An Investigation of Basic Probability Operations Using AND and OR Operations

by

Joannie Tsoi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2007

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The purpose of this thesis was to observe people's abilities to compute probability problems at a fundamental level. The problems in this study were presented in an abstract format to ensure non-ambiguity in its interpretation. The study was administered to university level students. The focus was to determine people's ability to answer probability questions that combined the probabilities of two single events, using the most basic types of operations: AND and OR. The study found that while most people were able to compute AND type probability questions, most had trouble with OR operations. Of special interest was a switching strategy that was employed in computing OR operations as the probability of a single event varied. The study also revealed that statistically sophisticated people were able to adopt the mindset of people who were statistically naïve. Further research is required in order to develop a better framework in understanding people's logical process of computing these basic probability questions as well as its application to our everyday lives.

Acknowledgements

I would first like to thank my advisor, Dr. Frank Safayeni, for his continual guidance in composing this thesis starting from the development of the idea to the finishing product. The accomplishment of this thesis would not have been possible without his expertise and advice. Lynne Wight has also been a tremendous help in guiding me through the process of thesis submission.

My two readers, Dr. Scott Jeffrey and Dr. Rob Duimering, have been generous in taking the time to read my thesis and in providing recommendations to enhance the quality of this paper.

Finally, I would like to thank my family and friends for their ongoing encouragement and support in writing this thesis. Thank you Bikin for reading my first draft and being there for me at all times. I especially want to thank Estelle for taking the time to help proofread and edit this thesis.

Table of Contents

List of Figures

List of Tables

1 Introduction

An important factor in misjudgment is a misperception of the question being asked. A tentative conclusion is that considerable advantage in comprehension may be realized by careful attention to the way in which questions are posed" (p. 55).

- (Borgida & DeBono, 1985; Morier & Borgida, 1984; Nisbett, Krantz, Jepson & Kunda, 1983), rephrased by Garfield and Ahlgren (1988)

Mathematics is an abstract concept. It took mankind a considerable amount of time to develop these concepts and eventually invented operations with short-hands and numerous computational possibilities (Bronowski, 1976). Even within simple mathematical operations, there are varying levels of abstraction. For example, the concept of addition may be easy to grasp because we can apply it to our daily lives. We can see 3 apples, 4 oranges and count to a total of 7 fruits. Multiplication is further abstracted from addition and may be harder to visualize. For example, it may be easy to picture 4 pairs of socks, resulting in 8 socks in total (4x2). However, multiplying larger numbers introduce rapid increases, causing difficulties with mapping these numbers onto something concrete. For example, imagine multiplying 248 by 734 which equals 182032. How does one visualize such an operation? Donaldson (1978, 1993) as cited by Stanovich (2003) emphasized the abstraction of mathematics by saying "what is involved in the mind's movement from 'seven fishes' to 'seven' is abstraction indeed, but it is more: it is a dramatic decontextualization. In the contexts of our ordinary life, we have to deal with quantities of fishes but we never encounter seven" (p. 337). Furthermore, we live in a world that is rich in context and mathematics is often encapsulated in linguistic expressions in which we have to decipher. This causes obstacles in formulating problems and carrying out the appropriate calculations in order to make proper decisions in our daily lives. Osherson, Shafir and Smith (1994) noted that “Probabilistic reasoning poses difficult computational challenges. If probabilities must be distributed over the sentences of an expressive language, these difficulties can become insurmountable” (p. 301).

In studying probabilistic reasoning, researchers often posed problems in the form of an anecdotal story. Problems presented in such a fashion had often been criticized because they were encased in a contextual manner, leading to possible misinterpretation and confusion. Cognitive researchers approached an aspect of heuristics and biases under a mathematical approach but with contextual information. To properly understand people’s logic behind probabilistic reasoning, this thesis posed basic probabilistic questions abstracted from contextual information. This established whether it is the mathematical concepts itself that people have difficulty with, or the inability to extract the appropriate information to solve probability questions. Furthermore, to determine the fundamental level of people’s ability to solve these problems, questions with less abstract concepts were asked. Basic combination of independent events using AND and OR operations should be studied as compared to more complex problems, such as conditional probability involving Bayes’ theorem. The purpose of this thesis was to discover the extent of people’s probabilistic reasoning in a mathematical sense without context and the influence of people’s intuitions on such calculations. Many claimed that concepts of probabilities were hard to grasp, but could not provide the reasoning for this. This thesis attempted to find out whether or not probabilistic reasoning was non-intuitive even at a fundamental level. This was achieved by posing basic AND and OR probabilistic problems with different ranges of the single probabilistic event and examining people's approaches to solving such problems. Studying people's thought processes in calculating fundamental probabilities could be beneficial in constructing improved methods of educational material in teaching basic probabilities.

2 A Review of Probabilistic Judgment in Literature

2.1 Heuristics and Biases of Probabilistic Judgment

A large body of research had been focused on the heuristics and biases of decision making by humans on subjective probability problems. Researchers claimed that people's judgments were misguided by representativeness, causality and attribution, and the availability of information that people had on certain problems (Gilovich, Griffin & Kahneman, 2002; Kahneman, Slovic & Tversky, 1982). The studies that were designed to illustrate these errors were often achieved by providing people with problems that were 'dressed up' with a background story. This led to the heuristics and biases found in probabilistic reasoning. For example, the Linda problem was often used to demonstrate the representativeness effect. The description of Linda was as follows:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and

> also participated in antinuclear demonstrations (Tversky & Kahneman, 1982, p. 92).

Several outcomes were then presented to subjects to which they had to rank from most probable to least probable. Many fell under the *conjunction fallacy* in which they believed Linda was more likely a bank teller *and* a feminist, rather than just a bank teller (Kahneman & Frederick, 2002). Another example was the Tom W. study. Tom was described as a stereotypical engineer. When presented with a base rate probability composing of engineers and lawyers (30% vs. 70%), most ignored the base rate of 70% chance that Tom was a lawyer. Instead, they determined that Tom had a higher probability of being an engineer. This was known as the *base rate neglect*. Studies showing similar concepts were conducted by many others including the disjunction fallacy (Bar-Hillel & Neter, 1993), and research on base-rate neglect displayed by health professionals when assessing patients' mammography information (Casscells, Schoenberger, & Grayboys, 1978). From these findings, Kahneman et al., (1982) concluded that intuitions should be regarded "with proper suspicion" and that people should "replace impression formation by computation whenever possible" (p. 31).

Recently, challenges have been raised against the aforementioned studies. One of the criticisms was raised by Moldoveanu and Langer (2002). Commenting on the experiments that involved participants in making decisions under uncertainties, they pointed out that:

> Making "correct" judgments or choices from the information that was given to the participants depends on an unspoken set of assumptions about that information, which may or may not be justified. It is not clear, therefore, that deviations from such "correct" answers instantiate cognitive "errors"; they may

> alternatively instantiate divergences in the set of assumptions that people make about the decision scenario. If the latter is true, then the decision scenarios that are associated with the "cognitive biases" literature are unspecified, and conclusions about the "incompetence" of decision makers vis-à-vis the axioms of calculus of probability may be premature (p. 373).

Dulany and Hilton (1991), Hilton and Slugoski (2001) also questioned the interpretation of the wording used by researchers in conducting these experiments. Many others have wondered whether or not psychologists have been applying the inappropriate normative model in observing these probability judgments (e.g. Margolis, 1987; Messer & Griggs, 1993; Levi, 1983; Ayton & Hardman, 1997) as summarized by Stanovich and West (2002). These criticisms prompted psychologists to further explore these heuristics and to provide reasoning and justification of the biases found in probability judgments. Kahneman and Frederick (2002) proposed that people employ a dual process model in which they labeled as *System 1* and *System 2* when making decisions. They determined that System 1 was an intuitive system in our cognitive process that allowed us to arrive at answers effortlessly and rapidly. System 2 was a reflective system that monitored System 1 to ensure its correctness. Therefore, it was established that factors that affected System 2 such as statistical sophistication, intelligence, manipulations of attention, etc., were responsible for eliminating biases. Another widely discussed method that influenced statistical reasoning beyond Bayesian inference was the concept of *Frequency Formatting*. Gigerenzer and Hoffrage (1995) conducted a series of experiments which showed that using natural numbers instead of percentages improved people's Bayesian inference. They also quoted others who had similar findings by stating that "Fiedler (1988) and Hertwig and Gigerenzer (1995) showed that the *conjunction fallacy* (Tversky & Kahneman, 1983) in the Linda problem and similar problems largely disappeared when questions were changed form probability to frequency format" (p. 698). Cosmides and Tooby (1996) also supported the frequency view from their result as seen from the following excerpt:

> Frequentist problems elicit Bayesian reasoning. This finding adds to the growing body of literature that shows that many cognitive biases in statistical reasoning disappear, and good statistical reasoning reappears, when problems are posed in frequentist terms (p. 62).

The Mental Model theory was another model that allowed individuals to infer probability of events without having the familiarity of probability calculus (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999). This model will be discussed in more detail in Section 5.4.3 - Discussion on Logical Thinking.

Although initial studies on probabilistic judgment revealed some interesting findings, the fact remained that the problems used by such studies were very context oriented and affected the direction of people's logical argument. Despite various attempts by researchers to explain the reasoning behind heuristics and biases, many were still using anecdotal problems to serve as the basis of analysis. Some have argued that the way people solve problems or use certain cognitive schemata was affected by how the problem was presented to them (e.g. Gigerenzer & Hoffrage, 1995; Girotto & Gonzalez, 2001; Moldoveanu & Langer, 2002; Fox & Levav, 2004). To establish how people truly arrive at probabilistic computations, it is essential to comprehend their level of mathematical skills by abstracting the problems and presenting it to them in a clear manner. This is similar to the way a mathematician would view a problem. Obtain the basic information and ignore information that is deemed irrelevant. Instead of using the contextual problems to explain and develop mental models, it would be beneficial to reformat these problems without the ‘fluff’ and investigate the results and potential errors that could emerge.

2.2 Various Research on Probability Judgment in Cognitive Sciences

Other research had also been carried out to study different types of probabilistic reasoning. Shimojo and Ichikawa (1989) studied variations of the “3 prisoners problem”, which dealt with prior probabilities. They discovered that people used certain subjective theorems, such as the *case theorem*, *ratio theorem*, and *invariant theorem,* to attempt to solve the problem. Support theory was used to explain the concept of subadditivity (Tversky & Koehler, 2002). The authors claimed that people formed impressions by making a selective global view. When people ‘unpack’ information, the sum of the exclusive components exceeded that of the implicit global impression. Macchi, Osherson, and Krantz (1999) further extended this theory by conducting studies that demonstrated superadditivity. Other researchers focused on aspects of conditional probabilities using Wason’s selection task as the basis to develop models and applying them to syllogistic reasoning (i.e. Oaksford & Chater, 2001; Fiddick, Cosmides & Tooby, 2000). Fox and Rottenstreich (2003) studied how people partitioned a problem under different conditions when making judgments under uncertainty. Fox and

Levav (2004) proposed a 'partition-edit-count' method that people applied to solve conditional probability problems by using questions similar to the Monty Hall problem and Wason's selection tasks. Probabilistic reasoning had also generated interest in the field of Artificial Intelligence. For example, Osherson et al. (1994) developed a method to extract a coherent body of judgment in probability that is consistent with a person's intuition. This approach was suitable for developing artificial intelligence systems.

Despite the various portrayal of study, very few in the cognitive field had explicitly looked at simpler probability operations to determine people's comprehension level of these mathematical problems. It might be true that computational limitation and performance errors alone may not account for deviation from normative models when making probability judgments (Stanovich & West, 2002). However, observing computation errors in solving basic level probability problems might serve as a basis to better understand people's cognitive models. It might also contribute to existing models in probabilistic reasoning.

