

**DECIDING JUSTICE: UNDERSTANDING AND DEVELOPING MORAL DECISION
MAKING***Daniel Marangoni*

Oklahoma State University

Francis Grabowski

Rogers State University

Abstract: This paper looks at the role that intellectual virtues such as intellectual patience and open-mindedness might play on decision making. While the idea that decision making that allows justice to flourish is relatively linear and objective, the process proves to be more complicated. Taking the idea of justice in a very literal sense, this paper discusses research that used participants as jurors in a hypothetical court case. Building early research in the area, modifications were made to a jury experiment to test how order effects and multiple judgments influence decisions and how intellectual virtue can improve the process.

Keywords: Virtue, cognition, education.

Introduction

Enacting justice through the decision making process is fundamentally important to a health society. In a very literal sense, the idea of justice through decision making is seen regularly across the world in the legal system. Judges, jurors, mediators, and the like are tasked with listening to information – usually from multiple perspectives – and making a decision. And while the decision making process is extremely complicated, the ultimate decision appears relatively simple: a defendant is either innocent or guilty.

McKenzie et al. (2002) examined how hearing both sides of a criminal case impacted believe of guilt. Participants were asked to imagine that they were jurors assembled for a trial. They were provided with pieces of information from the defense and the prosecution which they responded to by providing a numerical representation of their confidence in the guilt of the defendant. Scores were given at the start of the case, after hearing the first piece of evidence, and after hearing the second piece of evidence. Although each participant ultimately heard the same information, the order in which the information was provided caused scores to

vary greatly. The impact of order effects on the determination of guilt in a court case is profoundly important. Even when the same information is provided to all participants, the order that the information is presented could impact interpretation.

This seems to be a break down in what is considered justice. Once the facts are presented, then it should be the facts that determine whether a crime was committed or not. McKenzie et al. (2002) showed that a straight line to just decision making breaks down even from something as simple as changing the order of information. It would be nice if the barriers to justice in the decision making process were identified and mitigated.

In a sense, societies that educate their children attempt this. There is at the heart of education that children can and ought to be brought up to understand particular things about the world. For public education especially, there is often an explicit concern for citizenship and lawfulness. Broadly speaking, education is interested in identifying the barriers to justice and mitigating their influence. This paper is part of a larger research campaign to conceptualize intellectual virtue as aspects of cognition. Specifically, it works to understand how intellectual virtues play a role in cognition. The implications for the research are contextualized within an educational paradigm. In this paper, the researchers conceptualize how intellectual virtue and cognition overlap and what the overlap means for decision making. A variation of the McKenzie et al. (2002) study was used as the mechanism to explore this.

Intellectual Virtue

A virtue, as it is conceived within the Aristotelian tradition, is an abiding disposition acquired through consistent, voluntary action that enables a person to achieve more readily certain ends that would advance the well-being of himself or others. Therefore, a virtuous person is someone who does not merely do things well, but rather does them well often and without effort. A virtuous person, to put it bluntly, knows how to handle himself. Take, for instance, a firefighter who stands before a burning building. Trapped inside are families needing rescue. Most people would be overcome by fear and uncertainty. A skilled firefighter, however, has what it takes to get the job done; for not only is he able to regulate his fear, but he also has the practical wisdom to know how to carry out the rescue mission. Utter fearlessness would cause the firefighter to run recklessly into the building endangering himself and his fellow firefighters. Excessive fearfulness would have the opposite result: he would freeze in terror or perhaps run away. By contrast, what we want from a firefighter is courage—a healthy amount of fear that would prevent him from making a rash decision, yet

the ability to suppress enough fear to save those trapped inside. Courage, then, may be thought of a mean or midpoint between two extremes: one of rashness (which involves having a deficiency of fear) and one of cowardice (which involves having an excess of fear).

Along similar lines other moral virtues may be understood. Simply take some emotion or activity subject to excess and deficiency. The virtue lies in the Goldilocks zone: somewhere between the "too much" and the "too little." Temperance, the virtue pertaining to satisfaction of appetites, lies between inhibition (not enjoying pleasure enough) and overindulgence (enjoying pleasure too much). Generosity, the virtue pertaining to the giving of gifts, lies between miserliness (not giving enough gifts) and overindulgence (giving too many gifts). But note well – it is not merely the amount that counts; it is also how the emotion is felt or activity is performed. Is it done under the right circumstances, with the right intent, with respect to the right people, and so on? A courageous person knows how much fear to show, but he also knows *when* to show it, *where* to show it, and *how* to show it. The same goes for anger. There are, of course, times when people get angry at things or at times that they should not. But to show the right amount of anger, under the right circumstances, towards the right people – that is a sign of a virtuous good temperament.

