Author: K. B. van Boven (5930189)
Supervisor: Dr. R. W. F. Nouwen
Second reviewer: Dr. S. F. Donker
15 ECTS

27-01-2020

THE USE OF SEMANTIC MODIFIERS IN THE ASIAN DISEASE PROBLEM

HOW SMALL WORDS COULD HAVE A BIG IMPACT ON DECISION-MAKING

Universiteit Utrecht

BACHELOR THESIS UNIVERSITY UTRECHT
LIBERAL ARTS AND SCIENCES
MAJOR COGNITIVE AND NEUROBIOLOGICAL PSYCHOLOGY
FACULTY HUMANITIES

Table of contents

1. Introduction

Throughout every day of their entire lives, people are faced with an enormous amount of decisions. Be it little decisions like what drink to get with their breakfast or decisions with bigger implications, like whether or not a job offer is accepted, decision-making is a big part of the human life. Human decision-making has been studied intensely across many different disciplines, from behavioural psychology and sociology to mathematics and economics. The abundance of interest and research regarding this subject has not led to a consensus across all disciplines: there are countless theories, accounts and opinions on how decisions are made, many of which are unable to co-exist. Within the field of cognitive (neuro)psychology, decision-making has been studied intensely for years (see Crozier, Ranyard, & Svenson (1997) for an overview). Decision-making is an important part of cognition and, like with any cognitive task, theorized to mainly be executed in the brain. Therefore, it is a worthy subject within the cognitive neuropsychology. Having an idea about how decisions are made by humans is also of major significance in the study of artificial intelligence and modelling machines that are able to mimic this process. Learning about human decision-making and how to remake it in artificial intelligence gives way for many opportunities to make not only more accurate predictions for human behaviour, but also for ‘smarter’ and more advanced artificial intelligence.

One of the most interesting fields within human decision-making is the decision people make when it is not clear which options are more beneficial for them, called risky choices. Specifically, this research focuses on risky choices that are presented to people through written language. The written risky choice experiment used to test this is the classic Asian Disease Problem from Tversky & Kahneman (1981), a modified version of which is used in the current research. More information on the origin and details of this test will be discussed in the next section.

Influencing the decisions that are made in these risky choices is possible through different ways of stating the options or choices for people. The fact that decisions can so seemingly easily be influenced through language shows the importance of linguistic nuances and researches such as the current one. One of the most well-known and used theories on the topic of influencing decisions that people make is the so-called standard framing effect. The framing effect shows that the way options are formulated can influence how people make decisions. In short, this means that positive framing (stating gains) tends to lead to risk-avoidance and negative framing (stating losses) to risk-seeking. As discussed in more detail below, even just adding small modifying words can affect which decisions people are more likely to make.

For instance, Mandel (2001) stated that adding small words (i.e. semantic modifiers or slack regulators) is thought to make a significant difference in influencing decisions that people make, but the exact role that these modifiers play is still a large topic of discussion. It is theorized that using either a lower-bounded modifier such as *'at least'* or, conversely, an upper-bounded modifier like *'at most'* leads to a different interpretation of an otherwise identical sentence. *'At least'* followed by a numeric quantifier (a number, i.e. 200) often leads to an interpretation of *'200 or more'*, whereas *'at most'* has an adverse effect and will often lead to an interpretation like *'200 or less'* (Geurts & Nouwen, 2007; Coppock & Brochhagen, 2013; Blok, 2015). Once again, more detailed explanation on this will follow in section **2.3. Modifiers**.

The goal of this present research is to further examine the role of sematic modifiers, in particular the role *'at least'*, *'exactly'* and *'approximately'* play (the latter two are also known as slack regulators). These three modifiers have been chosen to, on one hand, be able to compare results with the existing framework within the field, but also to compare the influence of the boundedness of a modifier (lower vs. bilateral) and the acceptable range implied by the modifier (wide vs. precise). To reach this goal, the following research question has been formulated: How do the presence of semantic modifiers and slack regulators influence decisions made in the Asian Disease Problem? Another question is how and if the semantic modifiers and slack regulators interfere with the standard framing effect, an effect that will be tested on its' own as well. To come to an answer, Mandel's 2013 study (refer to section **2. Previous research** for more in-depth information) has been replicated and modified. This has been done because of the importance of this particular experiment and the ability to easily modify it to fit for the questions that are attempted to be answered in the current research.

The standard framing effect, along with the details of the Asian Disease Problem and background information leading up to the present research, will first be discussed. Afterwards, existing research on the role of semantic modifiers and slack regulators in risky choices such as the Asian Disease Problem will be put forward. Then, the methodology of the present research and how it has been constructed will be stated. Before going to the results, hypotheses will be constructed based on the theoretical background on the subject. The results of this research along with its' strong points and shortcomings will be stated and compared with the hypotheses, to see whether or not the expectations were accurate. Finally, the implications of the results for both the existing theoretical body of work and future research will be stated.

2. Previous research

2.1. Risky choices

One of the oldest and most influential theories on human decision-making is the expected-utility theory, originating from Bernoulli's analysis of the St. Petersburg paradox (1738), which serves as a descriptive model for human behaviour (Karmarkar, 1978). The expected-utility theory states that individuals rank different potential outcomes on the expected utility that they are thought to have. The rational decision made is the one that the has the largest utility for the individual (Fishburn, 1970; Machina, 2008). The value of this theory and its' models in the realm of economic decision-making has been great and plenty of research has been done on the back of the original theory. Friedman and Savage (1948) showed that consumers pick the alternative with the largest utility if there is no factor of risk involved, whereas when risk is involved, some different variables appear to enter the equation. Individuals weigh the risk they perceive to be there against the expected utility they foresee getting from making certain decisions. The expected-utility theory is, despite its' undeniable influence, not scientifically falsifiable. Expected utility is a subjective measure that is personal and different to every individual and therefore immeasurable. Possible scientific work regarding risky choices therefore mainly involves manipulation of the way choices are presented, which can be a way to influence the decision individuals are more likely to make.

