Sets

Evers, E.R.K.

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Sets

Ellen R. K. Evers

SETS

How the organization of stimuli affects judgments & choice

Proefschrift

ter verkrijging van
de graad van doctor aan Tilburg University
op gezag van de rector magnificus, prof. dr. Ph. Eijlander,

in het openbaar te verdedigen ten overstaan van
een door het college voor promoties aangewezen commissie
in de aula van de Universiteit
op vrijdag 13 februari 2015 om 14:15 uur

door

Ellen Riemke Katrien Evers
geboren op 26 december 1982 te Tilburg.

Promotor:	Prof. dr. Marcel Zeelenberg
Copromotor:	Dr. Yoel Inbar
Promotiecommissie:	Prof. dr. George Loewenstein
	Prof. dr. Barry Schwartz
	Prof. dr. Wilco van Dijk
	Prof. dr. Gideon Keren
	Prof. dr. Rik Pieters
	Dr. Bram van den Bergh
	Dr. Marijke van Putten

Table of Contents

CHAPTER 1

Introduction

In daily life, we are often confronted with sets. Think for example about shopping in the supermarket where goods are organized in groups of similar products, or playing a computer game in which achievements and rewards are organized in sets. Sometimes these sets are created by a manufacturer to be collected and used as a set, for example collectibles or boxed record-sets. Other times, manufacturers just create items that are not necessarily intended to use as a set but still have some overarching theme, like books by the same author, albums of the same band, or Starbucks mugs with a different illustration for each city. Finally, people may create their own sets. They may for example bring home an individual magnet on each trip they take, and display them together on a fridge. In such situations, items that were never intended to be used together are still perceived by the consumer as a set.[1]

Because sets are so prevalent in daily life, it is important to know if and how our behavior is influenced by groups of items being perceived as part of a set. Based on both psychological theory as well as everyday observations, it is likely that behavior is affected by the fact that items can be seen as (part of a) set. It is therefore surprising that virtually no research has been conducted to try and find out whether and, if so, how sets influence decision making.

Observational

As already mentioned, sets seem to be everywhere. Manufacturers produce their products in ways that make these products seem to belong together, having a consistent design (for example; see Figure 1.1).

People also create sets themselves, with an estimate of one in three Americans being a collector of items according to a certain theme (Belk, Wallendorf, Sherry, Holbrook, & Roberts, 1988). Furthermore, supermarkets and certain manufacturers of foods provide consumers with freebies that are structured around a common theme and that can be collected. This often dramatically increases sales, with the freebies becoming a hype resulting

1 Sets do not necessarily have to contain material products, but can also represent experiences (e.g., different rollercoaster rides) or accomplishments (e.g., medals for sports events). The latter are very prominent in computer-games where gamers can collect virtual trophies, or badges.

Figure 1.1: Sets of products having a consistent design

in remarkably unusual behavior. For example, in 2009, the Albert Heijn-supermarket in the Netherlands gave away a pack of 5 stickers with pictures of soccer-players for every 5€ spent. Similar pictures can be bought from the famous sticker-book company Panini for approximately 12 cents each. Because so many people started collecting these stickers and seem to be willing to break several social norms just to acquire these stickers (e.g., harass shoppers that exit the shop), the Albert Heijn felt forced to use crush barriers in front of their stores to keep customers from harassing and assaulting each other (Trouw 2009a, see also Figure 1.2). While these stickers were distributed actual fights broke out, kids stole them from each other, from their teachers, and some were even officially prosecuted for these crimes (see for example; Algemeen Dagblad 2009; Te Pas, 2009; Trouw, 2009a; Trouw 2009b). Why were people willing to put in this much effort, and even violence to acquire these stickers when similar stickers could be bought for a few cents? Few would agree that something worth 12 cents is worth fighting or stealing for.

But even the buying of Panini-stickers is difficult to reconcile with rational decision making as assumed in normative theories of choice. Recently a group of mathematicians have calculated the costs of one full album when the person collecting the stickers uses the most optimal completion strategy (Sardy & Velenik, 2010). In this optimal case, in which it is assumed that the collector has 9 friends using the same strategy and willing to trade, the cheapest for which the set can be completed is around $135, a hefty sum for a small booklet with 660 pictures. It seems unlikely that all these consumers would, in advance, indicate that a full book of stickers is worth $135 to them.