Intellectual virtues are open to the same analysis. Being tenacious, for instance, is a virtue. It involves knowing when to give ground and how much ground to give. Giving too much ground is a sign of intellectual cowardice or a weak will, whereas not giving enough ground is a sign of stubbornness. Being intellectually patient is also a virtue. When someone is patient – in the intellectual sense of the word – he knows when to and when not to render a judgment. In other words, he will suspend judgment for as long as he needs to suspend it, but he will render a judgment when it is called for. Consider a scientist engaged in research. The data collection is moving along slowly but surely. He is not yet in a position to announce any new discoveries, but he knows that the longer he waits, the more likely that other scientists will beat him to the punch. Part of him wants to make an announcement before he has sufficient proof. He knows, however, that doing so would be purely self-serving and a clear violation of intellectual integrity. Then again, were he a dithering, doubtful scientist, he would risk not making any discoveries at all for fear of being wrong. The patient scientist manages to balance these two extremes: he withholds judgment until sufficient evidence is gathered, but once it is gathered he quickly and confidently renders his judgment. Patience, therefore, requires a sense of timing: the scientist needs to know when to render judgment. It also requires a sense of measure: just how much evidence is enough on which to base his judgment? Finally, it requires a fair amount of strong-mindedness so that professional and

intellectual pressures do not cause the scientist to vacillate. We want our scientists to stand firm in their beliefs and not to be blown this way or that by factors internal or external to their research.

Quantum Cognition

The philosophical idea of intellectual patience, or the ability to appropriately delay judgment as it will be examined here, has particularly interesting connections to cognition – particularly to a relatively new way of thinking about cognition called quantum cognition. Busemeyer and Bruza (2012) say that “quantum theory is a new theory for constructing probabilistic and dynamic systems” and that there is a meaningful use to applying this to cognition and decision making. Quantum probability has proved to be a strong mathematical predictor of events that violate the assumptions of classic probability. For example, classic probability is commutative: it states that $AB=BA$. And yet, the existence of order effects has troubled researchers for decades. Research has shown that when people are presented with the same information in different orders, they come to different conclusions about that information. Quantum probability can account for this. Mathematically, it has been a robust predictor of probability; and as such, it demonstrated itself to be a valuable framework for cognitive mapping.

Prediction is not the only value of quantum theory or even quantum models of cognition. This paper is not concerned with mathematical models of decision making – that is, it is not concerned with predicting the exact values of decisions. But, the ability for quantum models of cognition and decision to make predication has philosophical and practical value. To make its predictions, it uses particular assumptions. Of particular interest to this study is the assumption that stating a belief creates rather than describes a belief.

While classic models of cognition think of belief in definite states at any given time, a quantum model works differently by acknowledging that that a person can be in an indefinite state regarding their belief. In court, jurors need to make decisions about the guilt or innocence of the defendant. A quantum model of decision making where indefinite states exist, then, cannot assume that your belief is either in a guilty or not guilty state. Rather, you “may be in indefinite state that allows both of these definite states to have potential” (p. 2). Heisenberg (1958) defined these potentials as *amplitudes*. While the amplitudes might increase such that the potential for a juror to claim the guilt of a defendant over his innocence, the juror remains suspended over the possibility of both guilt and innocence. Once the juror makes a decision, he projects his believe to either guilt or innocence. Before that, he waved

with potentials (amplitudes), but did not commit. In quantum mechanics, this movement is represented by a wave. In quantum decision, this is experienced as “conflict, ambiguity, confusion, and uncertainty” (p. 3). When the juror does make a judgment regarding the guilt or innocence of the defendant, quantum cognition describes this as “conflict resolution, decision, and certainty” (p. 3).