2.2. Framing effect

One of the most interesting, well-documented and relevant ways to influence decisions that individuals make without actually altering their options is framing. Framing refers to the way in which a message is framed and how this affects the perception of this message, and it has been shown to be effective across many fields. Framing refers to either station options in terms of gains (positive framing) or in terms of losses (negative framing). Objectively equivalent but varying descriptions of the same problem lead to different responses in decision-making, a result shown in domains such as (but not limited to) cognition, psycholinguistics and business (Levin, Schneider, & Gaeth, 1998). One of the most influential experiments with regards to framing was first constructed by Tversky and Kahneman (1981) and repeated many times after: the Asian Disease Problem (ADP). In the ADP, the context given to participants is as followed:

> Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the program are as follows.

After reading this problem, individuals get presented one of two different sets of programs. One group (consisting of half the subjects) are shown the following:

> If Program A is adopted 200 people will be saved.
>
> If Program B is adopted there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.
>
> Which of the two programs would you favor? (Tversky and Kahneman ,1981, p. 453).

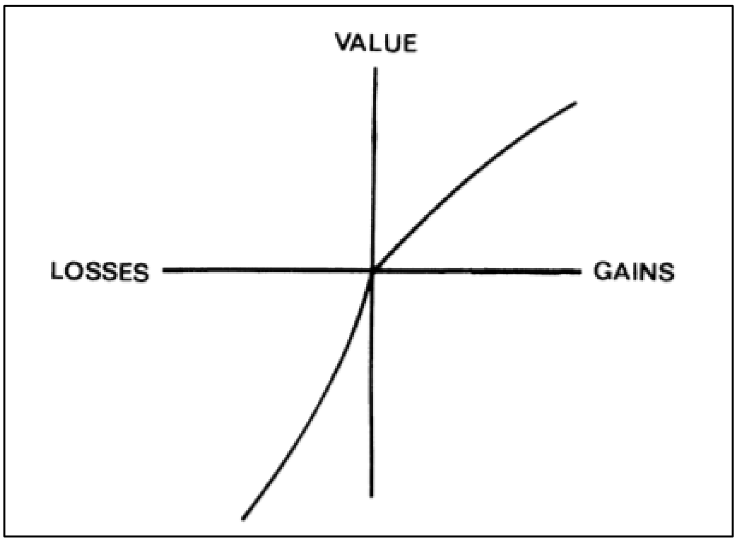
The other group (consisting of the other half of the subjects) are shown these options:

> If Program C is adopted 400 people will die.
>
> If Program D is adopted there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.
>
> Which of the two programs would you favor? (Tversky and Kahneman, 1981, p. 453).

The only obvious difference between the problems is the way that outcomes are described, using number of lives saved (positive frame) in program A&B and number of lives lost (negative frame) in program C&D. Programs A and C as well as programs B and D are objectively equivalent in their outcome, but nevertheless program A was shown to be chosen 72% of the time whereas program C was only picked in 22% of the instances (Tversky & Kahneman, 1981). This shows a large effect of the way the programs were framed on the decisions that were made. This difference between the ratio in which programs A and B and programs C and D were picked showed a large problem with the classical utility theory of decision-making, where effectively equivalent options are proposed to lead to identical decisions (Geurts, 2013). This phenomenon, where "*different, but logically equivalent phrases cause individuals to alter their preferences (…) typically involves casting the same information in either a positive or a negative light* (Druckman, 2004, p.671), is called the framing effect. A standard risky-choice framing effect occurs when risks are avoided in positive frames and risks are taken in negative frames.

Participants showed to be risk-averse when dealing with gains and risk-seeking when losses were emphasized (Lakshminarayanan, Chen, & Santos, 2010). This effect has been observed outside of contained experiments: there is a reluctancy to sell declined assets and a tendency to sell when assets have appreciated on the stock-market (Odean, 1998). Furthermore, house sellers tend to leave their houses on sale longer when the market price is below their buying price, where houses with a market price higher than the market price tend to get sold quicker (Genesove & Mayer, 2001). This translates to framing in the ADP because secure options (selling assets or houses) are chosen more frequently when the frame is positive and gains are emphasised (appreciated assets or a high price compared to the market), whereas risky choices (keeping assets or houses) are chosen more frequently when the frame is negative and losses are emphasized (declined assets and a relatively low price compared to the market). The fact that the asymmetric risk-preferences have been replicated in different high-stake scenarios suggests that this phenomenon is present in important settings in real-life scenarios. It is therefore an important field of research for explaining, understanding and influencing human decision-making.

In an attempt to explain the inconsistencies with the popular expected utility theory and the difference in choices between the positive and negative frames, a new theory was formed. Multiple effects are thought to be inconsistent with the utility theory: people tend to underestimate probable outcomes in comparison with certain outcomes, which Kahneman and Tversky (1979) call the certainty effect. This certainty effect leads to risk aversion in choices with sure gains and risk seeking in choices with sure losses, consistent with the aforementioned studies and theories. Furthermore, people tend to disregard properties that are shared by all potential options, leading to inconsistent decisions when the same choice is presented in different forms. This is called the isolation effect (Kahneman & Tversky, 1979). Both of these effects are taken into consideration when making a



descriptive model of real-life choices called the prospect theory (see *Figure 1*.). Instead of assigning value to the final state, it is assigned to relative gains and losses, and probabilities are replaced with decision weights. Decision weights make sure that the subjective probabilities used in the utility theory, which slant outcomes in the direction of safe decisions, are replaced (Whitaker & Fellner, 1966).

In the model, the losses are steeper than the gains, which means that a relatively smaller loss can account for a higher absolute value for an individual than a bigger gain can do. The figure is concave for gains and convex for losses, meaning that the gain or loss in value is relatively bigger with respectively smaller gains or losses. In the context of the ADP, the theory predicts higher absolute value in negative frames than in positive frames, therefore expecting relatively more risk-avoidance in negative frames than risk-seeking in positive frames. This expectation will be tested in the present experiment by comparing choices made in negative frames to those in positive frames. If relatively more people choose the uncertain option in the positive frame than they do certain options in the negative frame, then the results are consistent with this Prospect Theory Model.

