Assessment and resolution of difficult choices when deciding among alternatives

Stefan Gelissen

The work described in this thesis has been carried out under the auspices of the J.F. Schouten School for User-System Interaction Research, and was funded by SOBU.

Printing: Universiteitsdrukkerij Technische Universiteit Eindhoven

Cover: Bart Heijman Producties

An electronic copy of this thesis is available from the site of the Eindhoven University Library in PDF format (http://www.tue.nl/bib).

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Gelissen, Stefan

Assessment and resolution of difficult choices when deciding among alternatives / by Stefan Gelissen. – Eindhoven: Technische Universiteit Eindhoven, 2007. – Proefschrift. –

ISBN 978-90-386-0946-1
NUR 775

Keywords: Decision difficulty / Random choice procedure / Choice deferral / Reason-based choice / Regret / Trade off difficulty

Assessment and resolution of difficult choices when deciding among alternatives

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Eindhoven,
op gezag van de Rector Magnificus, prof.dr.ir. C.J. van Duijn,
voor een commissie aangewezen door het College voor Promoties
in het openbaar te verdedigen
op donderdag 3 mei 2007 om 16.00 uur

door

Stephanus Leonardus Maria Gerardus Gelissen

geboren te Weert

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. G.B. Keren

Copromotor:
prof.dr. Y. Ganzach

In memory of my father, Jac Gelissen

Contents

1

General introduction

Introduction

People are faced with making an enormous number of decisions during their lives. Some of these decisions are relatively easy, whereas others are experienced as rather difficult. What makes some decisions easy, whereas other are experienced, or perceived, as difficult? Easy decisions mostly involve alternatives of which one option is (much) more attractive than the other (i.e., one dominates the other alternative). Arguably, a decision involving a dominating alternative may not be considered a choice problem at all, because the decision maker will lack the experience of an internal conflict (Delquié, 2003). Difficult decisions, on the other hand, suggest that the alternatives under consideration are close in attractiveness (Festinger, 1964; Liberman & Förster, 2006). These difficult decisions are often accompanied with a discomforting feeling caused by an internal conflict, which signifies that the decision maker equally (dis)likes the available alternatives, and hence is indecisive about which alternative to choose. Decision makers may behave differently when trying to resolve this conflict. Some focus mainly on the most important (prominent) attribute that characterizes the available alternatives (e.g., Slovic, 1975; Tversky, Sattath & Slovic, 1988). To illustrate, consider choosing between equally attractive cars, defined on several attributes (e.g., price, color, gas mileage, etc.), of which one car is superior on some of these attributes (e.g., price), and inferior on the remaining ones (e.g., gas mileage). Difficulty in choosing stems from the fact that the superior attributes will evoke a preference for the car above, yet the inferior attributes will lead to a preference for the other car(s). This internal conflict may be resolved by choosing the car that is, according to the decision maker, superior on the most important feature (e.g., price), and rejecting the other car(s) by diminishing the importance of the remaining attributes (e.g., gas mileage).

Another way of resolving a difficult conflict between equally attractive options is avoiding choice (e.g., Dhar, 1997; Tversky & Shafir, 1992). For instance, the economist Thomas Schelling once decided to buy an encyclopedia for his children at a bookstore (Tversky & Shafir, 1992). This bookstore confronted Schelling with a difficult choice, since the two encyclopedias it sold were equally attractive to him. As a result, he bought neither of them. Yet, he argued, had the bookstore only sold one of the two encyclopedias, his internal conflict would not exist, and he would therefore have ended up with buying an encyclopedia for his children.

It might not always be possible to avoid difficult decisions as Schelling did. Choosing is sometimes inevitable in, say, custody disputes with equal parental fitness. In these cases a decision has to be made, as delaying a decision could harm the child (e.g., it is not beneficial to the child to have lengthy emotional disputes in a court). Elster (1989) suggested that for deciding in custody disputes with equal parental fitness, it may be appropriate to use a random procedure (e.g., a coin toss). Yet, people often strongly resent to decide by the toss of a coin. For instance, in a dispute between grandparents and their son-in-

law, both parties could not agree on with whom the two (grand)daughters would spend their Christmas ("Judge disciplined for coin flip", 2003). During the court hearing, the judge decided to toss a coin, which was won by the son-in-law. Yet, the grandparents filed a complaint, telling that their confidence in the courts was destroyed, and that the judge made a mockery of the hearing. As a result, the judge was reprimanded by the Supreme Court, which argued that tossing a coin was inappropriate, since judges are supposed to make decisions based on facts, law and best interest of the children, not on chance. This example, where a coin toss is used, by no means implies that people are repulsive towards each and every way of random selection. Possibly, when resolving a difficult decision, they resent other random procedures to a lesser degree, such as drawing lots in a lottery.

Why study choice difficulty?

According to normative theories, choice difficulty is not supposed to influence decision making (Shafir & LeBoeuf, 2004; Tversky & Shafir, 1992). These normative theories prescribe which decisions should be made, such that decision makers behave optimally given their goal, often defined as maximization of utility (Coombs, Dawes & Tversky, 1970). Following normative accounts, a large difference in attractiveness yields no internal conflict, and deciding which option to choose should be easy. Similarly, alternatives close in attractiveness will not evoke an internal conflict, since it matters little which of these equally attractive options is obtained. Hence, a choice between alternatives close in attractiveness should, as with alternatives differing substantially in attractiveness, be easy. In short, normative choice theories disregard that, irrespective of large or small differences in attractiveness, the decision maker experiences any internal conflict or choice difficulty.

Descriptive choice theories, on the other hand, are concerned with how choices are actually made (Coombs, Dawes & Tversky, 1970). Relying on research that investigates choice difficulty, descriptive accounts claim that the internal conflict *will* influence the decision maker's behavior. Accordingly, studying choice difficulty is important for several reasons (e.g., Beattie & Barlas, 2001; Delquié, 2003). First, experiencing choice difficulty gives rise to coping behavior, such as deferring choice (Tversky & Shafir, 1992; Dhar, 1997), or employing a random choice procedure (Elster, 1989). For instance, having difficulty in deciding among conflicting alternatives influences the purchase decisions of consumers, with deferral as a possible response. To illustrate, people will more likely defer choosing when confronted with a difficult decision among a popular SONY player for $99 and a top-of-the line AIWA player of $159, than when the SONY is the only available player, the latter which does not represent a conflict situation (Tversky & Shafir, 1992).

Second, difficult decisions may get decision makers more engaged in the decision task. The greater the experienced conflict, the more deliberation about the decision problem will take place. This deliberation concerns comparing the

alternatives along their attributes, and contemplating about the consequences of choosing one or the other option (Scholten & Sherman, 2006). A consequence of this deliberation is that a decision maker relies less on suboptimal cognitive shortcuts (e.g., choosing the option that is superior on the most prominent attribute), and instead, tries to make the trade off between the attributes involved (Delquié, 2003).

In short, although normative choice theories undermine the importance of choice difficulty on decision making, descriptive theories accept that difficulty plays a role in decision making, which is observable during making decisions (e.g., use of choice strategies), and in subsequent behavior (e.g., deferral).

What makes a decision difficult?

There are several factors that might make a decision difficult. One may, though, distinguish between factors that are external, and others that are internal to the alternatives under considerations. External factors have mainly to do with the given circumstances under which the alternatives are presented. These circumstances may refer to constraints provided by the consumer market, amount of knowledge of the decision maker, and the state of the world. The internal factors, on the other hand, attribute difficulty to characteristics of the alternatives themselves. The experienced difficulty, however, may stem not from solely external or internal factors, but often will result from a combination of a number of both types of factors. Furthermore, the external and internal factors may represent no real dichotomy. That is, it is possible that some factors defy a straightforward categorization, as they may fall in between these dichotomous factors.

External factors

Uncertainty

The existence of uncertainty influences decision difficulty, often expressed by the probability whether the expected outcome of a given alternative will be obtained. The experienced uncertainty may stem from two sources (Coombs, Dawes & Tversky, 1970). First, a decision maker may have incomplete knowledge about the world. For instance, it is not always clear how a future state or event will look like (e.g., will it rain or not). If a decision outcome (e.g., whether or not to carry an umbrella) depends on such states or events, deciding may be difficult. Second, it is possible that one lacks knowledge about oneself. A decision maker might be unsure which alternative will be most satisfying (even if the external consequences are known for sure). For instance, having to choose between pizza toppings can be difficult, since it is not straightforward which one will be most enjoyable at a particular diner.

Importance

The more important the possible outcomes (consequences) of the alternatives are, the more difficult the decision will be (Festinger, 1957, 1964).

Undoubtedly, most people would agree that choosing between cars will be more difficult than choosing between ice creams (of course, depending on the goals one has in mind, which explains why children would argue that choosing between ice creams is much more difficult). Presumably, importance may change as a function of temporal distance, since the value of outcomes are discounted as temporal distance increases (e.g., Trope & Liberman, 2003). Hence, a decision in the near future may be perceived as more important, compared to the same decision taken in the distant future. As a consequence, a decision maker may experience the decision as more difficult in the near than in the distant future.

Justification

People can be held accountable for their actions, and hence they might be asked to justify their decisions towards others (Tetlock, 1997). This external justification influences choice difficulty (Zhang & Mittal, 2005), because in most cases, a decision maker will be uncertain about the preferences of others (e.g., their superiors). Yet, a decision maker will try to choose the alternative that is most justifiable, which might be difficult, since it is often uncertain how others will trade off attributes, and what weights they assign to the attributes involved (Simonson, 1989).

Irreversibility

Not being able to reverse a decision, once it has been made, will increase difficulty (Festinger, 1957). The source of difficulty with decisions that can not be changed or revoked seems to be potential regret (Anderson, 2003), which manifests itself if the chosen alternative ends up being worse than the unchosen alternative(s) (Zeelenberg, Van Dijk, Manstead, & Van Der Pligt, 2000).

Number of available alternatives

The larger the choice set, the more difficult choosing is perceived to be (Hendrick, Mills & Kiesler, 1968). For instance, Iyengar and Lepper (2000) demonstrated that choosing between 30 different flavors of chocolates was experienced as more difficult than choosing between 6 different flavors, and furthermore, resulted in more dissatisfaction and regret regarding the chosen alternative.

Time pressure

Decision makers are more likely to use non-compensatory (i.e., focusing only on the most important attribute, and avoiding difficult trade off comparisons) than compensatory (i.e., choosing by making trade offs between all attributes involved) decision strategies when they are under time pressure (Dhar & Nowlis, 1999). Because non-compensatory strategies will simplify choice (Payne, Bettman & Johnson, 1993), decision difficulty reduces under time pressure.

Internal factors

Attractiveness of alternatives

Attractive alternatives are referred to as being better compared with some reference point, whereas unattractive alternatives are worse than some reference (Hsee & Leclerc, 1998; Zhang & Mittal, 2005). Consider a consumer who decides to buy a new mp3 player, and can choose between two available brands. If the two brands have a longer battery life time, and larger memory capacity than the player which the consumer currently owns (i.e., reference), then the two players are attractive to the consumer. Likewise, if the two brands have shorter battery times, and smaller memories, then the consumer will perceive them as unattractive. Chatterjee and Heath (1996) demonstrated that a decision among unattractive alternatives will be experienced as more difficult than a choice between attractive alternatives.

Multi attribute alternatives

Alternatives are often characterized by two or more attributes (or alternatively: dimensions). A mp3 player defined by "battery life time" and "memory size" is an example of a multi attribute alternative. It is possible that a choice involves options that are superior on some of these attributes, while being inferior on the other attributes. Choosing between these alternatives may require a trade off between the attributes' values.

Decisions, involving multi attribute alternatives, are difficult because people find these trade offs hard to make (Beattie & Barlas, 2001). Specifically, a trade off requires that one attribute (e.g., battery life) has to be compared to another, rather incomparable, attribute (e.g., memory size). Hence, when deciding among two mp3 players, a decision maker has to address the question of how to trade off, say, a 7 hours longer battery life to 512 MB less memory size. A trade off involving these incommensurable attributes will not be easy.

In addition, choice difficulty increases with alternatives having larger differences in their attribute values (i.e. larger trade off size) (Chatterjee & Heath, 1996). Thus, deciding among two players with a difference of 14 hours in battery life time, and 1024 MB in memory size, is experienced as more difficult than deciding among players with a difference of 7 hours and 512 MB. Closer inspection suggests that such larger trade offs imply larger sacrifices. That is, commitment to an alternative not only endows the decision maker with a large advantage (e.g., 1024 MB more memory size), it also signifies that a large advantage (e.g., 14 hours more battery life) is foregone with rejecting the other alternative. Supposedly, these larger sacrifices are the source of the experienced difficulty. Notwithstanding, deciding among alternatives with small trade off sizes may also be experienced as difficult, because it may be difficult to construct compelling reasons for justifying the decision maker's preference (Scholten & Sherman, 2006).

Besides variation in trade off size, the number of distinctive attributes, on which alternatives are characterized, may vary too. The more attributes that are involved in a choice problem, the more trade off comparisons are required. As

a result of the increase in the number of trade offs, a decision will be perceived as more difficult (Dhar, 1997; Hendrick, Mills & Kiesler, 1968).

The attributes that are involved in a choice problem, though, may not always be of equal importance. For instance, the differential importance between "battery life time" and "memory size" will be smaller, compared to the differential importance among "battery life time" and "dimensions of the players". Decisions among attributes having a smaller differential importance is found to be more difficult than among attributes with a larger differential importance (Scholten & Sherman, 2006).

Finally, some attributes may evoke a high level of (negative) emotion. These attributes mostly pertain to valued goals (e.g., "safety" and "health"), and decisions involving these emotion-laden attributes imply a trade off, in which the valued attributes might have to be sacrificed. Usually, these sacrifices are experienced as threatening, and will lead to an increasing level of choice difficulty (Luce, Payne & Bettman, 2001). For instance, a trade off involving "safety" and "money" in a car purchase will be more difficult than a trade off among "comfort" and "money".

Uni-dimensional alternatives

Although most choice problems will involve alternatives defined on two or more attributes, alternatives may sometimes be characterized by a single (abstract) dimension. Consider again the mp3 players described by "battery life time", "memory size", "weight", etc. It is possible to summarize all these characteristics by one overall rating. For instance, consumer magazines may use a rating scale, ranging from 1 (poor) to 9 (excellent), to inform consumers about two players. Different magazines, however, may judge the two players differently, as in the following example.

	mp3 player A	mp3 player B
Magazine X	5.9	7.7
Magazine Y	8.0	6.2

Despite the different ratings, assigned by the two magazines, the two players' average ratings are identical, namely 6.95. Since both players' averages are identical, and each magazine's ratings favor a different player, it is not obvious which of the two players a decision maker should prefer, and thus may yield an internal conflict.

How does a decision maker resolve a conflict that involves uni-dimensional alternatives? It is unlikely that deciding among uni-dimensional options requires a trade off between attributes, because there is only one (summary) dimension on which the alternatives are judged. Furthermore, due to the uni-dimensional character, it is impossible to resolve difficult decisions by focusing on the most important attribute. Moreover, because of the impossibility to emphasize the most important attribute, it may not be easy to construct compelling reasons for choice justification. This difficulty, in constructing compelling reasons, may explain why a choice between uni-dimensional

alternatives is perceived as difficult. Yet, these uni-dimensional options are relatively overlooked in judgment and decision making problems, and it remains unclear why, and under which circumstances, choosing between them is difficult.

Dichotomous alternatives

The attributes that constitute multi attribute alternatives may not always have a (continuous) dimensional character. Sometimes, an alternative includes a feature that the other alternative has not, that is, a unique feature. For example, a mp3 player may come with or without a "FM radio", or a "time and alarm function". It is not inconceivable to represent these features dimensionally, but in essence, these dimensions have a dichotomous (yes – no) character.

Evidently, decision makers predominantly focus on continuous rather than dichotomous dimensions in choice and judgment tasks. For instance, Slovic and MacPhillamy (1974) had participants compare two students' test results, and indicate which one was more likely to have a higher Grade Point Average (GPA). The respective test results were from a test taken by both students (representing a continuous dimension), and from a test that was unique for each student. The GPA predictions suggested that the continuous dimension received more weight from the participants than the unique dimensions. Supposedly, the continuous dimension facilitated a comparison between the students (i.e., easy to find out that one alternative exceeds the other on the specific test), while the results from the unique tests hindered a direct comparison (i.e., difficult to compare the two tests directly, and hence which test exceeds the other). As a consequence of the facilitated comparison with continuous dimensions, justifying one's preference will more readily focus on these dimensions (Markman & Medin, 1995). At the same time, constructing compelling reasons, that pertain to dichotomous (unique) dimensions, will be rather difficult, which, in turn, may increase the experienced choice difficulty.

For the same reasons, choosing between equally attractive alternatives from different product categories (e.g., a trip to China or a notebook) may be more difficult than choosing between alternatives belonging to the same category (e.g., two notebooks). Because a comparison between products from different categories mainly focuses on (non-comparable) dichotomous attributes (Johnson, 1984), it might be difficult to construct compelling reasons that (directly) relate to these attributes.

What makes studying choice difficulty difficult?

All of us are confronted with difficult decisions in our lives. It is not always clear, however, what people exactly mean when they say they experience a difficult decision, and what the underlying causes are of this experienced difficulty. We know, from experience, that participants often explain a choice between two small trade off size alternatives to be difficult, by telling that the differences in attribute values were small. This explanation, however, is simply a re-description of the alternatives involved, and conceals what actually made

the decision difficult (e.g., “it was difficult to construct compelling reasons”, or “the alternatives were equally attractive”).

Scientifically, the concept of “decision difficulty” does not have a standard formalization or operationalization, nor is there agreement about what the concept refers to (Hastie, 2001). For instance, it could be that the more elementary information processes are needed, the more difficult the decision will be. If this is the case, a measure of difficulty could be the number of processes involved, or the processing capacity required. A second (alternative) measure of difficulty, may be decision makers' subjective evaluation of the decision’s difficulty. Third, working backwards from a correct choice alternative (e.g., the one with the highest expected utility), may indicate the conditions and situations under which decision makers make errors, such as not taking into account information important for making the correct decision. The number of errors will, presumably, correspond with decision difficulty, and thus may be employed as yet another measure of difficulty. To conclude, there are several ways by which decision difficulty can be defined, leading to different conceptions of how it could be measured.

Measuring difficulty can thus be problematic, among other things, because there is no one, well defined, measurement standard (Tversky & Shafir, 1992). Besides the examples described above, the degree of difficulty seems also to correspond to, and thus can be measured by, the need to defer decisions (e.g., Tversky & Shafir, 1992; Dhar, 1997), and decision time (e.g., Hendrick, Mills & Kiesler, 1968). An important difference between the available measures, though, is that they assess decision difficulty in either an indirect or direct way, of which deferral is an example of the former, and subjective evaluation of the latter.

Direct measures

Choice difficulty may be examined by having participants themselves evaluate the extent to which they feel a decision is difficult. By itself, this procedure only points to the presence of difficulty, yet it conceals the source of the difficulty, and therefore does not explain why choosing was difficult. The source of the difficulty, though, might be revealed by asking participants why they thought choosing was experienced as difficult. Such explanations, however, are prone to introspection, a method that is problematic given the (controversial) assumption that participants have no direct access to their cognitive processes (Nisbett & Wilson, 1977). Therefore, the direct procedure is often used in conjunction with other measures, such as regret (Chatterjee & Heath, 1996; Iyengar & Lepper, 2000). A combination of these two measures, demonstrates that increasing (subjective) difficulty corresponds with increasing potential regret (Chatterjee & Heath, 1996).

Indirect measures

Difficult decisions bring along behavioral consequences, such as deferral, random selection, and longer decision times. These behavioral consequences, in turn, can be employed as (indirect) measures of choice difficulty. The following section, focuses on two such indirect measures, namely deferral and random selection.

Deferral

When decisions are difficult, people often display the tendency to defer choosing, thereby running the risk of foregoing available alternatives. Apparently, deferral can be used as a measure of difficulty, suggesting that the likelihood of not choosing increases with higher levels of experienced difficulty. As a measure of difficulty, however, deferral is not specific with regard to the source from which difficulty emanated. To illustrate, Dhar (1997) argued that the cause of deferral is preference uncertainty. Specifically, the smaller the decision maker's difference in preference among alternatives, the greater the tendency of not to choose. This explanation, nevertheless, is unspecific about the exact source from which the preference uncertainty originated, since this could be due to either external or internal factors, or any interaction between these factors. Tversky and Shafir (1992), on the other hand, were more specific about why people defer choosing, and suggested that deferral is the consequence of a specific (internal) factor, namely that one is not able to make difficult trade offs. According to Dhar (1997), though, this trade off explanation can not account for the result that participants are equally likely to defer choosing in case of small and large trade off sizes. After all, the trade off conflict emerging from smaller trade offs is easier to resolve (Chatterjee and Heath, 1996), and therefore should result in less deferral. An account solely based on trade off difficulty, thus seems inconsistent with the result that people often persist to defer, even when choosing between small trade off alternatives. The difficulty with small trade off alternatives, nevertheless, can be explained by arguing that people lack compelling reasons for choice justification, once they traded off the attributes involved (Scholten & Sherman, 2006). This account, of why choosing is difficult with smaller trade off sizes, may also explain the willingness to defer. Note that with regard to the source of the difficulty, this latter explanation is less shallow than the explanation based on preference uncertainty by Dhar (arguing that deferral is best explained, by suggesting that the alternatives were close in attractiveness). Furthermore, it shows that deferral in not sensitive to only one, but multiple factors, which makes it an impure measure of selection difficulty (Anderson, 2003).

Besides choice deferral when alternatives are equally attractive, decision avoidance may manifest itself when someone avoids choosing an alterative (i.e., omits action), subsequent to missing a similar, more attractive, option. For instance, participants who already bypassed the opportunity to buy a $40 ski pass, are less likely to buy a $90 ski pass (even if this represents a 10% discount of the usual price of $100), compared to participants who missed a $80 ski pass (Tykocinski, Pittman & Tuttle, 1995). This type of decision

avoidance, however, concerns a decision that involves rejecting (avoiding) an alternative which is advantageous ($90) compared to a reference point ($100). In other words, it deals with a choice between non-equally attractive options, of which the reference option (i.e., inaction) is preferred. For that reason, it can be distinguished from choice deferral, which is used for resolving a difficult choice between close, equally preferred, alternatives.

Random selection

Resolving difficult decisions by random selection is less accepted, and therefore less prevalent, than deferral. People seem reluctant, for instance, to toss a coin when confronted with a difficult decision (Elster, 1989), especially when the consequences of a decision are important to the decision maker. In soccer, for example, a coin toss is used for deciding which team will kick off, and which direction the teams will play in each half of the match. Supposedly, the outcome of this coin toss will have some influence on the match, but will not decide who will win the match, which diminishes the importance of the coin toss at the beginning of the match. At the European Championships in 1968, however, the semi final between Italy and the Soviet Union ended with 0-0, and a coin toss was used for deciding which of the two teams would reach the final ("Coin flipping", 2006). The coin toss was won by Italy, who later became European champions. Resistance, against having a coin making the important decision of who will win the match, may have resulted in the introduction of penalty shoot-outs, which are nowadays common at soccer championships (though these too may have a strong stochastic element).