2.3 Probabilistic Intuition and Mathematical Education

Research in mathematical education had been mainly focused on improving students' understanding of probabilities. Many claimed that the difficulty in teaching probabilities was students' inability to develop the intuition and fundamental ideas behind these problems. Garfield and Ahlgren (1988) suggested the following to explain students' inability to learn: students have an underlying difficulty with rational numbers and proportional reasoning, students' experiences and views conflicting with probability concepts and students' distaste for probability due to the abstract training they received for these problems. Since mathematical education researchers tended to focus their efforts on younger children, studies on probabilistic reasoning incorporated more basic types of probability problems when compared to cognitive research. Topics revolving around research in this field included but were not limited to: probabilistic intuition and its evolution with age (Fischbein & Schnarch, 1997), the effect of teaching probability concepts on students' probabilistic intuitions (Fischbein & Gazit, 1984), a study on intuitive interference in probabilistic reasoning using Reaction-Time tests (Babai, Brecher, Stavy & Tirosh, 2006), and inconsistencies about students' probabilistic reasoning due to conflicting frameworks (Konald, Pollatsek, Well, Lohmeier & Lipson, 1993). Fischebein and Grossman (1997) conducted studies concerning schemata and intuitions in combinatorial reasoning using basic probability concepts such as permutations, arrangement without replacements, arrangement with replacements, and combinations. It was concluded that people's guesses were not 'blind' and were expressed in a manner related to the combinatorial schema. The idea behind this thesis was to study people's probabilistic abilities at an even more fundamental level than the concepts covered by Fischbein and Grossman.

Inspired by previous research, the purpose of this thesis was to observe people's abilities to compute probability problems. It was important to ensure that people who were deemed to have an adequate background in probabilities would be able to solve probability questions at a fundamental level. The problems should be abstracted from 'story-like' context, leaving no ambiguity in its interpretation. The focus was to determine people's ability to answer probability questions that combined two single events, using the most basic types of operations: AND and OR. No research had been explicitly carried out in looking at the simple operation of combining two sets of data. Assuming that most people at the university level would be able to compute the probability of a single event, AND and OR operations were often one of the most fundamental concepts learned in probability theory. Due to the emphasis that had been placed on frequency formatting, the problems were also presented with 'natural' numbers instead of percentages. The next section will state the initial predictions and observations made prior to the study. This will be followed by a description of how the study was carried out. Analyses of the results are then discussed, followed by general discussions and concluding remarks.

3 Initial Predictions and Observations

The idea behind this thesis was to determine whether people would be able to carry out basic probability problems using AND and OR operations. It was suspected that while most people would be able to compute AND type probability problems, even university students with a technical background might encounter some difficulty with basic OR type probability questions because they are more complex.

The purpose of having different probabilities for the single events (0.2, 0.4, 0.6, and 0.8) in the AND and OR probability operation problems was to observe the effect of the initial probability on the actual answer to the problem. This initial probability (the probability of a single event/jar) will be referred to as the 'anchor.' It was predicted that most people would be able to arrive at the probability of a single event correctly; that is, they would be able to compute the probability of obtaining a red marble from each jar (they only needed to compute one as the two jars were identical in this study). This reference point (the anchor) would be used to determine the combined probability. It was suspected that students would make adjustments in the correct direction (direction toward the actual answer) from the anchor for both AND and OR operations. The correct direction for AND would be movement below the anchor, and the correct direction for OR operation would be movement above the anchor. The predicted rationale of how students would arrive at the combined probability is explained below.

<u>For AND operation:</u>

The student can determine the anchor, or the probability of choosing a red marble from a single jar. However, if he/she has difficulty merging the probabilities of choosing a red marble from both jar 1 and jar 2, he/she may be able to logically deduce the probability of choosing a red marble from both jars is lower than the single jar case. The probability of choosing a red marble from both jars 1 and 2 must be less than the anchor. No prediction was made in determining exactly *how much* lower people will adjust from the anchor.

For OR operation:

The student can determine the anchor, or the probability of choosing a red marble from a single jar. However, if he/she has difficulty merging the probabilities of choosing a red marble from at least one of the jars (from two jars), he/she may be able to logically deduce the chance of choosing a red marble from at least one of the jars is better than a single jar. The probability of choosing a red marble from at least one jar out of two must be greater than the anchor. Again, no prediction was made in determining exactly *how much* higher people will adjust from the anchor.

The two explanations above led to the following ***Directional Hypothesis***:

> People are able to determine the probability of a single event and make adjustments from this probability to arrive at a combined probability involving either AND or OR operations. They will move in a direction below the anchor for AND operations and move in a direction above the anchor for OR operations.

It was predicted that students would make insufficient adjustments from the anchor leading to an under estimation of OR operations and over estimation of AND operations. The reason behind insufficient adjustment might have been due to the result of multiplying decimal numbers. For example, when running simulations, people often make assumptions about their models. If these assumptions were independent and each of them was 90% accurate, the model might appear to be reasonable. However, it would only take 6 of these assumptions for the model to become inaccurate causing simulation results to be no better than a 50-50 chance ($0.9^6 = 0.53$) (Poile, 2004). The non-intuitiveness of this phenomenon is due to people’s familiarity of thinking that multiplication causes numbers to increase. Fischbein (1999) observed that people “have gotten used to this belief from their childhood when they operated only with natural numbers (for which these beliefs are correct). Later on, even after learning the notion of rational numbers (that is, including also sub-unitary fractions) they continue to hold the same belief – which obviously, does not correspond any more” (p. 29). Since multiplication itself is an abstract concept, people would have even more difficulties in rationalizing multiplication of decimals. The calculation of two independent events with AND operation entailed the multiplication of the probability of the events. Thus, multiplying two decimal numbers would cause the resulting probability to decrease at an unexpectedly rapid rate. One of the methods of determining the probability of the OR operations (in our case) was:

1 – P (picking a blue marble from Jar 1 and Jar 2)

Since the probability of picking 2 blue marbles (i.e. P(*x*)) would be surprisingly small (due to multiplication of decimals), the equation of: *1 –P(x)*¸ where P(*x*) is a small decimal number between 0 and 1, would upshot the probability of picking at least a red marble from both jars to be larger than anticipated. As a result, the unintuitive rapid rate of change from multiplying decimals may lead to overestimations in AND operations and underestimations in the OR operations. The actual equations for solving the AND and OR probabilities are shown in Figure 1.

AND operation:

P(red from both jars) = P(red from jar 1)*P(red from jar 2)

OR operation (2 methods):

(1) P(red from at least 1 jar) = P(red from jar 1)*P(red from jar 2) + P(red from jar 1)*P(blue from jar 2) + P(blue from jar 1)*P(red from jar 2)

(2) P(red from at least 1 jar) = 1 – [P(blue from jar 1)*P(blue from jar 2)]

Figure 1: Equations for AND and OR Operations

As a secondary analysis, different probabilities of the single events (0.2, 0.4, 0.6, and 0.8) allow us to study the possible effects of whether these different anchor values would cause people to vary the amount of adjustments from the anchor. For example, in the probability problems with OR operations, students may be more comfortable with adjusting a greater amount away from the anchor when the anchor is 0.2. Compared to an anchor of 0.8, students may be more conservative and adjust less because 0.8 is much closer to 1 than 0.2 (assuming students are making an adjustment above the anchor for OR operations). Refer to Section 5.1.3 for a more detailed explanation of the *varying adjustment* theory.

Originally, the study was directed at using a simpler model to determine people's intuitive concepts of AND and OR probabilistic problems. Some of the predictions made were based on this model. The initial model questioned whether people would be able to intuitively adjust in the correct direction in combining two probabilities given the probability of a single event. This model was more concerned with people's *intuitive movement* and *direction* relative to the initial given probability. An instrumental study on graduate students was conducted in this manner. It was discovered that most of these students had prior probabilistic training and their approaches in solving these basic probability problems were very methodological instead of being based on intuition. The study was tailored to be more compatible to the population sample and will be described in the next section.

4 Methodology

4.1 Subjects

A total of 246 students from the University of Waterloo participated in this study. The students were comprised of undergraduates from second to fourth year in either engineering or math, with a small percentage in arts and sciences. The students were split into eight groups. The first four groups (groups 1 to 4) had 25, 23, 25, and 24 students respectively. The remaining four groups (groups 5 to 8) had of 34, 40, 40, and 35 students respectively. 84% of these students had received previous instruction in probability, either at a university level or at a high school level.

4.2 Procedure

Online questionnaires were administered to the students using UW-ACE, a web-based course management system. Prior to the questionnaires, the students were informed of the purpose of the study. They were allowed to use calculators to solve the problems, but were told to compute the answer on their own, without referencing other sources. They were rewarded a 1% bonus mark for their efforts. There were eight different sets of questionnaires delivered to the corresponding eight groups of students.

An example showing the first and seventh sets out of eight sets of the questionnaire can be found in Appendix A and Appendix B. The purposes of the first two questions were to determine the field of study of the student and whether they had prior probability education. Next, a probability question dealing with either AND or OR operation, followed by a question asking for the student's confidence in their answers were administered. The student was then asked to provide an answer to the same probability question in the role of someone who had not been educated in probabilities. This served as a precaution to the study as it allowed for the examination of students' intuitive approach to the problem despite knowing the logical computation. The student was also asked to provide an explanation of the logic used to arrive at his/her answer.

The first four sets of questionnaires tested the ability of students to solve basic AND type probability questions. The question was structured so that students had to determine the probability of two identical events (each presented as probability of choosing a red marble in a jar) by combining the two events using the AND operation. An example of the AND type question with probability of the single event being 0.2 can be found in Appendix A (question #3). The probabilities of the events were varied with the values 0.2, 0.4, 0.6, and 0.8 making up the four sets of questionnaires. The other four sets of questionnaires tested the ability of students to solve basic OR type probability questions. Again, the question was structured so that students had to determine the probability of two events with identical probabilities by combining the two events using the OR operation. The probabilities of the choosing a red marble from each jar were again varied with the values 0.2, 0.4, 0.6, and 0.8. The seventh set out of the eight sets of questionnaire found in Appendix B is an example of the OR type questionnaire with the probability of the single event being 0.6. The only difference between the AND and OR type questionnaires was question #3 (see Appendix A and Appendix B). It should be noted that the number of students that answered OR type probability questions was deliberately set to be larger than the number of students that answered AND type probability questions, as it was predicted that a larger number of students would be able to carry out AND probability operations correctly. It would be more interesting to study the effects of OR probability operations with a larger sample size.

5 Results and Discussions

5.1 Global Analysis of Data

It was assumed that at least some of the students would be able to carry out basic probability calculations since the study was given out to second to fourth year university level students. It was of interest to examine people's intuitions on probabilistic reasoning and whether or not these intuitions matched the proposed directional hypothesis. For those who were unable to carry out the computation to the probability problems, their intuitive reasoning could be examined from their actual answers. For those who were able to compute the answers, they were asked to take on the role of a statistically naïve individual to allowed for the examination of their 'intuitive' method of computing probabilities. Analysis of data was split into two parts. For people who were able to answer the initial probability question correctly, their answer to the same question while assuming the role of someone who has not had probability training was considered. This group was termed as the 'correct' group. For people who did not answer the initial question correctly, their answer for the initial question was considered. This group was termed as the 'incorrect' group. The reason for dividing the analysis between the 'correct' and 'incorrect' groups was to observe the intuitive thought-processes of the statistically sophisticated compared with the statistically naïve. A global analysis combining data from the 'correct' and 'incorrect' groups was obtained from the results. Analysis showed that the data supported the directional hypothesis.