The notion that people violate rationality by choosing differently between extensionally equivalent options is based on the assumptions that programs A and C are equivalent (Geurts, 2013; Mandel, 2013). However, Frisch (1993) showed that 41% of the participants that demonstrated the framing effect did not agree that program A (*"200 people will be saved")* is identical to program C (*"400 people will die"*). Whereas the risky programs (B & D) are completely specified, the sure options are ambiguous. Participants might think that at least some of the 400 people unmentioned in program A have a chance of survival and some of the 200 people unmentioned in program C may not be saved (Mandel, 2001). To test this notion, Kühberger (1995) made an unambiguous version of the ADP, where (for instance) participants will read that 200 people will be saved and 400 people will not be saved in program A. In the ambiguous version, the findings of Tversky and Kahneman (1981) were replicated, but the framing effect was not found when the ambiguity in the programs was resolved. The ambiguity of numeric quantifiers is a topic of controversy within the field of risky choices, a controversy that the present research attempts to make clearer.

2.3. Modifiers

Numeric quantifiers, such as 200 and 400 in the ADP, are not judged as being exact quantities by most readers. Multiple theories based on the semantics of numeric quantifiers have been created, which have significant differences between them. Some claim that numeric quantifiers have lower-bound semantics, meaning individuals read '200 people' as being '*at least* 200 people'. In this set of beliefs, bilateral interpretation is thought to arise by coupling the natural lower-bound semantics of the numeric quantifier with an upper-bound modifier such as *'at most'* (Horn, 1992). This theory is however highly controversial, since people tend to interpret numeric quantifiers coupled with upper-bound modifiers as being a maximum value, where *'at most 200 people'* is interpreted as *'200 people or less'* (Blok, 2015). Another theory states that numeric quantifiers have bilateral semantics naturally, but they could be interpreted as unilateral depending on pragmatic inferences (Breheny, 2008), meaning context is used to understand the meaning beyond the literal interpretation (Spotorno et al., 2015). Despite their differences, all disagree with the notion that people always interpret numeric quantifiers as their exact values, confirming the ambiguity present in the original ADP.

The readers' interpretation of numeric quantifiers can be influenced by the use of scalar modifiers such as '*at least', 'at most',* or *'more than'*. The sentence "*Fred had at least three beers"* is extensionally equivalent to "*Fred had more than two beers"* on the first look, both sentences state that the amount of beers that Fred had was three or more. However, readers will likely expect different amounts of alcohol to be consumed when reading the first sentence compared to the second one, due to the varying implications that the different modifiers possess (Geurts & Nouwen, 2007; Coppock & Brochhagen, 2013). Another way to modify the interpretations of the reader is by controlling the amount of accepted deviation from the numeric quantifier used.

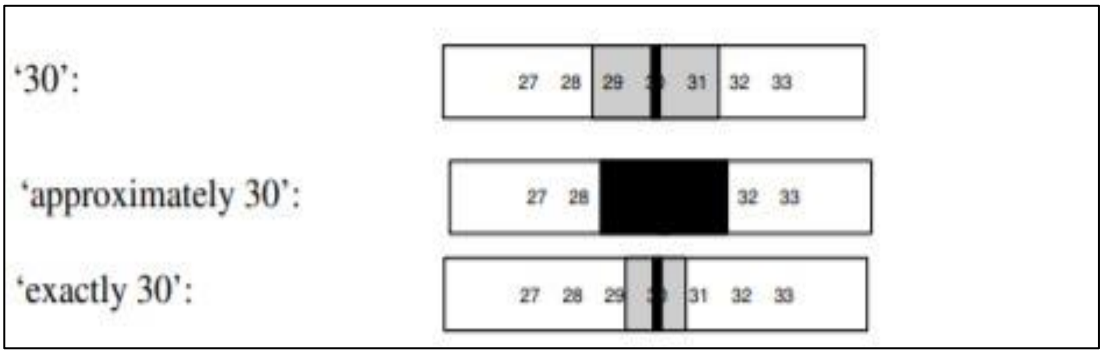


Figure 2. Acceptable values with and without slack regulators (Sauerland & Stateva, 2007, p.236).

This can be done using what Lasersohn (1999) calls slack regulators, words that "*help determine just how close to the truth is 'close enough'"* (p.525). The range in which *'close enough'* applies is called the pragmatic halo. Two of the most commonly used examples of slack regulators are '*exactly'* and '*approximately'.* Whereas exactly implies that the only values very close to the numeric value are acceptable, approximately does the opposite and widens the range of acceptable values (see *Figure 2.*). In this image, the black area marks values that are seen as true, the grey area marks the area the pragmatic halo and the white area marks false values.

The use of these modifiers can have significant influence on the decision-making process, as Mandel showed in his 2013 research. He modified the ADP using the terms *'at least'* and *'exactly'.* The first experiment showed that individuals exposed to the lower-bounded modifier *'at least'* were more likely to exhibit the standard framing effect, in which certainty is more preferred in the positive frame than the negative frame, than individuals exposed to bilateral slack regulator *'exactly'.* The group exposed to *'exactly'* made consistent choices more than twice as often as the group exposed to *'at least'* (73,3% compared to 31,7%, p.5). In a follow-up, a third group was added that was not exposed to any modifiers. The results showed that without modification, a medium framing effect was found whereas a lower-bounded, unilateral modifier (*at least*) made the framing effect larger and a bilateral modifier (*exactly*) made the framing effect nonsignificant. The size of the framing effect refers to how large the tendency was to pick risk-seeking in the negative frame when compared to the positive frame: a large effect means a lot more people tend to pick the certain option in the positive frame and the uncertain option in the negative frame than is the case with a medium and especially a small framing effect. Another experiment was ran with modifiers *'at least', 'exactly,* and *'at most',* which showed that framing effects were non-significant in the *'exactly'* group. Interesting about this research is that *'at least'* showed significant positive framing effects in the interpretation of the certain option and slightly negative framing effects in the uncertain option, whereas *'at most'* showed a reverse pattern (see *Figure 3*). This shows the large influence that modifiers can have on the interpretation of numeric quantifiers, an influence that this research is trying to clarify.