Coin tossing, as measure of difficulty, seems to be overlooked in choice difficulty research. The reason why coin tossing received no systematic attention might be twofold. First, the prevalence of coin tossing, as a way of resolving difficult decisions, is low. That a coin toss is more of a curiosity, rather than regularity, might explain why it is seldom used as a measurement of difficulty. Second, selection by a coin toss is often interpreted as if it does not matter to the decision maker which alternative is obtained (or that choosing was easy). In other words, tossing a coin can signal that it is unimportant which choice is made, or alternatively, to which of the available alternatives decision makers will commit themselves.

Despite the lack of attention, coin tossing can be used as a measurement tool of difficulty. For instance, it is possible to ask, in which of a number of distinctive choice problems a decision maker is more likely to accept a coin toss for resolving a difficult decision. Such difficult decisions imply that it is important which choice is made, and using the coin toss as measurement might not only indicate the degree of difficulty, but also reveal particular sources underlying the experienced difficulty, for instance, lacking compelling reasons for choice justification (Elster, 1989).

Aims of this thesis

The present thesis centers around two indirect measures of choice difficulty, namely deferring choice and coin tossing, and focuses on internal factors that influence difficulty.

In Chapter 2, both measurements are applied in the context of multi attribute choice problems that vary in trade off size. It presents experiments, of which the results are indicative of the circumstances (i.e., large or small trade off size) under which deferral and coin tossing may be employed. The chapter further discusses what sources of difficulty deferral and coin tossing bring to light, when they are applied to resolve difficult decisions among alternatives that vary in trade off size.

Chapter 3 shows how different display organizations, or framings, of identical uni-dimensional choice problems influence the likelihood to toss a coin. It further offers an explanation of how choice difficulty might change with these different display organizations.

Chapter 4 is devoted to choice problems involving dichotomous (yes-no) attributes. Although previous studies (Markman & Medin, 1995) demonstrated that it is more difficult to justify choice by emphasizing dichotomous rather than continuous attributes, it is unexplored by which means (i.e., deferral or coin tossing) people are willing to resolve a difficult decision involving these attributes. The chapter presents experiments that investigate the likelihood to defer choosing or tossing a coin, in the vicinity of both types of attributes, and further explores from which sources the difficulty with these attributes stems.

The fifth and final chapter discusses the experiments presented in this thesis, and draws conclusions about using deferral and coin tossing for measuring choice difficulty.

2

Resolving difficult choices by decision deferral and coin tossing as a function of trade off size

Abstract

This chapter explores to what extent "choice deferral" and "coin tossing" are acceptable as means for resolving difficult choices, when trade off sizes are varied. Empirical evidence is presented, suggesting that choice deferral is accepted irrespective of the size of the trade off. Coin tossing, on the other hand, is predominantly accepted in the context of smaller trade off sizes. This observed difference in acceptability is explained by assuming that difficulty, as a function of trade off size, is accounted for by different underlying sources. Accordingly, it was found that, as an explanation for decision difficulty, (1) anticipated regret was more likely used for larger trade off sizes, (2) lacking compelling reasons for smaller trade off sizes, whereas (3) difficulty in trading off pros and cons was used as an explanation, irrespective the size of the trade off. Argued is that the observed correlation between the two means of resolving difficulty (deferral and coin tossing) and the different sources underlying difficulty, suggests that these two means can be employed as measurement tools for revealing possible sources underlying the experienced choice difficulty, when trade off size is varied.

People often find it difficult to choose between two (or more) options. One factor influencing choice difficulty is trading off options' advantages against disadvantages. Imagine, for instance, a student who has to choose between two apartments. Although one apartment is larger than the other, it is also located further away from the university. For deciding among the two alternatives, the student might have to trade off the advantages and disadvantages associated with each option. Presumably, this trade off process depicts a commitment to one room, resulting in obtaining its advantages (e.g., larger room size), but also forces the student to accept its disadvantages (i.e., longer commuting time). At the same time, commitment unequivocally leads to rejection of the other room. This rejection results in avoiding the unchosen room's disadvantages (smaller room size), but more importantly, also in forgoing its advantages (shorter commuting time).

Before final commitment, the student may be confronted with opposite forces originating from the advantages and disadvantages of both rooms. These forces lead to the experience of a discomforting conflict (e.g., Festinger, 1957). Such conflict exists both between, as well as within alternatives (Fischer, Luce & Jia, 2000), and resolving it may be difficult. Moreover, Chatterjee and Heath (1996) have shown that by enlarging the trade off size (i.e., larger attribute differences), the choice between two alternatives becomes even more difficult. Supposedly, by increasing the differences on the attribute values, the decision maker becomes more reluctant to sacrifice the advantages of the un-chosen alternative. Evidently, larger sacrifices yield a greater conflict (Scholten & Sherman, 2006), which renders a decision more difficult.

A consequence of increasing choice difficulty is that people are more likely to defer choosing between alternatives. For instance, Tversky and Shafir (1992) demonstrated that when options were equally attractive, the tendency to defer was greater than when one option dominated the other. In the latter case, choosing will be experienced as easy, whereas in the former, people may experience a difficult to resolve conflict. Dhar (1997) examined deferral under conditions of either large or small attribute differences, and observed similar percentages of people deferring choice in both situations. This result suggests that choosing between small trade off options may be as difficult as choosing between large trade off options, in the sense that they have similar behavioral consequences (i.e., yield similar percentages of decision deferral).

Another way of resolving a conflict, emanating from a decision among equally attractive alternatives, is by employing a random device (Elster, 1989). Imagine two friends considering whether they should have dinner at an Italian or Greek restaurant. If both restaurants are equally attractive, they may as well decide by the toss of a coin. Nonetheless, people often resent the use of random selection (in particular when the decisional consequences are important to them). Consequently, deciding by the flip of a coin is rare. A possible explanation for this is that people may not like giving up decisional control, since loss of control may induce undesired feelings such as pessimism and distress (Skinner, 1996). Furthermore, random selection does not provide compelling reasons why one prefers a chosen alternative, whereas decision

makers often strongly rely on decisive arguments that resolve a conflict and will justify their preference (Shafir, Simonson & Tversky, 1993). However, when a decision maker is (and remains) indecisive about what to do, flipping a coin might be an acceptable way of resolving a difficult decision. Moreover, deferral is not always possible. For instance, when no other alternatives are available, and a decision has to be made fast, tossing a coin may be an acceptable solution.

Once the decision maker decides to select at random, it is important to distinguish between two situations. One, in which the decision maker is *indifferent* between the different options (e.g., going for dinner to a French or an Italian restaurant), and one does not really care which option is obtained. Under such circumstances, however, it is unlikely that a decision maker will experience the choice as difficult. Intuitively, a difficult choice stems from an act of deliberate choice, which is unnecessary if one is indifferent as to which alternative is obtained. On the other hand, random selection may also be appropriate in situations in which it *does* matter which alternative is obtained. Under such circumstances, the experience of a difficult choice may arise if, after some deliberation, the alternatives imply a conflict. If a decision maker can not resolve the conflict, and thus remains indecisive, she or he may find it acceptable to transfer the burden of choosing to a coin.

The purpose of the first experiment was to examine whether participants would, besides deferral, accept random choice as a means of resolving a decision conflict. Specifically, I presented simultaneously two pairs of alternatives, with either a large or small trade off size (cf. Dhar 1997). Two persons (decision makers) were introduced, each of whom had to choose between one of the respective pairs of options. Both persons claimed that they experienced great difficulty while making their choice. The consequences of this statement, regarding choice difficulty, are twofold. First, the statement conveys the information that it *does* matter, to the persons in question, to which alternative they commit themselves. Apparently, they carefully examined the available options, figured out that they would like to obtain both of them, and as a consequence, could not decide which alternative they preferred. Second, Liberman and Förster (2006) demonstrated that when a person is said to experience a difficult decision, people will draw the inference that to this person the provided alternatives were close in attractiveness.

Experiment 1

Method

Participants

A total of 130 students from Fontys University of Professional Education Eindhoven were recruited and paid 3 Euros for their participation. The experiment was part of a few unrelated judgment and decision making experiments that lasted for approximately 20 minutes.

Materials

Two pairs of rooms were constructed, defined on two attributes, namely commuting time and room size. The trade off size of the two attributes was varied and was large for one pair (rooms A and B), and small for the other (rooms C and D) (see Figure 1).

Pretest

Both pairs of rooms were pre-tested to determine whether varying trade off size resulted in different preferences within both pairs. Particularly, considering both pairs, rooms A and C may be regarded as comparable, since both are superior on commuting time and inferior on room size compared to rooms B and D, respectively. If, for instance, rooms A and C are preferred for their advantage of a shorter commuting time, then varying the trade off size is not expected to yield a difference in preference such that, for example, room A is preferred to B by 90%, while C is preferred to D by 60%. With such a difference in preference between both pairs, choosing between A and B might be easier than between C and D, since A dominates B more than C dominates D. Hence, when a difference in preference between both pairs is found, then choice difficulty might not be comparable for both pairs.

	A	**B**
Commuting time	10 minutes	40 minutes
Room size	14 m^2	26 m^2

	C	**D**
Commuting time	20 minutes	30 minutes
Room size	19 m^2	21 m^2

Figure 1. Overview of the pairs of rooms. Note that the trade off size for rooms A and B is larger compared to that of rooms C and D.

To test whether there was any difference in preference between both pairs, an independent group of participants either chose between rooms A and B (n = 30), or between rooms C and D (n = 32). The results of the pretest indicated that variation in trade off size did not change preference for an alternative (i.e., preference for the room with the larger room size was very similar for both pairs; Room A = 50%, Room C = 53%; χ^2 [1] = 0.06, p = .80).

After choosing between the pairs in the pre-test, participants rated choice difficulty on a six-point scale, ranging from 0 (not at all difficult) to 5 (very difficult). Results showed that difficulty was judged to be significantly higher when choosing between large (M = 4.53, SD = 2.79) compared to small (M =

3.03, *SD* = 2.69) trade off options, t (60) = 2.32, p = .02. The latter finding is a replication of Chatterjee and Heath (1996), who found that choice difficulty increased when trade off size was enlarged.

Procedure

The experiment was conducted on a laptop and participants performed the task according to their own pace.

The two pairs of rooms were presented simultaneously on the laptop screen. Both pairs were part of a cover story describing two persons, one who had to choose between rooms A and B (i.e., large trade off), and one who had to choose between rooms C and D (i.e., small trade off). The cover story did not refer to the pairs in terms of their trade off size.

A prior manipulation check was conducted, to ensure that participants noticed the difference in trade off size between the two pairs. To this end, half of the participants were asked to indicate for which pair the attribute values of both room size and commuting time were more dissimilar (implying which pair had a larger trade off size). The other half had to indicate for which pair the attribute values were more similar (implying which pair had a smaller trade off size).

Subsequently, all participants were told that the two persons, described in the cover story, experienced great difficulty in choosing between the respective pairs of rooms. Participants were assigned to one of two conditions. Participants in the "Defer" condition had to indicate which of the two persons (i.e., the one choosing between pair AB, or the one choosing between pair CD) was more likely to defer choosing, given the experienced difficulty. Participants in the "Coin" condition indicated which of the two persons was more likely to flip a coin in order to resolve the difficult decision. Order of presentation of both pairs of rooms on the screen was counterbalanced.

Results

Eleven participants were omitted from further analyses because, following the manipulation check, they perceived the rooms with the large trade off as more similar or vice versa. The number of participants judging the choice difficulty to be resolved by deferral or by coin flipping, as a function of trade off size, are portrayed in Table 1.

In the Defer condition, significantly more participants thought that the person deciding among rooms with a large trade off would defer choosing, compared with the person choosing between rooms with the small trade off, p = .006 by a binomial test. In the Coin condition, the reverse result was obtained: more participants indicated that the person choosing in the small trade off condition would flip a coin, compared with the person choosing in the large trade off condition, $p < .001$.

A discussion of these results is postponed to the discussion following Experiment 2.

Table 1
Number of participants judging choice difficulty to be resolved by deferral or by coin flipping, as a function of trade off size, in Experiments 1 and 2.

	Trade off size	
Difficulty resolved by	Large	Small
Rooms (Experiment 1)		
Deferral	41	19
Coin flipping	13	46
Printers (Experiment 2)		
Deferral	28	31
Coin flipping	8	48

Experiment 2

The aim of Experiment 2 was to replicate the results of the first experiment with different stimuli, namely printers.

Method

Participants

A total of 120 students from Tilburg University participated in the experiment and received 5 euros. The experiment was part of a few unrelated judgment and decision making experiments that lasted for approximately 30 minutes.

Design and procedure

Two pairs of printers were presented simultaneously on a laptop screen (see Figure 2). Similar to the first experiment, participants were assigned to either the Defer or Coin condition. Order of presentation of both pairs on the screen was counterbalanced.

Pretest

As in Experiment 1, the stimuli were pre-tested with an independent group of participants, to determine whether varying trade off size would result in a difference in preference between the two options of a pair. Approximately half of the participants (n = 32) chose between the printers having a large trade off (i.e., printers A and B), whereas the other participants (n = 30) chose between the printers with the small trade off (i.e., printers C and D). The results of the pretest showed that variation in trade off size did not change the alternatives' attractiveness (i.e., preference for the printer with the lower printing costs was similar for both pairs; Printer A = 84%, Printer C = 83%; χ^2 [1] = 0.01, p = .91).

Following the choice task, participants in the pretest were asked to rate choice difficulty on a six-point scale, ranging from 0 (not at all difficult) to 5

(very difficult). Congruent with Experiment 1, choice difficulty was higher when choosing between large (M = 3.72, SD = 2.32) compared to small (M = 3.13, SD = 1.78) trade off options, but non significant, $t(60) = 1.14$, $p = .26$.

	A	**B**
Printing costs	2.8 eurocent/page	5.2 eurocent/page
Printing speed	7 pages/minute	15 pages/minute

	C	**D**
Printing costs	3.7 eurocent/page	4.3 eurocent/page
Printing speed	10 pages/minute	12 pages/minute

Figure 2. Overview of the pairs of printers. Note that the trade off size for printers A and B is larger compared to that of printers C and D.

Results

Six participants were omitted from further analyses because, following the manipulation check, they perceived the printers with the larger trade off size as more similar or vice versa. The number of participants judging the choice difficulty to be resolved by deferral or by coin flipping, as a function of trade off size, are portrayed in the lower part of Table 1.

Unlike in Experiment 1, no significant difference between the large and small trade off size alternatives was found in the Defer condition, $p = .80$. In the Coin condition, significantly more participants thought that the person choosing between printers with a small trade off size would opt for flipping a coin, compared to the person choosing between the large trade off size options, $p < .001$. This latter result is congruent with, and further strengthens, the result obtained in Experiment 1.

Discussion

The first two experiments demonstrated, unequivocally, that a difficult decision was more likely to be resolved by a coin flip when trade off size was small rather than large. In other words, the acceptability of a coin flip was sensitive to the size of the trade off. With regard to deferral, however, the results were mixed. When facing a difficult choice between rooms (Exp. 1), deferral was more likely to occur when trade off size was large. Yet, deferral with printers (Exp. 2) showed no sensitivity to trade off size, which is consistent with Dhar (1997) who found similar percentages of deferral for both large and small trade off sizes.

Why is coin tossing sensitive to trade off size, whereas choice deferral might not be? One possibility is that random selection and deferral are applied for

resolving different types of choice difficulty. For instance, Scholten and Sherman (2006) identified two sources of difficulty, supposedly, contingent upon trade off size. Specifically, the experienced difficulty with larger trade off sizes is related to the sacrifices to be incurred by forgoing the advantages of the un-chosen alternatives, whereas difficulty with smaller trade off sizes stems from lacking reasons for choice justification. Indeed, the observed difference in sensitivity to trade off size between coin tossing and deferral, may be explained by introducing (and identifying) distinctive sources underlying difficulty that vary as a function of trade off size.

Elster (1989), for instance, suggested that people are willing to choose at random (e.g., by flipping a coin), when they have no compelling reasons for selecting an alternative. Accordingly, a coin toss is more likely to be accepted with small trade off size alternatives, because it will be difficult to construct compelling reasons for choice justification. Arguably, with larger trade off sizes, strong arguments in favor of an alternative can be made in case more weight is assigned to one of the alternatives' (prominent) dimensions. Concerning small trade offs, due to the small difference in attribute values, the effect of increased attribute weights on decisive arguments might be attenuated. For instance, consider a choice between printers characterized by printing costs and printing speed. A difference in printing costs of 3.7 and 4.3 eurocent/page, respectively, might be too small to use it as a convincing argument for choice justification, even if the importance of this attribute is strongly emphasized.

Regarding choice deferral, as a function of trade off size, Tversky and Shafir (1992) argued that people tend to defer their choice when they do not know how to trade off advantages against disadvantages. Yet, this concept of trade off difficulty may appear to be not well-defined. For instance, it is not clear whether it implies that difficulty stems from sacrifices to be incurred by trading off pros and cons, or that subsequent to such trade off, compelling reasons for justification are difficult to construct. Following Scholten and Sherman (2006), trade off difficulty indeed may encompass both difficulty stemming from sacrifices (large trade offs), as well as difficulty that is caused by lacking compelling reasons (small trade offs). Accordingly, trade off difficulty should be experienced irrespective of trade off size, and hence deferral should be accepted equally likely for both large and small trade offs.

The following experiment was designed to test whether choice difficulty, that depends on trade off size, can be better accounted for by the fact that it is hard (i) to justify one's choice by compelling reasons, or (ii) to trade off costs and benefits. It was predicted that choosing between small trade off size alternatives was difficult because of lacking compelling reasons for choice justification. The difficulty of trading off pros and cons was hypothesized to account for the choice difficulty experienced with both large and small trade offs.

Experiment 3

Method

Participants

A total of 145 students from Utrecht University were recruited and paid 3 Euros for their participation. The experiment was part of a few unrelated judgment and decision making experiments that lasted for approximately 20 minutes.

Design and procedure

The same two pairs of rooms as in Experiment 1 were presented simultaneously on the screen. The same manipulation, used in Experiment 1, was conducted prior to the main task, to ensure that participants noticed the difference in trade off size between the two pairs.

Participants were assigned to one of two conditions. Participants in the "Trade off" condition indicated which of two persons was more likely to explain the experienced difficulty by referring to difficulty in trading off the attributes of both alternatives of a pair. Participants in the "Reasons" condition indicated who, of the two persons, would most likely explain difficulty by referring to the difficulty in generating compelling reasons for the choice. Order of presentation of both pairs on the laptop screen was counterbalanced.

Table 2
The number of participants explaining choice difficulty by referring to trade off difficulty or lacking compelling reasons for choice justification (Exp. 3 and 4), as a function of trade off size.

	Trade off size	
Explanation	Large	Small
Rooms (Experiment 3)		
Trade off	39	34
Reasons	16	47
Printers (Experiment 4)		
Trade off	30	32
Reasons	11	58

Results

Nine participants were omitted from further analyses because, following the manipulation check, they perceived the rooms with the larger trade off size as more similar or vice versa. The number of participants explaining choice difficulty by referring to trade off difficulty or lacking compelling reasons for choice justification, as a function of trade off size, are presented in Table 2.

In the Trade off condition, explaining difficulty by referring to trade off resulted in no significant difference between the pairs, $p = .64$. In the Reasons condition, however, significantly more participants thought that the person choosing between the rooms having a small trade off would explain difficulty by referring to reasons for justification, as compared to the person choosing between the rooms with the large trade off, $p < .001$. Both results, in the Trade off and Reasons conditions, are consistent with the predictions.

Experiment 4

The aim of Experiment 4 was to replicate the results of the previous experiment, using different stimuli (printers).

Method

Participants

A total of 146 students from Tilburg University participated in the experiment and received 5 Euros. The present study was part of a number of unrelated judgment and decision making experiments.

Design and procedure

The experiment was identical to Experiment 3, except for replacing the stimuli with the two pairs of printers employed in Experiment 2. Participants were assigned to one of two conditions: the Trade off or Reasons condition. Order of presentation of both pairs on the laptop screen was counterbalanced.

Results

Fifteen participants were omitted from further analyses because, following the manipulation check, they perceived the printers with the larger trade off size as more similar or vice versa.

The number of participants explaining choice difficulty by referring to trade off difficulty or lacking compelling reasons for justification, are presented in the lower part of Table 2. The results of the present experiment replicated those of Experiment 3. Particularly, in the Trade off condition, no significant difference between the pairs was found, $p = .90$. In the Reasons condition, significantly more participants thought that the person choosing between the printers having a small trade off would explain difficulty by referring to reasons for justification, as compared to the person choosing between the printers with the large trade off, $p < .001$.

Discussion

Experiments 3 and 4 suggest that lacking compelling reasons for choice justification was more likely to be used as an explanation for choice difficulty with small rather than large trade off alternatives. Trade off difficulty, on the other hand, was equally likely to be used as an explanation for alternatives with

large, as well as small trade off. This latter finding is compatible with Scholten and Sherman (2006), who claimed that a decision among alternatives involving attributes, begins with identifying that these attributes require a trade off. They further proposed that trade off identification yields the experience of a conflict, regardless of trade off size. This may explain why in Experiments 3 and 4 of the present chapter, trade off difficulty was used as an explanation for the difficulty experienced with both large and small trade offs. According to Scholten and Sherman (2006), a decision maker will, subsequent to this trade off identification, investigate exactly how much of each attribute has to be traded off against how much of the other attribute(s). They argue that for small trade off sizes this investigation leads to the realization that choosing is difficult because it is hard to find good reasons for choice justification, the latter which is congruent with the findings in the Reasons conditions of Experiments 3 and 4.

In sum, difficulty stemming from lacking convincing arguments for choice justification is related to the size of the trade off (i.e., small trade off size), whereas the difficulty experienced in trading off pros and cons is rather insensitive to trade off size. Analogous to the sensitivity of decisive reasons to trade off size, resolving a difficult decision by the toss of a coin is mainly accepted with small trade off sizes. In that regard, despite that people often resent a coin toss (Elster, 1989), they might accept coin flipping for resolving a conflict in case they fail to construct convincing reasons for choice justification. On the other hand, deferral seems, similar to trade off difficulty, to be insensitive to the size of the trade off. Deferral, thus, may be adopted when choosing is difficult because of trade off difficulty (both large and small trade off size), as well as lacking compelling reasons for choice justification (small trade off size).

The aim of the following experiment was to test directly the relation between the two means for resolving a difficult decision (coin toss vs. deferral) and the source of the difficulty (trade off vs. reasons for justifying choice). Specifically, participants had to indicate whether a person would accept (i) a coin toss, or (ii) deferral, because it was either difficult to (i) construct compelling reasons for choice justification, or (ii) trade off advantages against disadvantages, respectively.

Experiment 5

Method

Participants

A total of 146 students from Tilburg University participated in the experiment and received 5 Euros. The present study was part of a number of unrelated judgment and decision making experiments.

Design and procedure

Stimuli were the pairs of rooms employed in Experiment 1. Unlike in experiment 1, however, each pair was shown on the laptop screen separately. Thus, participants were presented with either a large or a small trade off size pair. For rooms with a large trade off size, attributes values were explicitly described as differing considerably from each other. For rooms with a small trade off size, the attribute values were explicitly described as differing only slightly from each other.

Each pair of rooms was embedded in a cover story describing a person who experienced great difficulty while choosing between the two rooms. One group of participants (n = 74) was told that, as a result of the experienced difficulty, the person decided to defer the choice. The other group of participants (n = 72) was told that the person decided to flip a coin as a consequence of the experienced difficulty. Participants in both conditions were subsequently asked whether they thought the person would defer choosing / flip a coin, because it is hard to (i) trade off the advantages and disadvantages of both rooms against each other, or (ii) justify a choice for one or the other room with compelling reasons.