5.1.1 Frequency Analysis for AND Operations

For AND operations, using frequency analysis from SAS, it was found that only 11.7% of the students produced an answer bigger than the anchor. The remaining 54.26 % and 34.04% of the students produced answers that were either the same as or smaller than the anchor respectively. The differences between these numbers were found to be significant ($p < .0001$). Refer to Table 1 for a more detailed breakdown of the percentage frequency for each of the anchor.

Table 1: AND Frequency % for Both ‘correct’ and ‘incorrect’ Groups

Anchor	Same as anchor (%)	Smaller than anchor (%)	Bigger than anchor (%)
0.2	33.4	45.8	20.8
0.4	54.6	31.8	13.6
0.6	54.2	33.3	12.5
0.8	75	25	0
Overall %	**54.3**	**34**	**11.7**

Note that a large number of students arrived at an answer that was the same as the anchor for AND probability operations. The logical explanation for this answer can be found in the discussion of Section 5.4.1, Table 14: Common Mistakes for AND Operation. Similar results were obtained when the data was split into ‘correct’ and ‘incorrect’ groups. For the ‘correct’ group, 11.6% of the students had an answer above the anchor, 59.4% had an answer the same as the anchor, and 20% had an answer below the anchor. For the ‘incorrect’ group, 12% of the students had an answer above the anchor, 40% had an answer the same as the anchor, and 48% had an answer below the anchor.

From Table 1, it could be seen that people who calculated answers both below and above the anchor decreased 20% as the anchor increased from 0.2 and 0.8. Instead, a much larger proportion of people (> 40%) computed answers that were the same as the anchors for the anchor change of 0.2 to 0.8. Currently, no explanation could be given for the relationship between the increasing initial anchor value and the increasing number of answers that are the same as the anchor. Perhaps this relationship is due to some unusual statistical variation from the small sample sizes collected for the study for the AND operation. To demonstrate the size of the samples obtained for the AND operation, Table 2 below shows the actual frequency (instead of percentages) of the answers that were the same as, below, and above the anchor.

Table 2: The Sample Sizes of AND Operation

Anchor	**Same as anchor**	**Smaller than anchor**	**Bigger than anchor**
0.2	8	11	5
0.4	12	7	3
0.6	13	8	3
0.8	18	6	0
Overall	**51**	**32**	**11**

As seen, sample sizes were quite small as many of the cells were less than 10. Unusual statistical variation could very well contribute to the large increase of people that arrived at an answer equivalent to the anchor as the anchor values changed from 0.2 to 0.8. Further research needs to be conducted to study the correlation (if it exists) between initial anchor value and calculations resulting in answers being the same as the anchor.

To better compare the percentages of people who calculated answers that were either smaller or bigger than the anchor, frequency analysis was conducted with the exclusion of anchor value answers. This allowed for a more detailed examination of people's direction of movement from the anchor. From Table 3, it could be seen that a larger percentage of people were consistently computing answers below, rather than above the anchor for each of the anchor values. The overall result was that 74.4% had answers below the anchor compared to only 25.6% above the anchor. The p-value was 0.0014, which showed significant difference between the answers below versus the answers above the anchor.

Table 3: AND Frequency Excluding 'same as anchor' Answers

Anchor	**Smaller than anchor (%)**	**Bigger than anchor (%)**
0.2	68.7	31.3
0.4	70	30
0.6	72.3	27.3
0.8	100	0
Overall %	**74.4**	**25.6**

The significant portion of people who computed answers below the anchor compared to the people with answers above the anchor provided statistical support for the directional hypothesis.

5.1.2 Frequency Analysis for OR Operations

The frequency analysis of the combined data of ‘correct’ and ‘incorrect’ groups for OR operations revealed that 20.8% of the students produced an answer smaller than the anchor. The other 47.7 % and 31.5 % of the students produced answers the same as and greater than the anchor respectively. The differences between these numbers were found to be significant ($p = 0.0003$). Refer to Table 4 for a more detailed breakdown of the percentage frequency for each of the anchor.

Table 4: OR Frequency for Both 'correct' and 'incorrect' Groups

Anchor	Same as anchor (%)	Smaller than anchor (%)	Bigger than anchor (%)
0.2	41.2	14.7	44.1
0.4	30	22.5	47.5
0.6	57.5	22.5	20
0.8	62.85	22.85	14.3
Overall %	**47.7**	**20.8**	**31.5**

Similar results were obtained when data was split into 'correct' and 'incorrect' groups. For the 'correct' group, 18.2% of the students had an answer below the anchor; 45.4% had an answer the same as the anchor; and 36.4% had an answer above the anchor. For the 'incorrect' group, 22.9% of the students had an answer below the anchor; 49.4% had an answer the same as the anchor; and 27.7% had an answer greater than the anchor.

Unlike AND operations, there was no drastic change in the percentages of people that computed answers the same as the anchor as the initial anchor value increased. However, there was a drop in the number of people that calculated answers above the anchor. A more detailed investigation of people's direction of movement from the anchor was conducted with the exclusion of anchor value answers. From Table 5, it could be seen that a higher percentage of students arrived at an answer above the anchor for questions with low initial anchors (0.2, 0.4), while this percentage dropped considerably (> 20%) for the higher anchor questions (0.6, 0.8). With an initial anchor value of 0.2, the percentage of people that had answers above the anchor was 50% higher than people with answers below the anchor. For anchor values of 0.6 and 0.8, however, there were less people from the ‘bigger than anchor’ group compared to the number of people from the ‘smaller than anchor’ group.

Table 5: OR Frequency Excluding 'same as anchor' Answers

Anchor	**Smaller than anchor (%)**	**Bigger than anchor (%)**
0.2	25	75
0.4	32.1	67.9
0.6	52.9	47.1
0.8	61.5	38.5
Overall %	**39.7**	**60.3**

The overall result was that 39.7% of students had answers below the anchor compared to the 60.3% who had answers above the anchor. Although there were more people with answers above the anchor for the OR operation, there was very little evidence against the answers below and above the anchor being significantly different (p-value = 0.07). This result was caused by the ‘*switch’* of percentages between the ‘bigger than anchor’ and ‘smaller than anchor’ group from low initial anchor values (0.2, 0.4) to high initial anchor values (0.6, 0.8). This phenomenon will be discussed in Section 5.5: Changing Strategies in OR Operation.

5.1.3 Adjustments and the Effect of Anchor Values

Another point of interest from the initial predictions was whether or not the initial anchor value had any effect on the amount of adjustment from these anchors. This trend was analyzed by examining the distribution of the absolute value of adjustment away from each of the anchors. For example, if the anchor value was 0.2 and two of the answers given to the probability questions were 0.1 and 0.3, then the absolute value of adjustment for both of these answers would be 0.1. The smaller anchors were combined into a group of ‘lower’ anchors (0.2, 0.4) and the larger anchors were grouped as the ‘higher’ anchors (0.6, 0.8). The mean of the absolute value of adjustments was calculated for the lower and higher anchors for both AND and OR operations. For the AND operations, the mean of the adjustments of the lower anchors was 0.103, while the mean of the higher anchors was 0.129. Although these values were similar, the adjustment for lower anchors was slightly smaller than the higher anchors. Further analysis was conducted to compare the distribution of absolute adjustments between the lower and higher anchor values. The Wilconxon-Mann-Whitney test was applied on the data because data distribution showed strong evidence against normality. The p-value of the distribution comparing the adjustment of lower and higher anchor was 0.2939. This value indicted the insignificance of the differences between the distributions of the adjustments for the lower and higher anchors, despite of the slight variation in values observed from the mean of the adjustments between the lower and higher anchor values.

From the previous analysis, it was seen that people tended to adjust in a direction below the anchors for AND operations. As a result, there would be less room for adjustment for lower anchor values compared to higher anchor values. Refer to Figure 2 for a visual representation of the explanation.

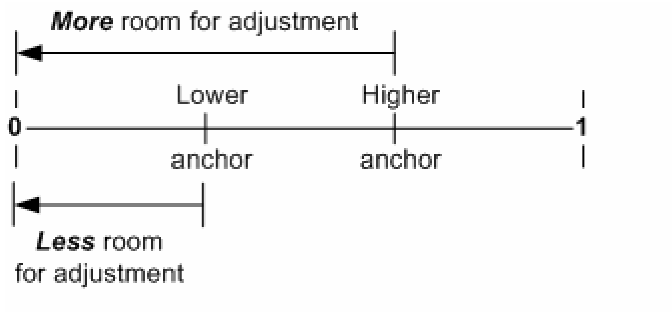

Figure 2: Adjustment for the AND Operations

Although there was a possibility that the amount of fluctuation of adjustments from the anchor was caused by the amount of space from the endpoint, it was determined that the distributions of the absolute value of adjustments between lower and higher anchors did not vary. Therefore, for AND operations, the initial anchor value did not have an effect on people's amount of adjustments from the anchor.

For the OR operations, the mean for the absolute value of adjustments of the lower anchor was 0.176, while the mean of the higher anchor was 0.126. Again, these values were similar, but in contrast to the AND operations, the mean of the lower anchor adjustments was slightly larger than the higher anchor adjustments. Analysis was conducted to compare the distribution of absolute adjustment between the lower and higher anchor values using the Wilconxon-Mann-Whitney test. The p-value of the distribution comparing the adjustment of lower and higher anchor for OR operation was 0.001. This showed a statistically significant difference between lower and higher anchors values on their absolute value of adjustments. Unlike the AND operations where people arrived at an answer by adjusting below the anchor, people generally moved in a direction above the anchor for OR operations. As the anchor values became larger in OR operations, there appeared to be less room for adjustment. It was found that higher anchor values instigated a reversal in people's direction of movement (moving below the anchor instead of above) for OR operations. Refer to Figure 3 for a visual representation of the explanation.

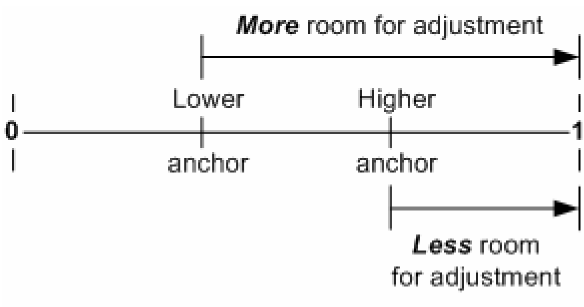

Figure 3: Adjustment for the OR Operations

Therefore, in contrast to the AND operation, the initial anchor value, which affected the amount of adjustment space to the endpoint, might be a factor in the rationalization of the smaller mean adjustment value for the higher anchors of the OR operation.

From the global data analysis, it was found that there was some statistical support for the directional hypothesis. Most people by large recognize the probability of a single event, the anchor. A large number of people from both AND and OR operations provided answers that were the same as the anchors. Furthermore, excluding those results, it was found that people were able to make adjustments in the 'appropriate' direction for the AND operations; their answers were below the anchor. This provided statistical support for the directional hypothesis. For OR operations, some support was shown for the directional hypothesis as there were a larger number of people who had answers above the anchor for the low anchor values of 0.2 and 0.8. However, it was found that a larger number of people had answers below the anchor for higher anchor values of 0.6 and 0.8. Initial anchor values also had an effect on the level of adjustments that people make from the anchor for OR operations, but had no effect on AND operations. Further investigation will have to be carried out in order to determine the validity of the 'room for adjustment' explanation on the actual amount that people adjust from the anchor. Also, the reasoning behind the relationship between anchor values and the amount of adjustments which varied for OR operations but not for AND operations still need to be determined.

Due to some unpredicted results from the data, it was found that the directional hypothesis only provided a narrow view of how people constructed basic probabilities. The data revealed that people applied models that were beyond the realm of logic provided by the directional hypothesis in computing basic AND and OR probability operations. These concepts will be discussed in the following sections.