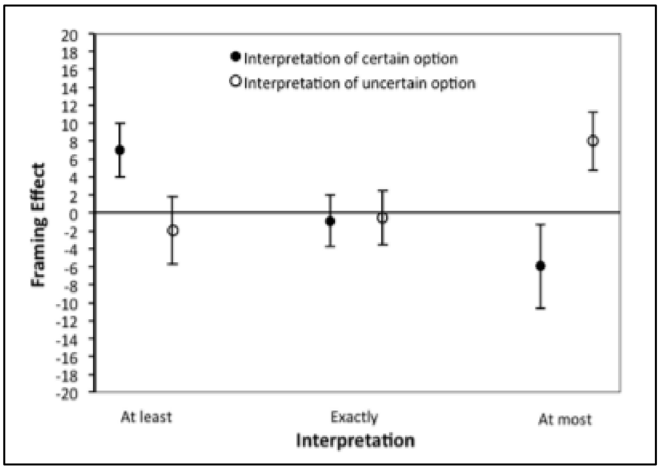


Figure 3. Framing effects with different modifiers (Mandel, 2013, p.8).

Plenty of research has been done in the past decennia to examine the standard framing effect in risky choices. Almost 40 years after the original ADP, it has been replicated numerous times with different variations throughout the years. The role that semantic modifiers play in affecting the standard framing effect and the decisions made in the ADP is something that has more recently become a topic of discussion, a discussion to which this research is set out to contribute. By slightly modifying both the contexts used and the modifiers added in Mandel's 2013 research, we attempt to answer the research question: How do the presence of semantic modifiers and slack regulators influence decisions made in the Asian Disease Problem? The modifiers used are the lower-bounded unilateral *'at least'*, the bilateral precise *'exactly'* and the bilateral vague *'approximately'*. The first two were also present in the Mandel (2013) research, and the aim is to see if the findings on those modifiers can be replicated. The modifier *'approximately'* was added to be able to compare two bilateral modifiers that have a converse effect on the pragmatic halo; *'exactly'* is thought to shrink this halo and *'approximately'* is theorized to broaden it (Lasersohn (1999); see *Figure 2*).

3. Methodology

3.1. Participants

Using Amazon Mechanical Turk (MTurk), a crowdsourcing marketplace, 61 participants were recruited and paid to complete the experiment. Only their primary language is asked, age and gender are unknown and unimportant for this research. Participants that did not answer English as their primary language have been excluded from the results to avoid having language-proficiency level as an unwanted variable, leaving the sample with 60 participants. There were also four filler questions per person to make sure individuals take the experiment seriously and read all the questions. For instance, one of these fillers was constructed like this:

> Due to an open gate, a farm lost 500 chickens. Two plans can be chosen to retrieve them.
>
> If plan A is chosen, it is certain that 20 chickens will be returned safely
>
> If plan B is chosen, there is a 95% possibility that all chickens will be returned safely and a 5% possibility that none of the chickens will return
>
> Which plan would you choose?

The answer with the highest utility here is clearly plan B. A score lower than 75% on these four fillers lead to exclusion. As a result, five participants were excluded from the sample, which left the sample at 55 participants. All filler questions were extremely similar to the example above and intentionally made to look like the target items, to minimize the chance that subjects figured out what the intention of the experiment was while making it. For the complete list of filler questions, refer to **Appendix 1**.

3.2. Design

To tackle the research questions stated in the introduction, an experiment has been constructed based on the original ADP (Tversky & Kahneman, 1981) and Mandel's (2013) modified version of the same problem. Three different variations of the ADP have been used for this research, each of which consists of the lives of 600 people being threatened by either natural disaster, illness or a hostage situation during a war. The reason for constructing two additional contexts, besides the disease-ridden region used in Mandel's 2013 research, is so that subjects can be exposed to multiple risky choices without literally repeating the same questions to them. This minimizes the chance that people will just gloss over a question because it seems familiar and maximizes the amount of observations we are able to get from each subject. The full contexts and options presented to the subjects can be found in **Appendix 2.**

After being shown the context, participants are then shown two options, one of which they need to choose: option A which will certainly secure 200 people and option B which has a 1/3 chance of saving everyone and a 2/3 chance of having everyone die. There are two variables: the polarity and the modifier that is present in the options. Polarity can either be positive (i.e. *200 people will be saved*) or negative (i.e. *400 people will die*). There are three different modifiers, a lower-bound unilateral semantic modifier *at least* and the bilateral slack regulators *exactly* and *approximately.* The options presented to subjects will look as follows (the example shows a positive frame, for the full research refer to **Appendix 2**):

If option A is chosen, it is certain that [exactly/at least/approximately] 200 people will survive

If option B is chosen, there is a 1/3 possibility that all 600 people will survive and a 2/3 possibility that nobody will survive

Every subject sees every version of the ADP twice in different polarities, interchanged with the fillers that have an obvious answer with the highest utility. Every ADP with the same quantifier is shown in both a positive and a negative frame to make comparing reactions to the same quantifier in different frames possible. This ensures that every subject contributes one observation for every condition ('*at least', 'exactly'* and *'approximately'*). The results are analysed to see whether there is an effect of the semantic modifiers and slack regulators on decision-making in risky choices. The existing framework on the standard framing effect will be tested along with the role that modifiers and slack regulators play.

4. Predictions

The semantic modifier *'at least'* is expected to increase the size of the framing effect, which means there will be a relatively larger abundance for uncertain options in the negative frame and certain options in the positive frame when compared to what the results would be without *'at least'*. Because of the lower-bounded nature of the modifier, people have the opportunity to mentally make the number of lives saved in certain option in the positive frame bigger than the numeric quantifier 200, whereas the uncertain option is unaffected. The negative frame however leaves a maximum of 200 people saved in the certain option, whereas the uncertain option gives an expectation of saving precisely 200 lives (1/3 chance to save 600 people). Furthermore, earlier research by Mandel (2013) showed that *'at least'* stimulated the standard framing effect, which is expected to be found again in this experiment.

The slack regulator *'exactly'* is expected to show the opposite pattern: including this word in the options is thought to work against the standard framing effect and promote consistent choices. This is because it's bilateral and its' precise nature does not leave much room for individual interpretation (see also *Figure 2*), therefore making the positive and negative frame extensionally equivalent without any ambiguity. This is based on the description invariance principle, which states that an individual's choices among a set of options should not vary because of different framing when the different frames describe identical options (Arrow, 1982; Tversky & Kahneman, 1986), as is the case with slack regulator *'exactly'*. Once again, Mandel has tested and confirmed this expectation in his 2013 research, and it is expected that similar results will come out of this experiment.

