

Running Head: SOCRATIC NOTE TAKING

### Abstract

The notion of Socratic Note Taking (SNT) is introduced to enhance students' learning from assigned readings. SNT features students asking questions and answering their own questions while doing the readings. To test the effectiveness of SNT, half the students from two sections of a philosophy course were assigned SNT on alternating weeks. Quizzes each week alternated between the two classes as either high or low stakes in a counterbalanced format. The design was a 2 (Notes: SNT or not) x 2 (Stakes: high or low) x 2 (Replication: first or second replication of a Notes x Stakes cell) within-participants factorial. On 10-point quizzes, SNT made an average difference of 1.22 points (more than a letter grade). In effect size terms that take error variance into account,  $\eta_p^2 = .43$ . Furthermore, the results indicate that SNT is particularly effective with weaker students, e.g., we found a nearly 3-point increase on 10-point quizzes for the weakest students.

Keywords: Socratic Note Taking; study compliance problem; motivation problem; reading compliance problem, quizzing

The Socratic Note Taking Technique: Addressing the problem  
of students not engaging with assigned readings before class

Mark Walker

David Trafimow

Jamie Bronstein

Our research addresses a familiar problem: many of our students do not read or study the assigned materials before class. We have developed a technique, The Socratic Note Taking Technique (hereafter SNT), which shows promise in mitigating this problem. A classroom experiment and subsequent survey demonstrates that SNT better addresses these problems than quizzing.

**The Reading Compliance problem**

The problem of students not engaging with assigned readings can be broken down into three sub-problems: the ‘reading compliance problem’, the ‘study compliance problem’, and the ‘motivation problem’.

The ‘reading compliance problem’, as we shall term it, refers to the fact that voluminous research across a wide array of disciplines shows that many students, often most, do not read the assigned texts on time. For example, Clump, Bauer, and Breadley (2004) found that about 27% of psychology undergraduate students read the assigned material before class. Professor Howard (2004) found that only 40% of his sociology students completed the assigned readings prior to class. Similar findings are reported from studies involving students studying philosophy (Brost et al., 2006), business (Artis, 2008; Starcher et al., 2011), science (Henderson et al., 2006; Jensen et

al., 2008), communication studies (Peterson, 2006), and education (Carney et al., 2008).

Evidence suggests a long term trend for an increase in the severity of the reading compliance problem (Burchfield et al., 2000; Clump et al., 2004; Sappington et al., 2002).

### **Study Compliance Problem**

The reading compliance problem is part of a larger problem, which we will refer to as the ‘study compliance problem’. Consider that in order to fully grasp the assigned readings, it is necessary but not sufficient for students to read. As Roberts and Roberts describe it, when given a reading assignment “. . . some students feel they have met their obligation if they have forced their eyes to “touch” (in appropriate sequence) each word on the assigned pages” (Roberts et al., 2008, p. 125). The fact that mere reading is not sufficient is confirmed when we reflect on why we ask students to engage with materials before class. Perhaps the most commonly cited reason is that it is not possible to cover all the necessary material during class (Ryan, 2006). Other reasons include a deeper understanding of the course material, and better class participation (Gurung, 2003; Narloch et al., 2006). All these goals are enhanced when students have seriously engaged with the material prior to class.

It is worth saying a bit more about the reading and study compliance problems. The term ‘read’ can be understood in a wide and narrow sense. In the narrow sense, ‘read’ means simply having read, in the manner noted above, where one’s eyes have passed over each word of the text. In the wide sense it means reading *and* studying the assigned readings. Other ways of expressing the wide sense of ‘read’ are suggested by the phrases that students should “read for comprehension”, “critically read”, or “engage with the readings.” Some illustration of the

difference can be found in a familiar example: a student comes to your office for help because he has failed the first two quizzes. When asked, he is quite adamant that he did the assigned readings. The obvious follow-up questions are to enquire about his study strategies: Did he take notes? Did he highlight the text? Did he quiz himself on the material? To these questions, the student has a pretty good response: “The syllabus says to read the assigned text, it says nothing about highlighting, note taking, or self-quizzing.” In writing on our syllabi that students should read the assigned text, many of us are thinking of the term ‘read’ in the wide sense, whereas some students—perhaps most—understand it in the narrow sense. It would behoove all of us to use ‘study’ on our syllabi and other course materials where we mean ‘read’ in the wide sense, to avoid such ambiguity.

In what follows, we will use ‘read’ in the narrow sense and ‘study’ to indicate the broad sense. It is perhaps worth mentioning that it would be wrong to think there is some sort of bright-line between reading and studying; rather, we should think of reading and studying as points on a continuum between merely passing one’s eyes over the material to diligently applying proven study techniques.

Given that we hope students will study (as opposed to merely read) the assigned materials, the question naturally arises as to what study techniques students should use to improve learning. A recent meta-survey of literature on effective study techniques rated ten different techniques in terms of their overall utility (Dunlosky et als., 2013).<sup>1</sup> Five techniques,

---

<sup>1</sup> The authors provide a brief description (p. 6) of the ten techniques as follows:

rereading, highlighting/underlining, summarization, keyword mnemonic, and imagery for text, were rated low in utility. Three techniques, elaborative interrogation, self-explanation, and interleaved practice, earned a moderate rating in utility. Practice testing and distributed practice were the only two techniques to earn a high overall utility rating. As these researchers note, ample research indicates that the most common study techniques employed by students are highlighting of texts and rereading. Unfortunately, research indicates that these are the least effective, hence their low utility scores (Dunlosky et al., 2013). So, the study compliance problem encompasses both the fact that students may merely read as opposed to study the assigned materials, and if students study, they will often employ the least effective study techniques.

### **Motivation Problem**

The motivation problem is how to get students to read and use effective study techniques with the assigned material before class. The exasperation some faculty feel about this problem is evident in Professor Maryellen Weimer's introduction to a special report on getting students to read what is assigned:

- 
1. Elaborative interrogation: Generating an explanation for why an explicitly stated fact or concept is true
  2. Self-explanation: Explaining how new information is related to known information, or explaining steps taken during problem solving
  3. Summarization: Writing summaries (of various lengths) of to-be-learned texts
  4. Highlighting/underlining: Marking potentially important portions of to-be-learned materials while reading
  5. Keyword mnemonic: Using keywords and mental imagery to associate verbal materials
  6. Imagery for text: Attempting to form mental images of text materials while reading or listening
  7. Rereading: Restudying text material again after an initial reading
  8. Practice testing: Self-testing or taking practice tests over to-be-learned material
  9. Distributed practice: Implementing a schedule of practice that spreads out study activities over time
  10. Interleaved practice: Implementing a schedule of practice that mixes different kinds of problems, or a schedule of study that mixes different kinds of material, within a single study session

Getting students to take their reading assignments seriously is a constant battle. Even syllabus language just short of death threats, firmly stated admonitions regularly delivered in class, and the unannounced pop quiz slapped on desks when nobody answers questions about the reading don't necessarily change student behaviors or attitudes (Weimer, 2010, p. 2).