5.2 The Differences Between AND and OR Operations

Initially, it was predicted that most people would be able to carry out AND type probability operations, but would have more difficulties with OR type probability operations. Looking at the results from the collected data, this prediction held true. For AND type operations, 74.2% of the people were able to answer the question correctly, when the four different anchor values of AND operations were combined. Amongst them, 86.6% of the people had been trained in probabilities before either at a high school or university level. A specific breakdown of the percentages of people that answered the AND probability questions correctly and the percentages of people with prior probability training for each anchor value are shown in Table 6.

Table 6: Prior Probability Training vs. Correct Answers for AND Operation

Anchor value	% of correct answers	% with prior probability training
0.2	84%	92%
0.4	70%	87%
0.6	76%	88%
0.8	67%	79%

Although a majority of people was able to answer AND type probability questions correctly, it could be seen that the percentage of people with prior probability training was higher.

For OR type probability questions, data showed that less than 50% of the people were able to answer these questions correctly. Combining all of the questions from the different anchor values, 45.2% of the people were able to obtain the correct answer. Out of these people, 82.2% had received prior training in probabilities. A breakdown of the percentages of people that answered the OR probability questions correctly and the percentages of people with prior probability training for each anchor value are shown in Table 7.

Table 7: Prior Probability Training vs. Correct Answers for OR Operation

Anchor value	% of correct answers	% with prior probability training
0.2	44%	82%
0.4	34%	76%
0.6	58%	85%
0.8	44%	86%

Again, the percentages of people with prior probability training were higher than the percentages of people who were able to answer these probability questions. Unlike the AND cases, these percentages were higher by a considerable amount (at least 35%). It should be noted that in both AND and OR cases, the students' logical explanations were verified to ensure that the correct answers were indeed 'correct' and followed the appropriate mathematical calculations.

The effect of prior probabilistic training and people's ability to compute AND and OR operations was further analyzed in Table 8 and Table 9. For AND operations from Table 8, it was seen that prior training in probabilities had a positive effect on knowledge. While 0% of the people who have not had training answered the questions correctly, a large majority of the people, over 80% and above who had training answered the questions correctly for all the anchors.

Table 8: The Effect of Prior Training for AND Operation

Anchor value	**% correct with prior training**	**% correct without prior training**
0.2	91.3%	0%
0.4	80%	0%
0.6	86.4%	0%
0.8	84.2%	0%

Table 9: The Effect of Prior Training for OR Operation

Anchor value	**% correct with prior training**	**% correct without prior training**
0.2	53.1%	14.3%
0.4	45.2%	0%
0.6	68.6%	0%
0.8	51.6%	20%

The effect of prior probabilistic training and people's ability to compute OR operations was analyzed in Table 9. For OR operations, prior training in probabilities also had a positive effect on knowledge, but the results were not as drastic as it was for AND operations. Surprisingly, for anchors 0.2 and 0.6, 14.3% and 20% of the people who had no prior training were able to compute OR operations correctly. Only around half the people who had prior training were able to answer OR operation problems correctly for all the anchors. From Table 8 and Table 9, it was seen that prior probability training improved one's abilities to calculate basic probabilistic operations. This training proved to be more effective for AND operations compared to OR operations.

Another difference between AND and OR operations was the propensity to overestimate answers for AND operations and underestimate answers for OR operations. This trend is evident when the mean of the answers is obtained from the 'correct' and 'incorrect' groups for each of the anchors in Table 10 and Table 11. The similarity in the mean values for both 'correct' and 'incorrect' groups showed that people who were statistically advance applied intuitive approaches to the problems that were comparable to the statistically naïve in their tendency to overestimate AND operations and underestimate OR operations.

Table 10: Mean Values of 'correct' and 'incorrect' Groups for AND Operation

Anchor	Correct estimation	Incorrect answer	Actual answer
0.2	0.19	0.19	0.04
0.4	0.42	0.32	0.16
0.6	0.56	0.53	0.36
0.8	0.78	0.73	0.64

Table 11: Mean Values of 'correct' and 'incorrect' Groups for OR Operation

Anchor	Correct estimation	Incorrect answer	Actual answer
0.2	0.28	0.30	0.36
0.4	0.55	0.52	0.64
0.6	0.61	0.61	0.84
0.8	0.73	0.75	0.96

For the AND operations, the mean of all the answers from the ‘correct’ group and the mean of all the answers from the ‘incorrect’ group were both larger than the actual answers for each of the anchors. For OR operations, both the ‘correct’ and ‘incorrect’ groups presented answers with means that were smaller than the actual answers for each of the anchors. From the initial predictions, it was assumed that people overestimated AND operations and underestimated OR operations because of insufficient adjustments made from the initial anchor value provided that they were following the directional hypothesis. However, the logic behind overestimations and underestimations was not due to the unexpected consequence of multiplying decimal numbers (theorized as the cause of insufficient adjustment). As mentioned in Section 5.1.2, people did not follow the directional hypothesis for the higher anchor values for OR operations; they did not adjust in the correct direction from the anchor. This occurrence was also observed from the mean of people’s answers in Table 11. The mean values of both the ‘correct’ and the ‘incorrect’ groups’ answers showed a significant amount of underestimation compared to the actual answer as these mean values were below the anchor, moving in the wrong direction for the anchor value of 0.8 for the OR operation.

The study revealed that a majority of people was able to carry out basic AND type probability questions but had more difficulty with basic OR type probability questions. While only a few people with prior probability training were unable to compute basic AND type probability questions, a high number of people had forgotten how to calculate basic OR type probability problems even with prior probability training. Although prior probabilistic training was beneficial for both AND and OR operations, further analysis revealed that such education proved to be more useful for AND operations than for OR operations. Even though it was predicted that more people would be unable to carry out OR probability operations, it was surprising to find such a large percentage of students unable to do so, especially at a second year university level or higher. Another interesting discovery was that while people showed propensity to overestimate AND operations and underestimate OR operations, the reasoning behind this occurrence was not caused by insufficient adjustment from the anchor. To obtain a better understanding into students' reasoning and approaches to these problems, their logical explanations will be reviewed in Section 5.4.

5.3 A Comparison Between the 'Correct' and 'Incorrect' Groups

The purpose of dividing the analysis between the 'correct' and 'incorrect' groups was to better examine people's intuitive thought process when calculating AND and OR probability questions. It would be acceptable to simply study people who made mistakes in their calculations and analyze the type of approaches they were using. However, curiosity also surfaced in investigating how someone with the knowledge to properly formulate an answer would go about answering such problems based on their 'gut feelings'. A comparison between the ‘correct’ group’s estimation of the answer by assuming a role of a person with no probability training and the ‘incorrect’ group’s actual answers revealed that a large number of people who were able to compute probabilities correctly had a good model of the ‘incorrect’ method of operation. Table 10 (p. 30) from Section 5.2 shows the mean of the correct estimates for each of the four anchor values (single event probability) compared to the mean of the incorrect answers for the AND operations. The OR operation means between correct estimates and incorrect answers are found in Table 11 (p. 31) from Section 5.2.

From the tables, it was clear that the mean of the correct estimations and incorrect answers were astoundingly close. This was especially true for the OR operations. Comparing the mean values for the AND operations might not be reliable as the sample sizes were quite small for some of the groups. For example, the incorrect sample size for the four anchors ranged from four to eight people. However, looking at the mean served only as an initial step. It hinted that even if people knew how to carry out AND and OR probability operations with an objective formula, they somehow retained the knowledge or intuition of calculating them incorrectly.

A more statistically sound method of looking at the ‘correct’ and ‘incorrect’ groups was to compare their distributions for the anchor values of 0.2, 0.4, 0.6, and 0.8 for both AND and OR operations. Due to the small sample sizes, it was found that these distributions were not normally distributed. Using the *proc univariate* function in SAS, it was found that for AND operations, the distributions showed either moderate or strong evidence against normality (with all p < = 0.0047). For OR operations, all distributions showed strong evidence against normality (all p <=0.003). The Wilcoxon-Mann-Whitney test was used on the data as it did not require the assumption of normally distributed interval variables. Table 12 displays the p-values comparing the answers between the 'correct' and 'incorrect' group for each of the anchors of the AND and OR operations.

Table 12: P-values for Comparing 'correct' and 'incorrect' Distributions

Anchors	**P-value**	**Difference Between Correct and Incorrect Groups**
AND 0.2	0.8729	- answers not statistically different
AND 0.4	0.0789	- little evidence against answers being statistically different
AND 0.6	0.4881	- answers not statistically different
AND 0.8	0.0668	- little evidence against answers being statistically different
Anchors	**P-value**	**Difference Between Correct and Incorrect Groups**
OR 0.2	0.7098	- answers not statistically different
OR 0.4	0.4184	- answers not statistically different
OR 0.6	0.7960	- answers not statistically different
OR 0.8	0.5256	- answers not statistically different

The test revealed that for AND operations, the answers between the 'correct' and 'incorrect' groups were either not statistically different (for the anchors 0.2, and 0.6) or there was little evidence against the answers of 'correct' and 'incorrect' groups being statistically different (for the anchors of 0.4 and 0.8). For OR operations, all of the answers between the 'correct' and 'incorrect' group for the four anchor values (0.2, 0.4, 0.6, and 0.8) were not found to be statistically different. Even without statistical calculations, it was clear that the distributions for the 'correct' and 'incorrect' groups were extremely similar as shown from the example in

Figure 4. This Figure displays the OR distribution with an anchor of 0.6 for both the ‘correct’ and ‘incorrect’ groups.

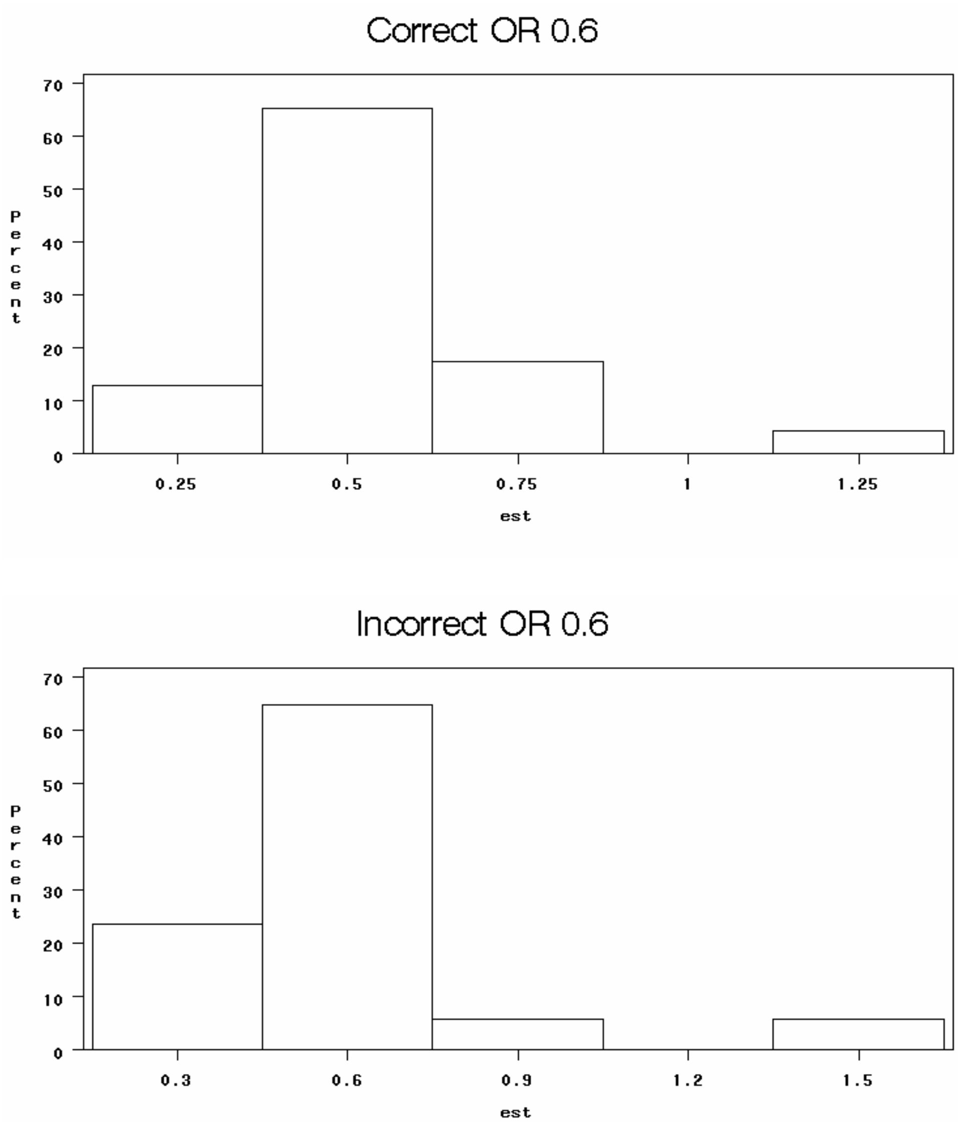

Figure 4: Answer Distributions in OR Operation Between ‘correct’ and ‘incorrect’ Groups