The last variable is the slack regulator *'approximately'*, which is similar to *'exactly'* but leaves more room for personal interpretation. Because of its vagueness and therefore wide pragmatic halo (Lasersohn, 1999), it is expected to promote the standard framing effect among choice-makers. It is however bilateral, which means that personal interpretation can make options look more appealing or more gruesome in both frames. Therefore, the expectation is that there is more standard choice framing when compared to *'exactly'*, but less than *'at least'*, since the latter can only make the certain option in the positive frame more positive and the same option in the negative frame more negative.

5. Results

As can be seen in *Table 1.*, people were more likely to choose the certain option in the positive frame than in the negative frame (66,7% of all choices within the positive frame versus 42,3% of all choices within the negative frame). To evaluate whether or not the response pattern was related to the polarization, a Pearson's chi-square test of contingencies (with $\alpha = .05$) was conducted. The chi-square test was statistically significant, $\chi^2(1, N = 336) = 20.17$, $p < .001$, although the association between distribution of choices and polarization was quite small, $\phi = .25$.

	CERTAIN OPTION	(%)	**UNCERTAIN OPTION**	(%)
NEGATIVE FRAME	71	42.3	97	57.7
POSITIVE FRAME	112	66,7	56	33.3
TOTAL	183	54.5	153	45.5

Table 1. Distribution of choices between frames

To analyse the observed effects of the modifiers, the certain options were coded as a score of 1 and uncertain options as -1 in the target items, replicating Mandel's 2013 analysis. Framing effects were then measured by subtracting the value of the choice in the negative frame from the choice in the positive frame. A value of 2 reveals a standard framing effect, whereas 0 signals consistent choices and -2 represents reversed framing effect. A higher mean score therefore translates to more occurrences of the standard framing effect. The mean score for choices with modifier '*at least'* was .3636 (95% CI [.0866, .6407]), modifier '*approximately'* showed a mean of .4727 (95% CI [.1981, .7473]) and choices with modifier *'exactly'* showed a mean score of .6909 (95% CI [.4314, .9504]).

The observed effects per modifier are shown as percentages in *Table 2.* and as counts in *Figure 4.*. As the table shows, the standard framing effect was more common with bilateral modifiers *'approximately'* and *'exactly'* than with unilateral lower-bounded modifier *'at least'*. To test whether the modifiers had a significant effect on the framing effects shown (or not shown), a Pearson's chi-square test of contingencies (with $\alpha = .05$) was conducted. The chi-square test was not statistically significant, $\chi^2(4, N=165) = 4.12$, $p = 0.39$.

	AT LEAST (%) [95% CI]	**APPROXIMATELY (%) [95% CI]**	**EXACTLY (%) [95% CI]**	**OVERALL (%) [95% CI]**
STANDARD FRAMING EFFECT	23.6 [12.4, 34.8]	27.3 [15.5, 39.1]	34.5 [21.9, 47.1]	28.5 [21.6, 35.4]
CONSISTENT CHOICES	71.0 [59.0, 83.0]	69.1 [56.9, 81.3]	65.4 [52.8, 78.0]	68.5 [61.4, 75.6]
REVERSE FRAMING EFFECT	5.5 [0.0, 11.5]	3.6 [0.0, 8.5]	0.0 [0.0, 0.0]	3.0 [0.4, 5.6]

Table 2. Observed effects per modifier

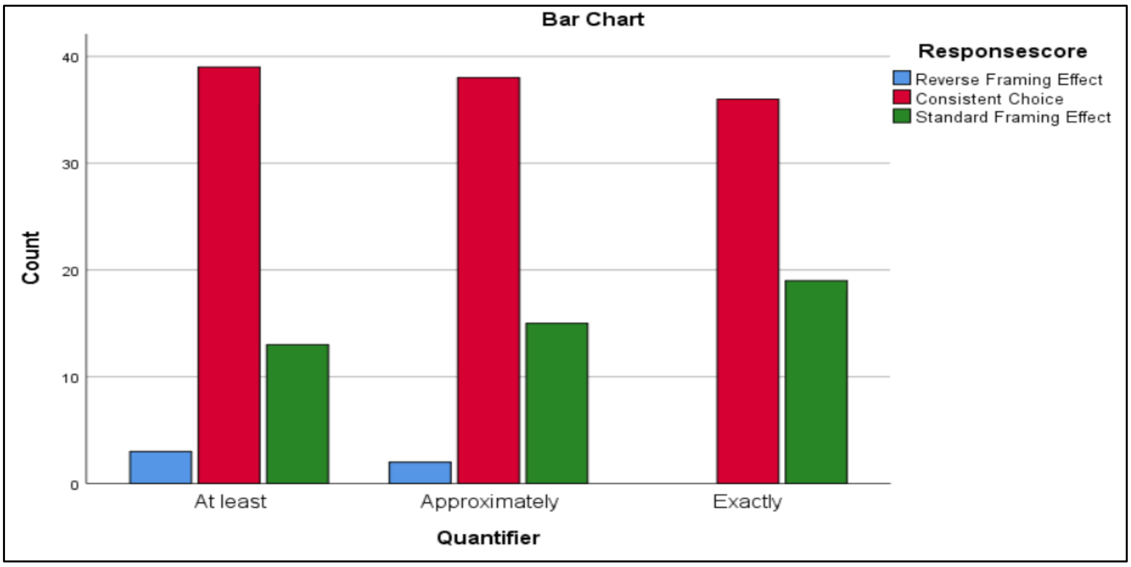


Figure 4. Counts of observed effects per modifier

6. Discussion

As mentioned beforehand, one of the main goals of this experiment was to attempt to replicate earlier findings (Lakshminarayanan, Chen, & Santos, 2010; Mandel, 2013; Tversky & Kahneman, 1981; see section **2. Previous research**) regarding risk-aversion in the negative frame and risk-seeking in the positive frame, also known as the standard framing effect in risky choices. The effect of the polarization of the options, whether results were expressed in lives saved (positive frame) or lives lost (negative frame), on the choices made (certain or uncertain) was analysed. As can be concluded from *Table 1.*, the standard framing effect that was shown in previous research was replicated in the current experiment. This means that people were more likely to exhibit risk-seeking when options are framed negatively and more risk-averse when they were framed positively. The standard framing effect did show to be present in this research with statistical significance, which proves consistency with the replicated Mandel study (2013). The strength of the effect however was considerably weaker than the experiment it was based on. The ϕ (mean square contingency coefficient) of the current research was 0.25, whereas Mandel (2013) came to a much stronger $\phi = 0.45$. This means that despite being consistent with the Mandel research, the current research showed to be less sensitive than its' predecessor. The reasons for the difference between both studies will be further examined in the current section.