The fact that we recognize but often fail to motivate students to read and study leads to a familiar catch-22: since many of our students do not come to class prepared, we spend valuable class time covering the assigned readings. Consequently, students quickly learn to anticipate that the assigned readings will be covered in class, and so are less likely to engage with the readings (Brost et al., 2006; Weimer, 2010).

The motivation problem is further compounded when we consider that motivation to read the assigned material is not necessarily sufficient to get students to use effective study habits. In other words, motivating reading is not the same thing as motivating the adoption of effective study techniques. Consider, for example, the self-report of Young, a highly motivated medical student who went from near the bottom of his cohort to near the top by using a study technique, retrieval practice, which involves self-quizzing:

It makes you uncomfortable at first. If you stop and rehearse what you're reading and quiz yourself on it, it just takes a lot longer. If you have a test coming up in a

week and so much to cover, slowing down makes you pretty nervous... You just have to trust the process, that was really the biggest hurdle for me, was to get myself to trust it. And it ended up working out really well for me (Brown et al., 2014, pp. 213-214).

If a highly motivated medical student is reluctant to use non-standard study techniques, we have reason to worry that average students will not use non-standard study techniques, even if their advantages are demonstrated. The literature confirms this worry: students will almost invariably revert to study techniques that they are more familiar with and which they find less onerous (McCabe, 2011).

### **Quizzing**

Although there are dozens, perhaps hundreds, of suggestions in the literature about how to address the reading and study compliance problems, in terms of popularity, frequent quizzing appears to be the most popular. In part, the popularity of quizzing can be inferred from the disparity in the literature on quizzing versus other methods. There is a substantial body of research that suggests quizzes will motivate more students to do the reading assignments when they know they must take a quiz on the readings (Howard, 2004; Clump et al., 2004; Fernald, 2004; Leeming, 2002; Johnson et. al, 2009; Narloch et al., 2006; Starcher et al., 2011).

Furthermore, discussions of competitor techniques in the literature are often framed specifically against quizzing. So for example, researchers have investigated the effectiveness of learning logs

versus quizzing (Carney et al., 2008), reading responses versus quizzing (Roberts et al., 2008), and reading questions versus quizzing (Henderson et al., 2006).

While we are not aware of any survey of the most popular techniques across disciplines to address the motivational problem, Starcher and Proffitt's (2011) survey of business faculty is revealing. In their invitation to comment on the motivation problem, 26 out of 85 suggestions they received involved quizzing (both in and out of class). Chapter summaries came a distant second, with only seven out of 85 suggestions involving some sort of written summary.

Interestingly, there is reason to suspect that Starcher and Proffitt's survey may underestimate the use of quizzing by business faculty. The prompt for their survey was two opened ended questions:

1. "What can I do to encourage students to prepare for class by completing their assigned textbook reading?"
2. "I know the use of the "pop quiz" or chapter quiz is one approach. I'm looking for additional ways to encourage students to read their textbooks" (Starcher et al., 2011, p. 400).

In light of the second question, the possibility of underreporting exists at two stages. The authors received about a 15% response rate to their survey. Those who use quizzing may have been less likely to respond at all, given that a foci of the research appears to be alternatives to quizzing.



And of those who responded, it may be that they failed to mention that they employ quizzing, since, again, a foci of the research appears to be *alternatives* to quizzing.

The available evidence, then, suggests that quizzing is the “technique to beat” in terms of how frequently it is used. Quizzing has a number of drawbacks; here we will mention three.

First, it is no secret that generally, students do not like frequent quizzing. Students often view quizzes as punitive (Connor-Greene, 2000). For many professors, this can lead to the worry of negative effects on teaching evaluations (Redding, 1998). It has been suggested that the logical implications of worrying about negative feedback inevitably leads to what many might think of as an unreasonable conclusion, “Faculty who reject quizzing on the basis of students’ ill will may want to reconsider the practice of giving exams on the same rationale” (Sappington et al., 2002, p. 274). For those with tenure, it may be tempting to be a bit sanguine here. But we should not forget that in many institutions, a majority of professors lack the security of tenure as a bulwark against unpopular teaching measures, no matter how effective they are. And, in fact, the objection misses the mark. Reluctance to use quizzing because of its unpopularity with students does not imply that one has the desire *never* to institute unpopular measures like exams. Rather, a more charitable understanding is that some professors may be reluctant to institute measures that are *particularly* unpopular. If exams are the norm, and frequent quizzes are the exception at one’s institution, then the worry that quizzing may be particularly unpopular with students cannot be so easily dismissed.

Second, quizzing has been criticized for encouraging, or at least requiring, only “surface learning” (Henderson et al., 2006; Roberts et al., 2008). Merely quizzing students about the bold

definitions in a textbook to see if they have read the textbook can encourage a false view about what we aim for in education (Henderson et al., 2006). While it is possible to ask written or multiple-choice questions that go beyond mere surface learning, it is very difficult to write questions at the appropriate level (Henderson et al., 2006). Take a simple example. Suppose we hope that students who seriously study the assigned readings will understand that there are no entailments between concepts X and Y. Suppose we ask a question to this effect on a quiz and the vast majority of students get it wrong. There are a number of possible explanations for this. They did not understand either concept X or Y, or they wrongly suppose X entails Y, or Y entails X. We might break this down into multiple questions. We could write questions about whether they have memorized the definitions for concepts X and Y (mere surface learning questions), and then ask separate questions about whether X entails Y, and whether Y entails X. In which case, one question has now morphed into four. If there are, say, six concept dyads that we hope students will understand from the reading, a short quiz at the beginning of class will quickly morph into something more like an exam. It is no wonder, then, that so many quizzes check only for surface learning.

Third, there is scant evidence that quizzing helps with the study compliance problem as opposed to the reading compliance problem. In the literature we have reviewed, the researchers examine the effect of quizzing where the baseline condition is that most students do not even read the text. As noted, much of the literature supports the claim that judicious use of frequent quizzing can improve at least reading compliance. To the best of our knowledge, the research has not systematically investigated whether quizzing can affect study compliance. Indeed, the

previous complaint, that quizzing encourages only surface learning, can be seen as acknowledgement that faculty are concerned with quizzing encouraging mere reading, rather than serious study.

### Socratic Note Taking

The SNT assignment requires students to take notes on the assigned reading using a three column format: page numbers, student-generated question, and answer. The excerpt below is from an instructor-generated example provided to students. (The complete prompt can be found in Appendix A.)