In sum, trade off size (large vs. small) and way of resolving choice difficulty (deferral vs. coin flipping) were manipulated in a 2 X 2 between-subjects design. The order of the two answer options (explanation by trade off difficulty or lack of reasons) was counterbalanced.

Results

The results of Experiment 5 are presented in Table 3.

Table 3
Results of Experiment 5: given that difficulty was resolved by either deferral or coin flipping, which type of justification (trade off or reasons) was more likely?

	Deferral		Coin flipping	
Trade off size	Trade off	Reasons	Trade off	Reasons
Large	27	11	28	9
Small	19	17	17	18

Deferral

I first examine the responses of participants who were told that the difficult decision was resolved by deferral. The majority of participants exposed to a large trade off, thought that it was more likely that someone would defer because it is hard to trade off advantages against disadvantages, p = .01. In the small trade off condition, however, deferral was considered to be equally likely

a result of trade off difficulty or lacking compelling reasons for choice justification, $p = .89$.

Coin flipping

The results of resolving choice difficulty by flipping a coin were congruent with those obtained with deferral. Specifically, in the large trade off condition, participants thought that it was more likely that someone would toss a coin because it was hard to trade off of the advantages and disadvantages of both rooms, $p = .003$. In the small trade off condition, it was equally likely that flipping a coin was a result of either trade off difficulty or lacking compelling reasons for choice justification, $p = 1.00$.

Discussion

The results of Experiment 5 showed that acceptability to resolve a difficult decision with large trade off sizes by deferral, is mainly based on the difficulty of trading off advantages and disadvantages. The acceptability to resolve a difficult decision with small trade off size by deferral, is equally likely to be based on the difficulty of trading off advantages and disadvantages or the inability to find good reasons. Thus, deferral is, from a descriptive viewpoint, an adequate means for resolving choice difficulty stemming from either trade off difficulty, as well as lacking convincing arguments for choice justification.

Coin flipping, as expected, was considered acceptable with small trade off sizes in case difficulty originated from lacking compelling reasons for choice justification. Yet, a coin toss was also accepted with small trade off sizes when it was difficult to trade off advantages against disadvantages. Thus, it appears that a conflict caused by a difficult trade off allows for a coin toss.

Regarding a coin toss with large trade off size alternatives, the first two experiments demonstrated that it is improbable that a coin toss will be accepted for resolving a difficult decision among large trade off size alternatives. The design of the present experiment, however, contained a condition in which a difficult decision, stemming from a large trade off, was resolved by a coin toss. Although improbable, this condition demonstrated, once again, that the difficulty with a large trade off size originated not from lacking compelling reasons for choice justification, but from the difficulty in trading off pros and cons.

What are the reasons for the decrease in accepting a coin toss with larger trade off sizes? Scholten and Sherman (2006) argued that a decision maker will first identify that a decision requires a trade off among attributes. This trade off identification is followed by scrutinizing how much of an attribute has to be traded against how much of another attribute. The result of this inspection is that a larger trade off signifies that a larger sacrifice is to be incurred. This sacrifice refers to forgoing a considerable advantage associated with the un-chosen alternative. Intuitively, a larger sacrifice will yield the experience of (potential) regret to a person if the chosen alternative ends up being worse than the un-chosen one. This anticipated regret may be avoided when a decision maker defers the choice. Yet, tossing a coin inevitably commits the decision

maker to one of the alternatives under consideration. Because a decision maker may want to avoid the anticipated regret, the reluctance to toss a coin will increase as trade off gets larger.

The following experiment examined whether regret, as an explanation for decision difficulty, was more likely with larger rather than smaller trade offs.

Experiment 6

Method

Participants

A total of 144 students from Radboud University in Nijmegen were recruited on campus and paid 5 Euros for their participation. The present study was part of a number of unrelated judgment and decision making experiments.

Procedure and design

The same two pairs of rooms and printers from the first two experiments were employed. The two pairs, of either rooms or printers, appeared simultaneously on the laptop screen. The manipulation check used in Experiment 1 was conducted to ensure that participants noticed the difference in trade off size between the pairs.

Subsequently, two persons were introduced who had to choose between the respective alternatives of a pair. One person had to choose between large trade off size alternatives, the other between small trade off size alternatives. Both persons were said to experience great difficulty while choosing. Participants had to indicate which of the two was more likely to explain the difficulty by arguing that if the chosen alternative turned out to be poor, it will lead to regret of not having chosen the other alternative.

Participants were presented with either rooms or printers. The order of presentation of the two pairs (either large or small trade off) was counterbalanced.

Results

Eleven participants were omitted from further analyses because, following the manipulation check, they perceived the pair with the large trade off size as more similar or vice versa. The number of participants who thought that choice difficulty was explained by (anticipated) regret, as a function of trade off size, are presented in Table 4.

The results show that, for both rooms and printers, the likelihood to use anticipated regret as an explanation for decision difficulty increased as trade off size increased. Aggregating the results (Tversky & Kahneman, 1971), showed that more participants (79 out of 133) thought that anticipated regret was more likely used as an explanation for the choice difficulty experienced with large trade off size pairs, $p = .04$.

Table 4
Results of Experiment 6: the number of participants explaining choice difficulty by referring to regret, as a function of trade off size.

	Trade off size	
Regret as explanation	Large	Small
Rooms	40	25
Printers	39	29

Discussion

The results of Experiment 6 suggest that regret was more likely to be used as an explanation for choice difficulty when choosing between large trade off size pairs. This, supposedly, may explain the reluctance to toss a coin with larger trade offs. Because larger trade offs imply that more sacrifices have to be incurred, a decision maker may regret the obtained alternative, once this ends up being worse than the other available alternative(s). This potential regret is inevitable in case of random assignment. Yet, the sacrifices are less considerable with smaller trade offs. Accordingly, anticipated regret is less associated with smaller trade offs, and the toss of a coin (i.e., random assignment) will become a more acceptable way of resolving a difficult decision.

The previous experiments in this chapter, suggested that deferral and coin flipping are possible means for resolving difficult decisions. Yet, whether or not these two methods are adopted, would depend on trade off size. Table 5 summarizes the conditions under which both means are considered acceptable for resolving a difficult decision.

Coin flipping is more likely to be accepted for resolving a difficult decision among small rather than large trade off alternatives. With small trade off sizes, the difficulty mainly stems from lacking compelling reasons for choice justification and the difficulty in trading off pros and cons, and coin flipping is acceptable for resolving difficulty originating from both these sources. As discussed above, anticipated regret seems to be a potential reason for not accepting a coin toss with larger trade off sizes.

Deferral, in contrast, may be adopted irrespective of trade off size. Deferring choice is used as a way of resolving conflict when it is difficult to trade off pros and cons. Besides trade off difficulty, however, the intention of deferral is avoiding anticipated regret when the alternatives' trade off size is large. Difficulty in constructing compelling reasons for choice justification may lead to deferral when the trade off size is small.

Table 5
Conditions under which deferral and coin flipping are accepted for resolving a difficult decision among multi attribute alternatives, as a function of trade off size.

	Trade off size	
Difficulty resolved by:	Large	Small
Deferral	- Trade off difficulty - Anticipated regret	- Trade off difficulty - Lacking reasons
Coin flipping	X	- Trade off difficulty - Lacking reasons

If a decision maker has to choose between either deferral or coin flipping, are both ways for resolving a difficult decision adopted to the same extent? The previous experiments in this chapter *separately* examined the likelihood that deferral, or coin flipping, will be adopted for resolving a difficult decision, given trade off size. In the following experiment the two means were presented *simultaneously* to participants, such that they had to choose between deferral and coin flipping for resolving a difficult decision among two alternatives.

Participants were presented with either a large or small trade off pair, instead of the simultaneous presentation of the pairs, used in the previous experiments (with exception of Experiment 5). Presenting large and small trade off pairs *separately* should extend the previous experiments for two reasons. First, separate and joint (simultaneous) presentation may not always yield comparable results (Hsee, 1996). Second, when presenting the large and small trade off pairs jointly, instead of separately, small trade off options may look like compromise options in comparison to the large trade off options. Simonson (1989) showed that people prefer compromise options, with attribute values falling in between more extreme options. Note that when large and small trade off pairs are presented jointly, the large trade off options have extreme attribute values, whereas the small trade off options have values that fall exactly in between these extremes (see Figure 3).

In that respect, the small trade off options constitute compromise options. A consequence of the simultaneous presentation of both pairs, however, may have been that participants assumed that besides large trade off options, the small trade off (compromise) options were also available on the market. An explanation for deferral with large trade off options, therefore, could be that participants assumed that in addition to large trade off pairs, desirable compromise options were also available. Deferral in case of large trade off options presumably gives them the chance for further search and ultimately lead them to these compromise options (i.e., small trade off options). When presented separately, participants are not aware of the small trade off options

that might be available on the market, which could serve as compromise options.

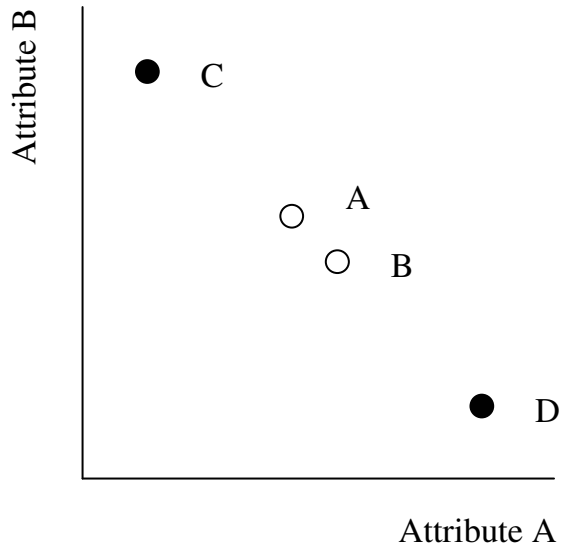

Figure 3. Example of two large (C and D) and small (A and B) trade off options. Options C and D have extreme values on both attributes, whereas A and B have values falling in between these extremes. Since A and B values' are midway those of C and D, the small trade off options may be interpreted as compromise options.

Experiment 7

Method

Participants

A total of 264 students from Fontys University of Professional Education Eindhoven participated in the experiment and received either 4 or 5 Euros. The present study was part of a number of unrelated judgment and decision making experiments.

Design and procedure

The same pairs of rooms and printers from the first two experiments were employed. Each participant was presented with one pair of either rooms or printers, having either a large or small trade off size. The attribute values of a large trade off size pair were described to the participants as differing substantially from each other. Conversely, the attribute values of a small trade off size pair were described as differing little from each other.

A cover story described a person who was said to experience great decision difficulty when choosing between the respective two options of a pair. Participants subsequently indicated whether the person, given the experienced decision difficulty, is more likely to resolve difficulty by (i) deferral, or (ii) flipping a coin.

Results

The number of participants judging choice difficulty to be resolved by deferral or by coin flipping, as a function of scenario and trade off size, are portrayed in Table 6.

Table 6
The number of participants in Experiment 7, judging the choice difficulty to be resolved by deferral or by coin flipping, as a function of trade off size.

	Difficulty resolved by:	
Trade off size	Deferral	Coin flipping
Rooms (n = 135)		
Large	57	12
Small	44	22
Printers (n = 129)		
Large	52	10
Small	43	24

Rooms

Irrespective of trade off size, 75% of the participants (101 out of 135) indicated that experienced decision difficulty was more likely to be resolved by deferral, compared with 25% (34/135) who thought that difficulty would more likely be resolved by a coin flip. A significant interaction was found between trade off size and manner of resolving difficulty, demonstrating an increase of coin flipping with small trade off sizes, χ^2 (1) = 4.55, p = .03.

Printers

The result with the printer pairs were congruent with those of the rooms. Specifically, irrespective of trade off size, 73% (95/129) of the participants thought that the experienced difficulty was more likely resolved by deferral, compared with 28% (34/129) who thought that difficulty would more likely be resolved by coin flipping. A significant interaction was found between trade off size and way of resolving difficulty, χ^2 (1) = 6.43, p = .01.

Discussion

Two conclusions can be drawn from Experiment 7. First, deferral is overall more accepted than coin flipping for resolving choice difficulty. Second, the means of resolving choice difficulty interacts with trade off size.

Separate presentation of the large and small trade off size pairs yielded results congruent with presenting the pairs jointly. Specifically, in case of simultaneous presentation, coin flipping was more likely employed for resolving difficulty induced by small trade off size pairs. Similarly, when the pairs were presented separately, an increase of coin flipping, as a way of resolving choice difficulty, was found for small trade off options.

In case of separate presentation, however, deferral is overall more preferred as a way of resolving a difficult choice. Yet, in case of small trade off options, coin flipping may increasingly replace deferral as means for resolving a difficult choice. Still, overall, these results demonstrate that decision makers may resent a coin toss (compared to deferral), which again confirms that this means is rarely employed for resolving a difficult decision (Elster, 1989).

General discussion

When a person faces a choice between alternatives defined on several attributes, deciding may require trading off costs against benefits. The experience of a difficult decision may arise both when the size of the trade off is large or small. That is, subsequent to the (primary) recognition of a difficult trade off, two (second-order) sources underlying the experienced difficulty may be identified (Scholten & Sherman, 2006). First, larger trade off sizes imply that significant sacrifices are to be incurred. These substantial sacrifices may lead to regret, once a chosen alternative turns out to be worse than the rejected one. Second, the difficulty with smaller trade off sizes emanates from lacking compelling reasons for choice justification.

Choice deferral or random selection (e.g., by a coin toss) are possible means for resolving difficult decisions. Deferral may be adopted both when a decision maker is anticipating regret (large trade off size) as well as lacking compelling reasons for choice justification (small trade off size). Consequently, deferral is equally likely to occur for large and small trade off sizes. Coin flipping, on the other hand, is an acceptable way of resolving difficulty when no compelling reasons for choice justification exist. Accordingly, people are more willing to toss a coin when a choice involves small trade off size alternatives. The reluctance of tossing a coin with large trade off sizes is further due to the regret anticipated. Intuitively, people will not toss a coin if they think the outcome might induce anticipated regret.

Both deferral and coin flipping can thus be viewed as means for resolving difficult choices. Besides their function with regard to resolution, they might also be used as a measurement instrument indicating the degree of experienced difficulty. For instance, if one alternative dominates the other, a decision maker is not likely to accept choice deferral, or alternatively, a coin toss. Yet, the

same decision maker is more likely to accept one of these means in case of equally attractive alternatives. It is important to note, however, that if both are employed as a measure of difficulty, it does not imply that a decision maker would apply these methods in reality. As measurement tools, the two investigate acceptance, and not the active use of deferral or coin tossing. In that sense, deferral or coin tossing only serve as a measure of difficulty.

Besides measuring the degree of experienced difficulty, can deferral and coin tossing reveal the underlying source(s) of the difficulty (e.g., anticipated regret or lacking compelling reasons)? Deferral seems ambiguous with regard to which factor caused the experienced difficulty (Anderson, 2003). Dhar (1997), for instance, avoided the problem of identifying the source from which difficulty originated, by introducing the concept of preference uncertainty. He argued that equally attractive alternatives induce uncertainty about which alternative to prefer. As a result, the decision to defer increases compared to when the alternatives would differ in attractiveness. Notice, though, that the concept of preference uncertainty is ambiguous, in the sense that it remains mute about whether difficulty stems from anticipated regret or lack of compelling reasons for choice justification.

Coin flipping may be more informative regarding the source of the difficulty. As a means for resolving difficult decisions, accepting random selection increases with smaller trade off sizes. With larger trade off sizes, a decision maker resents tossing a coin because of the sacrifices to be incurred, which may yield the experience of (anticipated) regret. This feeling of anticipated regret can not be avoided if the coin randomly assigns the decision maker to one of the available alternatives. As a measurement tool, therefore, whether or not a coin flip is accepted, may indicate from which source the difficulty originated, namely lacking compelling reasons for choice justification (accepting a coin toss) or anticipated regret (not accepting a coin toss). To illustrate, consider a person who is faced with a difficult decision among alternatives, defined on two distinctive attributes. Suppose you do not know the size of the trade off between these attributes. The decision to defer will not inform you whether the person experiences difficulty stemming from anticipated regret or lacking reasons. The decision to toss a coin, however, suggests that the person is lacking compelling reasons for choice justification, whereas the decision not to toss a coin suggests that anticipated regret is considerable.

When choice deferral and coin flipping are employed as measurement tools they, presumably, reveal source(s) that caused the choice to be difficult. Yet, in case the two are used in a context where alternatives vary in trade off size (i.e., present chapter's experimental designs), how likely is it that the experienced difficulty was stemming from anticipated regret or lacking compelling reasons, and not from some other underlying source(s)? Put differently, if deferral or coin tossing are used as measurement tools, when trade off size is varied, how valid is the conclusion that factors related to trade off size (Scholten & Sherman, 2006), and not some other external or internal factors (Chapter 1), caused the decision to be difficult? As an example, imagine a person who is

wearing an umbrella. What inference(s), about why the person wears an umbrella, are plausible? Most likely it will be raining, but alternative causes – say, sun burn, or sprinklers in a garden– may not be excluded. In other words, when is it plausible to conclude that a specific antecedent, and not some other factor(s), is most likely causing the observed behavior?

All experiments in the present chapter (with exception of Experiment 5 and 7) presented simultaneously two pairs of alternatives to the participants. These pairs' alternatives were defined on two identical attributes (e.g., room size and commuting time), who's attribute values were varied in trade off size. By presenting the pairs jointly, factors that may be characterized as external to the alternatives (Chapter 1), such as number of choice alternatives in a choice set, or importance of the alternatives to the decision maker, were similar (i.e., constant) for both pairs. Also, internal factors, such as number of attributes involved, or attributes' emotional level, were identical for both pairs. Because these (external and internal) factors were constant when trade off size changed, they can be excluded as possible sources that explain the difference in difficulty between the pairs. The only factors underlying difficulty that, presumably, varied between the presented pairs, were those contingent upon trade off size (i.e., anticipated regret and lacking compelling reasons for choice justification). Consequently, these factors should be considered in explaining why the decision among the pairs was perceived as difficult. In other words, when deferral and coin flipping are used as measurement tools, in the context of joint presentation of pairs with different trade off sizes, it will be valid to conclude that these two reveal sources of difficulty that are related to trade off size.

The observation that people are more likely to accept a coin toss with small trade off sizes, may also result from the alternatives being very similar (and hard to discriminate from each other). One could argue that, therefore, it does not matter which alternative is obtained, and hence explains why random choice (by flipping a coin) is acceptable with small trade offs. Could it be that such indifference between similar alternatives, instead of lacking compelling reasons, makes that someone is willing to toss a coin? In the experiments described in the present chapter, however, the person choosing between the small trade off alternatives was said to experience great difficulty. It seems rather implausible that one will characterize a choice as being difficult, if it does not matter which alternative is obtained. Moreover, when someone claims that choosing is very difficult, it is more likely that it mattered, to this person, which alternative was obtained and, as a result, she or he supposedly went through a choice process in which attributes were traded off. The outcome of this choice process was, presumably, a difficult to resolve conflict because the person could not find compelling reasons for justification. For that reason, using coin tossing as a measurement tool, in the context of a difficult choice between small trade off alternatives, is supposed to reveal that the difficulty was stemming from lacking compelling reasons for choice justification (and not indifference).

Besides (indirect) measures of difficulty as deferral and coin tossing, it is also possible to measure difficulty directly. For instance, Chatterjee and Heath (1996) employed a rating scale on which participants had to indicate how difficult they thought choosing between alternatives was. The direct measure shows an increase in difficulty with larger trade off sizes. By itself, however, this measure is not informative about the source of the difficulty. Although it correlates with measures of regret (Chatterjee & Heath, 1996; Iyengar & Lepper, 2000), compared to deferral and random selection, the direct measure seems to be rather insensitive in situations where difficulty stems from lacking compelling reasons for choice justification (i.e., small trade offs). Yet, this insensitivity to the difficulty caused by a lack of decisive arguments, may also be explained by arguing that, to decision makers, the sacrifices to be incurred (large trade offs) contribute, more than lacking compelling reasons (small trade off size), in the experience of choice difficulty. This emphasis on sacrifices by decision makers, in turn, may explain why the direct measure yields an increase in difficulty with larger trade off sizes.

3

Resolving difficult choices between uni-dimensional alternatives by a coin flip: The effect of different display organizations on choice difficulty and the construction of compelling reasons

Abstract

This chapter considers difficult decisions among alternatives defined on a single dimension (e.g., job applicants evaluated by two judges on one and the same rating scale), and investigates the acceptability of coin flipping in relation to lacking compelling reasons for choice justification. A coin toss is more likely employed for resolving a difficult decision, when the decision maker finds it difficult to construct decisive arguments for justification. These decisive reasons, supposedly, refer to the discrepancies in the alternatives' values on the single dimension (i.e., observable difference in, for instance, the ratings of the two judges for each individual job candidate) (Experiment 1). That is, a decision maker may have stronger reasons in favor of an alternative, if the displayed discrepancy of an alternative on the single dimension is smaller compared to that of another alternative. The apparent discrepancies, however, were expected to be less noticeable when a different display organization, of exactly the same values, was applied. It is hypothesized that particular display organizations will effect the likelihood of coming up with decisive reasons, and in turn, would effect the extent to which coin tossing would be accepted. Empirical evidence is presented supporting this hypothesis (Experiments 2 and 3).

Decision makers often try to justify their preference for a chosen option by invoking convincing and persuasive reasons. Shafir, Simonson, and Tversky (1993), for instance, proposed that a choice between alternatives is often guided by reasons for or against the respective alternatives. Decision makers will seek and construct such reasons, which they subsequently use to justify their preference to themselves and to others. In some cases, it will be easy to come up with convincing reasons. Consider, for example, a choice between two similar job candidates, one who has both better social skills and is far more motivated than the other. A preference for the former, dominating, candidate is easy to justify ("compared to the other candidate, she has both better social skills and is more motivated"). Now imagine a close decision among two equally attractive candidates: one who has better social skills, and the other who is more motivated. In this case, there is no dominating candidate. Choosing between them, presumably, requires a trade off among attributes, as well as finding compelling reasons to justify a preference for one or the other candidate. For these close decisions, though, such trade off and the finding of decisive arguments, will often be perceived as difficult. Yet, when forced to choose between such equally attractive alternatives, people feel the need to come up with some convincing reasons that would justify their preference, even when such reasons do not exist (Elster, 1989). For instance, Slovic (1975) presented participants with pairs of alternatives, differing on two dimensions. Participants first equated (i.e., matched) the two alternatives and, at a later stage, chose which of the two equally attractive alternatives they preferred. Since the participants matched the two alternatives, so that they would be equally attractive to them, one would expect that both options would be equally preferred by the participants. However, their choices deviated systematically from this expectation: participants chose the option that was superior on, what they thought to be, the more important (prominent) dimension. Slovic explained this systematic deviation by arguing that choices in favor of the option superior on the more important dimension are easier to justify.