People were relatively more likely to pick risk-avoidance (certain options A or C) in the positive frame than they were risk-seeking (picking uncertain options B or D) in the negative frame (see *Table 1.*). This contradicts the Prospect Theory Model (Tversky & Kahneman, 1979), which predicts higher absolute value in negative frames than in positive frames. Based on this theory, it is to be expected that people will be more risk-avoiding in the negative frame than risk-seeking in the positive frame, an expectation not met by the current research. The table also shows a majority of certain options were picked over uncertain choices, consistent with the certainty effect of Tversky and Kahneman (1979). People tend to underestimate probable outcomes in comparison with certain outcomes according to the certainty effect, which could be the reason why more decisions fell towards the certain option instead of the uncertain options. However, this could also be caused by other variables such as the different modifiers used.

The second goal of this experiment was to test the influence that the modifiers *'at least'*, *'exactly'* and *'approximately'* might have on the framing effect and decision-making in risky choice situations. Contrasting with Mandel's 2013 experiment, no significant effect of the modifiers was found on the framing effects shown. Albeit non-significant, the fact that more standard framing effects where shown in the *'exactly'* condition compared to those in the *'at least'* condition is surprising when comparing it to Mandel's 2013 research (see *Table 2*. and *Figure 4.*). We also calculated the average score of the different quantifiers, where a higher score corresponds with a bigger tendency towards the framing effect. The condition *'exactly'* (M = 0.6909) shows to be more sensitive to the framing effect than the *'approximately*' condition (M = 0.4727) and especially the *'at least'* condition (M = 0.3636).

As mentioned in the section **4. Predictions**, based on Mandel's results (66.7% framing effects in the *'at least'* condition and 21.7% in the *'exactly'* condition), a reverse pattern than shown in this experiment was expected. The expectation that more framing effects would be shown in the *'at least'* condition when compared to the *'exactly'* condition was based on the fact that *'at least'* has a larger pragmatic halo and therefore leaves more room for personal interpretation than *'exactly'* does. This means that subjects could make the positive frame for *'at least'* seem more desirable than the same option with modifier *'exactly'*, whereas the negative frame, when following the same logic, will seem less desirable with *'at least'* than the same option in the *'exactly'* condition. This was theorized to lead to relatively more risk-aversion in positive frames and more risk-seeking in negative frames. The '*approximately'* condition showed to be between the other conditions in terms of standard framing effect shown, as was expected beforehand. However, once again no statistical significance was found in these results and we can therefore not say that the modifiers had an effect on the framing effects shown. Confidence intervals were added in an attempt to generalize findings to the population, where we can say with 95% confidence that values will fall between those noted in *Table 2.*. To try and explain the different outcomes between the studies, a critical look will be taken on the differences in methodology and design between them.

6.1. Contexts

One of the most significant differences between the current study and earlier work on the subject was the use of different contexts (hostage situation, disease and natural disaster; see **Appendix 2**). Therefore, we decided to test the influence that the use of different contexts had on the decisions made by the subjects and how quickly they made these decisions. To find out whether or not the chosen context of the ADP had influence on the decisions made, the answers given were compared between the three different contexts of the target items. Both the contexts natural disaster and hostage situation had an identical 54.5%/45.5% split in favour of the certain option. The context disease had a slightly different split of 51.8%/48.2%

in favour of the certain option. This means there are no statistically significant influences from the contexts on the given answers, which was the goal when creating the different contexts: being able to introduce extensionally equivalent problems multiple times without literally repeating questions.

To further ensure that the use of different contexts did not have an effect on the difference in outcome between Mandel's 2013 experiment and the current one, the time taken to react was compared between the different contexts. The average time taken to respond, measured in milliseconds, did not seem to differ between the contexts disease (M = 8894, SD = 6012), hostage situation (M = 10079, SD = 11210) and natural disaster (M = 10379, SD = 8249). However, because of the online nature of the test, response time is not the most trustworthy variable. It is easy to walk away from the computer while taking the test, because there is nobody present that is controlling the subject taking the test. To correct for outliers, all values that differ more than 2 standard deviations (SD = 10080) from the mean reaction time overall in target items (M = 9456) were taken out. Once every reaction time higher than 29615 was taken out of the data set, we once again tested the means between the contexts.

The contexts hostage situation (M = 7898, SD = 4374), disease (M = 8225, SD = 5557) and natural disaster (M = 9151, SD = 6048) were compared using a one-way ANOVA. Examination of the Kolmogorov-Smirnov statistics and histograms indicate that the assumption of normality was supported. Levene's test for equality was statistically non-significant, indicating homogeneity of variances, F (2, 326) = 4.250, p = .015. The ANOVA indicated that there was no statistically significant effect of context on reaction time, F (2, 325) = 1.605, p = .202, partial η^2 = .01. We can therefore not conclude that the given context of target items had a significant effect on either the decisions made (as explained in the paragraph before this one) or time taken to make the decisions (as explained in this paragraph). This means that the use of different contexts to maximize exposure to target items per subject did not have an effect that could explain the inconsistencies between the current research and Mandel's 2013 research.

6.2. Sample

Another inconsistency between the experiments was the way in which the test was taken by subjects. In the current experiment, 61 people (of which 55 were actually analysed) got to fill out the experiment on their phone without any sort of supervision or authority being present. In Mandel's 2013 experiment, the 120 people answered the questions on a MacBook computer in a supervised laboratory (p.4). The absence of supervision in the current study could explain some of the differences in outcome, since the subjects were less controlled in the current study than in Mandel's. We can not know for certain that subjects made the survey individually, whereas having a supervisor present and a fixed location for the experiment means there is control over the environment. We did however use control items with an obvious answer to weed out the subjects that did not read the questions correctly or did not understand it completely (see **Appendix 1**). Furthermore, since Mandel's supervisor did not say or do anything (as far as is noted in Mandel's paper), it should not have an effect on the answers given by the subjects. The mere presence of a supervisor could influence the way that subjects pick their options in the experiment, but we are not able to say whether it did or not with any certainty.