<b>Textbook Page Numbers</b>	<b>Question</b> (Questions should focus on concepts and arguments.)	<b>Answer</b>
132-133	What does the “Questioner” believe?	<ul style="list-style-type: none"> <li>• People should not believe things on faith</li> <li>• Believing things on faith is dangerous</li> </ul>
133	What is faith?	<ul style="list-style-type: none"> <li>• To believe something on faith is to believe something even though you lack evidence for the belief</li> </ul>

The prompt includes a brief description of Socrates’ pedagogy which, famously, involved asking students a series of questions. Students are asked to formulate questions about the reading of the sort that they imagine a good teacher like Socrates would ask. The rationale for the assignment is as follows:

1. SNT was designed consistent with our knowledge about effective study techniques.

Decades of research confirms that student-generated questions shows improved learning as compared with merely reading target material (Dee-Lucas et al., 1980; Donaldson et al., 1980; Foos, 1989; Foos et al., 1994; Nairne et al., 1987; Slamecka et al., 1978; Van Blerkom, 2011; Van Blerkom et al., 2006). Recall that Dunlosky et al. (2013) found that only practice testing and distributed testing earned a high utility ranking among the ten different study techniques they discovered. As they define it, practice testing involves self-testing or taking practice tests on material to be learned. Distributed practice involves engaging with the material to be learned over an interval of time. SNT is consistent with this research. SNT offers a ready format for students to self-test. Students were told that they could use this assignment to study by covering up the third column and seeing if they could recall the answers to their questions. SNT is consistent with distributed practice, as the SNT assignment was due online the day before class. Students were also required to bring a hardcopy to class, which gave them a chance to review the material before taking a quiz.

2. SNT was designed with the goal of ease of mastery by students. Approximately ten minutes of class time during the first week, along with the prompt included in Appendix A, was the extent of priming students had for the assignment.
3. SNT was designed with the aim of permitting students to see a strong correlation between effort and earned grade. The first author of this paper has had a number of conversations with low achieving students who cite their reason for not doing well on quizzes and

exams as being due to the fact that they are “not good test-takers”. When asked why they did not prepare for quizzes and exams, they often said they expended minimal effort because it would have no effect: Since they are not good test-takers, there is no point in devoting time to studying. The implicit theory for many students seems to be of the entity versus incremental mindset about test-taking (Elliot et al., 2013; Grant et al., 2003). Part of the motivation problem associated with the entity mindset is circumvented by giving students credit simply for their study effort. It was emphasized in class several times that SNT was graded mostly on effort and every student in the class, with sufficient effort, should be able to score in the A range on the assignment. Most students who turned in the assignment scored in the 80 to 100% range. The most common reason for scoring lower was because the assignments were incomplete, for example, some students submitted notes for only half the chapter. For students who scored in the 80 to 90% range, the most common reason was because students neglected to include a question/answer dealing with an important concept or premise/conclusion. The initial efforts were graded quite leniently. We gradually raised the grading bar for SNT as students gained familiarity with the study of philosophy and with SNT.

4. SNT was designed with ease-of-grading in mind. Each SNT submitted by students averaged approximately 700 words in length. While this sounds relatively onerous in terms of grading effort, the consistency of the SNT format sped up the grading. (Grading-time is considered further in the discussion section). Having students include the page number is particularly helpful in this connection. If an important argument or concept is

introduced on a certain page in the text, it is a relatively simple matter to find the corresponding part of the students' submission.

The intent behind SNT was to have students ask expository questions about the text: the answers they asked should be found in the textbook. A few students spontaneously began asking critical and evaluative questions. For example, some students asked questions such as, "Does the author make a good argument?", and then proceeded to criticize the reading in the answer column. During the study, in order not to add any confounding variables, these additional efforts were neither rewarded nor penalized. Nevertheless, we can envision the possibility of incorporating critical questions into SNT. For example, one might ask students to add one or two critical questions in a separate section at the end of their assignments. One possible benefit is that it would emphasize the separate skills of understanding an author's position, as opposed to evaluating an author's position. A second possible benefit is that these questions could be used as a catalyst for small group or class discussion.

In the pilot study (see below) insufficient directions were provided as to what constituted a good question. Students tended to ask very broad questions with answers that covered two or more pages. This worked against the idea that asking students to formulate questions is a way for them to intellectually "work" the material, rather than the more passive approach of simply taking notes. In the main study, the directions were revised to indicate that in the typical case, a question of the right generality can be answered in one or two sentences. Even with these revised directions, students still tended to err on the side of asking questions that are too broad. Written comments on individual assignments brought this deficit to students' attention. In addition, this

problem was discussed by the instructor at the beginning of class when the first three SNT assignments were returned. The quality of the SNT assignments improved over the course of the semester. We believe this was in part from the feedback on the assignment, as well as students honing their ability to identify philosophical concepts, premises, and conclusions in the readings.

### **Hypotheses**

The main purpose of our study was to test the effectiveness of SNT as measured by quiz performance. The study involved a 2 x 2 x 2 design: SNT versus no SNT, and a high stakes/low stakes quiz, with replication. Students taking a particular quiz as a low stakes quiz were told that they would receive full credit for the quiz merely for taking the quiz. Students taking a particular quiz as a high stakes quiz were told that the grade on the quiz would count towards their final grade.

As discussed above, previous research indicates that students commonly employ some of the least effective study techniques. In addition, based on the foregoing comments about our rationale for SNT, we were optimistic that it would work, and with a large effect size too. Hence, our main hypothesis was that requiring students to employ SNT should yield increased learning:

**H1:** Students assigned to SNT will demonstrate increased learning relative to when they are not assigned SNT.

The foregoing review of the literature on quizzing suggests that it influences student learning. If the reason for that influence is that students put forth effort to obtain good grades, then manipulating the stakes involved should affect quiz performance. Specifically, when the stakes are low (students receive full points regardless of performance), students should perform less

well than when the stakes are high (points depend on performance). However, if the size of the main effect of high or low stakes turns out to be substantially less than the size of the main effect for SNT, this would help demonstrate the potential value of SNT. Thus, Hypothesis 2 is as follows:

**H2:** Students should perform better on high stakes quizzes than on low stakes quizzes. We also considered the possibility of an interaction whereby the combination of SNT and high stakes would lead to a larger than additive (interactive) increase in quiz performance.

We predicted (but hoped otherwise), that students would not employ SNT when it was not required. This prediction was based on the aforementioned research which suggests that students are reluctant to change their study behaviors:

**H3:** Students who are not assigned SNT and have evidence that SNT will increase their quiz grade will not employ SNT as a study technique to increase their quiz grade.