There are other ways, besides focusing on the most important dimension, to create convincing arguments in favor of an alternative. Festinger (1964) proposed that difficult decisions among equally attractive alternatives are resolved by spreading apart their attractiveness. This "spreading", of a priori equally attractive options, takes place once a decision is made, and changes are made such that the preferred option gains in attractiveness, while the rejected alternative's attractiveness decreases. Festinger reasoned that, as a consequence of this spreading of attractiveness, decisions are easier to justify in retrospect.

Justifying a choice by compelling arguments, thus, seems to be a reasonable way of resolving a difficult decision involving close alternatives. Choosing the alternative superior on the more prominent dimension (Slovic, 1975), or spreading apart the attractiveness of the available alternatives (Festinger 1964), may provide the decision maker with such compelling reasons. In the absence of convincing arguments, however, a difficult decision may be resolved by deferring the choice between alternatives (Tversky & Shafir, 1992; Dhar, 1997), with the hope of finding a better alternative. Someone who opts for

deferral, however, runs the risk of foregoing currently available alternatives (that may not be available later). In circumventing this loss of available options, a difficult decision among equally attractive alternatives might be resolved in yet another way, namely by tossing a coin.

When flipping a coin, a decision in favor of one or the other alternative is not determined by compelling reasons. Instead, the coin decides arbitrarily which alternative is obtained. Hence, if a decision maker fails to find compelling reasons in favor of one of the available options, flipping a coin might be an acceptable means for resolving the experienced difficulty. Yet, when confronted with a difficult choice, decision makers often dislike to use a random process, such as coin tossing, mainly because it implies renunciation of control, which may lead to undesired feelings, such as pessimism and distress (Skinner, 1996). Moreover, Elster (1989) explains this aversion by arguing that people want to have their choices determined predominantly by compelling reasons. He further contends that random choice brings along the possibility that they become assigned to an option that they, once obtained, might regret. Elster claims that, motivated to reduce this anticipated regret, decision makers try hard to come up with compelling reasons clearly favoring one of the available alternatives. Evidently, these feelings of anticipated regret intensify when the size of differences in attribute values increases (Chatterjee & Heath, 1996). At the same time, arguments in favor of an alternative, become stronger with increasing differences in attribute values (Chapter 2; Scholten & Sherman, 2006).

Overall, these latter findings are congruent with Elster's argument, namely that difficult choices imply a relation between potential regret and the need to come up with compelling reasons. Furthermore, aversion to a coin flip is more likely when compelling arguments can be made in favor of an option. In that case, increasing trade off would result in a declining preference for using a coin to replace the decision maker. Smaller trade offs, on the other hand, for which it seems to be rather difficult to come up with convincing reasons, should elicit an increasing acceptance for a coin toss. Support for this relation between trade off size, compelling arguments, and the willingness to flip a coin was shown in Chapter 2.

Multi versus uni-dimensional decision problems

Decision problems predominantly involve alternatives characterized by two or more distinctive attributes, that are represented dimensionally (e.g., job candidates described by motivation and social skills). Typically, a difficult choice implies some trade off between these dimensions, resulting in a conflict. Providing compelling reasons by, for instance, emphasizing the importance of one of the dimensions, may resolve this trade off conflict. Sometimes, however, people are confronted with stimuli described on only one dimension, for example, job candidates receiving (overall) ratings from two independent judges (see Figure 1). Compared to stimuli characterized on various distinctive dimensions, these uni-dimensional stimuli are defined on several dimensions that are converged on, and summarized by, a single value. Due to this uni-

dimensional character, it is rather impossible to resolve a difficult conflict by emphasizing the importance of one of the attributes, simply because there is only one attribute. However, because people are often in need for seeking reasons that justify their decisions, one may wonder how they will justify a difficult choice between equally attractive uni-dimensional stimuli. Further, if they are able to find compelling reasons, is it, as with multi-dimensional stimuli, easier to generate compelling reasons when the differences on the single dimension are large, rather than small? In addition, are decision makers more likely to toss a coin when the differences on the single dimension are small?

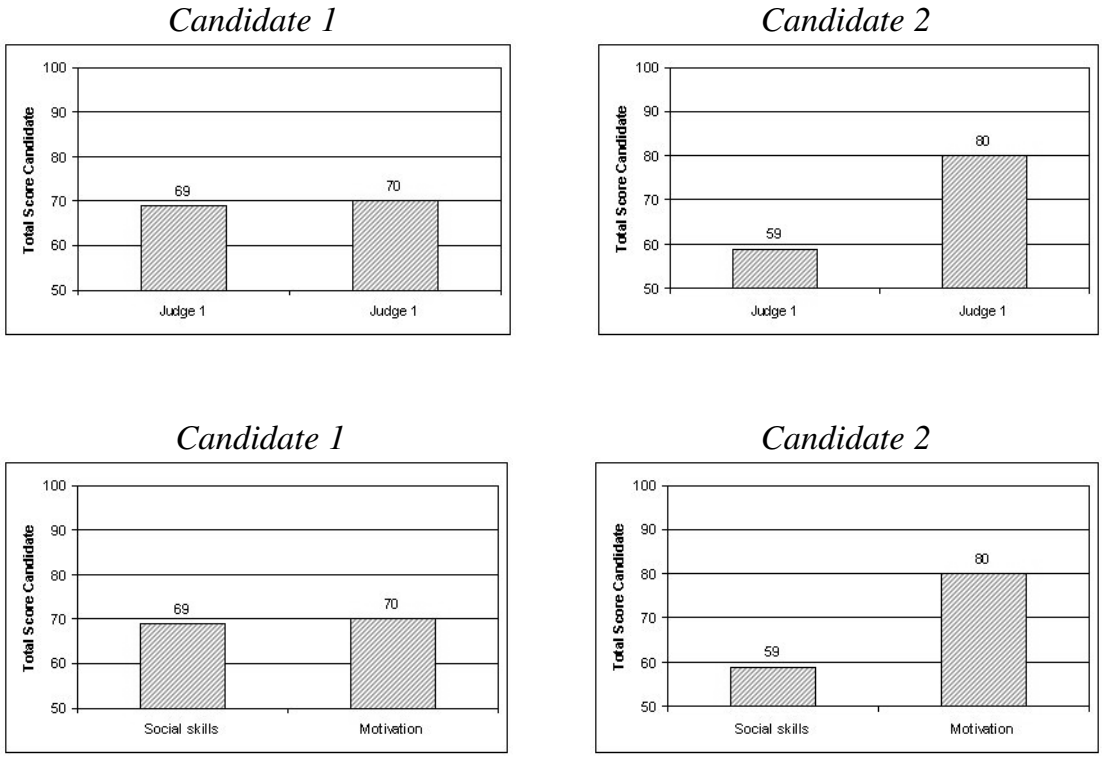

Figure 1. Ratings given to job candidates, being either uni-dimensional (top) or multi-dimensional (bottom), having either low (left) or high (right) variance.

A prerequisite for a difficult choice between uni-dimensional stimuli, however, is that alternatives would initially be viewed as equally attractive (e.g., Tversky & Shafir, 1992). Therefore, the following experiment was designed to test equal attractiveness among the alternatives. Scenarios were employed in which two job candidates were rated by a committee consisting of two independent judges. The variance of the judges' ratings was manipulated (cf. Ganzach, 1996), such that they either varied substantially (high variance) or were more or less similar (low variance). The purpose of this manipulation

was to create uni-dimensional stimuli that, at first sight, are analogous to multi-dimensional stimuli having either a large or small trade off size, respectively. Figure 1 displays examples of job candidates with high and low variability in their ratings.

Experiment 1

Method

Participants

A total of 71 students from Utrecht University were recruited on campus and paid € 5 for their participation. The present study was part of a number of unrelated judgment and decision making experiments.

Materials

Four pairs of uni-dimensional alternatives were constructed (see Table 1). Each pair of ratings was represented by two separate bar-graphs, displaying the job candidates' overall ratings as assigned by two independent judges. Ratings by the judges were either dissimilar, portraying high variance (right side of Figure 1) or similar, portraying low variance (left side of Figure 1). Averages of the assigned ratings of both candidates within a pair were identical, being 69.5 for one pair and 86.5 for the other pair.

The bar-graphs of both candidates appeared simultaneously on the screen of a laptop. A description was included, instructing participants to imagine they were in the role of a personnel officer, who had to decide which of the two candidates would be chosen at the final stage of a selection procedure. The two candidates were said to be interviewed and rated independently by two judges. The average ratings of both candidates were included in the description. It was explicitly emphasized that the two candidates' averages were identical.

Procedure

The experiment was conducted on a laptop. Participants were instructed to perform the task at their own pace.

Two different pairs of job candidates were presented (on separate screens) while performing the task. The two pairs were a combination of one low and one high variance pair, as displayed in Table 1. The average rating of this low and high variance pair was different. For approximately half of the participants (n = 36), the low variance pair consisted of two job candidates having an average rating of 69.5, whereas the high variance pair's average rating was 86.5 (i.e., pairs 1 and 4 in Table 1). The other participants (n = 35) were presented with pairs, of which the low variance pair had an average of 86.5, and the high variance pair an average of 69.5 (i.e., pairs 2 and 3 in Table 1). The two pairs were displayed one after the other, with no other tasks in between. The order of presentation of the low and high variance pairs was counterbalanced.

Table 1
Ratings assigned by the judges, to each of the candidates, in the low and high variance conditions.

		Candidate 1		Candidate 2	
	Variance	Judge 1	Judge 2	Judge 1	Judge 2
Average rating 69.5					
Pair 1	Low	69	70	71	68
Pair 2	High	59	80	77	62
Average rating 86.5					
Pair 3	Low	86	87	88	85
Pair 4	High	78	95	92	81

Participants were informed, prior to presentation, that information regarding two job candidates was going to be displayed simultaneously on a screen, and that their task was to rate both of these candidates on a six point scale, ranging from 0 (not at all attractive) to 5 (very attractive). After having rated the first pair, participants were told that they subsequently would have to rate a second pair. The instructions explicitly said that this second pair would be different from the first pair, without referring to the displayed size of the variance in the given ratings.

To assess whether participants noticed the differences in variance (i.e., low vs. high) between the two pairs, a manipulation check was conducted. Specifically, immediately after having rated the attractiveness of the candidates in a given pair, participants rated their similarity on a six point scale, ranging from 0 (not at all similar) to 5 (very similar). It was assumed that lower variance would correspond with a higher perceived similarity.

Results and discussion

Manipulation check

Consistent with the prediction, similarity was rated higher for low ($M = 3.96$, $SD = 0.98$) than for high variance pairs ($M = 3.00$, $SD = 1.16$), $t(70) = 6.53$, $p < .001$.

Attractiveness ratings

Table 2 portrays the attractiveness ratings.

Table 2
Mean attractiveness ratings for each candidate of a pair (standard deviations in parenthesis).

		Mean Attractiveness	
	Variance	Candidate 1	Candidate 2
Average rating 69.5			
Pair 1	Low	2.94 (0.95)	2.75 (0.87)
Pair 2	High	2.37 (0.91)	3.09 (0.78)
Average rating 86.5			
Pair 3	Low	3.74 (0.95)	3.31 (0.99)
Pair 4	High	3.03 (0.81)	3.58 (0.60)

In the condition in which the judges' average rating was 69.5, participants rated one of both candidates consistently as more attractive than the other. In the low variance condition, Candidate 1 was more attractive, $t(35) = 2.02$, $p = .05$. In the high variance condition, Candidate 2 was preferred, $t(34) = 3.05$, $p = .004$. Similar results were obtained when the judges' average ratings were 86.5 (low variance: $t[34] = 2.26$, $p = .03$; high variance: $t[35] = 3.16$, $p = .003$). Thus, one candidate, in each pair, was always rated as significantly more attractive.

The difference in attractiveness between candidates can be explained by focusing on the discrepancies in the ratings that were assigned by the judges. Specifically, of both candidates within a pair, the ratings of one candidate always exhibit less *discrepancy between judges* than those of the other. For instance, consider the high variance pair of candidates who received an average rating of 69.5 (i.e., Pair 2 in Table 1). The discrepancy between the judges for the first and second candidate were 21 (80 – 59) and 15 (77 – 62), respectively. The attractiveness ratings seem to reflect this difference in magnitude between the discrepancies: a candidate displaying less discrepancy was consistently rated as more attractive. Accordingly, for the candidates with a discrepancy in ratings of 21 and 15, respectively, the latter candidate was rated as being more attractive than the former.

In sum, Experiment 1 demonstrated that when candidates' ratings were displayed in bar-graphs, and one candidate had a smaller discrepancy in the assigned ratings than the other, then the two candidates were not judged as equally attractive. Evidently, the candidate associated with a smaller discrepancy will be the preferred one. Decision makers, in turn, may explain this dominance by invoking reasons referring to the discrepancy in the judges' ratings. For instance, less discrepancy in a candidate's ratings may be explained by arguing that this candidate was less ambiguous, and more reliable, while being interviewed by the judges. Hence, the candidate received more or less similar ratings from both judges. Once being guided in their decisions by such

compelling reasons, inferred from the displayed discrepancies, participants are likely to be reluctant to toss a coin when choosing between these candidates.

The above discrepancy in ratings, however, becomes less noticeable when the same ratings are displayed differently, as illustrated in the lower part of Figure 2. Specifically, in Experiment 1 the candidates' ratings were presented such that those of Judge 1 and Judge 2 were grouped together in a single bar-graph (see top of Figure 2). It was proposed that this kind of presentation, with ratings *grouped per candidate*, made the discrepancy between judges rather salient. The bottom of Figure 2, however, displays the same ratings, yet organized differently, namely *grouped per judge*. This latter kind of organization makes the discrepancy between judges less salient. Whereas grouping per candidate tends to emphasize the *discrepancy between judges*, grouping per judge shifts the focus towards the *discrepancy between candidates*. Furthermore, what stands out when ratings are grouped per judge, is that the magnitudes in discrepancy for both job candidates are identical. To illustrate, in Figure 2, the discrepancy of Judge 1 in the candidates' ratings is 18 (77 – 59), and similarly, the discrepancy for Judge 2 is 18 (80 – 62). Although the ratings grouped per judge show that Judge 1 gave somewhat lower ratings to both candidates than Judge 2, the discrepancies in the ratings, as given by both judges, were identical, namely 18.

May a different display organization influence the construction of compelling reasons and the likelihood to toss a coin? It was suggested that the construction of convincing reasons may be triggered by the difference in discrepancies, as reflected from both bar-graphs presented on the screen. Supposedly, the larger this difference, the more convincing the reasons are in support of a preferred alternative. The salience of a difference in discrepancy, however, is likely to depend on the display organization of the same ratings. When ratings are grouped per candidate, any difference in discrepancy (between judges) may appear to be rather prominent. The construction of compelling reasons may, therefore, not be difficult, and accordingly, a decision maker will be reluctant to toss a coin. In contrast, when ratings are grouped per judge, the discrepancies (between candidates) are identical, which will make it difficult to justify a preference for one or the other candidate with convincing reasons. As a consequence, the likelihood to toss a coin should increase when the ratings are grouped per judge, compared to when the grouping is per candidate.

Notwithstanding, for a direct comparison and evaluation of the two candidates, grouping per candidate might be a more natural, and hence a more suitable display organization, than having the same ratings grouped per judge. Yet, it is not unusual to receive the ratings of a candidate from two independent judges separately, such that these ratings are grouped per judge. Decision problems in which ratings are grouped per judge, therefore, may not be peculiar.

Experiment 2 employed the same stimuli used in the previous experiment, except that they were either grouped per candidate or per judge. The purpose of the experiment was to test how the two different display organizations would

influence a decision maker in resolving a difficult decision. Specifically, I examined whether grouping per judge would yield an increase in the likelihood to toss a coin compared to grouping per candidate.

In contrast to the first experiment, in which the choice between the candidates was made by participants themselves, the present experiment stated that the choice was to be done by another person. Attractiveness scores of Experiment 1 indicated that the job candidates were not rated equally attractive. Therefore, having participants themselves choose between the two candidates would, supposedly, result in one candidate dominating the other. Consequently, tossing a coin may seem as a rather implausible choice strategy to these participants. However, introducing a third person (e.g., a personnel officer) who claims that choosing between the candidates is very difficult, may circumvent the problem of dominance. A statement made by a third person regarding decision difficulty, would supposedly be interpreted as if the alternatives were close in attractiveness to this decision maker (a methodology adopted from Liberman & Förster, 2006), and thus may again render the toss of a coin as a plausible choice strategy.

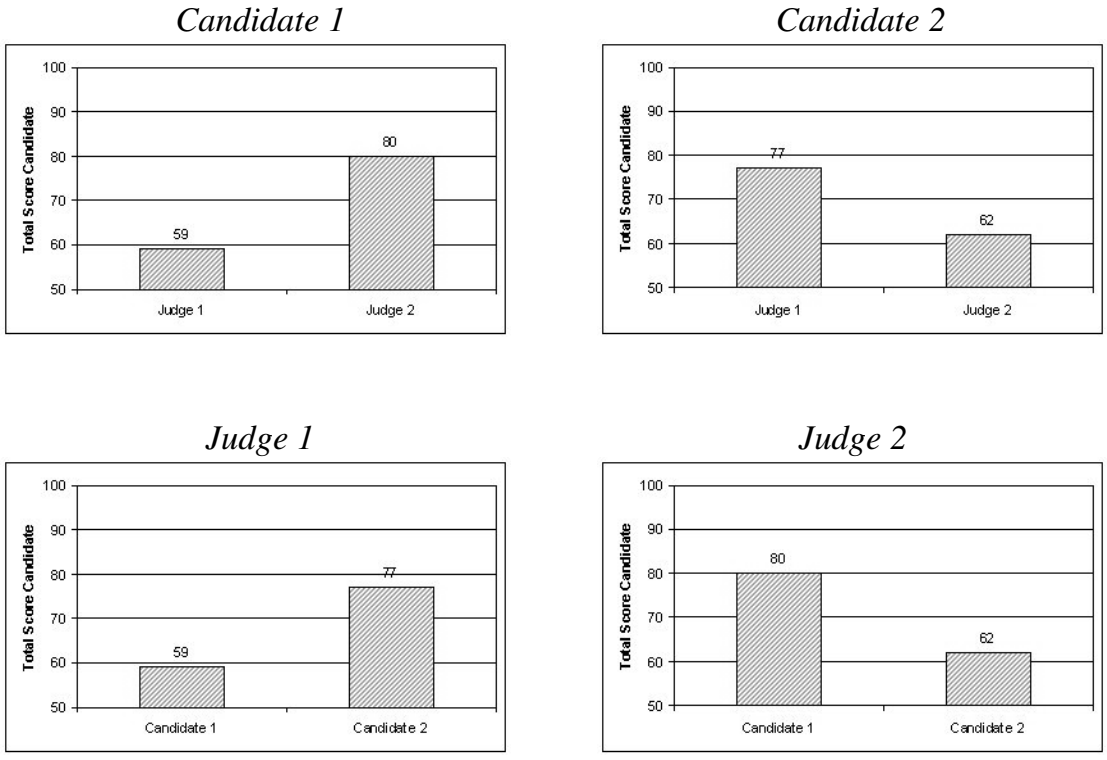

Figure 2. Different display organizations of the same information, with ratings either grouped per candidate (top) or per judge (bottom).

Experiment 2

Method

Participants

A total of 240 students from Tilburg University and Fontys University of Professional Education Eindhoven participated in the experiment and received € 5 for their participation. Participants performed a series of unrelated judgment and decision tasks, including the present experiment.

Materials

Employed were the low and high variance pairs with average ratings of 69.5 from the first experiment (see Table 1). For each pair, the candidates' ratings were presented by two bar-graphs, and displayed simultaneously on a laptop screen. The ratings were either grouped *per candidate* or *per judge* (see Figure 2). When grouped per candidate, each bar-graph consisted of the ratings given by the two judges to a single candidate. When grouped per judge, each bar-graph consisted of the ratings that a single judge gave to each of the two candidates.

A cover story described a personnel officer who had to choose between two job candidates. The two candidates were said to be rated independently by two judges, and both candidates' average rating was identical. The personnel officer claimed, after having considered these ratings, that choosing between the candidates is very difficult, thus conveying the information that, in his view, the alternatives were close in attractiveness (Liberman & Förster, 2006). Participants were subsequently told that one way of resolving this difficult choice is by tossing a coin.

Procedure

Participants performed the experiment on a laptop, according to their own pace. Grouping (*per candidate* or *per judge*) and variance (*low* or *high*) of the candidates' ratings were manipulated between subjects. Participants rated on a scale how likely (0 = not at all likely; 5 = very likely) they thought the personnel officer would toss a coin in order to resolve the difficult decision.

Results and discussion

The results are portrayed in Table 3.

Participants' ratings of the likelihood to toss a coin were submitted to a 2 (variance: low vs. high) x 2 (grouping: by candidate vs. by judge) analysis of variance. Grouping per judge was predicted to increase the likelihood for a coin toss.

Table 3
Mean judged likelihood to toss a coin (on a 0-5 scale) in Experiment 2, depending on variance and display organization.

	Grouped per candidate		Grouped per judge	
	n	*M (SD)*	*n*	*M (SD)*
Low variance	59	2.39 (1.31)	60	2.25 (1.47)
High variance	61	1.79 (1.40)	60	2.43 (1.38)

The analysis revealed no main effects of both variance (low vs. high), $F(1, 236) = 1.36$, $p = 0.25$, and grouping (per candidate vs. per judge), $F(1, 236) = 1.98$, $p = 0.16$. However, the interaction of variance x grouping was significant, $F(1, 236) = 4.78$, $p = 0.03$. To explain this interaction, it is helpful to focus on the discrepancies as displayed in the bar-graphs. For the low variance ratings grouped per candidate, the magnitude of the discrepancies (between judges) as displayed in the bar-graphs were 1 and 3, respectively. When these same ratings were grouped per judge, the magnitude of the discrepancies (between the candidates) were identical. Notice, however, that the distinctive groupings (candidate vs. judge) are not substantially different from each other with regard to the magnitude of the displayed discrepancies (i.e., the discrepancies were 1 and 3, respectively, when grouped per candidate, and 2 when grouped per judge, yielding a difference in magnitude of 2 (3 – 1) for ratings grouped per candidate, versus 0 (2 – 2) for ratings grouped per judge). For high variance ratings, in contrast, the two distinctive groupings displayed discrepancies that differed substantially in magnitude. When grouped per candidate, the discrepancies were 21 and 15, respectively, whereas grouping per judge yielded identical discrepancies of 18 (i.e., the difference in magnitude was 6 (21 – 15) for ratings grouped per candidate, versus 0 (18 – 18) for ratings grouped per judge). Assuming that tossing a coin is sensitive to the difference in magnitude between discrepancies, as displayed in the bar-graphs, then larger differences in the likelihood to toss a coin are expected for the different groupings of the high variance ratings, where substantial differences in magnitude between the distinctive groupings are observed (i.e., for high variance ratings, the difference in magnitude was 6 [grouped per candidate] vs. 0 [grouped per judge], compared to 2 [grouped per candidate] vs. 0 [grouped per judge] for the low variance ratings). In accordance with this prediction, there was a large difference in the judges' likelihood to toss a coin in the high variance condition, $t(119) = 2.55$, $p = .01$, and no difference in the low variance condition, $t(117) = 0.55$, $p = .59$.