This definitive state of decision appears uncontroversial in a classic sense. The classic view has been to state that people operate in a state of belief – they are happy or sad, or they believe that someone is innocent or guilty. But for Peres (1998), the idea that people are relatively indifferent to their potential belief states produces a provocative lesson – that taking a measurement of someone’s belief creates a belief rather than describes it. When the juror votes “innocent” or “guilty,” he creates a belief of innocence or guilt inside of himself. This idea, although simple, *is* quite provocative: it says that each time a person projects onto a belief; he creates in his mind a new state. Moving from state to state is not so simple. It is not so simple because judgments disturb each other and introduce uncertainty (Busemeyer and Bruza, 2012, p. 4). Moving from indefinite state to definite state in respect to a question asked changes subsequent responses. Asking someone if he likes dogs might gain an affirmative response, but if her were asked first if he were ever bitten by a dog, the answer might change. If the respondent generally likes dogs despite having been bitten by one as a child, he might be inclined to say that he like dogs. But if he says, “Yes, I have been bitten by a dog,” then his answer to the question, “Do you like dogs” might be less certain than when he had simply been asked about his attitude of dogs. Busemeyer and Bruza (2012) explain uncertainty by showing that “if the first question A produces an answer that creates a definite state with respect to that question, the state created by A may be indefinite with respect to a different question B” (p. 4).

The juror illustration is again helpful. If a juror is asked to make a decision about evidence from the perspective of the prosecution, the juror would be indifferent to a belief about information from the perspective of the defense. These two perspectives are simultaneously incompatible and so a decision about one precludes a decision about the other. It might, then, be “impossible to be in a definite state with respect to two different questions, because a definite state for one is an indefinite state for another” (Busemeyer & Bruza, 2012). In this case, the order that the information is presented would change the order that this indifference was made. In the example of the dog question, asking A then B would get a different response than B then A. Quantum cognition not only identifies this non-commutative nature of the question, but can explain it.

The Role of Virtue in Decision Making

From both a philosophical and quantum cognitive perspective, there are, presumably, advantages to delaying judgment can lead to more accurate decisions. If one is closed-minded and rushes to judgment, he forgoes the opportunity to reflect on a larger sampling of information and confuses the decision process by projecting belief onto multiple incompatible events. Talking about what he calls “open-mindedness,” Baehr (2011) describes its uniqueness by saying that it is distinguishable from other similar qualities in that “it involves a certain willingness and ability to transcend a particular cognitive standpoint, and to do so with the aim of ‘taking up’ or ‘taking seriously’ an alternative cognitive standpoint” (p. 156). Here the word “transcend” and the idea of being in superposition look very similar. Baehr (2011) goes on to say that open-mindedness is “largely or often a ‘facilitating’ virtue,” that allows the “mind to detach or remain detached from a default position or standpoint” (p. 157). This transcendence creates what he says is the “‘psychological space’ for other virtues and faculties to perform their respective functions.”

To be clear, the argument is not that open-mindedness is exactly what is at play with superposition and delaying judgment. First, from a philosophical perspective, the rich and beautiful discussion of virtues should not be ignored. That is to say, the treatment of the idea here is myopic for a particular purpose and draws only a little from the depths of the conversation. Engagement with the literature of intellectual virtues is a wonderful endeavor. Nevertheless, there is a curious connection between the ideas. Without fully expounding on the intricacies of the connections, the point is to note that quantum cognition can provide a psychological description of what philosophers use open-mindedness and other virtues to describe.

So, both the idea of open-mindedness and superposition (and what belief projection entails) implies that suspending judgment until an appropriate time can improve the accuracy of the judgment. Rather than being influenced by a default position or even by the last piece of information, delaying judgment to gather and think about more information could be helpful. The idea of a juror in a court case is particularly compelling here. Does making frequent judgments about the guilt of a defendant influence one’s ultimate decision?

Initial Experiment

Method

An initial exploratory experiment was set up to see test the effect of information order presentation and how the number of judgments impacted final decisions. Participants were

identified (n=44) and randomly assigned to be in one of four groups. Participants acted as jurors in a fabricated court trial and were asked to score their perception of the defendant's guilt. All participants received the same information. In this experiment, the information was alternating evidence from either the prosecution or defense. While all participants received the same information, the groups differed in two important ways: the information was presented in two different ways, and the number of times that participants had to give a score of guilt differed. Of the four groups, two groups received evidence from the defense last and two groups received evidence from the prosecution last. Also, of the four groups, participants either gave a score of guilt after each new piece of information (four times total), or only once at the end.

Each participant was provided a booklet that contained instructions and then four pages of information. The pages contained alternating information that either ended with defense or prosecution evidence, and had places for scores at either the bottom of each page or only once at the end. The score of guilt was a number from zero to twenty where zero was absolutely not guilty, ten was neutral, and twenty was absolutely guilty. The spectrum of numbers in between could be used to represent variations from the extreme scores. So, the groups were: prosecution evidence last with four scores (Prosecution_Last_Quad), prosecution evidence last with one score at the end (Prosecution_Last_Once), defense evidence last with four scores (Defense_Last_Quad), and defense evidence last with one score at the end (Defense_Last_Once). Participants were not aware that there were different groups. The four booklets were presented in equal numbers, and stacked in alternating order. Participants were given the next booklet in the stack when he or she agreed to participate. The 44 participants were equally divided so that each group had n=11 participants.