Finally, we considered whether SNT might help students to differing degrees. In part, a prediction along this line makes sense because the worse students perform in the no notes condition, the more SNT has the capability of improving their performance. If, as we hypothesize, low performance is due in large measure to using ineffective study techniques rather than, say, some innate inability to take quizzes, then we should expect a large negative correlation between performance in the no notes condition and improvement due to SNT.



**H4:** There will be a large negative correlation between student scores in the no notes cells and student scores in the SNT cells.

### Method

Before reporting on the experiment, we present a quick description of a prior pilot study. In the pilot study, 67 students taking American History took quizzes, with half of them using SNT for odd numbered quizzes and half of them using SNT for even numbered quizzes. Thus, all students used SNT for half of the quizzes and did not use SNT for half of the quizzes. There was a statistically significant effect whereby students scored more points for the SNT quizzes ( $M = 5.43$ ) than for the other quizzes ( $M = 5.00$ ),  $F(1, 66) = 829$ ,  $p < .01$ ,  $\eta_p^2 = .11$ .

However, there were two limitations that plagued the pilot study. First, the grading standard for the SNT was insufficiently rigorous. This problem was addressed in the main study. Second, the explanation of what constitutes good questions for a SNT assignment was improved for the main study. Despite these limitations, the result was sufficiently promising to justify an improved effort and a more complex experimental design in the main study.

### Participants

Students taking two sections of introductory philosophy (PHIL 101G, “The Art of Wondering”), in the Fall Semester of 2016 participated in the study. Permission to use the students’ data was obtained at the beginning of the semester and only students who gave permission were included in the analyses ( $n = 83$  and  $n = 75$ ). Participants in each class were divided into groups, depending on the first letter of their last names: A-L and M-Z. The two groups in the two sections combined to form a total of four groups. The reason for dividing

participants in this way was so they could be rotated through the conditions to be described presently. We only included students in the main experiment who completed all eight quizzes ( $n = 80$ ).

### **Design**

The design was a 2 (Notes: SNT or not) x 2 (Stakes: high or low) x 2 (Replication: first or second replication of a Notes x Stakes cell) within-participants factorial. That is, all students were exposed to all cells of the design. Each of the four groups proceeded through the eight cells of the design in a different order, to control for serial position effects.

### **Procedure**

All students took an initial quiz with SNT as a practice trial that was not analyzed. Subsequently, all students took eight quizzes, each of which represented a cell of the design. As described above, although all the groups took the tests the same days and in the same order, they were exposed to the four cells in different orders. Consequently, each student was exposed to two replications of each of the following cells: SNT-high stakes, SNT-low stakes, no notes-high stakes, and no notes-low stakes.

All students were informed at the beginning of the semester (on the syllabus and verbally), which quizzes would be under which conditions for them. As noted above, for the high stakes conditions, the students' actual scores on the quiz were recorded as part of their term grade. In the low stakes condition, students were told that they would be given a 10 out of 10 on the quiz, so long as they took the quiz. These quizzes were graded and given back to the students. Students could see, for example, that although they only got (say) four answers correct,

still a 10 out of 10 was recorded on the quiz and in the grade book. Students were told that their main incentive to try on the low stakes quiz was to help them prepare for the midterm and the final exam. Their performance on the low stakes quiz would give them some indication of how much studying they would need to do on particular course readings.

Finally, a survey was distributed the second last week of class to all students present. Only those students who had given permission at the beginning of the semester were included in the analysis ( $n = 120$ ). The results will be explained in the context of the findings.

## Results

Consistent with the design, the data were analyzed as a 2 (Notes: SNT or not) x 2 (Stakes: high or low) x 2 (Replication: first or second replication of a Notes x Stakes cell) within-participants factorial. Consistent with our main hypothesis (H1), there was a statistically significant main effect for the Notes factor. Participants performed much better in the SNT than in the no notes condition ( $M = 6.49$  and  $M = 5.27$ ),  $F(1, 79) = 59.21$ ,  $p < .001$ ,  $\eta_p^2 = .43$ . Before continuing, we call the reader's attention to the large partial eta-square statistic that indicates an abnormally large effect size.

We also tested the main effect for the Stakes factor but, in fact, the mean in the high stakes condition ( $M = 5.89$ ) differed very little from the mean in the low stakes condition ( $M = 5.87$ ),  $F(1, 79) < 1$ ,  $\eta_p^2 = .00$ . Thus, contrary to a secondary hypothesis (H2), our findings indicate that the stakes were irrelevant to quiz performance. In addition, there was not a statistically significant Notes x Stakes interaction,  $F(1, 79) < 1$ ,  $\eta_p^2 = .02$ . Finally, there was a theoretically uninteresting Replication main effect; students performed better in the first

replication than in the second ( $M = 6.19$  and  $M = 5.58$ ),  $F(1, 79) = 21.78, p < .001, \eta_p^2 = .22$ . We considered the possibility that students learned to use SNT even when not required by the teacher by the time of the second replication. However, although there was a small but statistically significant Notes x Replication interaction [ $F(1, 79) = 4.54, p < .05, \eta_p^2 = .05$ ], it was not in the correct direction to support this possibility because the mean performance in the no notes conditions actually decreased, rather than increased, for the second replication ( $M = 5.43$  and  $M = 5.11$ ). This provided some confirming evidence for (H3): students did not start using SNT when it was not assigned. (Additional evidence from the survey data is discussed below.) Results are summarized in Table 1. [Insert Table 1 here].

We hypothesized that SNT works particularly well for the students who perform comparatively poorly without notes (H4). To test this hypothesis, we averaged students' scores in the SNT condition to obtain an average level of SNT performance; we averaged students' scores in the no notes condition to obtain an average level of non-SNT performance; and we computed a difference score between these two averages. Note that all of these averages were for *each* student, so that each student had a baseline non-SNT performance as well as a difference score, indicating the effect of SNT for that student. According to H4, there should be a negative correlation between baseline non-SNT performance and difference scores. That is, the worse the baseline non-SNT performance is, the greater the difference score should be. In fact, the correlation was strongly negative,  $r = -.62, p < .001$ . The scatter plot of Fig. 1 shows the improvement for individual students. [Insert Fig. 1 here.]