Alternatively, the decreased likelihood to toss a coin in the high variance condition with the ratings *grouped per candidate* can be explained by arguing that one of the candidates in this pair received one rating (59) that was below a neutral reference point (i.e., 60 according to Dutch standards), and one rating (80) that was above the reference point, whereas the ratings of the other candidate in the same pair were always higher than this neutral reference point

(62 and 77, respectively). Supposedly, this reference point helps to evaluate ratings, such that any ratings above the reference are considered to be good, and any ratings below the reference are interpreted as bad (Hsee, Loewenstein, Blount & Bazerman, 1999). To prevent the possibility that in the high variance condition the (poor) candidate, with one rating below reference point, gets assigned to the job, the personnel officer might be reluctant to toss a coin (note, though, that this reference point explanation can not account for the higher likelihood to toss a coin when the same ratings were grouped per judge, suggesting that the increased likelihood for this grouping was due to lacking compelling reasons for choice justification). In the low variance condition, with ratings *grouped per candidate*, however, none of the two candidates received ratings below the neutral reference point. Since both candidates' ratings are well above 60, the personnel officer might be willing to run the risk of tossing a coin. To conclude, the obtained difference in the likelihood to toss a coin between the high and low variance condition, when solely focusing on the ratings *grouped per candidate*, could be explained by arguing that in the former condition, one candidate's rating was below the reference, whereas in the latter condition all candidates' ratings were above the reference point.

The issue concerning ratings below and above a neutral reference point was addressed in the following experiment, which alternatively, employed a distinctly different context. Specifically, the context described an organizer of a music festival, who was said to have to choose between two music bands. The organizer used the number of hits on the internet, as produced by two search engines, as a measure of bands' popularity (see Figure 3). Compared to the judges' ratings in the first two experiments, the number of hits on the internet has no neutral reference point that can, subsequently, be used for evaluation. As a consequence, a decrease in the likelihood to toss a coin in the high, compared to the low, variance condition can not any longer be explained by arguing that someone does not want to run the risk of getting randomly assigned to an alternative that contains a (bad) value below a reference.

Additionally, in the following experiment, the number of hits were displayed not in bar-graphs, but in tables. The use of tables is useful to generalize the findings of the previous experiment.

For the following experiment, it was predicted that the displayed discrepancy in ratings would be less noticeable if the number of hits were grouped per search engine (than per band). Hence, the likelihood to toss a coin was predicted to be higher when number of hits were grouped per search engine. Because the displayed discrepancy was, presumably, more salient with grouping the hits per band, it should be easier for the organizer of the festival to construct compelling reasons to justify his choice. As a consequence, the likelihood to toss a coin was predicted to be lower with the number of hits grouped per band.

Experiment 3

Method

Participants

A total of 293 students from Tilburg University and Fontys University of Professional Education Eindhoven were recruited on campus and paid € 4 for their participation. The present study was part of a number of unrelated judgment and decision making experiments.

Low variance pairs

Grouped per Band

Band A	*Band B*
Search Engine X: 71	Search Engine X: 67
Search Engine Y: 69	Search Engine Y: 73

Grouped per Search Engine

Search Engine X	*Search Engine Y*
Band A: 71	Band A: 69
Band B: 67	Band B: 73

High variance pairs

Grouped per Band

Band A	*Band B*
Search Engine X: 96	Search Engine X: 49
Search Engine Y: 44	Search Engine Y: 91

Grouped per Search Engine

Search Engine X	*Search Engine Y*
Band A: 96	Band A: 44
Band B: 49	Band B: 91

Figure 3. Different display organizations of the number of hits for two music bands, as produced by two search engines on the internet, with the respective number of hits either grouped per band or per search engine.

Materials

Low and high variance pairs, describing two bands' number of hits by two search engines were constructed (see Figure 3). The average number of hits for each band was 70. The number of hits were presented in two tables, and both tables were displayed simultaneously on a laptop screen. The values in the tables were either grouped *per band* or *per search engine*. When grouped per

band, each table consisted of the number of hits that the two search engines produced for a single band. When grouped per search engine, each table consisted of the hits that a single search engine found for each of the two bands.

A cover story described an organizer of a music festival who had to choose between two music bands. The organizer used the number of hits, reported by two search engines, as a measure of the bands' popularity. From the engines' search results, it appeared that the average number of hits for each band were identical.

The organizer claimed, after having considered the hits by both engines, that choosing between the bands is very difficult, which supposedly conveyed the information that, in his view, the bands were close in attractiveness. Participants were subsequently told that one way of resolving this difficult choice is by tossing a coin.

Procedure

Participants performed the experiment on a laptop and according to their own pace. Grouping (*per band* or *per search engine*) and variance (*low* or *high*) of the bands' number of hits were manipulated between subjects. Participants rated, on a 6 point scale, how likely (0 = not at all likely; 5 = very likely) they thought the organizer would toss a coin in order to resolve the difficult decision.

Subsequent to this task, participants were asked to choose which of the two bands they themselves would prefer. The reason behind asking participants for their own preference was twofold. First, compatible with Experiment 1, is there a preference for the band displaying a smaller discrepancy in the number of hits? Second, might this preference change with different display organizations (i.e., grouped per band vs. search engine)?

Results and discussion

The results are displayed in Table 4.

Table 4
Mean judged likelihood to toss a coin (on a 0-5 scale) in Experiment 3, depending on variance and display organization.

	Grouped per band		Grouped per search engine	
	n	*M (SD)*	*n*	*M (SD)*
Low variance	71	2.70 (1.75)	74	3.18 (1.45)
High variance	77	2.79 (1.57)	71	3.03 (1.29)

A 2 (variance: low vs. high) x 2 (grouping: by band vs. by search engine) analysis of variance revealed a main effect of grouping, indicating that grouping by search engine yielded a higher likelihood to toss a coin (M = 3.10, SD = 1.37) than grouping per band (M = 2.75, SD = 1.65), $F(1, 289) = 3.95$, p = .05. The analysis revealed no main effect of variance (low vs. high), $F(1, 289) = 0.03$, p = .87, nor an interaction of variance x grouping, $F(1, 289) = 0.44$, p = .51.

Congruent with Experiment 2, these results demonstrated that the two different display organizations (i.e., grouped per band vs. search engine) influenced a decision maker in resolving decision difficulty. In particular, the likelihood to toss a coin was assessed to be higher in case the displayed discrepancy in the number of hits was less noticeable (grouped per search engine), compared to a display organization were a discrepancy in hits was more salient (grouped per band). This result suggests that for the former it was more difficult to construct convincing reasons for choice justification than for the latter display organization.

Might these results imply that, for the participants themselves, the alternatives also become more equally attractive when the number of hits were grouped per search engine, instead of per band? Choice results, representing the preferences of the participants, depending on whether the number of hits were grouped per band or per search engine, are summarized in Table 5.

Table 5
Number of participants choosing the music band having either a small or large discrepancy in the search engines' results, as a function of grouping and variance.

	Preference band	
	Small discrepancy	Large discrepancy
Low variance		
Grouped per band	48	23
Grouped per search engine	49	25
High variance		
Grouped per band	37	40
Grouped per search engine	30	41

Across both types of grouping, 97 participants (67%) in the low variance condition favored the alternative with the small discrepancy, compared to 48 who favored the large discrepancy alternative. For the high variance condition, 67 participants (45%) preferred the small discrepancy, whereas 81 chose in favor of the large discrepancy. Thus, with regard to small and large discrepancy

alternatives, the participants' preferences were reversed when the low and high variance conditions are compared, χ^2 [1] = 13.90, $p < .001$. Supposedly, when variance is low, participants favor the alternative displaying a smaller discrepancy (which is congruent with the attractiveness ratings obtained in Experiment 1). Yet, when variance increases, the alternative with the highest number of hits (96) gets more attractive. In other words, when variance is substantial, it seems that the attractiveness of an alternative also depends on the extremeness of the values it received (i.e., highest overall rating), which explains the increasing preference for the alternative displaying the large discrepancy.

Might participants their own preference change with different display organizations (i.e., grouped per band vs. search engine), that is, do alternatives get closer in attractiveness when they are grouped per search engine? For low variance alternatives, the preferences did not change as a function of display organization, χ^2 [1] = 0.03, $p = .86$. Similarly, no change in preference was found for high variance alternatives, χ^2 [1] = 0.50, $p = .50$. These latter findings imply that the alternatives do not become closer in attractiveness, when the display organization changes from *grouped per band* to *grouped per search engine*. Moreover, these results suggest that choosing may get more difficult in case the discrepancy is less noticeable (i.e., grouped per search engine), not because the options become closer in attractiveness, but more likely, because it is harder to construct compelling reasons for justification.

General discussion and conclusion

In order to resolve a difficult choice, decision makers try to come up with compelling reasons to justify their preference for the chosen alternative. In the absence of such reasons, however, decision makers may accept a random process (e.g., tossing a coin), that will replace them in making the decision. Even though decision makers are often averse to have a random process being in charge of their decisions, they may rely on such a process when resolving a difficult decision for which they have trouble with seeking and constructing compelling reasons for justification.

The present study examined how different display organizations, of the same choice options, effected the construction of compelling reasons. The degree to which a difficult choice was resolved by convincing reasons, was measured by the likelihood to toss a coin. It was conjectured that the more difficult it is to construct reasons for justifying a preference, the more likely a decision maker will accept a coin toss. To investigate this alleged relation, choice alternatives were employed that were defined by two values, with each alternative's values displaying some discrepancy (i.e., the difference between its two values). A choice problem consisted of two of these alternatives, introduced by another person, who claimed that choosing between these alternatives was a difficult task. Results showed that a difficult decision was less likely to be resolved by a coin toss, when the two alternatives' respective discrepancies were different from each other. Further support for the suggested relation between these

discrepancies and the likelihood to toss a coin, was found when the display of the same choice problem was reorganized. A consequence of this change in display organization, was that any noticeable difference in discrepancy between two alternatives disappeared, and accordingly, results showed an increase in the likelihood to accept a coin toss. The increased likelihood to accept a coin toss, was interpreted as implying that it was more difficult to construct convincing reasons after reorganizing the same choice problem, which made that any difference in discrepancy was less noticeable.

Description invariance vs. different display organizations

Different display organizations of the same alternatives thus seem to influence both the construction of compelling reasons, and the associated likelihood to accept a coin toss. Following normative considerations, though, a difference in display organization should not have any effect. This is reminiscent of the concept of description invariance, which requires that changing the representation of a choice problem should have no effect on the *preference* between alternatives (Tversky & Kahneman, 1986). Yet, violations of description invariance are often observed with framing effects. For instance, people prefer "75% lean" over "25% fat" ground beef (Levin & Gaeth, 1988), and "95% fat-free" over "5% fat" yoghurt (Sanford, Fay, Stewart & Moxey, 2002), despite the fact that these descriptions are logically equivalent. Notice, however, that the present study differs from these framing experiments in two ways.

First, while framing yields logically equivalent descriptions of the same choice problem, the information presented in each of these descriptions is not identical (i.e., "75% lean" and "25% fat" are not *literally* identical descriptions). The present study manipulated the display organization of the choice problems, which also yielded different descriptions, but in contrast with framing experiments, the information presented with each different organization was *literally* identical (i.e., candidates' ratings assigned by the judges were identical, regardless whether they were grouped per candidate or per judge). Yet, framing experiments and the present study have in common that each description shifts the focus of attention towards different information. For instance, a cup will more likely be described as half full, if it was previously *empty*, and half empty if it was previously *full* (McKenzie & Nelson, 2003). These logically equivalent frames, thus, focus on different kinds of reference point information (i.e., empty or full cup). Similarly, different display organizations seem to focus on different kinds of information. When the ratings in Experiment 2 were grouped per candidate, the focus was on the discrepancy between judges. In contrast, when the same ratings were grouped per judge, the focus was on the discrepancy between candidates. Thus, different display organizations focus on different dimensions of the same ratings.

Second, the present studies differ from traditional framing experiments in that violations of description invariance concern changes in *preferences* that occur with different descriptions of the same choice problem. The present study, however, mainly focused on how different descriptions of the same

choice problem effected a specific property of the *decision process*, namely decision difficulty. Importantly, the preferences were *not* expected to change along with different descriptions. For this reason, the cover story, that accompanied each choice problem, involved a person who said that choosing between the two alternatives was difficult. This, supposedly, conveyed the information that the alternatives were close in attractiveness to the person who had to choose between them (Liberman & Förster, 2006). Consequently, it can be assumed that, despite the difference in display organization, the alternatives were always approximately equally attractive, and thus no change in preference occurred.

Alternative explanations for the increased likelihood to toss a coin

When the same choice problem is portrayed in different displays, are there other factors, besides difficulty in constructing compelling reasons, that explain the likelihood to toss a coin? Schkade and Kleinmuntz (1994) also examined how different organizations of a choice problem effected decision process measures. For instance, they had identical information either organized by attribute (for each attribute, the values of all alternatives were presented together) or by alternative (for each alternative, the values of all attributes were presented together). Their results showed that different organizations influenced choice process measures, such as effort expenditure, but resulted in no changes regarding preference. These findings may raise the question whether the likelihood to toss a coin is related to the amount of effort expended, instead of generating compelling reasons. May a decision maker decide to toss a coin as soon as a lot of effort is required for choosing between two alternatives? For instance, imagine someone confronted with a difficult decision among two high variance alternatives that are grouped per judge (as in bottom of Figure 2). For ease of comparison between the two candidates, it might be useful to mentally transform the choice problem such that the alternatives become grouped per candidate (see top of Figure 2). However, this transformation will require mental effort, and by tossing a coin, a decision maker may save this effort. The relation between mental effort and tossing a coin, may thus explain the observed difference in the likelihood to toss a coin for high variance scores, and would predict a similar difference for the low variance scores. Notwithstanding, results show that no such difference was observed with low variance scores (see Table 3), which makes the relation between effort expenditure and the likelihood to toss a coin rather implausible.

A decrease in the likelihood to toss a coin can also be explained by an increase of (anticipated) regret. This feeling of regret occurs when, in retrospect, a chosen alternative turns out to be worse than the un-chosen one(s). Intuitively, people will eschew tossing a coin if substantial regret may be expected. Thus, was an increase in (anticipated) regret the reason that participants were less likely to accept a coin toss when a display organization emphasized discrepancies between the ratings? The answer to this question seems to be negative. A different display organization may emphasize different dimensions of the same information (e.g., the discrepancy between candidates

vs. that between judges as in Experiment 2), the information itself (i.e., respective values of the alternatives under consideration) remains identical in a different display organization. Presumably, decision makers focus on this latter information, concerning values, when predicting whether they will be anticipating regret or not. That is, although it may be difficult to construct compelling reasons for justifying a choice (given a display organization), whether or not a choice outcome yields regret is based on the values displayed, since these values determine if a chosen alternative turns out to be worse than the un-chosen one(s). Thus, instead of anticipated regret, it is more likely that the decision to toss a coin, in the context of a particular display organization, is based on lacking compelling reasons for choice justification.

Alternative measures of difficulty

The present paper offers a measure of difficulty (i.e., coin toss acceptability) that may differ from the measures that are usually employed (e.g., time to reach a decision: see Scholten & Sherman [2006] for an overview). Tossing a coin, however, may be particularly useful for measuring difficulty related to the construction of compelling reasons. Another measure of difficulty, that indicates that difficulty might stem from lacking convincing arguments for justification, is choice deferral (Anderson, 2003). What distinguishes these two measures, however, is their sensitivity to anticipated regret (Chapter 2). Yet, it was argued above that potential regret may not have influenced choice difficulty, when changing display organizations. Therefore, choice deferral and coin flipping are predicted to yield similar results, when employed for measuring difficulty in the context of different display organizations. Results from pilot work on choice deferral, employing the scenario of Experiment 3 (i.e., popularity of two music bands as measured by the number of hits produced by two search engines), support this prediction. Despite a small sample size (n = 89) in this pilot, the results suggested that participants were more likely to accept deferral in case the discrepancy between alternatives was less noticeable (i.e., when the number of hits were grouped per search engine), which is congruent with the results obtained for coin flipping. Nevertheless, the latter conclusion is preliminary due to the small sample size employed in the pilot experiment, and therefore further research investigating choice deferral is needed.

4

The resolution and assessment of choice difficulty when choosing between dichotomous and continuous choice pairs

Abstract

This chapter discusses decision difficulty when choosing between alternatives that consist of either dichotomous or continuous attributes. Whereas the latter attributes are represented dimensionally, the former are defined as features that an alternative possesses or not (i.e., yes – no attribute values). Empirical evidence is presented indicating that, compared to deciding among continuous alternatives, a difficult choice between dichotomous alternatives is more likely explained by lacking decisive reasons for choice justification (Experiment 1), and is more likely resolved by the toss of a coin (Experiment 2). Furthermore, difficult decisions involving dichotomous alternatives, are found to be more likely resolved by deferring choice, while their difficulty is more likely explained by anticipated regret, as well as the difficulty in trading off pros and cons (Experiment 3). The implications of the observed correlation between means of resolving difficulty (choice deferral and coin flipping) and underlying sources of decision difficulty (lacking convincing arguments, anticipated regret, and trade off difficulty) are discussed. In particular, which sources of difficulty may be measured by choice deferral and coin tossing, when applied to dichotomous and continuous alternatives?

During their lives, people are confronted with numerous decisions, involving different sorts of alternatives. For instance, they may have to choose between two ice creams, a trip to China or Peru, and whether or not to accept a job offer. Intuitively, the diversity in choice alternatives seems to be enormous. Despite this diversity, however, many choice alternatives are represented in a similar way, namely by attributes. Choosing between such alternatives boils down to some process in which the alternatives are compared, yielding a preference for one of the respective alternatives. This comparison between alternatives, though, may consider each alternative as a whole (i.e., alternative based choice), but often will focus on a limited number of attributes (due to limited information processing capacity), which are subsequently compared (i.e., attribute based comparison) (e.g., Payne, 1976).

Alternatives, that are compared in a choice process, are sometimes characterized by a single attribute. When deciding among these single-attribute alternatives (e.g., a large and a small pizza), most people will prefer the alternative that is superior on the dimension that represents the singular attribute in question (i.e., the larger pizza, unless one is on a diet). A single attribute, however, may not always be represented dimensionally. It is also possible that alternatives differ on a single attribute in terms of a dichotomy, such that some possess this attribute, whereas other do not (e.g., identical cars equipped with air conditioning or not). The extent to which this attribute is desired, supposedly, will determine the decision maker's choice.

Choice alternatives, however, are more often characterized by two or more distinctive attributes. When facing a choice between two rooms, for instance, a student may have to decide among multi attribute alternatives varying, say, on commuting time and room size. If examination of these attributes reveals that one room is superior on both attributes, then choosing is easy – the dominating alternative will be chosen. Yet, a situation that involves a dominant alternative may not always be interpreted as a choice, since it lacks the experience of a conflict (Delquié, 2003). It is also possible, however, that neither room dominates the other. For example, suppose that one room has a larger size, whereas the other room's commuting time is shorter. A decision among these latter two rooms, arguably, yields the experience of an internal conflict once the student trades off the advantages, associated with both rooms, against each other. As a consequence of this conflict, the decision maker may call upon choice strategies, such as selecting the alternative that is superior on the most important dimension (Slovic, 1975; Tversky, 1972; Tversky, Sattath & Slovic, 1988).

Related to multi attribute alternatives are options defined by several values, converging on a single (often more abstract) dimension. A job officer choosing between two job candidates, with each candidate rated independently by two judges on an identical scale, ranging from 1 (poor) to 10 (excellent), is an example of a situation involving uni-dimensional alternatives. The single dimension on which the candidate is evaluated, supposedly, summarizes a number of distinctive qualities. Yet, comparison of the candidates focuses on only one (overall) dimension. A conflict in choosing between these candidates

arises if both candidates' average ratings are identical (e.g., 7.5), and additionally, one of these job candidates received ratings 7 and 8 from the respective judges, whereas the other received 9 and 6 from them. Which of the candidates is preferred by the job officer, will depend on the discrepancy between the judges' ratings (Chapter 3). Specifically, it appears that the candidate with the smaller discrepancy is more likely to be chosen. Thus, concerning the above two job candidates, the discrepancies are 1 (8-7) and 3 (9-6), respectively. A job officer, therefore, will most likely prefer the first candidate, since the variability in her scores is smaller compared to that of the other candidate. For explaining such a preference, the job officer may argue that such smaller variability suggests that the candidate made a more reliable impression towards both judges which, in turn, sounds as a convincing argument for choosing this candidate.

Different alternatives, though, may not always be defined by their respective values on continuous dimensions. Consider two identical cars that differ only in that one has air conditioning, whereas the other is equipped with air bags. When choosing between these cars, a direct comparison of the two alternatives on some (continuous) dimensions is no longer possible, since an alternative either will have a defining feature or not. It is still possible, though, to represent the attributes dimensionally, in the sense that they are dichotomous (i.e., these features may be represented by yes – no values on the dimension). Yet, the difficulty with these dichotomous attributes remains that they are not comparable (e.g., how should one compare the features air conditioning with air bags?). Decision makers, therefore, might instantiate an abstract (continuous) attribute (e.g., necessity) on which they can subsequently compare the alternatives (Markman & Medin, 1995).

Related to a decision among dichotomous alternatives, are choices between alternatives from different product categories, say, a choice between a television and a trip to Berlin. A decision among these alternatives may, similarly, yield a comparison pertaining to features that both alternatives do not have in common. Corresponding to a choice between dichotomous alternatives, alternatives from such different product categories are compared by evoking abstract attributes, like "necessity" and "importance", on which it is easier to compare these non-comparable alternatives (Johnson, 1984).

Comparability between alternatives and reason-based choice

It appears that there are several types of attributes on which alternatives can be compared. An important distinction, though, can be made between dichotomous and continuous attributes. In contrast with dichotomous attributes, it may be easier to compare alternatives on continuous attributes, given that these latter attributes do not require the construction of an abstract attribute. Yet, once a decision maker has figured out such an abstract attribute, deciding among dichotomous alternatives is easy, supposedly because one will prefer the alternative superior on this attribute.

When faced with a conflict between continuous attributes, a decision maker has to address the trade off between attributes. How exactly to compare these

attributes may be complicated, which makes the resolution of this conflict difficult. For instance, how should a student compare, and trade off, a 5 m^2 larger room size to a 25 minutes longer commuting time? Rather than combining these attributes into an abstract attribute, conflicts between continuous attributes are often resolved by focusing on the most important, more prominent, dimension (e.g., room size), and subsequently choosing the alternative superior on this dimension (Slovic, 1975).

When experiencing and resolving a difficult choice between alternatives, decision makers often seek and construct compelling reasons for choice justification (Shafir, Simonson & Tversky, 1993). The construction of convincing reasons, presumably, requires that a decision maker is able to compare the alternatives under consideration. Dichotomous alternatives, supposedly, hinder a direct comparison, and call for an abstract attribute (e.g., necessity) on which the non-comparable attributes are subsequently compared. Continuous alternatives, on the other hand, share attributes (e.g., room size) on which they are directly comparable. The construction of convincing arguments for continuous attributes will, as a consequence, be more straightforward (e.g., that one room has a larger size than the other is a decisive argument), and for that reason, justification might be easier if one uses what is common to both alternatives. For dichotomous attributes, on the other hand, constructing compelling reasons may be more demanding, and hence it is less likely that a decision maker will employ these dichotomous attributes for choice justification. Evidence, that justification based on dichotomous attributes is not likely, was provided by Markman and Medin (1995). They constructed alternatives (i.e., descriptions of video games) composed of both comparable and non-comparable attributes. The comparable attributes referred to dimensionally represented properties (e.g., two games that differ in the number of ballparks), whereas non-comparable attributes were defined by properties that only one of the alternatives possessed (e.g., one game has practice trials, whereas the other has not). Their results implied that justifications, of why one preferred an alternative, were more likely to include comparable (or what they called alignable) than non-comparable (non-alignable) attributes. Evidently, by referring to comparable rather than non-comparable attributes, it is easier to justify why one prefers an alternative.