Besides collectible stickers, manufacturers also seem to put in extra effort to provide consumers with other sets that they can collect and complete. Microsoft even forces its game-developers to include virtual badges that can be collected. No data about whether

Figure 1.2: Kids lining up behind crush barriers in the hopes of getting collectable soccer-stickers from regular shoppers at the Albert Heijn supermarket.

such attempts actually increase revenue is publicly available, but assuming that the overall aim of manufacturers is to make money and realizing that adding freebies costs money and time, the provision of collectable sets is likely to motivate people to consume more of their goods.

Finally, when looking at sets, there appears to be some common theme in their design. Most sets consist of items that share a clear similarity between all items in the set; they usually feature the same shape and design but are all different colors or are all the same shape and color-scheme but feature different images (see for example Figure 1.3). This could of course be a (remarkable) coincidence, but it could also be that sets adhering to such characteristics are preferred over other designs.

Theoretical

As already mentioned, there is no research looking directly at how people interact with sets, as far as I am aware. There is, however, limited research in domains that could possibly be generalized to behavior expressed by people when interacting with sets. First

Figure 1.3: Two sets of consumer products. On the left a set of Trexi toys that consists of items all having the same shape and design but different colors. On the right a set of Penguin books that are all the same shape and same colors, but with different designs.

of all, we know that when people are confronted with a large amount of stimuli, they automatically group and organize those stimuli according to certain principles (Koffka, 1935; Wertheimer, 1923). Furthermore, there is some evidence that people like some patterns (which are organizations of visual stimuli) more than others (see for example; Garner, 1970), though there is no consensus *why* some patterns are liked better than others (Garner, 1970; Glanzer & Clark, 1964; Palmer, 1991; Pomerantz & Lockhead, 1991). One common feature of patterns that are generally liked better, is that the organization of these patterns is less complex. They exhibit a very clear organization (for example see; Figure 1.4).

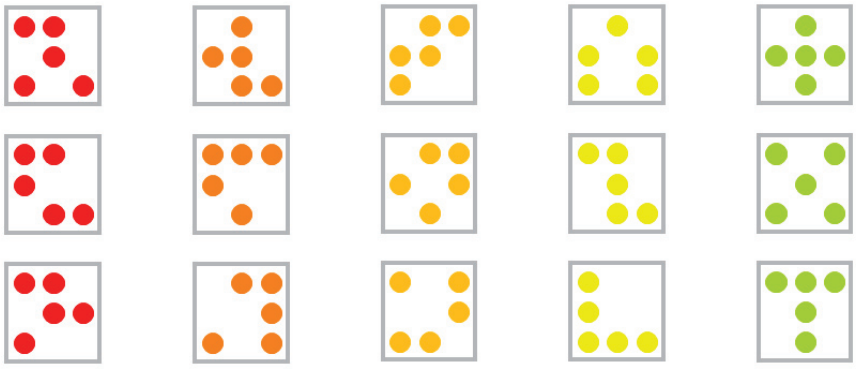


Figure 1.4: Examples of patterns taken from Garner & Clement (1963) that were rated on their "goodness". Colors reflect average rating with the patterns on the left rated lowest on pattern goodness and those on the right highest.

Furthermore, philosophers and psychologists have for a long time posited that people do, and should, generally prefer "data" (information, visual stimuli, etc.) that is low in complexity (see for example; Attneave, 1954; Chater, 1999; Chater & Vitanyi, 2003; Kintsch, 2012; Schmidhuber, 1997; Wertheimer, 1923).

Hypotheses

Accepting the premise that people generally prefer organizations that are low in complexity leads to the following broad hypotheses:

- The lower the complexity of a set of stimuli, the more people like this set.
- As a consequence, reducing the complexity of a set increases preference for this set.
- People are motivated to reduce the complexity of a set.