Looking at the logical explanations that people provided for their answers, a majority of the ‘correct’ group gave the same logical explanation as the ‘incorrect’ group for both AND and OR operations for various anchor values. This was especially true for OR operations as shown in Table 13. The percentages of the ‘correct’ group from Table 13 was calculated by combining all the students in the ‘correct’ group for all the cases of either the AND or OR operations in which the logical explanation of obtaining the answer was the same for both the ‘correct’ and ‘incorrect’ group. This number was then divided by the total number of students that were in the ‘correct’ group, producing the percentage value. Since a majority of the people was able to answer the AND questions correctly, it was more interesting to observe the results from the OR operation. The logical explanations for the students’ answers to these probability questions will be discussed in greater detail in Section 5.4: Calculation Logic.

Table 13: Percentage of 'correct' with Same Logical Explanation as 'incorrect'

	% in ‘correct’ group with the same logical explanation as the ‘incorrect’ group
AND operations	65.2%
OR operations	79.7%

The results showed that a large number of people who, despite being able to compute basic probabilities, were able to retain the knowledge of the incorrect computational method. This indicated a strong tendency for people to employ certain methodologies in calculating these probabilities. These methods have not been forgotten even after learning the correct technique of basic probability calculations. There might be some ‘intuitive’ logic that people apply to these probability questions when the appropriate mathematical method of calculation is not used or known. Research on probabilistic judgment claimed that people have a natural ability to solve probabilistic problems depending on how the problem is presented, such as using frequency formats (Gigerenzer & Hoffrage, 1995). However, from the findings of this analysis, instead of being naturally good statisticians, people naturally revert back to being in the mind frame of the statistically naïve.

5.4 Calculation Logic

The global analysis showed some statistical support for the directional hypothesis (refer to Section 5.1) and the presence of overestimation and underestimation for AND and OR operations respectively. However, the reasoning for arriving at the answers to these probability questions was completely different from the logic that was predicted in Section 3, which was based on a simpler intuition model. As predicted, a majority of people were able to identify the probability of 1 jar (or a single event) with little effort. The difficulty resided in combining probabilities of 2 jars. It seemed that when students encounter trouble in combining the probabilities, instead of making an attempt to estimate an answer through logical deduction, many of them went into a 'number crunching' mode. They would try to manipulate the given numbers using some sort of mathematical operation. This was different from studies on base-rate neglect, or conjunction fallacy (Kahneman et al., 1982), where people's *intuitions* misled them because of how the problems were presented (in a very contextual format). In this case, because the numbers drove the logic, people lost their gut feeling, or *intuition*, based on these calculations. This concept was neatly summarized by Garfield and Ahlgren (1988) as they stated that:

> Students often tend to respond to problems involving mathematics in general by falling into a “number crunching” mode, plugging quantities into a computational formula or procedure without forming an internal representation of the problem (Noddings, Gilber-MacMillan, & Lutz, 1980). They may be able to memorize the formulas and the steps to follow in familiar, well-defined problems but only seldom appear to get much sense of what the rationale is or how concepts can be applied in new situations (Chervany, Collier, Fienberg, & Johnson, 1977; Garfield 1981; Kempthorne, 1980). Within the conceptual underpinnings, the details they have learned or memorized, for whatever use they might be, therefore quickly fade (p. 46).

Another basis for the reliance of mathematical computation in determining the answers was due to the way the problem was presented to the students along with the targeted population chosen for this study. The population had been compromised by selecting people with prior probability training. An improved study on people’s *intuitions* of basic probability involving AND and OR operations would be to select participants without background in probability training and to frame the question without asking participants to *compute* the answer. Such a study might be better suited for the initial predictions made in this thesis. Further discussions and a sample question are found in Section 7. Surprisingly, the mathematical tendencies in calculating these probabilities were common as they were predicted by those who were able to compute the correct answer as shown in Section 5.2. Specific cases of people’s logic in arriving at their answers for the AND and OR probability questions is analyzed in the following section.

5.4.1 AND Operation Logic

It was shown in Section 5.2 that a majority of people (74.2%) were able to carry out AND probability operations. Refer to Figure 1 for the equation of how to solve the AND type probability questions. For the people who computed the AND probability questions incorrectly, some of the common mistakes made are summarized in Table 14.

Table 14: Common Mistakes for AND Operation

#	Answer compared to anchor	% of incorrect people with this logic	Operation logic
1)	Same	27.3%	- forms correct probability of 1 jar, disregards the need of combining with the probability of the other jar resulting in answer being the probability of just 1 jar (probability of both jars are identical)
2)	Smaller	18.2%	- forms correct probability of 1 jar, divide this probability by 2 to obtain result
3)	Same	9.1%	- combine the marbles from jar 1 and jar 2 together, formulate the probability of choosing a red marble using the combined result of both jars

These common mistakes were also identified by people who were able to answer the problems correctly (the correct group) with 27.5% using explanation 1) and 4.3% using explanation 2), and a rather large percentage of people (30%) using explanation 3). An interesting phenomenon found in the common mistakes was also observed in the global analysis. A significant number of people provided an answer that was the same as the anchor. In the case of the AND operation, instead of attempting to combine the probabilities of both jars together, people took the probability of a single event as their answer. It is uncertain why people chose to ignore the probability of the other jar. Perhaps people had the notion of combining the probabilities, but when uncertain of the method of operation, they reverted back to just using the probability of a single jar. It was difficult to obtain people's train of thought from examining their logical explanations as they were not revealing. Most of them provided explanations such as: *5 marbles in each jar, 3 are red, therefore 3/5 = 60%* (for an anchor of 0.6) or *total of 5 marbles with 4 of them being red, choose one and you have 4/5 chance of picking out red* (for an anchor of 0.8). A different approach to the current study will have to be prepared in order to collect the appropriate data for understanding the reasoning behind people's tendency to 'ignore one of the jars'.

Although most people were able to formulate the correct probability of choosing a red marble from one jar (a single event), 8.8% of the total people that answered the AND probability questions made several mistakes in formulating the probability of a single jar. These explanations were not grouped together as there were various ways in which people rationalized the probability of choosing a red marble from a single jar followed by different methods of combining the probabilities of jar 1 and jar 2 for AND operations. Aside from the combination logics (for two jars) displayed in Table 11 (ignoring 1 jar, dividing the probability by 2, combining marbles of 2 jars), the other combination logics were as follows:

- Multiplying the probabilities of the two jars together (this is the correct method of arriving at the answer for AND operations)
- Adding the probabilities of the two jars together

Out of the people who answered incorrectly, 22.7% of them formulated the wrong probability for a single jar, but used the correct operation (multiplying the probabilities of the two jars together) to calculate the combined AND operation.

Some of the methods in which people formulated the probability of obtaining a red marble from a single jar p(A) were as follows:

a) p(A) = 1/(total # of red marbles)
b) p(A) =1/2
c) p(A) = # of red marbles / # of blue marbles
d) p(A) = 1/(total # of marbles); regardless of the number of red marbles

As a comparison, the correct method of identifying the probability is: p(A) = r / n, where r is the total # of red marbles in the jar and n is the total # of marbles in the jar.

Attempts at estimating how people arrived at these probabilities were made. However, since most students did not provide an explanation of how they computed these probabilities, the assumptions might not be accurate. For a) and d), the students were expressing the probability as a $1/n$ ratio (where n was the total set of possibilities). Fischbein and Gazit (1984) found a similar misconception in their paper of studying the effect of education on probability intuition in children. The reason for having a '1' in the numerator was because "one extracts in fact, at each trial, a single marble" (Fischbein & Gazit, 1984, p. 9). The total set of possibilities (n) in a) was regarded as the total # of red marbles in the jar. In d), it was regarded as the total # of marbles (which was correct). It was challenging to pin down the exact reasoning of the logic behind b). An explanation could be because the outcome would be to either draw a red marble or not, resulting in a 50/50 chance. Fox and Rottenstreich (2003) found that people would often adopt a default case partition where an event target would either occur or not occur when trying to solve a probability query. There might be other reasoning behind the 1/2 probability formulation, but this could not be verified under the circumstances. For c), the person simply took the ratio of red marbles over the blue marbles.

5.4.2 OR Operation Logic

Unlike AND probability operations, a majority of people (54.8%) were unable to carry out OR probability operations correctly. Few of the people who computed the OR probability questions incorrectly made common mistakes that are summarized in Table 15.

Table 15: Common Mistakes for OR Operation

#	Answer compared to anchor	% of incorrect people with this logic	Operation logic
1)	Same	27.1%	- forms correct probability of 1 jar, disregards the need of combining with the probability of the other jar resulting in answer being the probability of just 1 jar (probability of both jars are identical)
2)	Same	22.9%	- combine the marbles from jar 1 and jar 2 together, formulate the probability of choosing a red marble using the combined result of both jars
3)	Bigger	22.9%	- forms correct probability of 1 jar, adds the probabilities of jars 1 and 2 together OR doubles the probability of 1 jar
4)	Smaller	10%	- forms correct probability of 1 jar, multiply the probabilities of jars 1 and 2 together (this is the correct method for the AND type probability question, but not OR probabilities)

Logic explanations 1) and 2) for OR operations (from Table 15) matched the logical explanations 1) and 3) for AND operations (from Table 14). It seemed that people who were unable to answer these questions correctly arrived at the same logic regardless of the probability operation being AND or OR type. Again, a large percentage of people arrived at an answer that was the same as the anchor for the OR operation. Even more surprising was that in the case of OR operations, not only would people ignore one of the jars, they also showed the tendency of combining the marbles from both jars into one. This also led to answer which was the same as the anchor. For both of these logics, calculating the probability became easier as one only had to formulate the probability of a single event. The common mistakes of the 'incorrect' group were also identified by people who were able to answer the problems correctly with 26.6% using explanation 1), 17.2% using explanation 2), 20.3% using explanation 3) and 9.4% using explanation 4).

Only 4.5% of the total people that answered OR probability questions of the people from the OR operation groups made mistakes in formulating the probability of choosing a red marble from a single jar. The logic behind incorrectly computing the single event probability was the same as the cases in the AND section (refer to Section 5.4.1) with the exception of b) (no one used the probability of drawing a red marble for 1 jar = 1/2 in the OR cases). Out of the people who answered incorrectly, only 7.1% of the people had logical explanations and performed mathematical operations that were close to the correct answer. These people either formed the incorrect probability of a single jar and performed the correct operation, or formed the correct probability of a single jar but erred in listing all the different cases of drawing a red marble. They would miss some of the cases or miss certain components in the cases. Table 16 illustrates some of these answers.