Markman and Medin explained their results by arguing that a comparison between alternatives is accomplished by an alignment process (e.g., Gentner & Markman, 1997), that yields a set of commonalities, and two kinds of differences: alignable, and non-alignable differences. The alignable differences are related to the commonalities, and arise when two alternatives have different values along a common dimension (e.g., two mp3 players that differ in memory size). The non-alignable differences, on the other hand, are not related to the commonalities, and arise when one alternative possesses a feature that the other alternative is lacking (e.g., only one of the two mp3 players comes with an AC power adapter). Furthermore, Gentner and Markman (1994) provided evidence that alignable differences are more salient than non-alignable in the comparison process. Since choice may be driven by this alignment process (Medin,

Goldstone & Markman, 1995), justifications are more likely to be focused on alignable (continuous, quantitative) rather than on non-alignable (dichotomous, qualitative) differences.

The present chapter focuses on difficult decisions that entail a comparison between either dichotomous or continuous attributes. Three experiments, contrasting both types of attributes, were designed to examine factors that might determine experienced difficulty (e.g., lacking decisive reasons for justifying choice), and the means by which difficulty is likely to be resolved (e.g., a coin toss). The results of these experiments suggest that these factors, causing choice difficulty, may coincide with specific means of resolving difficulty. The implications of these observed correlations will be discussed.

Experiment 1

The first experiment was designed to test whether it is more difficult to construct compelling reasons in case a close decision concerns alternatives, consisting solely of dichotomous attributes, or alternatively, continuous attributes. The experiment is similar to that reported by Markman and Medin (1995), yet differs from their study in two respects.

First, they employed alternatives that contained both comparable and non-comparable attributes at the same time. They explained why comparable attributes are more likely used for choice justification, by claiming that these attributes are more salient than the dichotomous attributes in the comparison process. The present experiment, however, employs alternatives that have either comparable (continuous) or non-comparable (dichotomous) attributes, but never simultaneously. With the latter alternatives, both types of attributes do not compete for saliency in the comparison process used for choosing between these alternatives. Therefore, it may be possible to test directly whether it is more difficult to construct compelling reasons for dichotomous rather than comparable attributes, without having to argue that one of the attributes is more salient in the comparison process.

Second, Markman and Medin asked participants for written justifications of their preferences, and subsequently checked whether they listed more comparable or non-comparable features. A problem with this procedure is that a salient attribute will be listed anyway, regardless whether it is comparable or non-comparable. Therefore, instead of having participants listing attributes, the following experiment directly asked whether it is more difficult to construct compelling reasons when faced with a decision involving either comparable or non-comparable attributes.

Method

Participants

A total of 135 students from Radboud University Nijmegen were recruited on campus and paid 5 Euros for their participation. The present study was part of a number of unrelated judgment and decision making tasks.

Materials

Descriptions of four pairs of mp3 players were constructed. Players in each pair were defined by two attributes (see Figure 1), having either a dichotomous (i.e., Pair D1 and D2) or a continuous (i.e., Pair C1 and C2) character. Subsequently, four stimulus sets were created. A set consisted of two pairs of mp3 players, of which one pair had dichotomous and the other continuous attributes (see Table 1).

Pair D1 (dichotomous attributes)

	mp3 player X	mp3 player Y
FM radio	No	Yes
Time / alarm function	Yes	No

Pair D2 (dichotomous attributes)

	mp3 player X	mp3 player Y
AC power adapter	Yes	No
Picture storage/viewer	No	Yes

Pair C1 (continuous attributes)

	mp3 player X	mp3 player Y
Weight	60 gr (ultra light)	130 gr (light)
Available colors	2	5

Pair C2 (continuous attributes)

	mp3 player X	mp3 player Y
Battery life time	15 hrs.	22 hrs.
Memory size	1024 MB	512 MB

Figure 1. Pairs of alternatives employed in the present studies, characterized by either dichotomous (Pair D1 and D2) or continuous (Pair C1 and C2) attributes.

Procedure

The experiment was conducted on a laptop, and participants were instructed to perform the task according to their own pace.

For each set, both pairs of mp3 players appeared simultaneously on the laptop screen. A cover story described two executives, each working for a different company. One executive had to choose, for his company, between the pair described by the dichotomous attributes, whereas the other had to choose, for his company, between the pair described by the continuous attributes. Both executives were said to experience great difficulty in choosing between their

respective pairs. This information about difficulty, supposedly, signified that for each executive, the mp3 players were assessed to be close to each other in attractiveness (Liberman & Förster, 2006). Subsequently, participants were informed that one of the two executives explained his experienced difficulty by telling that it was difficult to construct convincing reasons to justify his choice towards the company's management. Participants had to indicate which of the two executives (i.e., the one choosing between the options in the dichotomous pair, or the one choosing between the options in the continuous pair) was more likely to explain choice difficulty by referring to the difficulty in generating compelling reasons.

Each participant was presented with only one of the four stimulus sets (see Table 1). For each set, the order of presentation on the screen of the dichotomous and continuous pair was counterbalanced. The experiment was, thus, a 4 (stimulus set) x 2 (order) between-subjects design.

Results and discussion

The percentages of participants, for each set, who indicated that choice difficulty was more likely explained by referring to the difficulty in constructing compelling reasons for justification, are presented in Table 1.

Table 1
Percentages of participants for each stimulus set who indicated that, compared to continuous alternatives, difficult decisions among dichotomous alternatives are more likely explained by lacking compelling reasons.

	Lacking convincing reasons	
Stimulus set	*n*	Dichotomous
Pair D1 vs. C1	36	47%
Pair D1 vs. C2	35	80%
Pair D2 vs. C1	33	61%
Pair D2 vs. C2	31	65%

The results showed that, when faced with a difficult decision, it is more difficult to construct compelling reasons for dichotomous than continuous attributes. This was confirmed when the results of the four stimulus sets were aggregated, which revealed that 63% (85 out of 135) of all participants indicated that it was more difficult to construct compelling reasons for dichotomous pairs, $p = .003$ by a binomial test. This result is congruent with, and further strengthens, the finding by Markman and Medin (1995), which showed that dichotomous attributes were less likely to be used in justifying preferences.

Once a difficult decision can not be easily resolved by firm convincing reasons, one may consider reverting to a random procedure (Elster, 1989). Yet, people are often reluctant to decide randomly among alternatives. In the absence of convincing reasons, however, they may be willing to submit their decision (reluctantly though) to an appropriate random process. For instance, when faced with a difficult decision involving a trade off between (continuous) attributes, it is more likely that a decision maker refers to lacking convincing reasons when the differences between the attributes are small rather than large (Chapter 2; Scholten & Sherman, 2006). At the same time, people indicate that they are more likely to toss a coin in case a difficult decision involves small trade off alternatives (Chapter 2). Together, these findings suggest that when lacking decisive reasons, people may more likely accept the flip of a coin.

However, the extent to which people accept the toss of a coin, when the difficult decision involves dichotomous alternatives, is not clear. On the one hand, compared with continuous alternatives, difficult decisions among dichotomous alternatives are more likely explained by referring to the difficulty in constructing compelling reasons (Experiment 1). Accordingly, people should be more willing to toss a coin with difficult choices involving dichotomous alternatives. On the other hand, choosing may yield sacrifices, that are to be incurred, by foregoing the un-chosen alternative. These sacrifices may, presumably, be larger with dichotomous rather than continuous attributes. For instance, imagine deciding among two mp3 players defined on continuous attributes, such as battery life time and memory size. Suppose, further, that one of the players has a longer battery life (22 hrs.), and the other a larger memory size (1024 MB). A preference for a player superior on, say, battery life time, comes along with a sacrifice on the other attribute, namely memory size (on which the other alternative is superior). In comparison to the large (sacrificed) memory capacity of the rejected player, the preferred alternative also possesses some memory capacity, though to a lesser degree (e.g., 512 MB). Yet, when the players are characterized by dichotomous attributes, the sacrifices to be incurred might even be larger, because the attributes have an all-or-none character. Consider a choice between two mp3 players, one that comes with an AC power adapter, whereas the other player can store and view pictures. A preference for one player implies foregoing the feature associated with the other, in an all-or-none way. That is, choosing one player goes together with sacrificing all of the other player's feature. Although the sacrifice on memory size with the player superior on battery life was considerable, the chosen player still possesses some memory capacity. Thus, the sacrifice with continuous alternatives may be smaller compared to that experienced with the dichotomous alternatives, where choosing implies foregoing *everything* associated with the other player. Due to these larger sacrifices, associated with dichotomous attributes, a decision maker may be less likely to accept the toss of a coin, compared with the case of continuous alternatives. Random assignment in case of continuous attributes, after all, maintains some of the attribute on which a preferred alternative is inferior.

In sum, when contrasting dichotomous and continuous alternatives, it is not obvious whether one would more likely accept a coin toss with dichotomous pairs, because for these alternatives decision makers may be lacking compelling reasons for justifying their preference, or conversely, would more likely accept a coin toss with continuous alternatives, because with these alternatives the sacrifices to be incurred are supposed to be smaller. The following experiment examined whether a difficult decision among either dichotomous or continuous alternatives is more likely to be resolved by a coin toss.

Experiment 2

Method

Participants

A total of 117 students from the Fontys University of Professional Education Eindhoven participated in the experiment. The experiment was part of a series of unrelated judgment and decision experiments. Participants received € 5 for their participation.

Design and procedure

The same four stimulus sets of Experiment 1 were employed. Participants indicated which of the two executives - the one choosing between the dichotomous pair, or the one choosing between the continuous pair - would be more likely to flip a coin in order to resolve the difficult decision. Order of presentation of both pairs on the laptop screen was counterbalanced.

Results and discussion

Table 2 displays the percentages of participants who thought that, for which of the two executives, choice difficulty was more likely to be resolved by the toss of a coin. Aggregated over the four stimulus sets, more participants (70 out of 117, or 60%) indicated that the executive choosing between dichotomous alternatives would more likely resolve the difficulty by a coin flip, $p = .04$ by a binomial test.

This finding showed that coin tossing is more likely to be accepted when resolving a difficult choice that involves dichotomous rather than continuous attributes. Similarly, the first experiment demonstrated that decision difficulty with dichotomous pairs was more likely explained by lacking compelling reasons. Combined, these results suggest that accepting a coin toss coincides with lacking decisive reasons for choice justification.

Besides using a coin toss for conflict resolution, deferring to choose is another possible means for resolving difficult decisions (Tversky & Shafir, 1992). In case dichotomous and continuous alternatives are contrasted, to what extend will deferral be an acceptable way for resolving decision difficulty? Similar to coin tossing, choice deferral is used when a decision maker finds it

hard to justify a choice by convincing arguments (Anderson, 2003, Chapter 2). Accordingly, it can be predicted that the acceptance of deferral will be higher when resolving difficult decisions involving dichotomous rather than continuous pairs. Furthermore, choice deferral is also predicted to be higher with dichotomous pairs, under the assumption that avoiding to choose depends on (anticipated) regret (Anderson, 2003). Presumably, people are repulsive to toss a coin when anticipated regret is substantial. Instead, they will more likely decide to defer choosing under such circumstances. Substantial regret is, supposedly, experienced when deciding implies that large sacrifices are to be incurred. Moreover, it was hypothesized earlier that these sacrifices were expected to be larger for dichotomous choice alternatives. Therefore, anticipated regret is also expected to be larger with these latter choice options, and accordingly, choice deferral will be more likely for dichotomous than continuous pairs.

Table 2
Percentages of participants for each stimulus set who indicated that, compared to continuous alternatives, difficult decisions among dichotomous alternatives are more likely resolved by a coin toss.

	Resolution difficulty by coin toss	
Stimulus set	*n*	Dichotomous
Pair D1 vs. C1	28	46%
Pair D1 vs. C2	31	67%
Pair D2 vs. C1	30	60%
Pair D2 vs. C2	28	64%

Findings of Experiment 2, though, can be interpreted as being incongruent with the hypothesized larger sacrifices (and hence, larger anticipated regret) for dichotomous pairs. These results showed a higher acceptance of coin tossing with dichotomous alternatives, which suggests that the sacrifices to be incurred may not have been larger for dichotomous, but conversely, for the continuous pairs. Intuitively, people will not accept a coin toss if they anticipate large sacrifices to be incurred, which, in turn, may result in substantial regret if the chosen alternative turns out to be worse than the un-chosen one. This leads to the prediction that, because anticipated regret may have been larger with continuous attributes, deferral will be more likely for continuous rather than dichotomous pairs.

To conclude, it is not clear whether the tendency to defer is higher for dichotomous or continuous pairs. Choice deferral is more likely for the former alternatives, under the assumption that for this type of alternatives a difficult decision is more often explained by lacking compelling reasons for choice

justification. For continuous alternatives, choice deferral is more likely under the assumption that sacrifices may turn out to be larger with these attributes.

The aim of the following experiment was to examine empirically where the occurrence of choice deferral is higher: if a difficult choice is between dichotomous or continuous pairs. In addition, the experiment investigates for which of these two types of pairs, choice difficulty is more likely explained by anticipated regret. This measure of anticipated regret may provide insight as to why people opt for deferral. Finally, the experiment measures for which of these pairs choice difficulty is more likely explained by the difficulty in trading off attributes' pros and cons. This trade off difficulty, supposedly, is another reason for why people defer choosing (Tversky & Shafir, 1992).

Experiment 3

Method

Participants

A total of 293 students from Tilburg University and Fontys University of Professional Education Eindhoven were recruited on campus and paid 4 Euros for their participation. The present experiment was part of a number of unrelated judgment and decision making experiments, all performed on a laptop.

Design and procedure

The same four stimulus sets of Experiment 1 were employed. Participants were assigned to one of the following three conditions. In the "Deferral" condition, participants indicated which of the two executives (i.e., the one choosing between the dichotomous, or the continuous pair) would be more likely to defer choosing in order to resolve the difficult decision. In the "Regret" condition, participants indicated which of the two would more likely explain difficulty by arguing that if the chosen alternative turned out not to be good, it will lead to regret that the other alternative was not chosen. Finally, in the "Trade Off" condition, participants' task was to indicate who would be more likely to explain difficulty by arguing that it is hard to trade off the alternatives' attributes. Order of presentation of the pairs on the laptop screen was counterbalanced.

Results and discussion

Table 3 presents the percentages of respondents who indicated that choice difficulty was more likely to be resolved by deferral. Aggregated over the four stimulus sets, more participants (66 out of 95, or 70%) indicated that the executive choosing between dichotomous alternatives would more likely resolve the experienced difficulty by choice deferral, $p < .001$. In the Regret condition, the combined results showed that more participants (57 out of 93, or 61%) thought that the executive choosing between dichotomous pairs would

more likely explain difficulty by referring to anticipated regret, $p = .04$. Similarly, in the Trade off condition, collapsed over the four sets, more participants (65 out of 105, or 62%) thought that the executive choosing between dichotomous attributes would more likely explain the experienced difficulty by arguing that it is hard to trade off the attributes under consideration, $p = .02$.

The results showed that deferral is more likely accepted when choice alternatives consist of dichotomous attributes. Furthermore, this higher occurrence of deferral coincides with more anticipated regret and the experience of trade off difficulty, suggesting that these are possible antecedents of why someone would defer choosing.

Despite the observed correlation between deferral and anticipated regret, closer inspection of the results for the individual stimulus sets suggests that these two are associated with different stimulus properties. Specifically, with regard to potential regret, the two stimulus sets containing choice pair C2 (i.e., continuous attributes varying on "battery life time" and "memory size") yield results that seem to differ from the two sets that included pair C1 (i.e., continuous attributes varying on "weight" and "available colors"), in the sense that the latter are clearly in the direction of dichotomous attributes, whereas the former lack a clear direction (i.e., close to 50%). Moreover, after the two sets containing pair C2 were combined, and similarly, the two sets containing pair C1, a significant difference between them, regarding a preference for the dichotomous pair, was revealed, χ^2 (1) = 9.99, $p = .001$, supporting that the sets containing C2 or C1 yielded different results in case of anticipated regret. This observed difference, might be explained by arguing that anticipated regret, as an explanation for choice difficulty, focuses not only on the type of attributes involved (i.e. dichotomous vs. continuous), but also on the attributes that are traded off against each other.

Table 3

Percentages of participants for each stimulus set who indicated that, compared to continuous alternatives, difficult decisions among dichotomous alternatives are more likely resolved by deferral, are more likely explained by (anticipated) regret, and are more likely explained by trade off difficulty.

	Deferral		Anticipated regret		Trade off	
Stimulus set	*n*	Dichotomous	*n*	Dichotomous	*n*	Dichotomous
Pair D1 vs. C1	23	61%	24	67%	25	64%
Pair D1 vs. C2	20	70%	24	46%	28	46%
Pair D2 vs. C1	24	79%	21	90%	27	70%
Pair D2 vs. C2	28	68%	24	46%	25	67%

To illustrate, let us assume that anticipated regret is related to the potential sacrifices to be incurred. These sacrifices, because of their all-or-none character, might have been larger with dichotomous compared to continuous attributes. As a consequence, a more regrettable choice outcome is experienced when choosing between, say, a "FM radio" and "time and alarm function" than between alternatives that vary on "weight" and "available colors" (which, for example, can be observed for the Set consisting of D1 and C1, where 67% thought that difficulty is more likely explained by anticipated regret for the dichotomous pair).

At the same time, sacrifices seem to relate to the importance of the specific attributes. That is, the more important an attribute is considered to be, the more a decision maker will experience that a sacrifice is to be incurred by foregoing advantages associated with this important attribute. This explains why percentages close to 50% are observed when a Set includes C2 (instead of percentages clearly in the direction of dichotomous alternatives, which would be the case if anticipated regret is based solely on the all-or-none sacrifices associated with dichotomous attributes). Apparently, a substantial amount of the respondents thought that more anticipated regret was evoked by the attributes of pair C2 ("battery life time" vs. "memory size") compared to the attributes of pair D1 ("FM radio" vs. "time and alarm function") or D2 ("AC power adapter" vs. "picture storage/viewer"). Supposedly, to these participants, the attributes of pair C2 were of more importance than those of D1 or D2, making that for the former the sacrifices to be incurred are experienced as larger. Thus, to these participants it seems harder (and, eventually, more regrettable) to forgo a longer "battery life time" or larger "memory size" by rejecting one of two available players (Pair C2), than rejecting players that have, for instance, either a "FM radio" or "time / alarm function" (Pair D1). In short, whether or not a decision is explained by anticipated regret, is related to the sacrifices to be incurred, which, in turn, seem to depend on the type of attributes, as well as their relative importance.

On the other hand, accepting deferral may more likely focus, not on the importance of the attributes, but solely on the type of attributes (i.e., dichotomous vs. continuous). To illustrate, Table 3 shows that, irrespective of which pair is involved in a decision (i.e., dichotomous or continuous), all sets yield percentages higher for dichotomous pairs (i.e., all percentages were well above 60%). Furthermore, the pattern of percentages observed with deferral differs from that obtained with anticipated regret. Such different pattern might suggest that the antecedents of why one would opt for deferral, or should experience anticipated regret, are different. Presumably, if deferral and anticipated regret were determined by similar attribute properties, then it should be reasonable to expect similar patterns of percentages for these two. A different pattern for deferral and anticipated regret, though, may support the hypothesis that the two were based on different stimulus properties, namely that deferral is based solely on the type of attributes involved, whereas anticipated regret is based on both the type attributes involved and their relative importance.

To test whether a different pattern of results was obtained for deferral and anticipated regret, the sets containing either pairs C2 or C1 were collapsed and submitted to a log-linear analysis with condition (deferral vs. anticipated regret) and stimulus sets (combined sets containing either pair C1 or C2), and type of attribute (dichotomous vs. continuous) as factors. The significant interaction between these three factors, G^2 (4) = 11.70, p = .02, supported the observation that deferral and anticipated regret yielded a different pattern of results. Moreover, this difference implies that deferral and anticipated regret might depend on different stimulus properties.

In sum, anticipated regret may be caused by stimulus properties that differ from those that lead to deferral. As a consequence, anticipated regret may not (robustly) have influenced the decision whether or not to defer choosing. What factor, other than anticipated regret, may make a decision maker avoid choosing? Difficulty in trading off pros and cons is often suggested to influence deferral (Anderson, 2003; Tversky & Shafir, 1992), and congruent with this proposition was a non-significant interaction in the present experiment between the factors condition (deferral vs. trade off difficulty), stimulus sets (combined sets containing either pair C1 or C2), and type of attribute (dichotomous vs. continuous), G^2 (4) = 2.58, p = .63.

The absence of this interaction suggests that when deciding is difficult, both deferral and trade off difficulty seem to invoke a comparison that focuses on similar stimulus properties, namely attribute type (i.e., dichotomous vs. continuous attributes). This comparison faces the problem of how to trade off the attribute values associated with different attributes. When attributes have a dichotomous character (i.e., attribute values are yes or no), deciding among alternatives might be more difficult, because such a comparison pertains to features that the alternatives do not have in common. This rather difficult comparison makes, supposedly, that decision makers more likely defer choosing when deciding among dichotomous rather than continuous alternatives.

Finally, it is important to note that part of the present experiment's conclusions were based on the notion of "attribute importance". It is recommended, though, to measure this "attribute importance" more directly in future research, instead of deriving it from the results as they were obtained in Experiment 3. Such future research regarding attribute importance, supposedly, will further strengthen the conclusions as they were drawn above. Another shortcoming of the present experiment is the small number of participants employed for each stimulus set (see Table 3), which makes that the conclusions drawn from these results are rather tentative. Therefore, future research with more observations for each set is recommended.

General discussion

The main questions concerning choice difficulty raised, by the present chapter, were: which underlying sources of difficulty are more likely associated with either dichotomous or continuous alternatives, and what means for

resolving difficulty are more likely to be employed when contrasting these different types of alternatives. In answering these questions, empirical evidence was presented, suggesting that existing decision difficulty is more likely explained by lacking compelling reasons for justification, when choosing involves dichotomous rather than continuous alternatives (Experiment 1), and is more likely resolved by a coin toss for the former than the latter type of alternatives (Experiment 2). Furthermore, the difficulty stemming from dichotomous pairs is more likely resolved by deferring to choose, and is more likely explained by anticipated regret, as well as by difficulty in trading off pros and cons (Experiment 3).

To conclude, when contrasting dichotomous and continuous pairs, these results suggest that underlying sources of difficulty often correlate with the acceptability of a specific means for resolving the existing difficulty. With regard to dichotomous pairs, for example, an increase in lacking compelling reasons for choice justification was found to coincide with an increase in coin tossing for resolving the existing difficulty. What are possible implications of these uncovered correspondences, obtained in the present chapter? May these results, for instance, suggest that means of resolving difficulty can be used for revealing - and measuring - which sources caused a decision to be difficult?