So, with reference to table 2, looking at students scoring an average of two on a ten item multiple choice test in the non SNT cells, the negative correlation cited above leads to the prediction of an improvement of a whopping 2.88 for these students, for a total of 4.88. [Insert Table 2 here.] Interestingly, taking students with the highest baseline scores, the negative correlation leads to the prediction that SNT actually harms their performance. Whether SNT actually harms the performance of these students, or whether this is simply an artifact of regression to the mean, is not something that our study established.<sup>2</sup>

Finally, there are a few descriptive findings from the final survey that are of interest. First, only 5% of students stated that they used SNT for quizzes where it was not required. This is despite the fact that most of the students (82.5%) who filled out the final questionnaire disclosed they heard from the instructor that quiz scores were higher for students who did SNT than for students who did not. These results further support (H3). Our survey confirmed what others have found: underlining (approximately 74%), rereading (approximately 25%), and note taking in format other than SNT (approximately 21%), were the most common study techniques. (Students were permitted to select more than one technique, hence the percentages add up to greater than 100%.) Finally, there is evidence suggesting that students see SNT as preferable to

---

<sup>2</sup> Assuming there is some detriment to quiz performance for students with the highest baseline, it is fascinating to speculate why this might be. A helpful suggestion by an anonymous reviewer is that further studies of SNT might break down survey data by quiz performance level. This might provide some insight as to what the highest quiz performers do differently when not assigned SNT.

taking quizzes. To the proposition “It would be good for future classes to take one or two quizzes per week and eliminate the Socratic note assignment completely”, in a forced choice, 78.81% fell on the ‘disagree’ end of the continuum (Strongly disagree = 28, Disagree = 43, Somewhat Disagree = 22), while 21.19% sided with the ‘agree’ end of the continuum (Somewhat Agree = 13, Agree = 6, Strongly agree = 6). To the proposition “It would be good for future classes to do more Socratic note assignments worth more of the final grade and eliminate or severely reduce the number of quizzes”, in a forced choice, 14.28% sided with the disagree end of the continuum (Strongly Disagree = 5, Disagree = 5, Somewhat Disagree = 7), while 85.71% fell on the agree end (Somewhat Agree = 26, Agree = 36, Strongly Agree = 40).

### **Discussion**

In this section, we will discuss how the data bears on our hypotheses, and the utility of SNT versus high stakes quizzing. Although quizzing was held constant across the different cells of the study, the high stakes condition is of particular relevance as this is the most common form used in classrooms to address the study compliance problem.

The study strongly supported our main hypothesis (H1): we found an unusually large effect size for SNT versus no notes, confirming that SNT is more effective at addressing the study compliance problem as compared with students’ usual study techniques.

We were a little surprised to find no confirming evidence for the hypothesis of a main effect for the Stakes factor, or at least a Stakes x SNT interaction. Our initial thought was that students in the SNT low stakes condition would perform less well because they had no further motivation to study the material once they had completed SNT the night before. Students in the

high stakes/SNT cell, on a given week, had the extra incentive to study further (perhaps reviewing their SNT assignment) to perform better on the high stakes quiz. Our study did not look at how students used SNT once the assignment was finished, so this may be worth exploring further. The fact that there was no significant difference between the means in the no notes/high stakes cell, and the no notes/low stakes cell seems to run counter to the literature that suggests high stakes quizzes improve student performance. We are not sure what to make of this, but two points seem relevant. First, there is a relative paucity of classroom (as opposed to laboratory) studies on the effects of quizzing (Batsell Jr. et al., 2017). Second, all the studies we are aware of that investigate the effect of quizzing in classroom studies use a no quiz/high stakes quiz design, as opposed to our study which used a low stakes/high stakes quiz design. The no quiz/high stakes quiz design is often measured against performance on midterms or final exams. This opens the intriguing possibility that taking a quiz, rather than the threat to a student's grade, accounts for much of the perceived effect of quizzing. This is an avenue for further research.

As predicted by (H3), the vast majority of students did not use SNT when it was not assigned as a means to improve their quiz grades. Our study confirms this with two independent lines of evidence. First, we found that the difference between those in the SNT cells and those not in the SNT cells in a given week did not decline over the course of the study. A possibility we recognized in designing the experiment was that students might freely use SNT when not assigned, given that they would be told it increased their quiz grade (assuming that it did). The fact that students did not freely adopt SNT was further confirmed in our survey data, which showed that only 5% used SNT on a regular basis when it was not assigned. Our study strongly

suggests that knowledge of more effective study techniques is not sufficient to motivate many students to adopt such techniques in order to better prepare for quizzes. As professors, we found this particularly surprising (and somewhat distressing). The means for the quizzes in all cells were in the F and D range, suggesting that if low quiz grades were sufficiently motivating, then students would have adopted more promising study techniques.

Finally, our study provided strong support for the hypothesis that SNT would help students inversely, proportional to the degree that they performed less well in the no notes cells. Of course, a priori, we should expect to see some skewing towards the weaker students, since they have more room for improvement. Still, the size of the negative correlation we found,  $r = -.62, p < .00$ , is particularly encouraging. The large predicted improvement for low baseline quiz scores exceeded our wildest dreams. Not only is this a benefit for these students, but the whole class benefits because SNT reduces the achievement gap between the weakest and strongest students. Most of us would agree that teaching is easier and more effective to the extent that the disparity between the weakest and the strongest students is reduced. SNT achieves this goal.

As noted above, informal evidence suggests that many students attribute their poor quiz scores to an entity, rather than incremental, mindset about test-taking. Further study is warranted as to the prevalence of the entity versus incremental mindset with respect to test-taking, and the possibility of using SNT to disabuse students of the entity mindset.

In terms of its utility for improved learning, it will be helpful to contrast two extreme alternatives: the quiz format (daily or weekly quizzes on the assigned readings) versus weekly SNT. These are extreme in the sense that some blended format using both quizzes and SNT is



entirely possible (and indeed the experiment involved such a blended format). However, for the purposes of contrast, the extreme versions will best help frame our discussion.

Our survey shows that students overwhelmingly prefer SNT to taking high stakes quizzes. It seems reasonable to conclude that students find SNT less punitive than quizzing. Several students made remarks along those lines in the written comments section of the survey. As one student put it:

I would rather do the [S]ocratic notes than take a quiz. Not because I felt unprepared but because I felt stressed out if I was going to pass it or not. I couldn't really enjoy the class because I was too worried about my quiz grade. I learned through [S]ocratic note taking more than anything.

As indicated, when looking at increased learning, SNT wins hands down over the quiz format. Further research is warranted to see how SNT and quizzing might compare on some measures of learning independent of quizzing and SNT, such as essay writing, final exam performance, or quality of class discussion.

Two further advantages are worth noting: First, SNT makes for a different learning environment. There is evidence that suggests that high stakes quizzing might be counterproductive to fostering a learning environment. Hayek, Toma, Oberlé, and Butera (2014) showed that expectations of grading increased the preference effect—the preference people have for evidence that confirms rather than disconfirms their expectations—through the mechanism of increasing perceptions of competitive social comparison. Frequent quizzes are likely to increase

perceptions of competitive social comparison, which arguably is an undesired outcome with which to begin, with the additional undesired outcome of decreasing the extent to which students are willing to attend to evidence that disconfirms their expectations. Second, most instructors believe that there is intrinsic benefit simply in writing. The correlation between the number of pages undergraduates write and their critical thinking abilities has been extensively documented (Arum et al., 2011). So, other things being equal, a writing assignment is to be preferred to merely circling an answer.