Table 16: 'Incorrect' Group with Partially Correct Logic

Marbles of 1 jar	**Explanation**	**Comments**
1 red, 4 blue	P(red marble for each jar) = 1/5 P(at least 1 red marble) = 1/5+1/5+1/5*1/5	- forms correct probability for 1 jar - have the right cases but did not include probability of drawing the blue marble from the other jar
3 red, 2 blue	P(1 red from 1st jar, 1 red from 2nd jar) + P(1 red from 2nd jar, blue from other) + P(1 red from 1st jar, blue from other) = 2/3*1/2+ 1/3*1/2 +1/3*1/2	- forms incorrect probability for each jar - lists the right cases and operations
3 red, 2 blue	P(success) = P(success)*P(failure) + P(success)*P(success) = 0.6*0.4 + 0.6*0.6	- forms correct probability for each jar - missing one of cases (1 red and 1 blue has be counted twice)
4 red, 1blue	P(at least 1 red marble) = 1 – P(getting blue marbles from both jars) = 1 – 1/4	- forms incorrect probability of getting blue marbles from both jars - have the right idea *Note: This person did not know how to compute basic AND case probability he could not calculate P(getting both blue marbles)

Another interesting observation was made from the results of how people were computing these probabilities. It was initially predicted that student would be able to determine the probability of choosing a red marble from one jar, and also be able to logically deduce the chance of choosing a red marble from both jars (probability lower than a one jar because it is harder to obtain a red marble from two jars – AND case) or choosing a red marble from at least one of the two jars, its probability higher than one jar because chances are better for two jars – OR case. Such reasoning was true for a few cases. There were four people who correctly computed the OR probability questions and reasoned the same way when they had to estimate how a person without prior probability training would solve the question. However, there was only one person that showed the same reasoning as the initial prediction out of the people who answered the question incorrectly. The reason provided was: S*ince I get to pick from 2 jars, the probability will be higher than that of a single jar*. More people might have adopted this reasoning from the list of people who answered the question incorrectly. This will be further discussed in Section 5.5: Changing Strategies in OR Operations.

It could be argued that a few people who answered the AND type probability questions also had a similar type of logic as the initial prediction. Of all the people that answered incorrectly for the AND type questions, 18.2% used a logic which *formed the correct probability of a single event, and then divided this probability by 2 to determine the final result* (refer to Table 11). Although the students did not explicitly state that the probability would decrease if they were to obtain a red marble from *both* jar 1 and jar 2, they might have divided the probability of choosing a red marble from the single jar by 2 because the answer '*made sense*'. They rationalized that the resulting probability of choosing a marble from both jars *should* be smaller than that of a single jar.

5.4.3 Discussion on Logical Thinking

Judging by the amount of people who had prior probability training and the number of people that answered the questions from the study incorrectly (refer to Section 5.2), it could be concluded that computation of probability even at a basic level was not a trivial matter. Even if students had received previous training in probabilities, the concepts that they had been taught were forgotten.

From analyzing the results from the experiments, it seemed that most students that were unable to answer the probability questions correctly fell under the following categories:

i) Formulate the correct probability for the single event, but use a wrong operation to combine the two probabilities

ii) Formulate the incorrect probability for the single event, but use the correct operation (or logic that is inline with the correct operation) to combine the two probabilities

A majority of students provided an answer based on the logic from i). By presenting students with two sets of probabilities, most of them recognized the probability for each set, but had trouble using the correct method to combine them. As stated earlier, instead of estimating an answer based on the single event, they performed various mathematical operations on the two sets of probabilities in an attempt to combine them. The computation of the probability was two folds:

1) Formation of events that occur

2) Computation of events: either by removing an event, or computing the events together (through mathematical operation, or reorganizing event space)

This computation was common in both AND and OR cases. However, while this was acceptable for AND type operations, it was unsuitable for OR type probability operations. Thus, from the analysis in the previous sections, it was seen that a lot more people were unable to provide the correct answer for the OR type probability problems compared to the AND type probability problems.

For AND type probability questions, most people were able to compute the answer. In fact, even if they formed the wrong probability of a single event, they still used the right operation to combine the events (refer to previous page, item ii). Section 5.4.1 illustrated that these students actually made up 22.7% of the people who answered incorrectly. The basis behind why AND type probability questions were easier may be due to the following:

- AND type probability operation involving independent events has only one case
- AND type probability operation is coded into a simple formula (multiplication of all the independent events)

Apparently the concept of *multiplying the probabilities of each independent event to obtain the combined probability of all these events* was something that most people retained from prior probability training. Therefore, whenever people encounter independent events and need to compute their combined probabilities, it was trivial as long as they were able to calculate the probability of each independent event. The concept of multiplying independent probabilities may be considered intuitive because this knowledge seemed to be deeply embedded in their minds. However, if one took a step further and asked people to explain *why* the probabilities are multiplied together, it was uncertain whether or not people would be able to answer this question. For example, when asked what the probability of flipping two heads from a fair coin was, this method provided a solution of 1/2 x 1/2 = 1/4. This equation does not require any understanding of how to form the set for the combined event and allows one to do away without having to list all the separate cases {HH, HT, TH, TT}. The two fold operation of how people calculate probability as described above would be enough to compute the correct answer for basic AND type probability problems.

1) Formation of events that occur – *forms probability of each jar*
2) Computation of events– *compute probability of events together with multiplication*

For OR type questions, there was no longer a "magical formula" that encapsulated the solution into a single calculation. However, it seemed that people with prior probability training had forgotten this fact as they tried to apply an equation (various forms of simple arithmetic such as addition, multiplication, etc.) to compute the probability of drawing out *at least* one red marble from the two jars in a single step. The aforementioned two fold operation caused error in calculating probabilities for OR type questions because in the second step of computation, they: *a) remove an event (by ignoring the probability of one of the jars altogether), b) combine events together with mathematical operation, c) combine events together by reorganizing the sample space (combing the marbles of both jars and determining probability from that single sample space).* The trouble with using such a method was that people were considering the events of each jar separately and tried to combine them in a single step. This method was suitable for AND operations because they only needed to combine the event of each jar *once* as it accounted for only a single case of obtaining red marbles (the case of getting a red marble from both jars). For OR operations, people needed to combine the events of the two jars *several times* to account for the different cases of drawing out a red marble from either jar 1 or jar 2, or both. In other words, instead of only considering the events of two jars separately, looking at them simultaneously would assist people in listing out all the possibilities. Johnson-Laird and colleagues (1999) proposed that people who were naïve in probability reasoning applied mental models of possibilities that accounted for the true cases with each model being equally probable. Using the *proportionality principle*, people would compute the probability of a certain event as the proportion of models where that specific event occurred. For example, consider the following problem given in the paper:

> *There is a box in which there is a black marble, or a red marble, or both. Given the preceding assertion, according to you, what is the probability that there is a black marble with or without another marble* (Johnson-Laird et al., 1999, p. 70)?

According to these authors' theory, such a question would result in the following mental model (Johnson-Laird et al., 1999, p. 70):

Black	-
-	*Red*
Black	*Red*

where a model with different possibilities was captured in each row. With the mental model theory described above, the probability of a black marble with or without another marble would be 2/3. One of the main criticisms of the mental model was the assumption of a "canonical representation of the sample space for a given problem" (Fox & Levav, 2004, p. 637). Usually, the format of many problems concerning probability evaluations would not present the problem in a method that would identify all the individual cases. Fox and Levav (2004) highlighted the argument made by Lagnado and Sloman (in press) that there was no claim to the assumption in identifying where the fixed representations of these models were emerging from. The current study for an OR type problem in this thesis was stated in the following format:

> *There are 2 jars of marbles. Each jar has [1] red marbles and [4] blue marble. If you were to pick 1 marble out of the first jar and 1 marble out of second jar, what are the chances of getting at least 1 red marble?*

The numbers in the square brackets varied in the study for the OR type questions. The different cases needed to compute the probability of drawing the red marble were not explicitly listed out for the students. It was presented using the words *at least* which in this situation, represented as i) red from jar 1, ii) red from jar 2, and iii) red from both jars. There was very little evidence that people were able to formulate the individual cases for solving OR type probability questions. For instance, it was seen in section 5.4.2 that only 7.1% of the people who answered the OR question incorrectly came up with the idea to list all the relevant cases to solve this problem. Unlike the problems presented by Johnson-Laird et al. (1999), in which the context encouraged partition, leading to certain probability judgment, the logic in this study was not explicitly laid out for the students. The finding was that a majority of students were unable to construct these partitions. They displayed a tendency of viewing the two jars as separate entities and then attempted to quickly eliminate this partition using various methods to combine them into a single partition. The problem was that students needed to first consider the events of the two jars simultaneously to be able to list out all the cases: {RR, RB, BR, BB} where R is probability of choosing a red marble and B is the probability of choosing a blue marble, and the position they were in represented which the jar they were in (i.e. RB = red from jar 1, blue from jar 2). Only after listing down all the cases would it then be appropriate to consider the probability of each jar separately and combine them to determine the probability for each case where a red occurred {RR, RB, BR}. Finally, the separate cases have to be added together to arrive at the probability of drawing at least one red marble out of the two jars.

The calculation of the OR type probability questions was more complex than AND type probability questions as it accounted for several cases in the sample space. Determining the answer required more than applying a simple mathematical operation between the probabilities of two events. However, there was a *short-cut* method of approaching this type of problem that permitted the use of a simpler equation. Asking for the probability of *at least* one red marble could also be viewed as:

1 – Probability(drawing blue marbles from both jars 1 and jar 2)

Looking at the problem this way was much easier as this single equation enabled the determination of the answer without having to list all the separate cases. Of the people who computed the OR type probability questions correctly, 63.4% used the above formula. Only one person out of the people who computed the answer incorrectly attempted to use this operation. The logic behind this formula still required a different partitioning mechanism than the two-fold operation that the ‘incorrect’ group often used. To formulate this equation, people still had to consider two jars simultaneously in drawing the conclusion that the answer was merely the whole sample space minus the case where no red marbles were drawn from jars 1 and 2. Viewing the two jars as separate events and combining them would not be enough. Thus, both the basic AND and OR type probability problems have simple formulas that could be used to determine the answer. Unlike the AND operation, the OR type operation required a more in-depth understanding of the logic and the partitioning of the event space in order to be able derive the equation. Another interesting finding about the shortcut method of OR operation was that calculating the probability of both marbles being blue from jars 1 and 2 was just the AND type operation.

5.4.4 Comments on Frequency Formatting

Girotto and Gonzalez (2001) argued that regardless of whether a question was presented in probabilistic or frequency format, naïve individuals were able to solve these problems if the form of the question and the structure of the problem were framed in a certain way. For example, they elicited that people faired much better in a *two-step* conditional probability question than a *single-step* question that corresponded to the form of a usual conditional probability question. The two-step question was structured such that the conditional problem forced people to first count the chance of P(E) and then count the chance of P(H&E) for solving the problem of P(H|E) where P(H|E) = P(H&E)/P(E). In the current study, the problems presented the information with natural numbers (i.e. 3 red marbles and 2 blue marbles) as opposed to using a probabilistic format (i.e. 0.6 of the marbles are red). However, despite using these natural numbers, people were still making errors in their computation of AND and OR operations. Therefore, it seemed the natural frequency representation was not effective in this case. People were only able to perform better on the frequency formatted questions given by Gigerenzer and Hoffrage (1995) because the problems were framed in a way that guided people into structuring them correctly.