First, the observed correspondence between lacking compelling reasons and coin tossing is congruent with previous studies on choice difficulty. For example, in Chapter 2 it was demonstrated that, in comparison with large trade off pairs, for small trade off size alternatives (i.e., small differences between alternatives on the continuous attributes), a person is more likely to explain decision difficulty by arguing that it is hard to construct compelling reasons for justification. At the same time, the existing difficulty between these small trade off size alternatives is more likely resolved by tossing a coin. In general, thus, it seems that lacking decisive reasons coincides with coin flipping. This, in turn, suggests that when coin flipping is used as measurement tool for decision difficulty, it may reveal that a decision maker finds it difficult to construct compelling reasons.

Second, results discussed in the present chapter imply that for dichotomous pairs, a higher acceptability of coin tossing (Experiment 2) might coincide with higher anticipated regret (Experiment 3). At first sight, this correspondence seems to be rather inconceivable. Yet, empirical evidence obtained in Experiment 3, concerning the relation between choice deferral and anticipated regret, might explain why this rather peculiar correspondence was found. Specifically, the decision to defer and the experience of anticipated regret seem to depend on different properties of the choice alternatives. Arguably, whether or not a regrettable choice outcome is experienced, depends on the relative importance of the attributes involved, as well as the type of attribute (i.e., dichotomous vs. continuous). The decision whether or not to defer choosing, on the other hand, may focus more on the type - instead of the importance - of the attributes involved. In short, anticipated regret and deferral may be based on different stimulus properties, which explains why, in Experiment 3, they were found to be independent of each other.

Similar to deferral, the decision of whether or not to use a coin toss for resolving a difficult decision, may focus predominantly on the specific attribute type involved, rather than on the importance of these attributes. This (assumed) similarity between these two means of resolving difficulty, would further imply that anticipated regret (as underlying source of difficulty) and coin tossing are independent of each other in the context of dichotomous versus continuous alternatives. Moreover, this independence suggests that whether or not to accept a coin toss is not (robustly) influenced by anticipated regret. As a consequence, it will not be likely that when a coin toss is used as a measurement tool, it reliably indicates whether anticipated regret played a role in the experienced difficulty with dichotomous pairs.

Third, what underlying sources of difficulty may choice deferral measure, when taking into account the sources with which it was found to coincide? As already explained above, it may not be likely that deferral reliably measures whether anticipated regret caused the decision among dichotomous pairs to be difficult. On the other hand, as results of Experiment 3 seemed to imply, deferring to choose may be related to trade off difficulty. Furthermore, a decision maker who experiences that it is difficult to trade off pros and cons, may come to the conclusion that this difficulty is related to lacking compelling reasons for choice justification (Scholten & Sherman, 2006). This suggests that, when confronted with a difficult decision among either dichotomous or continuous pairs, deferral a measurement tool can, similar to coin tossing, be used for revealing whether on not someone might be lacking compelling reasons for choice justification.

Is it possible that choice deferral and coin flipping, in the context of dichotomous and continuous pairs, are influenced by, and therefore measure, some other external or internal factors (Chapter 1) that cause a decision to be difficult? It was conjectured that the pairs, employed in the present experiments, not only differed in terms of the type of alternatives involved (i.e., dichotomous versus continuous), but also with regard to the attributes' relative importance. The latter can be characterized as an (external) factor influencing choice difficulty, as it depends on the decision maker' evaluation of the attributes (Chapter 1). The findings obtained in the present chapter, however, suggest that the use of choice deferral and coin tossing, as means of conflict resolution, are not contingent upon this external factor. Therefore, in the present chapter's choice contexts, choice deferral and coin tossing are not supposed to measure the external factor of attribute importance.

Contrarily to attribute importance, other external (e.g., number of alternatives) and internal (e.g., number of attributes) factors did not, supposedly, vary between the pairs. Therefore, these latter factors can be excluded as antecedents of using choice deferral or coin flipping. For that reason, it seems valid to conclude that, in the context of dichotomous and continuous attributes, choice deferral and coin tossing measure two sources of difficulty, namely trade off difficulty and lack of compelling reasons for justification, and not some other external or internal factors.

Finally, why may choice deferral and coin flipping not be influenced by anticipated regret in the present experiments, whereas these two were found to be influenced by anticipated regret in the context of choice pairs with continuous attributes that varied in trade off size (Chapter 2)? Anticipated regret, experienced in the context of trade off size pairs, seems to be due to the substantial sacrifices that are to be incurred. The difference between pairs in their magnitude of the (to be incurred) sacrifices are, supposedly, more salient if both pairs are characterized on the same attributes (e.g., "room size" and "commuting time"), as in Chapter 2, and may be less obvious when the respective pairs are defined on different attributes (e.g., "FM radio" and "time / alarm function" versus "battery life time" and "memory size"). In effect, the difference between large and small sacrifices may have been more obvious for the trade off size pairs in Chapter 2. Due to this salient difference in sacrifices, anticipated regret might be more likely to influence the acceptability of using choice deferral or coin tossing.

On the other hand, if pairs consist of different attributes (as in the present chapter's experiments), the difference in magnitude of the (to be incurred) sacrifices between pairs may become less apparent. As a consequence, anticipated regret may not (robustly) influence the decision to defer or flip a coin, as was found in the present chapter.

5

General discussion

Making decisions can be difficult. Evidently, when two alternatives are equally attractive, deciding among them may evoke the experience of a difficult to resolve conflict. According to normative accounts, however, choosing between such equally attractive alternatives should not be difficult at all (Shafir & LeBoeuf, 2004). That is, since the alternatives are close in attractiveness, it matters little which one is obtained, and therefore no internal conflict should be experienced. Yet, contrary to normative accounts, decision makers often will *describe* these decisions as difficult, and maintain that choosing between such equally attractive alternatives *does* evoke the feeling of an internal conflict, not seldom accompanied by an affective reaction (Zajonc, 1980), or the notion that one is lacking decisive reasons for justifying one's preference (Shafir, Simonson & Tversky, 1993). As a result of this experienced difficulty, decision makers may rely on means for resolving the apparent conflict, such as choice deferral and random choice.

Previous research on decision difficulty mainly focused on decision deferral employed in choice contexts in which alternatives, defined on several continuous dimensions, were varied in trade off size (e.g., Dhar, 1997). The present thesis also investigated the acceptability of using a random device in these contexts, specifically coin tossing. Furthermore, the present thesis explored the acceptability of these two means for resolving difficulty in contexts on which previous research remained relatively mute, namely, choices in which alternatives were defined on a single dimension (Chapter 3), and where alternatives were characterized by dichotomous attributes (Chapter 4).

Besides their apparent use for conflict resolution, I suggested that choice deferral and coin tossing might also be utilized as measurement tools that provide insight into what caused a specific decision to be difficult. For that reason, possible sources underlying the experienced difficulty (e.g., lacking compelling reasons for choice justification and anticipated regret), in the different choice contexts, were assessed, and it was conjectured that the tendency to defer choosing and accepting a coin toss correlated with these particular sources underlying difficulty.

Findings, obtained in this thesis, yielded correlations implying that choice deferral and coin tossing are both related to difficulty stemming from trading off pros and cons and lacking compelling reasons for justification. In addition, choice deferral was found to correlate with anticipated regret, but only when choosing involved pairs that varied in trade off size (for a more extensive review of the findings, see the Summary section of this thesis).

In the remainder of this chapter, I first discuss the pros and cons of the experimental design, employed in most studies of this thesis. Second, the use of choice deferral or coin tossing as measurement tool is examined. Third, behavioral aspects of both deferral and coin tossing are explored, for instance, why the former is more often used for resolving choice difficulty than the latter in real-life situations. Finally, some shortcomings of deferral and coin flipping as measurement tools are discussed.

Discussion and underlying assumptions of the experimental design

The experimental design, employed in this thesis, investigated participants' thoughts (predictions) about what another person might do in order to resolve a difficult decision (e.g., the likelihood to toss a coin), or how another person would explain the experienced difficulty (e.g., by referring to lacking compelling reasons for choice justification). One may raise the question whether these participants' assessments, of other persons, are accurate? Are participants able to correctly predict other persons' behavior and the content of their mental states, when these other persons are confronted with a difficult decision?

For answering these questions, it is important to explicate two assumptions underlying the experimental design employed in the present thesis. First, the persons described in the scenarios were said to experience great decision difficulty. Liberman and Förster (2006) demonstrated that participants, who are told about someone else experiencing a difficult decision, will subsequently infer that to this other person the alternatives were close in attractiveness.

Second, for predicting the described person's behavior (e.g., tossing a coin) and mental state (e.g., "I am lacking compelling reasons for justification"), participants have to pretend that they themselves were confronted with a (difficult) decision among two equally attractive alternatives. It has often been demonstrated that, not only can people pretend that they are in the position of someone else, they are also capable of correctly predicting the other's mental state and (subsequent) behavior (see Theory of Mind literature, for instance, Leslie [1994]). Thus, with regard to the experiments in the present thesis, and the interpretation of their results, it can be (safely) assumed that participants were able to correctly predict the behavior and beliefs of the person(s) described in the choice problems.

The above assumptions, however, would not have been needed, in case participants themselves first equated (i.e., matched) alternatives, and subsequently, chose which of the two equally attractive alternatives they preferred. Yet, when having to choose between these previously matched alternatives, participants seem to focus mainly on the most important (i.e., prominent) attribute, and prefer the alternative that is superior on this attribute. Apparently, one of the (previously matched) alternatives will dominate the other in choice, supposedly, making the decision among them not difficult, and easy to justify with compelling reasons for justification.

Alternatively, it is possible to ask participants if they could pretend that they themselves would experience great difficulty in choosing between two presented alternatives. However, such instructions could yield nonsensical responses if, according to these participants, one alternative would dominate the other, representing no difficult choice at all. Similarly, these undesirable responses may be evoked when participants think that they themselves are not in the need of obtaining the product (e.g., a printer) described in the scenario.

The experimental design, employed in the present thesis, is supposed to prevent that one alternative dominates the other (i.e., the decision was said to be very difficult). Moreover, the particular design creates a situation in which

the participants' own assessment of the attractiveness does not matter. Specifically, if decision difficulty is attributed to another person, participants will infer that to *this other person* the alternatives were equally attractive, even if they themselves would think that one option is dominating. In this way, they are involved in a scenario that does not concern their own judgment about the attractiveness of the alternatives, but instead, represents a conflict situation where someone else thinks these alternatives *are* close in attractiveness. Finally, a scenario in which some other person has to choose between alternatives, will not lead to a situation in which participants will think about themselves "I do not want to obtain this product", and therefore may argue that they can not pretend that choosing between these alternatives would be difficult to them.

Deferral and coin flipping as measurement tools

The experiments in this thesis showed that both deferral and coin tossing coincided with particular sources underlying choice difficulty. Further, it was conjectured that these observed correlations suggest that these two means, could each be used for resolving difficulty as caused by these specific sources. For instance, deferral and coin tossing might appear to be acceptable means for resolving a conflict that stems from the difficulty in trading off attributes' pros and cons. Scholten and Sherman (2006) contend that the (primary) recognition of this trade off, among attributes, may be followed by the realization that one will be lacking strong arguments for justification. Again, for resolving this latter (second-order) type of conflict, it was argued that both deferral and coin flipping might be acceptable means of resolution.

The two are not acceptable to the same extent for resolving another type of (second-order) conflict, namely that associated with anticipated regret. This feeling of anticipated regret may arise when a trade off implies that substantial sacrifices are to be incurred. Regarding both choice deferral and coin flipping, the former makes a decision maker being able to avoid this anticipated regret. Coin flipping, on the other hand, yields in an inevitable commitment to one of the available alternatives. Consequently, the decision maker has to confront the anticipated regret, which one would rather prefer to shun. For that reason, it is likely that when difficulty stems from anticipated regret, choice deferral will be acceptable, whereas coin flipping will be avoided.

Is it also possible to view choice deferral and coin tossing differently, namely for revealing the circumstances under which deciding is difficult or, more specific, as measurement tools that reveal underlying sources of experienced difficulty? With regard to trade off difficulty, both of them may be equally accepted as means for resolving the experienced difficulty. Therefore, employing them as measurement tools, supposedly, reveals that the experienced difficulty might be caused by this specific source. The (second-order) conflict originating from lacking decisive reasons might, similar to trade off difficulty, also be resolved by both deferral and coin flipping. As such, the two may show, as measurement tools, that difficulty was caused by lacking compelling arguments.

Where the two of them differ, though, is in their acceptability when the (second-order) conflict related to anticipated regret plays a role in the experienced difficulty. What consequence might this difference have, if these two means were to be used for measuring anticipated regret? Consider, for instance, a choice context involving alternatives that vary in trade off size, and in which a decision among these alternatives is experienced as difficult. Under these circumstances, decision makers are willing to defer their decision regardless of whether the trade off size is large or small (Chapter 2; Dhar, 1997). Thus, choice deferral, when used as a measurement tool, would suggest that, irrespective of the size of the trade off, anticipated regret might be the underlying cause of choice difficulty. On the other hand, coin tossing is more likely accepted when trade off sizes are small, and will be avoided when anticipated regret is substantial. This avoidance, in turn, suggests that the anticipated regret must have been rather negligible with smaller trade offs. This latter conclusion, however, contradicts with the one that was previously drawn from using deferral as a measurement tool. In other words, results from coin flipping cast doubt on the conclusion that was drawn on the basis of deferral. Moreover, if choice deferral was used as the only measurement in this context, its conclusion regarding anticipated regret might have been misleading. For that reason, if the purpose is to reveal whether or not anticipated regret caused a decision to be difficult, it might be useful to use both measurement tools in combination.

In sum, it is possible to reveal underlying sources of decision difficulty (i.e., trade off difficulty and lacking compelling reasons) by employing choice deferral and coin tossing as measurement tools. Preferably, though, they should be used in combination, especially when difficulty is, supposedly, stemming from anticipated regret.

Deferral and coin tossing as behavioral aspects of decision making

Instead of resolving difficult choices themselves, decision makers may thus rely on means that resolve the difficulty for them, namely deferring to choose and random selection. Frequently, though, the two are not used to the same extent. Because people often find coin tossing repulsive, they more readily will defer their (difficult) decisions. This repulsiveness, associated with coin tossing, may stem from the (potential) commitment to one of the available alternatives, which is inevitable as soon the coin assigns the decision maker to one of the options under consideration. Deferral, on the other hand, does not lead to any commitment, at least not for the near future. In practice, however, deferral may lead to postponing the decision indefinitely, and to no commitment at all (Tversky & Shafir, 1992), making that the decision maker misses the available alternatives.

Decision makers may also prefer deferral over tossing a coin, because the latter brings along the feeling that one is not in control of the decision making process, and hence the feeling that one can not influence the outcome. This lack of control may lead to undesired feelings such as pessimism and distress (Skinner, 1996), which explains why coin tossing is usually avoided. The

decision process leading to deferral, may invoke the feeling (even if illusory) that the decision maker is in complete control of the situation. Yet, this does not imply that the decision process underlying deferral is necessary a deliberate one. The underlying source(s) that caused the decision to be difficult, for instance, may be unknown to the person who decides to defer.

Another reason why coin flipping occurs less often than deferral, is due to the possibility to call coin tossing a non-rational, or unreasonable, action to take when a decision maker is confronted with a difficult choice. Whether a deliberate act is rational or not, though, depends on the definition of the concept of "rationality". According to Kacelnik (2006), rationality in Philosophy and Psychology (PP-rationality), addresses the question of whether or not a person has used reasoning, and hence focuses on the process that leads to (observable) outcomes, and not on the outcomes themselves. In other words, behavior is called PP-rational if it emerges from a particular reasoning process, which is understood by entities such as "thoughts" and "beliefs". In contrast, a coin is deprived from any reasoning process, nor does it have "thoughts" and "beliefs". For that reason, having a coin in control of the decision making process may be perceived as non-rational, since a decision outcome was not (directly) based on any kind of reasoning process.

Following Kacelnik (2006), in contrast to PP-rationality, rationality in Economics (E-rationality) addresses the question whether a specific behavior results in maximized utility. This latter definition of rationality emphasizes observable behavior, namely the decision outcome, rather than the (reasoning) process that caused it. Following E-rationality, entities such as institutions, which have no "beliefs" and "thoughts" of themselves, can act in a rational way. Furthermore, the use of a random procedure, such as coin tossing, may also be congruent with an E-rational way of thinking, and accordingly, can be interpreted as a rational act. With regard to the latter, if a choice concerns two equally attractive alternatives (i.e., with identical utilities), then it does not matter which one is obtained, since both lead to maximal utility. Accordingly, a fair coin will assign the decision maker with equal chance to one of both choice outcomes, and hence will be consistent with an E-rational way of thinking.

In sum, depending on which definition is applied, the act of using a coin can be judged as either incompatible (PP-rationality) or compatible (E-rationality) with rational considerations. Often, though, the deliberate use of a coin as a decision instrument, is perceived as a non-reasonable (or non-rational) act, and therefore it seems more likely that PP-rationality is, rather than E-rationality, adopted as definition. Deferring to choose, on the other hand, may look as if it emerges from a deliberate reasoning process. Thus, according to PP-rationality, deferral will be perceived as a rational act, which may explain why deferral is more likely to be used than coin flipping when resolving a difficult decision.

With regard to the use of a coin toss, Elster (1989) proposed that employing a visibly arbitrary chance device for making a decision is often repulsive. If people do not like tossing a coin, then why should they accept it as a means for resolving a difficult decision? Moreover, is tossing a coin, as used in the present thesis, an (ecological) valid measurement tool if people, in real life,

resent using a coin toss? To answer these questions it is important to distinguish between (i) accepting a coin toss as a means for resolving decision difficulty, and (ii) accepting the outcome of the coin toss (i.e., commitment to an alternative). In real-life settings, people most likely agree upon using a coin in the former way, yet dislike the latter. That is, if people are confronted with a difficult decision, they may consider to use a coin toss for resolving a difficult choice. Yet, from the moment on the coin commits them to one of the available alternatives, they (ad hoc) reject using this random procedure, supposedly, because (only then) they realize that they are tied to the assignment determined by the outcome of the toss. Thus, in real-life situations, people will accept a coin flip for resolving choice difficulty, but resent (ad hoc) the commitment resulting from the outcome of the toss. However, because they (a priori) may accept a coin toss for resolving the experienced difficulty, coin flipping can be regarded as an ecologically valid measurement tool for decision difficulty.

Finally, concerning choice deferral, may this be called a "means for resolution", as the decision problem remains unsolved? Note that, although choice deferral does not resolve which alternative a decision maker will obtain, it will resolve (at least temporarily) the experienced difficulty involving a choice between close alternatives. Concerning the latter type of resolution, however, it is possible to distinguish between difficulty due to the choice outcome, and difficulty related to employing a solution procedure (strategy) (Zhang & Mittal, 2005).

Outcome difficulty will increase when the decision involves two unattractive, compared to attractive, alternatives (Chatterjee & Heath, 1996). Moreover, the tendency to defer choosing increases when the attractiveness of a choice set decreases (Dhar & Nowlis, 1999), suggesting that deferral is sensitive to outcome difficulty. Intuitively, coin flipping will show the opposite result, namely a (corresponding) decrease when the attractiveness of the alternatives decreases. In contrast to outcome difficulty, choice deferral and coin tossing may also be used when the experienced difficulty stems from employing a particular procedure (i.e., procedural difficulty), for instance, trading off attributes (Tversky & Shafir, 1992). This trade off difficulty, in turn, encompasses both difficulty stemming from substantial sacrifices to be incurred, as well as difficulty that is caused by lacking compelling reasons for choice justification (Scholten & Sherman, 2006).

In the present thesis, both choice deferral and coin flipping were, supposedly, employed for resolving procedural difficulty. Consider, for instance, Experiment 2 of Chapter 2, which demonstrated that deferral was equally likely for printers with large and small trade off sizes. Evidently, similar percentages of choice deferral suggest that there was no difference in (overall) attractiveness between choice pairs, that is, outcome difficulty was similar for both large and small trade off pairs. In contrast, the likelihood to toss a coin changed with trade off size, which could suggest that both pairs' outcome difficulty was indeed varying. Yet, since the deferral results already demonstrated that the outcome difficulty was similar for both pairs, it seems more plausible to assume that the pairs differed in the difficulty associated with

employing a (trade off) process (i.e., procedural difficulty), namely sacrifices to be incurred and lacking compelling reasons for justification. Moreover, it can be argued that choice deferral is employed for resolving difficulty stemming from sacrifices to be incurred, as well as lacking compelling reasons, which explains why deferral was found to be equally likely for large and small trade offs. Coin tossing, on the other hand, is exclusively employed in case the experienced difficulty is stemming from lacking decisive arguments for justification, which, in turn, explains why its acceptance was found to change with trade off size.

In sum, the change in likelihood to toss a coin, contingent upon trade off size, seems to indicate that procedural difficulty was varying (and being resolved), and furthermore, implies that deferral and coin tossing are employed for resolving procedural difficulty, instead of outcome difficulty, in the present thesis' experiments.

Shortcomings of deferral and coin tossing as measurement tools

Despite their purported use as measurement tools for revealing underlying sources of choice difficulty, both deferral and coin flipping display some shortcomings in their proficiency of what they tend to measure.

Related to the above discussed distinction between PP- and E-rationality, is the difficulty in discriminating between (i) the difficulty to generate compelling reasons, and (ii) the strength - or number - of satisfactory reasons for choice justification. The former, apparently, focuses on the process that leads to decisive arguments for justification (cf. PP-rationality), whereas the latter seems to emphasize the output of this process (cf. E-rationality). Although the employment of deferral and coin tossing as measurement tools might reveal that the decision maker is lacking compelling reasons, it will shed no light on whether the difficulty was stemming from the process for generating compelling reasons, or alternatively, the strength of the reasons.

Furthermore, these measurement tools also remain mute about whether a difficult choice process is followed by the decision makers' recognition that convincing reasons for justification are lacking or, conversely, the absence of these reasons for justification makes that deciding is difficult. The first approach presupposes that choosing begins with a process that yields a value reflecting the attractiveness of the alternatives involved. If these alternatives appear to be close in attractiveness, then one will experience a difficult choice. Subsequently, a decision maker may attempt to invoke information for choice justification and apply this, on an ad hoc basis, for resolving the difficult decision (Zajonc, 1980).

The second approach - arguing that lacking compelling reasons is what makes deciding difficult - suggests that deciding starts with constructing reasons in favor or against the alternatives. Subsequently, the generated reasons enter into the decision process, and a difficult choice signifies that these (a priori) reasons for justification are balanced. That is, the alternatives are equally attractive because the arguments are not decisive, indicating no clear

preference for one or the other alternative, which results in the experience of a difficult decision.

Conclusion

The present thesis has explored the acceptability of choice deferral and coin tossing in different choice contexts that confronted a decision maker with a difficult decision. Conjectured was that these two means for resolving decision difficulty, may coincide with particular underlying sources that caused a decision to be difficult (i.e., trade off difficulty, lacking compelling reasons for justification, and anticipated regret). As a result, it was argued that, choice deferral and coin tossing, both can be employed as measurement tools that may reveal the underlying sources leading to the experience of a difficult choice between alternatives.

References

Anderson, A.J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological bulletin, 126,* 130-167.

Beattie, J., & Barlas, S. (2001). Predicting perceived differences in tradeoff difficulty. In E.U. Weber, J. Baron, & G. Loomes (Eds.), *Conflict and tradeoffs in decision making* (pp. 25-64). Cambridge, UK: Cambridge university press.

