s misfortunes do not usually disturb me a great deal.
    A B C D E

15. If I’m sure I’m right about something, I don’t waste much time listening to other people’s arguments.
    A B C D E

16. After seeing a play or movie, I have felt as though I were one of the characters.
    A B C D E

17. Being in a tense emotional situation scares me.
    A B C D E

18. When I see someone being treated unfairly, I sometimes don’t feel very much pity for them.
    A B C D E
19. I am usually pretty effective in dealing with emergencies.
   A  B  C  D  E

20. I am often quite touched by things that I see happen.
   A  B  C  D  E

21. I believe that there are two sides to every question and try to look at them both.
   A  B  C  D  E

22. I would describe myself as a pretty soft-hearted person.
   A  B  C  D  E

23. When I watch a good movie, I can very easily put myself in the place of a leading character.
   A  B  C  D  E

24. I tend to lose control during emergencies.
   A  B  C  D  E

25. When I'm upset at someone, I usually try to “put myself in his shoes” for a while.
   A  B  C  D  E

26. When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me.
   A  B  C  D  E

27. When I see someone who badly needs help in an emergency, I go to pieces.
   A  B  C  D  E

28. Before criticizing somebody, I try to imagine how I would feel if I were in their place.
   A  B  C  D  E
APPENDIX D: Theories of Empathy Scale

Using the scale below, please indicate your agreement with each of the following statements.

1                         2                         3                         4                         5                         6                         7

STRONGLY DISAGREE                      STRONGLY AGREE

1. A person’s level of empathy is something very basic about them, and it can’t be changed much. ______

2. Whether a person is empathic or not is deeply ingrained in their personality. It cannot be changed very much. ______

3. People can’t really change how much empathy they tend to feel for others. Some people are very empathic and some aren’t and they can’t change that much. ______

4. No matter who somebody is, they can always change how empathic a person they are. ______

5. People can always change how much empathy they generally feel for others. ______

6. Anybody can change how empathic a person they are. ______