5.5 Changing Strategies in OR Operations

It was established that people had little difficulty with AND type questions because a majority of them knew that calculating the probability of two independent events involved multiplying the two probabilities together. Looking at the data gathered from the study, people also applied a similar approach to calculate OR type operations. However, they used the mathematical operation of *addition* instead of *multiplication*. From the data, it was seen that for the lower anchor values (0.2, 0.4), many people who computed the problems incorrectly added the probabilities of choosing a red marble from jar 1 and choosing a red marble from jar 2. The same answer resulted from *doubling* the probability of a single jar as well. For the anchors of 0.2 and 0.4 for OR type questions, the largest number of people (21.4%) provided the answer of either adding or doubling the probability of choosing a marble from the single jar. Perhaps people added the probabilities of choosing the red marble from the two jars or doubled the probability of a single jar was because they reasoned the same way as the initial prediction of this thesis: *because one can either pick a red marble from jar 1 or jar 2, the probability will be higher than the probability of a single jar*. (Also refer to Section 5.4.2). Adding the probabilities of jar 1 and jar 2 or doubling the probability of a jar gave a reasonable answer because:

- The resulting probability was larger than that of a single jar

- The resulting probability was within reasonable bounds (i.e. not exceeding 1)
- Able to use simple mathematical operation to compute result

For the anchor of 0.2, the difference between adding the two events and the actual answer was only off by 0.04 (0.4 compared to 0.36). A concept that explained the reasoning behind doubling the probability of choosing a marble from one of the jars was the idea of linearity. Van Dooren, De Bock, Depaepe, Janssens and Verschaffel (2003) introduced the idea of the 'illusion of linearity' and applied it to probabilistic reasoning. They believed that "students have a strong tendency to apply linear or proportional models anywhere, even in situations where they are not applicable" (p. 113). For example, given that the probability of obtaining two ones in rolling two fair dice is 1/36, people reasoned that the two dice have to be thrown 18 times to have a probability of 1/2 to get two ones at least once (Van Dooren et al., 2003). In the current study, if people fell under the illusion of linearity, they would reason that drawing a red marble twice (one from each jar) would double their chances. Following this logic, if three jars were presented to them and the question asked for the probability of drawing at least one red marble out of the three jars, they would triple the probability of a single jar (assuming probabilities of all jars were the same) to arrive at the answer. One of the explanations provided by a student from the OR type probability problem with anchor of 0.4 was: *The probability of getting a red marble from one jar is 2/5 and we are picking form two jars so the probability is twice that, therefore the answer is 80%.* This particular student seemed to have fallen into the trap of linearity. However, those students who added the probabilities of the two jars together were not using a linearity approach. If the problem provided different probabilities for the two events, then the illusion of linearity would disappear. For example, if the probabilities of drawing a red marble from jar 1 and jar 2 were different, addition *P(jar 1) + P(jar 2)* would result in a different answer than doubling the probability of one of the jars, *P(jar 1)*2*. Upon further analysis, it was found that only 4 people out of the 15 who had the same answers (either from adding or doubling the probabilities) for the lower anchor used the doubling method as their reasoning. Although the illusion of linearity may be used to explain errors in certain problems, it did not seem to be applicable for the OR type problems.

A shift in people's strategy occurred for the higher anchor values of 0.6 and 0.8. For these anchor values, the probability of drawing a red marble from a single jar was either 0.6, or 0.8. If one followed the above logic and added the probabilities for the two jars or doubled probability of one jar, the result would exceed 1. A majority of people recognized that a probability of > 1 was not possible. Whereas a large number of people used the *addition* reasoning for the anchors of 0.2 and 0.4, only one person provided this reasoning and arrived at an answer greater than 1, when the anchors were 0.6 and 0.8. There seemed to be a shift in people's logic when the answer no longer made sense. For the anchors of 0.6 and 0.8 for OR type questions, most people arrived at answers that were the same as the anchors, which was achieved by the following two strategies:

1) Formulating the probability of choosing a red marble of 1 jar, ignore the other jar

2) Combine the marbles of the 2 jars together; compute the probability of choosing a red marble from the new combination

These two strategies were shown in Table 15: Common Mistakes for OR Operation. Most people adapted to using 1) for the higher anchor values. Quite a few of the people used the method shown in 2) as well. When the probabilistic situation got too complicated or created a puzzling answer such as in the OR type questions with higher anchor values, people avoided computation altogether. Instead of attempting to use mathematical operation to combine the probabilities of each jar, people used the single event logic that was less complicated in nature. This created a strong tendency to avoid using the probability from the second jar or to mix the marbles together to form a single event. For example, one person attempted to use addition to combine the two events from each jar but reverted back to the probability of choosing a red marble from a single jar when the answer exceeded 1 as seen by his explanation:

> *Each jar has 3 red and 2 blue marbles. Total outcome for getting a red marble out of the first jar is 3/5. Getting a red out of the second is also 3/5. But this would make the entire probability (3/5+3/5 = 6/5) > 1. That is not possible. I am confused here. Answer is 3/5.*

It should be noted that strategies 1) and 2) were also used by a number of people when the anchor values were 0.2 and 0.4, but more people used the *addition* approach. Figure 5 provides a graphical representation that shows the shift in the people's strategies from the changing anchor values.

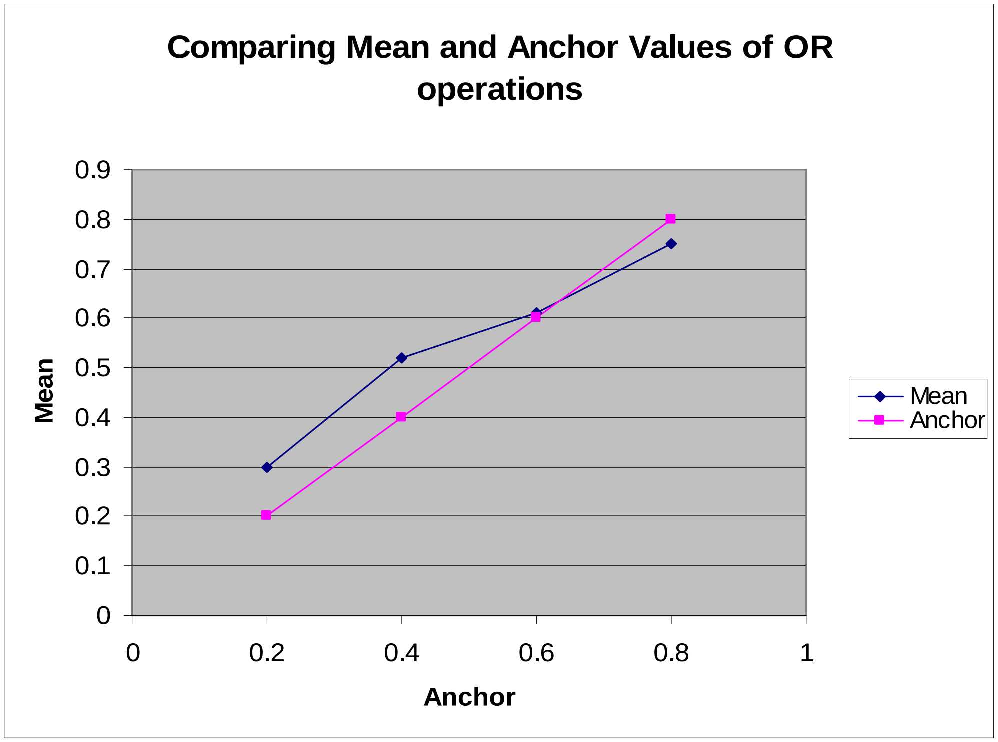

Figure 5: Graphical Representation of Strategy Shift

From Figure 5, it could be seen that for the lower anchors (0.2, 0.4), the means were at least 0.1 above the anchor values because people who were calculating their answers with the *addition reasoning* were pulling the mean higher. For larger anchors (0.6, 0.8), the means were a lot closer to the anchor. For the anchor value of 0.8, the mean was below the anchor because of a few exceptional cases that pulled the mean down.

6 General Discussions

Reiterating the prediction in this study, it was forecasted that people would be better at computing AND operations than OR operations. The reason for this was because OR operations are generally more complex to formulate and compute. In the real world however, there are various situations and decisions that involve OR scenarios. When considering several events at once, people generally want to determine the probability of the occurrence of *at least* one of the events rather than *all* of the events. Most people know that intuitively, the probability of at least one incident occurring is higher than the probability of all the incidents occurring. Therefore, they tend to align their thought process with decisions that yield a higher probability. Consider the situation of job hunting. When applying for jobs, a person will most likely judge his chances of getting at least one interview instead of judging his chances of getting interviews from all the jobs he applied for. Upon attending several interviews, one usually considers the chances of getting at least one offer rather than receiving an offer from all of the job interviews. There are many other OR type scenarios that people encounter in their daily lives. In academics for example, students tend to only study a subset of questions and hope the probability of at least one of the questions will be on the exam. In terms of finance, an angel investor would hope that at least one of the startup companies he had invested in would be successful.

At first, it may seem odd that people have difficulties or may demonstrate a lack of intuition with OR probability operations despite the degree of exposure they receive. However, it is important to consider what is actually occurring in people's thought processes when they encounter these OR type situations and whether these processes are in fact correct. From my experience, the following two facts are intuitive:

i) The probability of at least one event occurring is much higher than all events occurring

ii) The probability of at least one event occurring increases if the set of events itself increases (i.e. the more questions one studies for, the better the chance that one of them will be on the exam)

From the findings in this thesis however, when there is more than one event applied to OR type operations, people tend to view these events separately, rather than simultaneously, causing errors in their computation. Referring back to the job hunting example, how would a person process his chance of getting at least one job offer? He knows that the more interviews he has, the better the chance of getting a job offer. He may compute the probability of getting a job offer for each of the interviews separately based on how he conducted himself during the interview. The study in this thesis have shown that he would look at the probabilities of each interview individually to determine his chances of getting at least one job. That is, if he considers his chances of getting a job from three separate interviews are 0.1, 0.2, and 0.8, he may conclude that his chances of getting a job is 0.8, ignoring the probabilities of the other two interviews since they are lower. On the other hand, if the probability of getting a job from each interview are the same, say 0.2, he may conclude that his chances of getting a job is 0.6, the sum of all the probabilities! The data gathered from this thesis provided a good basis for understanding some simple heuristics that people apply to these problems. Despite encountering many OR type probabilities, the method people go about in estimating these probabilities are often incorrect.

7 Concluding Remarks

As stated in the beginning of the thesis, how a problem is framed is essential in conveying information and affecting how people understand a problem. In OR operations, it was seen that people have trouble with looking at events simultaneously, making the mistake of looking at them separately. This could be a result of one usually making comparisons between objects by looking at one object first, then the other, but not at the same time. It would be interesting to conduct further research to observe whether people will change their method of computation if their attention was directed to look at events simultaneously. This could be accomplished by changing the form of the questions in this study. Currently, the question provides the information for each jar and asks for the probability of a red marble after picking a marble out of each jar. The question may be altered to read 'reach both hands, one into each jar, to draw the two marbles' or by using a picture that conveys the same information (as shown in Figure 6). These strategies should enhance people's focus to consider the probabilities of the two jars simultaneously.