These ideas lead to more concrete hypotheses on how people interact with sets, of which we tested the following:

When owning a (part of a) set;

- If perceived as an incomplete set, people will be motivated to complete it.
- When adding items to this set, people will prefer items that simplify the set over other items.
- When adding items to this set, people will prefer items that keep the set simple over those that make the set more complex.

When choosing between sets:

- People will prefer a set that is objectively simpler over a more complex set.
- Providing information that simplifies a set increases preference for that set.

Set-Induced Violations of Rational Choice Theory

These hypotheses may not seem particularly surprising. However if supported by data, they violate several axioms of rational choice theory one of the most prevalent theories on how people make decisions. In its simplest form rational choice theory assumes that people, before taking any action, calculate the costs and benefits of taking the action and then act according to which action results in the highest value for themselves. In other words, in these models people are assumed to be utility maximizers. This implies that, unless we assume sets add some form of unknown utility, the existence of sets should not affect decision making in any way.

Outline of this Dissertation

In what follows, I present the first steps I have taken together with my collaborators to gain an understanding in when and how sets influence decision making. The first two empirical chapters focus on how the perception of a set influences our motivation to acquire items that are part of that set. In Chapter 2 we show that when people perceive to possess an almost complete set, they are strongly motivated to complete it paying extra for items that complete the set; a set-completion premium. This effect even emerges for things people would normally be aversive to such as gross pictures.

In Chapter 3, we test whether this effect also emerges in non-product domains such as virtual badges earned by playing computer-games that are particularly fleeting. Furthermore, we show that by creating subsets within larger sets of such virtual badges, motivation is increased even more because there are multiple opportunities for (sub)set-completion.

In Chapter 4 we go from acquisition to liking and preferences. In Chapter 4 we investigated whether the organization of a set affects how much people like this set. We find that people generally prefer sets that are either all-different or all-similar on all salient dimensions.

Finally, in Chapter 5 we test whether the preference for such structured sets results in violations of normative rational choice theory (they do) and show the perception of structure as the underlying mechanism.

The underlying process, future directions, relation to other psychological theories and practical implications are discussed in the general discussion. I have already empirically examined some of these relationships and implications, these studies are described in detail in the appendix.

Since none of this research could have come to light without my co-authors, all empirical chapters, as well as descriptions of studies I conducted with co-authors are written in “we-form”. However, the introduction and general discussion largely reflect my own thoughts about the causes and consequences of sets and are therefore writing in “I-form”.

CHAPTER 2

Set Completion Premium[2]

We investigate how the existence of sets influences the decisions of consumers. Across four experiments, we find that owning a large proportion of a set motivates consumers to complete it. We first demonstrate the existence of a set-completion premium in Experiments 1 and 2. Next we investigate the underlying process, and find that owning a large proportion of a set makes it more likely that the owned items are seen as an incomplete set which motivates consumers to complete it (Experiment 3). Finally, in Experiment 4, we find that these effects emerge even for sets of aversive stimuli.

2 This chapter is based on Evers, E. R. K., Zeelenberg, M., & Inbar, Y. (2014). *The set-completion premium*. Manuscript under review.