In keeping with McKenzie et al. (2002), weak information was a single bullet point of information and strong information was three bullet points of information. The prosecution provides strong evidence two times (SD1 and SD2), and the defense provides weak evidence once and strong evidence once (WD and SD). All participants received the following introduction:

The defendant on trial was arrested after police received an anonymous tip from a witness that he was the man who might have committed a warehouse burglary the night before. You have been selected to evaluate the evidence in the case as a juror in the trial.

Next, all participants received the following information alternating between prosecution and defense in the orders mentioned above.

Results

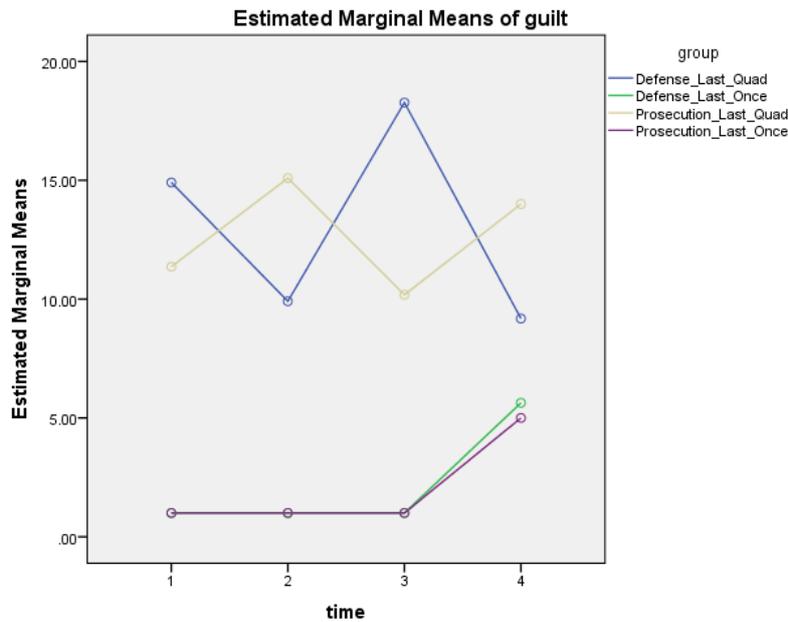
To test for the main effect that order of information had on the guilt score for each group, a one-way ANOVA was run to determine the differences in confidence of guilt at the end of the experiment—time 4. This test revealed that guilt scores were statistically significantly different for different groups, Welch's $F(5, 26.707) = 32.983$, $p < .000$. Welch's test was reviewed since the test of homogeneity of variance was violated in Levene's statistic. Since the Welch Anova was statistically significant ($p < .000$), the data was analyzed using the Games-Howell post hoc test.

Table 1

	N	Mean	Std. Dev.
Defense_Last_Once	11	5.00	1.55
Prosecution_Last_Once	11	5.63	4.15
Prosecution_Last_Quad	11	9.18	7.39
Defense_Last_Quad	11	14.00	5.69

Guilt scores increased from the lowest at Defense_Last_Once ($n=11$, $M=5$, $SD=1.549$) to the highest at Defense_Last_Quad ($n=11$, $M=14$, $SD=5.692$). The difference in the guilt scores for each group can be seen in Table 1. Games-Howell shows that an increase in mean score from Defense_Last_Once ($M=5$, $SD=1.55$) to Defense_Last_Quad ($M=14$, $SD=5.69$) of 9, $SE=1.78$ was statistically significant at $p=.003$. Comparison between these two groups was used since they represent the groups with the largest difference in mean score. Both groups that scored guilt only once has nearly identical mean scores with no significant difference between them ($p=.963$), while the mean scores in the quad groups fell on different sides of the guilt and innocence threshold (score of 10).

Discussion



The study showed that despite receiving the same information, the groups ended with statistically significantly different confidences in guilt at the end of the experiment. First, it is important to note that the information provided to the participants was intended to be stacked in favor of the defense. The strongest piece of prosecution evidence could be considered the DNA evidence, but this is resolved by the fact that the defendant has a twin brother—who would have the same DNA. Also, the brother gave the defendant the items from the scene of the crime, and the defendant can prove he was at home at the time of the crime. Consequently, guilt scores at time 4 that were under the guilt threshold of 10 were considered more accurate than scores above 10.