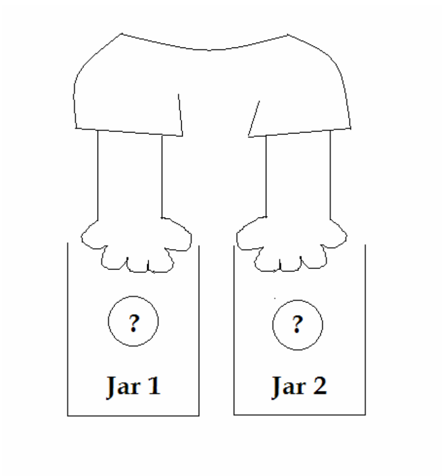

Figure 6: Drawing 2 Marbles Simultaneously

Studying how people calculate basic AND and OR probabilities brought about some fascinating findings, including people's intuitive ability to move in the correct direction from the anchor depending on the type (AND or OR) of problem presented; the ease of estimating other's common mistakes in calculating these problems for people who can compute the problems correctly; the tendency for people to only consider a single event when more than one event is given; and the switching of strategy that was employed when the probability of a single event of the OR type problem becomes too high. However, the study from this thesis only provided a glimpse into understanding people's logic in calculating these basic problems. Another approach in testing the directional hypothesis and the idea of making insufficient adjustments would be to conduct this study with a population that had no prior training in probabilities. The question could also be framed differently to eliminate the need for mathematical computations. For example, the anchor (probability of the single event) itself could be given to the participants and the question could simply ask the participants to indicate the direction of the combined probability of drawing marbles out of the two jars. A sample question with AND type operations is shown in Figure 7.

There are 2 jars of marbles. Each jar has 1 red marble and 4 blue marbles. The probability of getting a red marble out of one of the jars is 0.2 (1/5). This is shown on the line below. If you were to pick 1 marble out of the first jar and 1 marble out of second jar, what are the chances of getting a red marble out of both jars?

You do not need to compute the answer. Please mark a spot on the line below indicating the direction the answer would lie from the initial point and roughly how far it would be.

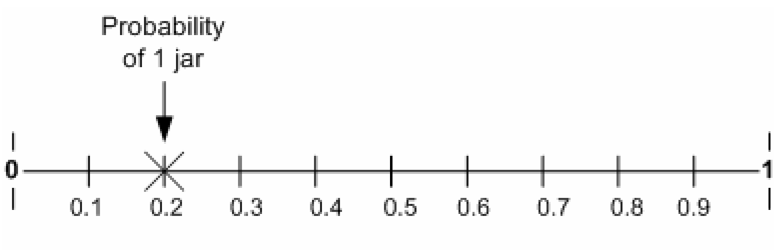

Figure 7: Sample Question of New Study

As indicated previously (Section 5.4 – Calculation Logic), this would be beneficial for studying people's intuition and how they handle basic probabilistic type questions. More analysis and studies will have to be conducted to further understand people's rationalities in calculating these probabilities to either build a better framework in developing cognitive models or to devise better methods to teach probabilistic reasoning.

The primary purpose of this thesis was to observe people's abilities to compute probabilities at a fundamental level using AND and OR operations. Encouraged by past research, attempts were made to construct probabilistic questions on an abstract mathematical level devoid of information irrelevant to the problem. People's logical approaches to these problems were analyzed on various levels. This allowed for a better understanding of people's logistical reasoning in calculating these fundamental probabilities and the common types of errors that were occurring. While few had difficulties with AND operations, many had trouble with OR operations. Although people's intuition provided a general guideline in these calculations, initial anchors on these problems seemed to affect people's answers. Given two or more events in a probabilistic question, more emphasis should to be placed in viewing these events simultaneously and constructing the different possible cases from these events. The ease at which the sophistically educated reverted back to the 'incorrect' methods of thinking could be used to proposed ways to devise problems in which OR type probabilities could become as intuitive as AND operations. This thesis revealed certain errors that people make in OR operations that could be applied to revise the current method of teaching probabilities at the basic level.

People already have so much trouble with OR operations without contextual information, it would be even more confusing to deal with the complex problems in a world full of context and information. Yet, it seemed that OR type probabilities are encountered frequently and are very applicable in our daily lives. In order to avoid making repeated mistakes in our decisions without being conscious of it, there's a need to devise better methods of teaching basic probabilities before advancing to the next level. This thesis provided a basis in developing better teaching methods in learning OR operations. Just as one cannot learn to leap before learning how to stand firmly, one cannot make certain decisions without having a good solid foundation of basic probabilistic concepts.

References

Ayton, P., & Hardman, D. (1997). Are two rationalities better than one? *Current Psychology of Cognition, 16*, 39-51.

Babai, R., Brecher, T., Stavy, R., & Tirosh, D. (2006). Intuitive interference in probabilistic reasoning. *International Journal of Science and Mathematics Education, 4*, 627-639.

Bar-Hillel, M., & Neter, E. (1993). How alike is it versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology, 65*, 1119-1131.

Borgida, E., & DeBono, K. (1985). *Hypothesis-testing and the role of expert knowledge.* Manuscript submitted for publication.

Bronowski, J. (1976). The ascent of man. London: British Broadcasting Corporation.

Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299*, 999-1000.

Chervany, N. L., Collier, R. D., Fienberg, S., & Johnson, P. (1977). A framework for the development of measurement instruments for evaluating the introductory statistics course. *American Statistician, 31*(1), 17-23.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*, 1-73.

Donaldson, M. (1978). Children's minds. London: Fontana Paperbacks.

Donaldson, M. (1993). Human minds: An exploration. New York: Viking Penguin.

Dulany, D. E., & Hilton, D. J. (1991). Conversational implicature, conscious representation, and the conjunction fallacy. *Social Cognition, 9,* 85-110.

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task. *Cognition, 77,* 1-79.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50,* 123-129.

Fischbein, E. (1999). Intuitions and schemata in mathematical reasoning. *Educational Studies in Mathematics, 38*, 11-50.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics, 15*, 1-24.

Fischbein, E., & Grossman, A. (1997). Schemata and intuitions in combinatorial reasoning. *Educational Studies in Mathematics, 34*, 27-47.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*, 96-105.

Fox, C. R., & Levav, J. (2004). Partition-edit-count: Naïve extensional reasoning in judgment of conditional probability. *Journal of Experimental Psychology: General, 133*(4), 626-642.

Fox, C. R., & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science, 14*, 195-200.

Garfield, J. (1981). An investigation of factors influencing student attainment of statistical competence. (Doctoral dissertation, University of Minnesota, 1981). *Dissertation Abstracts International, 42*, 5064A.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education*, *19,* 44-63.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684-704.

Gilovich. T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition, 78*, 247-276.

Hertwig, R., & Gigerenzer, G. (1995). *The chain of reasoning in the conjunction task.* Unpublished manuscript.

Hilton, D. J., & Slugoski, B.R. (2001). Conversational processes in reasoning and explanation. In *Blackwell handbook of social psychology. Vol. 1: Intraindividual processes,* (pp. 181-206). A. Tesser & N. Schwarz (Eds.). Oxford, U.K.: Blackwell.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J. (1999). Naïve probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62-88.

Kahneman, D., & Frederick., S. (2002). Representativeness revisited: Attribute substitution intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Kempthorne, O. (1980). The teaching of statistics: Content versus form. *American Statistician, 34*(1), 17-21.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education, 24,* 392-414.

Lagnado, D. A., & Sloman, S. A. (in press). Inside and outside probability judgment. In D. J. Koehler & N. Harvey (Eds.). *Blackwell handbook of judgment and decision making*. London: Blackwell.

Levi, I. (1983). Who commits the base rate fallacy? *Behavioral and Brain Sciences, 6*, 502-506.

Macchi, L., Osherson, D., & Krantz, D. H. (1999). A note on superadditive probability judgment. *Psychological Review, 106*, 210-214.

Margolis, H. (1987). *Patterns, thinking, and cognition.* Chicago: University of Chicago Press.

Messer, W. S., & Griggs, R. A. (1993). Another look at Linda. *Bullentin of the Psychonomic Society, 31*, 193-196.

Moldoveanu, M., & Langer, E. (2002). False memories of the future: A critique of the applications of probabilistic reasoning to the study of cognitive processes. *Psychological Review, 109,* 358-375.

Morier, K., & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin*, *10,* 243-252.

Nisbett, R., Krantz, D., Jepson, C., & Kunda, Z. (1983). The use of statistical thinking in everyday inductive reasoning. *Psychological Review, 90,* 339-363.

Noddings, N., Gilbert-MacMillan, K., & Lutz, S. (1980, April). *What does and individual gain in small group mathematical problem solving?* Paper presented at the meeting of the American Educational Research Association, Montreal.

Oaksford, M., & Chater, N. (2001). Human rationality and the psychology of reasoning: Where do we go from here? *British Journal of Psychology, 92*, 193-216.

Osherson, D., Shafir, E., & Smith, E. E. (1994). Extracting the coherent core of human probability judgement: a research program for cognitive psychology. *Cognition, 50*, 299-313.

Poile, R. C. (2004). Model validation in social science simulation. Master thesis, University of Waterloo.

Shimojo, S., & Ichikawa, S. (1989). Intuitive reasoning about probability: Theoretical and experimental analyses of the "problem of three prisoners". *Cognition, 32*, 1-24.

Stanovich, K. E. (2003). The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving. In J. E. Davidson & R.J. Sternberg (Eds.), *The psychology of problem solving* (pp. 291-342). New York: Cambridge University Press.

Stanovich, K. E., & West, R. F. (2002). Individual differences in reasoning: Implications for the rationality debate? In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1982). Judgments of and by representativenss. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293-315.

Tversky, A., & Koehler, D. J. (2002). Support theory: A nonextensional representation of subjective probability. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D., & Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics, 53*, 113-138.

Appendix A

Sample Survey Questionnaire – Set 1

Msci 211 Survey #1

The purpose of this survey is to study intuition and how people without any probability training estimate chance events. In this study, you will be asked to first solve a relatively simple probability question and then to estimate the response of a person without any knowledge in probability.

Please read each of the questions carefully. Answer each question to the best of your abilities. Your answer will not affect your mark in this course so please do the best you can without consulting other people or materials as that may affect the experiment. The use of a calculator is allowed for this experiment.

By completing this survey, you will receive 1% bonus mark for the course. Thank you for your time, your assistance is much appreciated.

1. Please specify your undergraduate program.

2. Have you ever had probability training before? If yes, please specify where you have received training (i.e. from a previous course in high school/university). It would also be helpful if you to write down the course name/number.

3. There are 2 jars of marbles. Each jar has 1 red marble and 4 blue marbles. If you were to pick 1 marble out of the first jar and 1 marble out of second jar, what are the chances of getting a red marble out of both jars?

4. Please give an explanation for your answer in Question 3.

5. How confident are you in your answer being correct in Question 3?

- Not certain at all
- Not very certain
- 50/50 (half a chance of being right, half a chance of being wrong)
- pretty certain
- 100% certain

6. Please do your best to guess how a person who has never had any probability theory training would answer Question 3 based on their impression or gut feeling.

7. Please give an explanation for your answer in Question 6.

Done

Appendix B

Sample Survey Questionnaire – Set 7

Msci 211 Survey #7

The purpose of this survey is to study intuition and how people without any probability training estimate chance events. In this study, you will be asked to first solve a relatively simple probability question and then to estimate the response of a person without any knowledge in probability.

Please read each of the questions carefully. Answer each question to the best of your abilities. Your answer will not affect your mark in this course so please do the best you can without consulting other people or materials as that may affect the experiment. The use of a calculator is allowed for this experiment.

By completing this survey, you will receive 1% bonus mark for the course. Thank you for your time, your assistance is much appreciated.

1. Please specify your undergraduate program.

2. Have you ever had probability training before? If yes, please specify where you have received training (i.e. from a previous course in high school/university). It would also be helpful if you to write down the course name/number.

3. There are 2 jars of marbles. Each jar has 3 red marbles and 2 blue marbles. If you were to pick 1 marble out of the first jar and 1 marble out of second jar, what are the chances of getting at least 1 red marble?

4. Please give an explanation for your answer in Question 3.

5. How confident are you in your answer being correct in Question 3?

- Not certain at all
- Not very certain
- 50/50 (half a chance of being right, half a chance of being wrong)
- pretty certain
- 100% certain

6. Please do your best to guess how a person who has never had any probability theory training would answer Question 3 based on their impression or gut feeling.

7. Please give an explanation for your answer in Question 6.

Done