Introduction

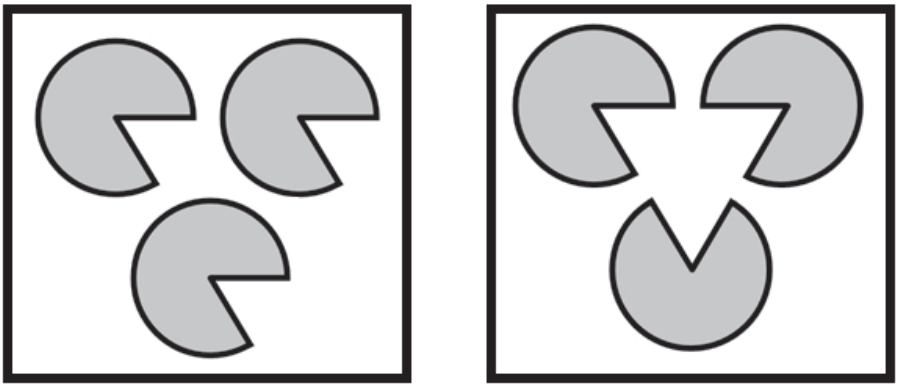
Products are often produced in sets. Sometimes these sets are intended for consumers to collect in their entirety, such as a specific predefined set of baseball cards or beanie babies. Other sets may have an overall theme even though they are not explicitly intended for collectors—for example, the Adidas sneaker collection inspired by Star Wars™ , or the many different Casio G-Shock™ watches. Researchers know quite a bit about the motivations and behavior of people who identify themselves as collectors (Belk, Wallendorf, Sherry, & Holbrook, 1991; McIntosh & Schmeichel, 2004), but almost nothing is known about how sets affect the behavior of ordinary consumers. Of course, research on assortment effects has investigated how consumers choose from sets (Kahn & Wansink, 2004). However, with a few exceptions (Evers, Inbar, & Zeelenberg, 2014; Mourey, Oyserman, & Yoon, 2013), researchers have not studied how owning part of a set affects subsequent choices. This is remarkable, because everyday observation suggests that people's purchasing decisions are often influenced by the sets to which products belong.

One of our colleagues was an avid reader of Nancy Drew mysteries as a child. At the time, she acquired all but three of the 25 original books manufactured. Despite not having looked at them in over 10 years, if she ever finds the three missing books she plans to buy them immediately—so that she can put them in her parents' basement with the other 22. Evidently, our colleague is motivated to complete her set of books even though she no longer has any interest in reading them. In other words, her reason for buying the three remaining books has nothing to do with their individual attributes—*except insofar as they belong to a nearly-complete set already in her possession*. We do not believe that this makes her unusual. Rather, we think that this story is an example of a general fact: Consumers often seem motivated to complete sets, even when this was not their initial goal—and even when the set-completing items do not seem to provide much intrinsic utility.

Why might this be the case? One part of the explanation is likely a well-known phenomenon called the "goal gradient": All else equal, animals (including people) work harder to reach a goal when they are closer to achieving it (Miller, 1944). With people, this happens even when the absolute distance to the goal is held constant but its subjective closeness varies. For example, in one study people were given customer loyalty cards on which they could collect stamps towards a free cup of coffee. In one condition people were given an empty card with room for ten stamps. In the other, they were given a card that had room for 12 stamps and had already been stamped twice. Even though the number of remaining stamps required for a free coffee was the same in both conditions, people given the pre-stamped card returned to buy coffee more often, presumably because they felt subjectively closer to a full card (Kivetz, Urminsky, & Zheng, 2006).

But the goal gradient effect alone is an insufficient explanation of ordinary consumers' motivation to complete sets (although it has been proposed as an explanation for the motivations of self-described collectors, (see; Carey, 2008; McIntosh & Schmeichel, 2004). The reason is that by definition, goal gradient effects operate *only when people already have a goal*. This does not seem to be the case for many ordinary consumers (such as our colleague described above) who have no particular goal in mind when they start buying items. Rather, in these cases it seems that the goal to complete the set *emerges* once one owns a certain proportion. In other words, there seems to be something about almost-complete sets that creates a completion goal, even when this goal was not there from the start. We propose that such a goal may arise as a consequence of the "gestalt" of nearly-complete sets.

Early research by the gestalt psychologists on perceptual grouping started with the question of why we perceive meaningful structures instead of mere collections of perceptual elements. In his pioneering work, Wertheimer (1923) showed that people automatically find unifying patterns ("gestalts") in groups of stimuli, and described several principles of how these gestalts are structured. Gestalts are perceived automatically, and ignoring them takes substantial effort. In some cases doing so seems nearly impossible—the structure of the individual stimuli almost "forces" an overarching global experience on the perceiver. A well-known example is the principle of closure (see Figure 2.1). Even though the stimuli in both images are partial circles, the "negative space" triangle in the image on the right is nearly impossible to ignore. Instead of merely seeing three individual partial circles, a global overarching structure becomes salient.



Three partial circles Three partial circles creating the Gestalt of a triangle

Figure 2.1: Illustration of the emergence of the Gestalt.