In terms of overall utility, the biggest drawback of the SNT assignment versus the quiz format is the time it takes to grade. We found that a quiz of ten questions can be graded by hand and entered into the grade book in about one minute. The Socratic assignment takes between two and three minutes on average to grade and enter. These small differences at the individual student level can quickly add up. With two large classes of 90 students, the difference is between about three hours of grading for a quiz versus six to nine hours of grading for the Socratic Notes assignment. While there is no denying the general point that written assignments often take longer to grade than multi-choice exams, several points can at least blunt the contrast.

To create a good multiple-choice quiz takes time. As noted above, it is very difficult to construct quizzes of the right difficulty, and quizzes may favor questions that demonstrate mere reading, rather than studying, of the material. If the additional time to construct a well-designed quiz (assuming this is even possible) is factored in, the discrepancy in time commitment closes. In fact, with smaller classes, the time commitment may actually favor SNT in some instances. The length of time to construct a quiz does not change much between a large and small class. If it

takes, say, an hour and a half to formulate, edit, and print a multiple-choice quiz for a class of 30 students, and a minute to grade and enter the grade, this represents about 120 minutes of labor. In contrast, once the original prompt is formulated, assigning SNT involves simply telling students what pages to cover. Assuming conservatively three minutes grading time per assignment, this involves 90 minutes of labor to grade 30 SNTs. The time advantage to SNT quickly disappears if one can reuse quiz questions from year to year, perhaps by having a large test bank.<sup>3</sup> And electronic grading of multiple-choice quizzes could further increase the advantage to quizzing. Any advantage for SNT quickly disappears when we turn to larger courses. So while there are certain circumstances where SNT may save time over quizzing—in smaller classes where new quiz questions must be generated with some frequency—in many situations, quizzing will have some advantage in terms of time.

There are a couple of ways that the time disadvantage of SNT might be reduced or neutralized. The first is to use undergraduates to assist with grading SNT assignments. This is possible because grading SNT is relatively straightforward. For our main experiment, a senior undergraduate philosophy major was employed. The undergraduate TA graded about 80% of all SNTs (approximately 70 per week on average) and quizzes for the two classes and was paid ten hours a week for this (and for attending lectures). On a per student basis, the TA's cost was calculated at about \$8 per student for the entire semester. The additional cost was justified on the

---

<sup>3</sup> We would like to thank an anonymous referee for pressing on this and several other issues related to the time investment in quizzing.

basis that it meant that students would write a lot more than they would have otherwise (our university has a university-wide push for requiring more writing). With the undergraduate TA help, the instructor for the two courses in the main experiment spent less time grading, in comparison with non SNT versions of the same courses.

A second possibility to decrease grading time for SNTs is in groups. Imagine, for example, a course with 11 SNT assignments on the weekly readings. The first two are graded with comments to make sure that students have good feedback on what is expected. The final nine assignments are graded in groups of three. Students are told that two of the assignments will be graded at five out of five if they meet some minimal requirements, e.g., handing in their assignments on time, following the format, meeting a minimal word count, etc., while the third assignment will be examined more closely and graded out of 20. Students will not be given advance notice of which assignment will be weighted higher and graded more rigorously, so they will have motivation to make a good effort on each assignment. Although this variant was not tried, it appears to have the potential to greatly reduce the time spent grading SNTs for the instructor or TA—perhaps cutting it in half—as compared with grading each SNT in the usual manner.

SNT is a tool for improving student learning. All tools have limitations: no tool works well in all situations. Our pilot study and main study show that SNT improves students' comprehension of the assigned readings. We have every reason to suppose that SNT should work across a variety of disciplines, beyond history and philosophy, as a tool to help students study assigned readings more effectively. Future research in other disciplines should look to confirm

this prediction. Indeed, a priori, it would be very surprising if SNT worked only in history and philosophy given that almost all study techniques work (or fail to work) across a wide range of disciplines (Dunlosky et al., 2013). Assuming this prediction is borne out, then the case will be complete that SNT should dethrone quizzing as the technique to beat when it comes to getting students to engage with the readings before class.

## References

- Artis, Andrew B. "Improving Marketing Students' Reading Comprehension with the SQ3R Method." *Journal of Marketing Education* vol. 30 no. 2 (2008): 130-137.
- Arum, Richard, and Josipa Roska. *Academically Adrift: Limited Learning on College Campuses*. Chicago: University of Chicago Press, 2011.
- Batsell Jr, W. Robert, Jennifer L. Perry, Elizabeth Hanley, and Autumn B. Hostetter. "Ecological Validity of the Testing Effect: The Use of Daily Quizzes in Introductory Psychology." *Teaching of Psychology* vol. 44 no. 1 (2017): 18-23.
- Brost, Brian D, and Karen A. Bradley. "Student Compliance with Assigned Reading: A Case Study." *Journal of Scholarship of Teaching and Learning* vol. 6 no. 2 (2006): 101-111.
- Brown, Peter C., Henry L. Roediger and Mark McDaniel. *Make It Stick*. Cambridge: Harvard University Press, 2014.
- Burchfield, Colin M., and John Sappington. "Compliance with required reading assignments." *Teaching of Psychology* vol. 27 no. 1 (2000): 58-60.
- Carney, Amy G., Sarah Winstead Fry, Rosaria V. Gabriele, and Michelle Ballard. "Reeling in the Big Fish: Changing Pedagogy to Encourage the Completion of Reading Assignments." *College Teaching* vol. 56 no. 4 (2008): 195-200.
- Clump, Michael A., Heather Bauer, and Catherine Bradley. "The Extent to Which Psychology Students Read Textbooks: A Multiple Class Analysis of Reading across the Psychology Curriculum." *Journal of Instructional Psychology* vol. 31 no. 3 (2004): 227-232.