6.3. Design

In an attempt to maximize the number of observations we could analyse with the relatively small sample available, the current experiment ended up showing the ADP to subjects in three different contexts (see **Appendix 2**), each context shown twice with varying polarization. This ensured that we were able to acquire three observations per subject, compared to just the one observation per subject in Mandel's 2013 experiment. The advantage of using multiple observations per subject is that we were able to get more data from our subjects, but a possible disadvantage could be that the subject gets used to the questions and therefore answers differently than they would have done when seeing the question originally. If an individual feels like he or she is subjected to the same question twice, they might choose the same option as they did the first time without properly thinking about it beforehand. The other side of this argument is that seeing the same type of question multiple times could increase an individual's understanding of the question asked, and therefore result in a more dependable answer than seeing a question just once or twice.

Different contexts for the ADP were created to minimize repetition and this did not appear to influence decisions made by subjects (as shown earlier in the segment **6.1. Contexts**), but it still could be of importance in creating the inconsistencies between both studies. The choice to maximize observations was made to ensure a bigger statistical power and to make the results more applicable to the general population, which it did (Martin & Bateson (1993): more data collection equals better statistical power). The increase in analysable data was prioritised over the potential risk that multiple observations could have an impact on decision-making in the experiment, a risk that was minimized through the use of different contexts.

Another significant difference between the current study and Mandel's 2013 study is the use of unrelated- or filler tasks. The current study intertwined target items with a few filler questions with obvious 'correct' answers (see **Appendix 2**), whereas the target items in Mandel's study were divided with multiple unrelated tasks that took approximately 40 minutes to complete (p.4). There is a large difference in time spent on non-target items between the studies and this could play a role in creating the inconsistent results when comparing the two. A longer time spent between target items means that subjects are less likely to remember their first decision when making the second one on the target item with varying polarization and modifiers. The fact that the current study did not allow for such extensive non-target tasks (mainly due to budgetary restriction) could both benefit and harm the trustworthiness of the results. On one side, rapidly answering the target items could mean that the items are understood better and therefore answers are more reliable. On the other hand, rapid repetition of similar questions could lead to complacency in the subjects and copying their earlier answers instead of thinking about every question independently. We can not conclude what the effect may be of the difference in time spent on non-target items, it could both promote or obstruct sensitivity of the experiment.

7. Implications for future research

Interesting to note is the average frequency of standard framing effects shown in people that exhibited it at least once. From the 47 instances where a subject showed the standard framing effect, 44.7% of those (21 instances) where distributed among just 12.7% (seven) of the subjects. These subjects showed the standard framing effects in every context possible, whereas 50.9% (28) of the subjects showed consistent choices in every context possible. This means that all the observed framing effects, standard or reversed, where observed in less than half of the subjects (49.1%). People that showed standard framing effects did it almost twice on average (1.88), and only 40% of this group showed standard framing effects once. We can therefore theorize that some people are very sensitive for the effects of framing whereas others are never influenced by them. Sensitivity to framing effects in groups has been studied before, without any major significant outcomes (Milch, Weber, Appelt, Handgraaf, & Krantz, 2009). In their 2008 study, Flachaire and Hollard did research regarding individual sensitivity to framing effects and found that even in socially homogenous population, there is substantial homogeneity in sensitivity to framing effects. An explanation regarding why this could be the case was never given in this paper. The open answer methodology they used, and the ambiguity of their classification does also make these results less trustworthy, and more research on this subject is needed to form a more complete image. It is nevertheless an interesting note within the field of framing effect- and decision-making studies.

Another relevant field of research within risky choices and specifically the ADP is the ambiguity of numeric quantifiers. As mentioned in the section **2.2. Framing effect**, Frisch (1993) showed that a significant part of his subjects that demonstrated the framing effect did not agree that programs A and C or B and D were identical. Because of the online method that the current research used to test its' subjects, there was no option to reach subjects and ask them about the experiment. It would have been interesting to see whether or not they felt like programs were equivalent, and whether or not the use of modifiers influenced their opinion on this. Reflection on the options available, along with making an unambiguous version of the ADP in the way Kühberger (1995) did, would make for a clearer picture on how numeric quantifiers are interpreted and how semantic modifiers can impact this. Further research is needed to clarify the ambiguity of numeric quantifiers and the role semantic modifiers can play in solving this perceived ambiguity.

8. Conclusion

The goal of this research was two-fold: firstly, to replicate tests where standard framing effect was present in risky choices and secondly, to further research the role of different modifiers on this effect. The standard framing effect proved to be consistent with earlier research, which means that people are more risk-seeking in a negative frame and more risk-averse in a positive frame. No such conclusion can be made on the role that unilateral lower-bound modifier *'at least'* and bilateral modifiers *'exactly'* and *'approximately'* play in the process of framing effects and decision-making. This lack of statistical significance can and has been explained through multiple possible factors, that likely all interact with each other and even more factors. We can say that the decision-making in risky choices of certain individuals can be influenced by changing the way in which possible outcomes are worded, be it in wins (lives saved in this experiment) or losses (lives lost in this experiment). More research is needed on the role of modifiers in this process to come to a more complete image, especially more excessive research on a bigger scale can give us more clarity. The study of modifiers in risky choices and, more specifically, the ADP is one that has potential to give us more information on human decision-making and how it can be influenced. This could in turn learn us more about how to mimic and predict human decision-making and possibly prevent certain unwanted processes within the decision-making. There is however still a long way to go, and more research is needed to get more of a grasp on the processes behind one of our biggest and ever-present cognitive challenges: decision-making.

Appendix 1: Filler questions[1]

Filler 1

An error at the bank has left 600 people without any money on their account. Two plans can be chosen to fix this error.

If plan A is chosen, it is certain that 600 people will get their money back

If plan B is chosen, there is a 10% possibility that all 600 people will get their money back plus an extra 3% and a 90% possibility that all the money is lost in administration

Which plan would you choose?

Filler 2

Due to an open gate, a farm lost 500 chickens. Two plans can be chosen to retrieve them.

If plan A is chosen, it is certain that 20 chickens will be returned safely

If plan B is chosen, there is a 95% possibility that all chickens will be returned safely and a 5% possibility that none of the chickens will return

Which plan would you choose?