First, the experiment worked test order effects. Defense_Last_Quad and Prosecution_Last_Quad were used to explore the possibility that the order of presentation could impact the confidence of guilt. The profile plot below visualizes the results: that in fact the order of presentation has significant impact on the confidence of guilt. At time 4, the group that heard defense evidence last had a mean guilt score that fell below 10 and the group that heard prosecution evidence last had a mean guilt score above 10. Since 10 served as a neutral value, it acts as the threshold between conviction and acquittal. That is, although both groups received the same information, the group that read defense evidence last would have acquitted the defendant while the group that read prosecution evidence last would have convicted him. So, while the statistical differences are interesting, the practical implications are profound. Justice, in this case, would be determined by the lawyer who had the last word.

Second, the experiment explored how the number of decisions impacts overall decision making. The presumption was that since judgments create rather than describe beliefs, there would be a difference in confidence of guilt if the number of explicit judgments changed. Again, the profile plot visually represents how this played out in the experiment. If participants only made one determination of guilt at the end, it did not matter in which order they received the information. If all the information was given before the decision, not only were participants likely to give nearly identical scores, these scores were well below the guilt threshold. Since the evidence was explicitly stacked to favor the defense, these low group means are considered highly accurate. A single decision at the end completely mitigated the effects of order and resulted in a highly accurate determination of innocence.

Second Experiment

Recording a guilt score after hearing all of the information in the case resulted in a better decision than recording guilt after each piece of information. This was controlled by the experiment, however. Participants were either asked to score four times, or were only given once chance to record. While this might have particular implications for controlled environments (such as with educators who can control the set-up of educational assessments), it does not give insight into how real jurors might act during a court case. During a trial, jurors are not explicitly asked to record their perception of guilt throughout the trial, but what prevents them from rushing to judgment?

This question is important in the discussion of intellectual virtue since those jurors who demonstrate intellectual open-mindedness or patience in hearing the information, they might self-regulate how often they project a belief during the case. Jurors that rush to judgment might look a lot like the participants who were asked to score guilt after each round of information. Those jurors that had open to and patient with new information might more closely resemble those participants in the experiment who only score guilt once at the end. The purpose of experiment two was to explore whether or not the decision process could be influenced with an intervention.

Method

Experiment two repeated the initial experiment by including the Prosecution_Last_Quad and Prosecution_Last_Once groups but added a new group. Since the discrepancy in scores between the two groups that received defense information last was smaller, these two groups were not re-used in experiment two. The new group was an

intervention group. So, experiment two had three groups all of which received prosecution information last. The intervention group was identical to Prosecution_Last_Quad except that it received a page of information at the start of the experiment that the other groups did not. The interventional group page read:

*Hearing both sides of a story can be complicated and can include a lot of confusing information.
 In this case, your opinion will determine if a man is guilty or innocent and so you should carefully read all of the information to help you make the best decision.
 To make the best decision, you should wait until after you have read all of the information before coming to a final conclusion.*

Experiment two utilized an online version of the booklet rather than a physical booklet. Participants were recruited through the distribution of the link to the online survey. The online survey was set up so that each visitor would be randomly given the survey for one of the three groups. The surveys posted online corresponded to the booklets of the original groups. Here, Prosecution_Last_Quad received the Control survey, Prosecution_Last_Once was given the Single survey, and the intervention group received the Intervention survey. Randomization was set up to distribute each survey in an equal number. This restricts pure randomization but closer mimicked the original randomization method of the initial experiment where an equal number of booklets were printed and passed out in random order.

Results

As with the initial experiment, mixed ANOVA was run to look at the change in guilt scores of each group over time.