- Connor-Greene, Patricia A. "Assessing and Promoting Student Learning: Blurring the Line between Teaching and Testing." *Teaching of Psychology* vol. 27 no. 2 (2000): 84-88.
- Dee-Lucas, Diana, and Francis J di Vesta. "Learner-Generated Organizational Aids: Effects on Learning from Text." *Journal of Educational Psychology* vol. 72 no. 3 (1980): 304-311.
- Donaldson, Wayne, and Michael Bass. "Relational Information and Memory for Problem Solutions." *Journal of Verbal Learning and Verbal Behavior* vol. 19 no. 1 (1980): 26-35.
- Dunlosky, John, Katherine A. Rawson, Elizabeth J. Marsh, Mitchell J. Nathan, and Daniel T. Willingham. "Improving Students' Learning with Effective Learning Techniques: Promising Directions from Cognitive and Educational Psychology." *Psychological Science in the Public Interest* vol. 14 no. 1 (2013): 4-58.
- Elliot, Andrew J., and Carol S. Dweck. *Handbook of Competence and Motivation*. New York: Guilford Publications, 2013.
- Fernald, Peter S. "The Monte Carlo Quiz: Encouraging Punctual Completion and Deep Processing of Assigned Readings." *College Teaching* vol. 52 no. 3 (2004): 95-99.
- Foos, Paul W. "Effects of Student-Written Questions on Student Test Performance." *Teaching of Psychology* vol. 16 no. 2 (1989): 77-78.
- , Joseph Mora, and Sharon Tkacz. "Student Study Techniques and the Generation Effect." *Journal of Educational Psychology* vol. 86 no. 4 (1994): 567-576.
- Grant, Heidi, and Carol S. Dweck. "Clarifying Achievement Goals and their Impact." *Journal of Personality and Social Psychology* vol. 85 no. 3 (2003): 541-553.

- Gurung, R. A. "Pedagogical Aids and Student Performance." *Teaching of Psychology* vol. 30 no. 2 (2003): 92-95.
- Hayek, Anne-Sophie, Claudia Toma, Dominique Oberlé, and Fabrizio Butera. "The Effect of Grades on the Preference Effect: Grading Reduces Consideration of Disconfirming Evidence." *Basic and Applied Social Psychology* vol. 36 no. 6 (2014): 544-522.
- Henderson, Charles, and Alvin Rosenthal. "Reading Questions." *Journal of College Science Teaching*, vol. 35 no. 7 (2006): 46-50.
- Howard, Jay R. "Just-in-time Teaching in Sociology; or How I Convinced my Students to Actually Read the Assignment." *Teaching Sociology* vol. 32 no. 4 (2004): 385-390.
- Jensen, Murray, and Randy Moore. "Reading Trade Books in a Freshman Biology Course." *The American Biology Teacher* vol. 70 no. 4 (2008): 206-210.
- Johnson, Bethany C., and Mark T. Kiviniemi. "The Effect of Online Chapter Quizzes on Exam Performance in an Undergraduate Social Psychology Course." *Teaching of Psychology* vol. 36 no. 1 (2009): 33-37.
- Leeming, Frank C. "The Exam-a-day Procedure Improves Performance in Psychology Classes." *Teaching of Psychology* vol. 29 no. 3 (2002): 210-212.
- McCabe, J. "Metacognitive Awareness of Learning Strategies in Undergraduates." *Memory and Cognition* vol. 39 no. 3 (2011): 462-476.
- Nairne, James S., and Robert L. Widner. "Generation Effects with Nonwords: The Role of Test Appropriateness." *Journal of Experimental Psychology: Learning, Memory, and Cognition* vol. 13 no. 1 (1987): 164-171.



Narloch, Roger, Calvin P. Garbin, and Kimberly D. Turnage. "Benefits of Pre-lecture Quizzes."

*Teaching of Psychology* vol. 33 no. 2 (2006): 109-112.

Peterson, Valerie V. "The Divide and Conquer Exercise: Finding the Essence of Difficult

Readings." *Communication Teacher* vol. 20 no.1 (2006): 14-17.

Redding, Richard E. "Students' Evaluations of Teaching Fuel Grade Inflation." *American*

*Psychologist* vol. 53 no. 11 (1998): 1227-1228.

Roberts, Judith C., and Keith A. Roberts, K. A. "Deep Reading, Cost/Benefit, and the

Construction of Meaning: Enhancing Reading Comprehension and Deep Learning in

Sociology Courses." *Teaching Sociology* vol. 36 no. 2 (2008):125-140.

Ryan, Tracey E. "Motivating Novice Students to Read their Textbooks." *Journal of Instructional*

*Psychology* vol. 33 no. 2 (2006): 135-141.

Sappington, John, Kimberly Kinsey, and Kirk Munsayac. "Two Studies of Reading Compliance

among College Students." *Teaching of Psychology* vol. 28 no. 4 (2002): 272-274.

Slamecka, Norman J., and Peter Graf. "The Generation Effect: Delineation of a Phenomenon."

*Journal of Experimental Psychology: Human Learning and Memory*, vol. 4 no. 6 (1978):  
592-604.

Starcher, Keith, and Dennis Proffitt. "Encouraging Students to Read: What Professors are (and

aren't) Doing about it." *International Journal of Teaching and Learning in Higher*

*Education* vol. 23 no. 3 (2011): 396-407.

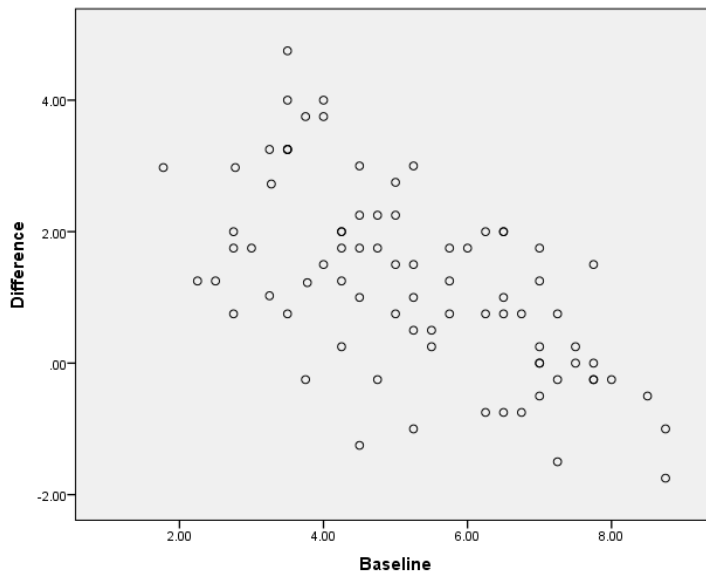
Van Blerkom, Dianna L. *College Study Skills: Becoming a Strategic Learner*. Cengage Learning,  
2011.

-----, Malcolm L. Van Blerkom and Sharon Bertsch. "Study Strategies and Generative Learning: What Works?" *Journal of College Reading and Learning* vol. 37 no. 1 (2006): 7-18.

Weimer, Maryellen. *11 Strategies for Getting Students to Read What's Assigned: Faculty Focus Special Report*. Madison, WI: Magna Publications, 2010.