Filler 3

An intern at the kitchen has dropped a box with 100 portions of lettuce. Two plans can be chosen to pick the lettuce up.

If plan A is chosen, it is certain that 90 portions will remain edible

If plan B is chosen, there is an 80% possibility that all the lettuce needs to be thrown out and a 20% possibility that 95 portions will remain edible

Which plan would you choose?

Filler 4

During a flood In the library, 400 books are damaged. Two plans can be chosen to recover these books.

If plan A is chosen, it is certain that 25 books will be recovered

If plan B is chosen, there is an 85% possibility that all books will be recovered and a 15% possibility that 10 books will be recovered

Which plan would you choose?

[1] The plans that are written in italics are the 'right' answer, pcking 2 or more 'wrong' plans leads to exclusion from the sample

Appendix 2: Contexts and options

Context 1: hostage situation

In a region of war, 600 people are held hostage. Two plans can be chosen to attempt to bring them to safer regions with the following outcomes.

Context 2: disease

A rare disease has broken out in a quarantined population of 600 people. Two plans can be chosen to tackle the disease with the following outcomes.

Context 3: natural disaster

A small community of 600 people is threatened by a natural disaster. Two plans can be chosen to bring people to safety with the following outcomes.

Positive polarization

If plan A is chosen, it is certain that *[exactly/at least/approximately]* 200 people will survive.

If plan B is chosen, there is a 1/3 possibility that all 600 people will survive and a 2/3 possibility that nobody will survive.

Negative polarization

If plan A is chosen, it is certain that *[exactly/at least/approximately]* 400 people will die.

If plan B is chosen, there is a 2/3 possibility that all 600 people will die and a 1/3 possibility that nobody will die.

References

Arrow, K. J. (1982). Risk perception in psychology and economics. *Economic Inquiry, 20,* 1-9.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae, 1738*(5), 175-192.

Blok, D. (2015). The semantics and pragmatics of directional numeral modifiers. *Proceedings of SALT, 25,* 471-490.

Breheny, R. (2008). A New Look at the Semantics and Pragmatics of Numerically Quantified Noun Phrases. *Journal of Semantics, 25*(2), 93-139.

Coppock, E., & Brochhagen, T. (2013). Raising and resolving issues with scalar modifiers. *Semantics & Pragmatics, 6*(3), 1-57.

Crozier, R., Ranyard, R., & Svenson, O. (1997). *Decision making: Cognitive models and explanations*. New York, NY: Routledge.

Druckman, J. N. (2004). Political Preference Formation: Competition, Deliberation, and the (Ir)relevance of Framing Effects. *American Political Science Review, 98*(4), 671-686.

Fishburn, P. C. (1970). *Utility Theory for Decision Making.* New York, NY: Wiley.

Flachaire, E., & Hollard, G. (2008). Individual sensitivity to framing effects. *Journal of Economic Behavior & Organization, 67*, 296-307.

Friedman, M., & Savage, L. J. (1948). The utility analysis if choices involving risk. *Journal of Political Economy, 56*(4), 279-304.

Frisch, D. (1993). Reasons for Framing Effects. *Organizational Behavior and Human Decision Processes, 54*, 399-429.

Genesove, D., & Mayer, C. (2001). Loss Aversion and Seller Behavior: Evidence from the Housing Market. *The Quarterly Journal of Economics, 116*(4), 1233-1260.

Geurts, B. (2013). Alternatives in Framing and Decision Making. *Mind & Language, 28*(1), 1-19.

Geurts, B., Nouwen, R. (2007). 'At least' et al.: The Semantics of Scalar Modifiers. *Language, 83*(3), 533-559.

Horn, L. R. (1992). The Said and the Unsaid. *Semantics and Linguistic Theory, 2,* 163-192.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis if Decision under Risk. *Econometrica, 47*(2), 263-292.

Karmaker, U. S. (1978). Subjectively Weighted Utility: A Descriptive Extension of the Expected Utility Model. *Organizational Behavior and Human Performance, 21,* 61-72.

Kühberger, A. (1995). The Framing of Decisions: A New Look at Old Problems. *Organizational Behavior and Human Decision Processes*, *62*(2), 230-240.

Lakshminarayanan, V. R., Chen, M. K., & Santos, L. R. (2010). The evolution of decision-making under risk: Framing effects in monkey risk preferences. *Journal of Experimental Social Psychology, 47,* 689-693.

Lasersohn, P. (1999). Pragmatic Halos. *Language, 75*(3), 522-551.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes*, *76*(2), 149-188.

Machina, M. J. (2008). Non-expected Utility Theory. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics* (pp. 1-14). London: Plagrave Macmillan.

Mandel, D. R. (2001). Gain-Loss Framing and Choice: Separating Outcome Formulations from Descriptor Formulations. *Organizational Behavior and Human Decision Processes, 85*(1), 56-76.

Mandel, D. R. (2013). Do Framing Effects Reveal Irrational Choice? *Journal of Experimental Psychology: General, 143*(3), 1-14.

Martin, P., & Bateson, P. (1993). *Measuring Behaviour: An Introductory Guide,* Cambridge: Cambridge University Press.

Milch, K. F., Weber, E. U., Appelt, K. C., Handgraaf, M. J. J., & Krantz, D. H. (2009). From individual preference construction to group decisions: Framing effects and group processes. *Organizational Behavior and Human Decision Processes, 108,* 242-255.

Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance, 53*(5), 1775-1798.

Sauerland, U., & Stateva, P. (2007). Scalar vs. Epistemic Vagueness: Evidence from Approximators. In M. Gibson & T. Friedman (eds.), *Semantics and Linguistic Theory (SALT) 27,* 228-245. New York: CLC Publications.

Spotorno, N., McMillan, C. T., Rascovsky, K., Irwin, D. J., Clark, R., & Grossman, M. (2015). Beyond words: Pragmatic inference in behavorial variant of frontotemporal degeneration. *Neuropsychologica, 75,* 556-564.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, 211*(4481), 453-458.

Tversky, A. & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59,* 251-278.

Whitaker, J. K., & Fellner, W. (1966). Probability and Profit. A Study of Economic Behavior Along Bayesian Lines. *The Economic Journal, 76*(301), 105-107.