Chatterjee, S., & Heath, T. B. (1996). Conflict and loss aversion in multiattribute choice: The effects of trade-off size and reference dependence on decision difficulty. *Organizational behavior and human decision processes, 67,* 144-155.

Coin flipping. (2006, August 14). In *Wikipedia, The Free Encyclopedia*. Retrieved 14:00, September 26, 2006, from http://en.wikipedia.org/w/index.php?title=Coin_flipping&oldid=69507668

Coombs, C.H., Dawes, R.M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction.* Englewood Cliffs: Prentice Hall.

Delquié, P. (2003). Optimal conflict in preference assessment. *Management science, 49,* 102-115.

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of consumer research, 24,* 215-231.

Dhar, R., & Nowlis, S.M. (1999). The effect of time pressure on consumer choice deferral. *Journal of consumer research, 25,* 369-384.

Elster, J. (1989). *Solomonic judgements: Studies in the limitations of rationality.* Cambridge, UK: Cambridge university press.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford university press.

Festinger, L. (1964). *Conflict, decision and dissonance.* Stanford, CA: Stanford university press.

Fischer, G.W., Luce, M.F., & Jia, J. (2000). Attribute conflict and preference uncertainty: Effects on judgment time and error. *Management science, 46,* 88-103.

Ganzach, Y. (1996). Preference reversals in equal-probability gambles: A case for anchoring and adjustment. *Journal of behavioral decision making, 9,* 96-109.

Gentner, D. & Markman, A.B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological science, 5,* 152-157.

Gentner, D., & Markman, A.B. (1997). Structure mapping in analogy and similarity. *American psychologist, 52,* 45-56.

Hastie, R. (2001). Problems for judgment and decision making. *Annual review of psychology, 52,* 653-683.

Hendrick, C., Mills, J., & Kiesler, C.A. (1968). Decision time as a function of the number and complexity of equally attractive alternatives. *Journal of personality and social psychology, 8,* 313-318.

Hsee, C.K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational behavior and human decision processes, 67,* 247-257.

Hsee, C.K., & Leclerc, F. (1998). Will products look more attractive when presented separately or together? *Journal of consumer research, 25,* 175-186.

Hsee, C.K., Loewenstein, G.F., Blount, S., & Bazerman, M.H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological bulletin, 125,* 576-590.

Iyengar, S.S., & Lepper, M.R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology, 79,* 995-1006.

Johnson, M.D. (1984). Consumer choice strategies for comparing non-comparable alternatives. *Journal of consumer research, 11,* 741-753.

Judge disciplined for coin flip. (2003, June 21). *Detroit Free Press*.

Kacelnik, A. (2006). Meanings of rationality. In M. Nudds, & S. Hurley (Eds.), *Rational animals?* (pp. 87-106). Oxford, UK: Oxford University Press.

Leslie, A.M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition, 50,* 211-238.

Levin, I. P., & Gaeth, G. J. (1988). How consumers are effected by the framing of attribute information before and after consuming the product. *Journal of consumer research, 15,* 374-378.

Liberman, N., & Förster, J. (2006). Inferences from decision difficulty. *Journal of experimental social psychology, 42,* 290-301.

Luce, M.F., Payne, J.W., & Bettman, J.R. (2001). The impact of emotional tradeoff difficulty on decision behavior. In E.U. Weber, J. Baron, & G. Loomes (Eds.), *Conflict and tradeoffs in decision making* (pp. 86-109). Cambridge, UK: Cambridge university press.

Markman, A.B., & Medin, D.L. (1995). Similarity and alignment in choice. *Organizational behavior and human decision processes, 63,* 117-130.

McKenzie, C.R.M., & Nelson, J.D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic bulletin & review, 10,* 596-602.

Medin, D.L., Goldstone, R.L., & Markman, A.B. (1995). Comparison and choice: Relations between similarity processes and decision processes. *Psychonomic bulletin & review, 2,* 1-19.

Nisbett, R.E., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review, 84,* 231-259.

Payne, J.W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational behavior and human performance, 16,* 366-387.

Payne, J.W., Bettman, J.R., & Johnson, E.J. (1993). *The adaptive decision maker.* Cambridge: Cambridge university press.

Sanford, A.J., Fay, N., Stewart, A., & Moxey, L. (2002). Perspective in statements of quantity, with implications for consumer psychology. *Psychological science, 13,* 130-134.

Schkade, D. A., & Kleinmuntz, D. N. (1994). Information displays and choice processes: Effects of organization, form, and sequence. *Organizational behavior and human decision processes, 57,* 319-337.

Scholten, M., & Sherman, S.J. (2006). Trade offs and theory: The double-mediation model. *Journal of experimental psychology: General, 135,* 237-261.

Shafir, E., & LeBoeuf, R.A. (2004). Context and conflict in multiattribute choice. In D.J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 341-359). Malden, MA: Blackwell publishing.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49,* 11-36.

Simonson, I. (1989) Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research, 17,* 150-162.

Skinner, E.A. (1996). A guide to constructs of control. *Journal of personality and social psychology, 71,* 549-570.

Slovic, P. (1975). Choice between equally attractive valued alternatives. *Journal of experimental psychology: Human perception and performance, 1,* 280-287.

Slovic, P., & MacPhillamy, D. (1974). Dimensional commensurability and cue utilization in comparative judgment. *Organizational behavior and human performance, 11,* 172-194.

Tetlock, P.E. (1997). An alternative metaphor in the study of judgment and choice: People as politicians. In W.M. Goldstein & R.M. Hogarth (Eds.), *Research on judgment and decision making* (pp. 657-680). Cambridge, NY: Cambridge university press.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological review, 110,* 403-421.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review, 76,* 31-48.

Tversky, A. (1977). Elimination by aspects: A theory of choice. *Psychological review, 79,* 281-299.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin, 2,* 105-110.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business, 59,* S252-S278.

Tversky, A., & Shafir, E. (1992). Choice under conflict: The dynamics of deferred choice. *Psychological science, 3,* 358-361.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological review, 95,* 371-384.

Tykocinski, O.E., Pittman, T.S., & Tuttle, E.S. (1995). Inaction inertia: Foregoing future benefits as a result of an initial failure to act. *Journal of personality and social psychology, 68,* 793-803.

Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American psychologist, 35,* 151-175.

Zeelenberg, M., Van Dijk, W.W., Manstead, A.S.R., & Van Der Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition and emotion, 14,* 521-541.

Zhang, Y., & Mittal, V. (2005). Decision difficulty: Effects of procedural and outcome accountability. *Journal of consumer research, 32,* 465-472.

Summary

Choosing can be difficult when deciding among equally attractive alternatives. Usually, though, people try to resolve such difficult decisions. As a consequence, they may consider avoiding to choose (deferral), or alternatively, use a randomizer (e.g., toss a coin). The use of these latter two means for resolving choice difficulty, might coincide with particular sources underlying the experienced difficulty (e.g., lacking compelling reasons for choice justification, and anticipated regret).

If a particular behavior, such as deferral and coin tossing, is found to be correlated with particular underlying sources of choice difficulty, it seems plausible to conjecture that this behavior may bring to light why the decision was difficult. In other words, these means for resolving difficulty can be employed as measurement tools that reveal particular underlying sources of difficulty.

Experiments, reported in this thesis, investigated to what extent choice deferral and coin tossing were used in different contexts concerning a choice between (equally attractive) alternatives. Furthermore, I explored why choosing in these contexts was experienced as difficult (i.e., what sources were underlying difficulty). Combined, the results are indicative of what sources, underlying decision difficulty, choice deferral and coin tossing are supposed to reveal as measurement tools.

Chapter 2 considered alternatives, defined on two attributes, that varied in trade off size. A choice involving alternatives, of which each alternative is superior on different attributes, often evokes the experience of a conflict if one attempts to trade off the alternatives' advantages and disadvantages. Such conflict may arise irrespective of the trade off size (i.e., the size of the differences between attributes). The resulting conflict is often difficult to resolve and, as a consequence, a decision maker may consider to rely on choice deferral and coin tossing. The acceptance of these two approaches, however, was shown to be conditional upon the size of the trade off. In particular, deferral is accepted regardless of the size of the trade off, whereas coin tossing is more likely accepted when trade off size is small.

The observed distinction between deferral and coin tossing, as a function of trade off size, might be explained by suggesting that the two are sensitive to different sources underlying choice difficulty. These sources, presumably, vary depending on the size of the trade off. Congruent with the purported different sources underlying difficulty, was empirical evidence demonstrating that for smaller trade offs, difficulty was more likely stemming from lacking compelling reasons for choice justification. For larger trade off sizes, on the other hand, difficulty was more likely explained by anticipated regret. Finally, difficulty stemming from trading off alternatives' pros and cons was not found to change when trade off size was varied.

Combined, these results suggest that choice deferral, that was found to be accepted irrespective of trade off size, may be used for resolving difficulty stemming from lacking compelling reasons for justification (small trade off

size), anticipated regret (large trade off size), and trade off difficulty (both large and small trade off size). The higher acceptability of coin tossing with small trade off sizes suggests that a coin toss is used when it is difficult to come up with decisive arguments for choice justification, or alternatively, when it is hard to trade off attributes. At the same time, coin flipping may not be acceptable if a regrettable choice outcome is anticipated, which explains why a coin toss is not accepted with large trade off sizes.

Chapter 3 focused on difficult decisions among alternatives described on a single dimension (e.g., job candidates evaluated by several judges using identical rating scales). Obviously, the difficulty between these uni-dimensional alternatives can not result from a trade off between several distinctive attributes. The preference for an alternative appeared to be based on the discrepancy between alternatives' values on the single dimension. In particular, when two alternatives were presented, and one alternative displayed a smaller discrepancy in its values than the other, then the former was likely to be preferred over the latter. Additionally, it was conjectured that convincing reasons for choice justification took into account the displayed difference in discrepancies between the alternatives. Congruent with this proposition was the finding that a different display organization of the same ratings, made the discrepancy between the alternatives on the single dimension less noticeable. This different display organization yielded an observable increase in the likelihood to toss a coin. Supposedly, when alternatives do not display a salient discrepancy on a single dimension, it is hard for a decision maker to construct compelling reasons for resolving the difficult choice. Similar to the findings of Chapter 2, these results demonstrated that tossing a coin may coincide with lacking compelling reasons.

Chapter 4 investigated difficult decisions involving alternatives defined by two distinctive attributes. These attributes were characterized as being either continuous or dichotomous, where continuous attributes are dimensionally represented (e.g., mp3 players that differ in memory size), whereas dichotomous (yes – no) attributes are defined as properties that one alternative possesses but is absent for the other (e.g., one mp3 player comes with an AC power adaptor, whereas the other not). The results obtained in this chapter implied that a difficult choice between alternatives consisting of dichotomous attributes is more likely, compared to continuous alternatives, to be resolved by (i) deferring the decision, and (ii) the flip of a coin. Similarly, the experienced difficulty with dichotomous attributes is more likely explained by (i) lacking convincing reasons for choice justification, (ii) anticipated regret, and (iii) difficulty in trading off pros and cons.

Results reported in Chapter 4 suggest that, with dichotomous alternatives, a difficult decision that was caused by the experience of anticipated regret, may be resolved not only by deferral, but also by the toss of a coin. It seems implausible, though, that a decision maker who anticipates regret would be willing to accept a coin toss for resolving the experienced difficulty. This inconceivable correlation was (tentatively) explained by arguing that accepting deferral or coin tossing, may not be (robustly) determined by anticipated regret.

Moreover, in the context of dichotomous attributes, the acceptability of these two means for resolving difficulty, may more likely depend on a difficult trade off between alternatives' pros and cons, and the lack of compelling reasons for justification.

In sum, the experiments presented in this thesis demonstrated that choice deferral and coin tossing coincided with particular underlying sources of choice difficulty. Specifically, both choice deferral and coin tossing were found to correlate with decision difficulty stemming from trading off attributes, and lacking compelling reasons for choice justification. Furthermore, in the context of choice pairs that were characterized by continuous attributes with a large trade off size (Chapter 2), choice deferral was found to coincide with anticipated regret.

To conclude, these observed correlations suggest that choice deferral and coin tossing are acceptable means for resolving choice difficulty stemming from particular underlying sources. Furthermore, these correlations imply that choice deferral and coin tossing can be employed as measurement tools that will reveal the specific underlying sources that made choosing between alternatives to be experienced as difficult.

Samenvatting

Kiezen tussen even aantrekkelijke alternatieven kan moeilijk zijn. Dikwijls proberen mensen deze moeilijke beslissingen op te lossen door de keuze te ontlopen (uitstellen) of door gebruik te maken van een random procedure, zoals een munt opgooien. Het toepassen van deze twee manieren voor het oplossen van een moeilijke keuze, hangt vaak samen met bepaalde oorzaken die verklaren waarom de keuze in eerste instantie moeilijk was (bijvoorbeeld, het ontbreken van overtuigende redenen die de keuze voor een alternatief rechtvaardigen òf de verwachting dat men spijt kan krijgen van een bepaalde keuze).

Stel dat een bepaald gedrag, zoals uitstellen of het opgooien van een munt, samenhangt met een bepaalde oorzaak (bijvoorbeeld, het ontbreken van overtuigende keuze-argumenten). Als dat het geval is, dan kan worden aangenomen dat dit gedrag, op zijn beurt, aan het licht brengt waarom het kiezen moeilijk was. Met andere woorden, deze manieren van gedrag kunnen wellicht ook gebruikt worden als meetinstrumenten die de oorzaak van de moeilijkheid blootleggen.

De experimenten in dit proefschrift bestuderen, in verschillende contexten, de mate waarin uitstellen en het opgooien van een munt worden toegepast bij het nemen van moeilijke beslissingen. Daarbij is ook onderzocht waarom kiezen als moeilijk wordt ervaren in deze contexten (dat wil zeggen, wat waren de oorzaken van de moeilijkheid). Samengenomen tonen de resultaten aan welke oorzaken van de keuzemoeilijkheid aan het licht worden gebracht, indien uitstellen en het opgooien van een munt gebruikt worden als meetinstrumenten.

Hoofdstuk 2 besteedde aandacht aan keuze-alternatieven die beschreven waren aan de hand van twee attributen die, op hun beurt, verschilden in grootte van trade-off (waarbij de grootte van de trade-off betrekking heeft op de omvang van de verschillen tussen de eigenschappen). Deze alternatieven waren zo samengesteld dat de één beter was dan de ander op één van de twee attributen en vice versa. Bij het kiezen tussen deze alternatieven is het mogelijk om de verschillende voor- en nadelen op de attributen (dat wil zeggen: de eigenschappen van de alternatieven) tegen elkaar af te wegen. Deze afweging van eigenschappen heeft vaak tot gevolg dat mensen een conflict ervaren. Zo'n conflict, dat kan voorkomen bij grote alsmede kleine trade-offs, is vaak moeilijk op te lossen, met als gevolg dat een persoon de keuze uitstelt of een munt opgooit. De bereidheid om zich te verlaten op één van beide manieren van conflict-oplossen, lijkt af te hangen van de grootte van de trade-off. Dat wil zeggen, uitstellen vindt plaats ongeacht de grootte van de trade-off, terwijl het opgooien van een munt meer waarschijnlijk is bij een kleine trade-off.

Het geobserveerde verschil in accepteerbaarheid - tussen uitstellen en het opgooien van een munt - kan verklaard worden door dat beide wellicht gevoelig zijn voor verschillende oorzaken die een keuze moeilijk maken. Deze oorzaken zijn, op hun beurt, mogelijk afhankelijk van de grootte van de trade-off. In overeenstemming met de gestelde afhankelijkheid tussen trade-off-grootte en soort van moeilijkheid, waren resultaten die aantoonden dat voor een

kleine trade-off het moeilijk was om overtuigende redenen te vinden voor het rechtvaardigen van een bepaalde keuze. Voor een grote trade-off, aan de andere kant, is kiezen moeilijk vanwege de spijt die men eventueel kan krijgen nadat voor een optie gekozen is. Ten slotte, moeilijkheid veroorzaakt door het afwegen van de eigenschappen lijkt plaats te vinden ongeacht de grootte van de trade-off.

Als deze resultaten worden gecombineerd dan schijnt het dat uitstel van een keuze (die plaatsvindt ongeacht de grootte van de trade-off) gebruikt kan worden om een moeilijke keuze op te lossen die wordt veroorzaakt door (i) een gebrek aan overtuigende keuze-argumenten (kleine trade-off), (ii) verwachte spijt (grote trade-off), en (iii) het afwegen van eigenschappen (zowel bij kleine als grote trade-off). Het opgooien van munt, dat alleen plaatsvindt bij een kleine trade-off, suggereert dat de munt toegepast wordt bij het oplossen van een moeilijke keuze waarbij overtuigende keuze-argumenten ontbreken, en waar eigenschappen moeilijk tegen elkaar af te wegen zijn. Verder lijkt het opgooien van een munt niet toegestaan indien achteraf een bepaalde keuze spijt tot gevolg heeft. Dit laatste verklaart waarom mensen niet bereid zijn een munt op te gooien bij grotere trade-offs.

Hoofdstuk 3 richtte zich op alternatieven die beschreven waren aan de hand van slechts één dimensie (bijvoorbeeld, twee sollicitanten die beoordeeld zijn door verschillende beoordelaars die ieder één en dezelfde meetschaal gebruikten). Het mag duidelijk zijn dat een moeilijke keuze tussen deze één-dimensionale alternatieven niet veroorzaakt wordt door het afwegen van verschillende eigenschappen. De voorkeur voor een alternatief lijkt, in dit geval, neer te komen op het verschil dat de verschillende alternatieven op de enkele dimensie hebben. Dat wil zeggen, als twee alternatieven worden aangeboden, en één vertoont een kleiner verschil in zijn schaalwaarden (de beoordelaars geven ieder beoordelingen die weinig van elkaar verschillen) dan de ander, dan hebben mensen een voorkeur voor het eerste alternatief. Overigens lijken keuzeredenen, om de gemaakte keuze mee te verantwoorden, te verwijzen naar het verschil in schaalwaarden tussen de alternatieven. Indien deze schaalwaarden op een andere manier worden gepresenteerd, waardoor dit verschil minder in het oog springt, dan hebben mensen waarschijnlijk meer moeite met het vinden van overtuigende keuze-argumenten. Dit duidt er op dat een (eventueel) zichtbaar verschil tussen de schaalwaarden gebruikt wordt voor het bedenken van overtuigende keuzeredenen. Bevindingen in dit hoofdstuk doen vermoeden dat een dergelijke presentatie van de schaalwaarden, waarbij de verschillen minder in het oog springen, ertoe leidt dat mensen sneller bereid zijn een munt op te gooien. Dit laatste maakt het aannemelijk dat als een verschil tussen de schaalwaarden moeilijk waarneembaar is, mensen het moeilijk vinden om overtuigende keuze-argumenten te genereren. Deze bevindingen laten opnieuw zien, in overeenstemming met de resultaten uit Hoofdstuk 2, dat het opgooien van een munt acceptabel is als mensen geen overtuigende keuzeredenen kunnen bedenken.

In Hoofdstuk 4 werden moeilijke keuzes tussen alternatieven, beschreven op twee verschillende soorten attributen, onderzocht. Deze attributen waren

continu of dichotoom van aard. Het eerste type attribuut wordt aan de hand van een dimensie beschreven (bijvoorbeeld, mp3 spelers die verschillen wat betreft hun geheugengrootte). Het tweede soort attribuut wordt beschouwd als een (ja-nee) eigenschap dat één alternatief bezit maar bij de ander afwezig is (bijvoorbeeld, één type mp3 speler wordt geleverd met een stroomadapter, maar de andere(n) niet). De resultaten in dit hoofdstuk tonen aan dat, vergeleken met continue alternatieven, een moeilijke keuze tussen opties met enkel dichotome attributen eerder opgelost wordt door (i) de keuze uit te stellen, en (ii) een munt op te gooien. Daarnaast wordt de moeilijkheid met dichotome attributen eerder toegeschreven aan (i) het ontbreken van overtuigende redenen om de uiteindelijke keuze mee te verantwoorden, (ii) spijt die eventueel verwacht wordt, en (iii) de afweging van de eigenschappen.

De resultaten in Hoofdstuk 4 doen vermoeden dat een moeilijke die veroorzaakt wordt door eventuele spijt achteraf, zowel opgelost kan worden door de keuze uit te stellen, alsook door het opgooien van een munt. Niettemin lijkt het onwaarschijnlijk dat iemand die een moeilijke keuze moet nemen, en daarbij achteraf spijt kan krijgen met een bepaalde keuze, bereid is een munt te gebruiken. Deze onwaarschijnlijke samenhang tussen het opgooien van een munt en eventuele spijt achteraf, die gesuggereerd wordt door de gevonden resultaten in dit hoofdstuk, kan verklaard worden door aan te nemen dat bij dichotome alternatieven het opgooien van een munt niet samenhangt met (of anders gezegd: niet veroorzaakt wordt door) de spijt die iemand achteraf kan verwachten. Bij het kiezen tussen dichotome eigenschappen lijkt het waarschijnlijker dat de ervaren moeilijkheid veroorzaakt wordt door het ontbreken van overtuigende redenen om de keuze mee te verantwoorden, en de moeilijke afweging tussen de betrokken eigenschappen.

De gevonden samenhangen in dit proefschrift suggereren dat het uitstellen van een keuze en het opgooien van een munt, beide manieren zijn om een moeilijke keuze mee op te lossen die teweeggebracht wordt door bepaalde oorzaken. Zowel het uitstellen als het opgooien van een munt kunnen in verband worden gebracht met keuzemoeilijkheid die veroorzaakt wordt door het afwegen van eigenschappen, alsook een gebrek aan overtuigende redenen waarmee de uiteindelijke keuze kan worden verantwoord. Indien keuze-opties beschreven worden aan de hand van attributen, waarbij sprake is van een aanzienlijke trade-off-grootte (Hoofdstuk 2), dan lijkt uitstellen samen te hangen met de spijt die iemand eventueel achteraf kan verwachten. Aan de hand van deze gevonden verbanden kan, ten slotte, worden gesteld dat de twee beschreven manieren, die gebruikt worden om een moeilijke keuze mee op te lossen, wellicht geschikt zijn als meetinstrumenten die de oorzaken van een moeilijke keuze aan het licht brengen.

Acknowledgements

This thesis has benefited from the discussions I had with my colleagues at the JDM research group and the J.F. Schouten school, both at the Eindhoven University of Technology. Thank you.

I owe special thanks to Gideon Keren for his supervision, and constructive comments during this research project. I would also like to thank Janneke Staaks and Marc van Buiten for joining me on the (4 year) tour along several universities in the Netherlands.

Finally, I would like to thank all my friends (you know who you are!) for being there. A special and warm thanks goes out to Jess and my mother for their constant and invaluable support.

Stefan

Curriculum vitae

17 March 1974	Born in Weert, The Netherlands.
1986 – 1993	VWO (pre university education), Bisschoppelijk College Echt.
1993 – 1997	Mechanical engineering, Fontys University of Professional Education, Eindhoven, B.Sc. degree.
1998 – 2002	Cognitive psychology, Maastricht University, faculty of psychology, M.Sc. degree (Cum Laude).
2002 (January – July)	Visiting scholar, Michigan State University, Lansing (MI), department of psychology.
2002 - 2006	Ph.D. student, J.F. Schouten school for User-System Interaction, Eindhoven University of Technology, faculty of Technology Management.