**Table 1.** *Quiz performance means and standard deviations (in parentheses) are represented as a function of the three within-participants factors: SNT (Notes or No Notes), Risk (High Risk or Low Risk), and Replication (Replication 1 or Replication 2).*

SNT	Replication 1		Replication 2	
	High Risk	Low Risk	High Risk	Low Risk
Notes	6.81 (1.64)	7.08 (2.12)	6.04 (2.02)	6.04 (2.04)
No Notes	5.38 (2.68)	5.34 (2.34)	5.48 (2.26)	4.89 (2.09)



**Figure 1.** Difference scores are represented along the vertical axis as a function of baseline scores along the horizontal axis.

**Table 2** shows the improvement on averages on a ten-item quiz:

Table 2. *Predicted improvement using SNT on baseline on 10-item quiz*

Baseline	2	3	4	5	6	7	8	9	10
Improvement	2.88	2.37	1.86	1.35	.84	.33	-0.18	-0.69	-1.2

## **Appendix A: Prompt provided to students in main study**

### **Socratic Note Taking**

The purpose of this assignment is to ensure that students have mastered the basics of taking notes on the assigned readings. Taking notes on class readings is a proven method to deepen students' understanding of course material, as opposed to merely reading the assigned text. Many educators believe a good set of course notes also makes studying for quizzes and exams easier and more effective.

The term 'Socratic note taking' is an homage to Socrates. Socrates is considered by many to be one of the greatest of the Ancient Greek philosophers. He was also a renowned educator famous for a method of teaching that involves asking students questions. This assignment asks you to play the role of both Socratic teacher and student. Your questions should focus on important concepts, information, explanations, and evidence. In other words, you should think about the sorts of questions a good teacher might ask.

In general, your questions should be answered by one or two sentences. If you find it takes more to answer your question, consider asking a more specific question. An exception is where answers are naturally grouped together. For example, if the textbook suggests that there are three possible causes for some phenomenon X, then it would be helpful to answer these together even though the answer is likely to exceed two sentences. In this instance, the question should indicate that a more extensive answer is required. For example, the question might ask: "What three possible causes have been cited to explain X?"

Using your completed notes as a study tool is straightforward. Simply cover up the answers in the right hand column and test yourself by asking yourself each question.

Sample Notes Assignment:

**Due Date:** August 31<sup>st</sup>. Submit your assignment to Canvas by 11:59 pm. Also, please bring a hardcopy to class Sept. 1<sup>st</sup>. (Late submissions will be accepted ONLY with a university recognized excuse, e.g., illness, absence for university event, etc.)

**Purpose:** The purpose of this assignment is to ensure that students have mastered the basics of taking notes on the assigned readings. Taking notes on class readings is a proven method to deepen students' understanding of course material, as opposed to merely reading the assigned text. A good set of course notes also makes studying for quizzes and exams easier, and increases the likelihood of getting a good grade.

**Assignment:** Take notes for reading 2.4 from the textbook. Your notes must follow EXACTLY the question/answer format of the sample provided. (See below). Failure to follow directions will result in penalties ranging from 25 to 100%. Assignments must be typed. Use complete sentences. Grammatically deficient assignments will have points deducted.

**Recommendation:** I recommend reading the chapter first. Next, read it again, taking notes for the assignment. You should allow yourself three to five hours for this assignment.

**Checklist:**

- ✓ Submit on Canvas by 11:59 pm, August 31
- ✓ Bring hardcopy to class Sept. 1<sup>st</sup>
- ✓ Name
- ✓ Date
- ✓ Title of Reading

- ✓ Notes are EXACTLY in the format provided in the example

(This example covers less than half a chapter. The expectation is that your notes will be about twice as long.)

Jane Doe

Aug. 30<sup>th</sup>, 2015

“Is Faith an Answer?”<sup>4</sup>

<b>Textbook Page Numbers</b>	<b>Question</b> (Questions should focus on concepts and arguments.)	<b>Answer</b>
132-133	What does the “Questioner” believe?	<ul style="list-style-type: none"> <li>• People should not believe things on faith.</li> <li>• Believing things on faith is dangerous.</li> </ul>
133	What is faith?	<ul style="list-style-type: none"> <li>• To believe something on faith is to believe something even though you lack evidence for the belief.</li> </ul>

---

<sup>4</sup> From the assigned textbook, Phil Washburn, *Philosophical Dilemmas* (New York: Oxford University Press, 2015).



133	What is a religious example of faith?	<ul style="list-style-type: none"> <li>• People say they believe in God even though they admit there is no evidence for the belief. This is to believe in God as a matter of faith.</li> </ul>
133	What is evidence?	<ul style="list-style-type: none"> <li>• To say there is evidence for something is to say that there are established facts that make the belief more probable.</li> </ul>
133	What example of evidence is offered in the textbook?	<ul style="list-style-type: none"> <li>• For example: if blood stains at the crime scene match the accused, this makes the guilt of the accused more probable.</li> </ul>
133	What is the connection between religious faith and evidence?	<ul style="list-style-type: none"> <li>• To believe in heaven on the basis of faith is to believe in heaven despite the fact that there is no evidence that makes heaven probable.</li> </ul>
133-134	What two explanations and examples does the textbook offer for faith?	<ul style="list-style-type: none"> <li>• One explanation is that faith in something comes from an authority.</li> <li>• For example, parents, priests, etc. may tell children to believe in God even though they cannot provide evidence that God exists. This is to have faith in God based on authority.</li> <li>• A second explanation as to why people have faith is for reasons of comfort.</li> <li>• For example, believing in an afterlife is comforting, so people have faith.</li> </ul>

134	Why does faith based on authority lead to bad consequences?	<ul style="list-style-type: none"> <li>• People who believe based on authority fail to think for themselves.</li> <li>• Seeking out evidence is harder, but leads to individual and societal happiness.</li> <li>• People who believe religious matters based on authority will tend to believe political matters based on authority. This puts too much power in the hands of political authorities.</li> </ul>
135	Why does choosing faith based on comfort lead to bad consequences?	<ul style="list-style-type: none"> <li>• Some people may find ideas comforting that are hurtful to others.</li> <li>• For example, the author imagines that it might be comforting to believe that men are superior to women, even though there is no evidence for this. This sort of sexism is bad for society.</li> <li>• Some may avoid taking responsibility based on the faith in their own innocence. This sets a bad example, and if everyone fails to take responsibility, if everyone claims faith in their own innocence, then society itself will be undermined.</li> </ul>
136	What are beliefs supposed to be about?	<ul style="list-style-type: none"> <li>• Beliefs are supposed to be about the world.</li> <li>• Evidence tells us how the world is, so our beliefs should be based on evidence, not faith.</li> </ul>