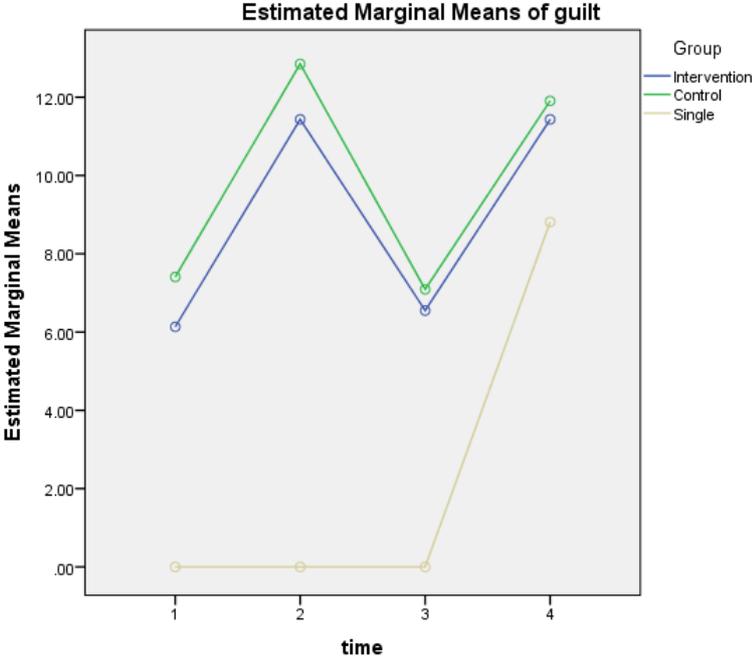
Table 2

	N	Mean	Std. Dev.
Single	63	8.8	5.3
Intervention	60	11.4	4.15
Control	54	11.9	5.12

There was a statistically significant interaction between the groups and time on guilt confidence, $F(5.278, 459.152)=27.185, p<.000$. With this interaction present, it was necessary to review simple main effects. Of particular interest is the simple main effect for group at time 4. A between subjects Anova (one-way ANOVA) was run for time 4. This test revealed that guilt scores were statistically significantly different for different groups, Welch's $F(2, 112.544)= 6.248, p < .000$. Welch's test was reviewed since the test of homogeneity of

variance was violated in Levene’s statistic. Since the Welch Anova was statistically significant ($p = .003$), the data was analyzed using the Games-Howell post hoc test. Guilt scores increased from the lowest at Single ($n=63, M=8.8, SD=5.3$) to the highest at Control ($n=54, M=11.9, SD=5.36$). The difference in the guilt scores for each group can be seen in Table 1. Games-Howell shows that the difference between the mean score of Single group was significantly different from that mean score of the Intervention group ($p=.008$) and the Control group ($p=.006$). While there was a difference in mean score between Intervention and Control, the difference was not statistically significant.

Discussion



The nature of the intervention was built on the idea that particular intellectual virtues would be helpful in this type of decision making process. Also, since intellectual virtues might have the potential to be cultivated, educational interventions such as this one might prove successful. And so, it does need to be noted that there was not a statistically significant difference between the end scores for the Control and Intervention group. Nevertheless, the profile plot below shows that the intervention group moved in predictable ways (according to the Control group) but according to a plot that is slightly shifted down.

Ultimately, the Intervention group means close resembled the group means of the Control group rather than that of the Single group. It is to be argued, however, that the shift in movement is important. It is important for two reasons. The intervention did impact decision

making. Second, the intervention was decidedly small in the grand scheme of things. The influence of order effect and number of decisions are powerful forces. We believe that an impact at all from such a weak intervention opens the possibility that more robust interventions could have greater impact.

Conclusion

Intellectual virtues have a rich philosophical tradition that can be a useful in the classroom. Many school children work through citizenship or values training along with their academic work. Teaching and developing intellectual virtues connects naturally with academic work. It places virtues like intellectual courage and intellectual patience at the heart of academic development. Teaching these virtues in the classroom also gives students the opportunity to practice them along with their studies. Intellectual virtues frequently have moral counterparts that might be strengthened through the development of the intellectual virtues. Students who learn about intellectual patience and practice it in the classroom might find themselves more patient in general.

This paper worked to show that intellectual open-mindedness or patience or any like virtues would be helpful in the decision making process. Waiting until hearing all of the information and making a single decision proved vastly more accurate than rushing to decision. The intellectual virtue of open-mindedness or patience encouraged this type of response to information. While the intervention only made a small impact in the decision making process, it did in fact, have an impact. Arguably, there would be value in more robust instruction and practice in intellectual virtues in the classroom as a means of mitigating the effects of order and multiple decisions on the decision making process.

References

- Baehr, J. (2011). *The inquiring mind: On intellectual virtues & virtue epistemology*. New York: Oxford University Press.
- Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge, UK: Cambridge University Press.
- Heisenberg, W. (1958). *Physics and philosophy: the revolution in modern science*. New York: Harper & Row.
- McKenzie, C.R.M., Lee, S. M., & Chen, K.K. (2002). When negative evidence increases confidence: Change in belief after hearing two sides of a dispute. *Journal of Behavioral Decision Making*, 15, 1-18.
- Perez, A. (1998). *Quantum theory: Concepts and methods*. Kluwer Academic.