We believe that something similar happens in the case of items that belong to a larger set. When the items owned constitute a small proportion of the whole, they are perceived as separate and individual (analogous to the circles in Figure 2.1 on the left). However, when these items constitute a large proportion of the whole, they may be seen less as individual items and more as a part of a set. Just as the perceptual experience of a triangle emerges from the three partial circles in Figure 2.1 on the right, the perceptual experience of a partial set may emerge from the individual items. Once these items are perceived as a set, the complete set becomes a natural reference point—and reference points often evoke a goal to attain them (Heath, Larrick, & Wu, 1999; Pope & Simonsohn, 2011). This explains why consumers may find themselves with a goal to complete sets despite having had no such goal when first item from the set was purchased. Once a large proportion of the set is owned, the items in one's possession start to look like a nearly-complete set. When this happens, the full set becomes one's reference point—prompting a focus on the missing items and motivating action to obtain them.

We test this account in seven experiments, of which four are reported in this manuscript and three are described in the appendix. In all experiments we examine whether consumers who possess partial sets are more motivated to obtain additional items when these items complete the set than when they do not (holding the actual number of items owned constant). In Experiment 2.3, we test whether consumers owning a nearly-complete set are in fact more likely to see the owned items as a set, rather than as individual items, and whether this motivates them to obtain the missing items. Finally, since this effect is expected to emerge purely from a drive for completion, we tested whether the same effects would be found using sets of clearly aversive stimuli.

Across the four experiments, we operationalize motivation in different ways to ensure generalizability: actual and hypothetical willingness to pay (Experiments 2.1 and 2.2-2.3, respectively) and effort spent to obtain missing items (Experiment 2.4). Furthermore, we test these effects both using desirable (Experiment 2.1-2.3) and undesirable (Experiment 2.4) goods.

Experiment 2.1

Experiment 2.1 was our first exploration of the set-completion premium. We expected that people would be willing to pay more for an item when it completed a set as compared to when it did not.

Method

During a week-long testing session[3], 137 students from Tilburg University (89 females, M_{age} = 21.3, SD = 3.0) were first introduced to the BDM procedure for eliciting willingness to pay (Becker, DeGroot, & Marschak, 1964) and told that they would be bidding on a product for real later in the experimental session. After an unrelated task, the experimenter came into each cubicle and gave participants three pens (blue, green, and black) of a set consisting either of four pens in total, or ten pens in total (see Figure 2.2).

Set-completion condition Control condition

Figure 2.2: Two sets of pens used in Experiment 2.1

Participants were told that these pens were theirs to keep and that they now had the chance to bid—with their own money—on the 4th (red) pen using the previously practiced BDM procedure. We expected that participants would be willing to pay more for the red pen when it was the last pen in a set of four, as compared when it was the fourth of a set of ten.

Results and Discussion

As expected, participants were willing to pay more for the red pen when it completed the set of pens (i.e., it was fourth of a set of four; M = 29.0 cents, SD = 45.9) as compared to when it did not complete the set (i.e., it was fourth of a set of ten; M = 16.1 cents, SD = 21.8, $F(1, 135) = 4.46$, $p = .04$, $d = 0.38$). Thus, participants were willing to pay more for an identical pen when this pen completed a set of pens as compared to when it did not. We conceptually replicated these findings in two studies described in the appendix as Experiment 7.1 and 7.2.

3 We ran the study in a weekly testing session which usually results in around 120 participants. Our stopping-rule was that data-collection would end at the end of the week.

Experiment 2.2

The set-completion account applies most clearly in cases where consumers own nearly all of a set, because it depends on seeing the items in one's possession as an incomplete set rather than as separate goods. To gain more insight into when exactly consumers shift from perceiving a group of individual objects to perceiving an incomplete set, we tested how motivation to complete sets varies with the proportion of the set owned. Experiment 2 included six conditions in which we systematically varied what proportion of a complete set participants owned, from a low of four of ten to a high of four of five.

Method

A total of 431 workers on Amazon.com's Mechanical Turk (mTurk)[4] were assigned to one of six conditions in which they were asked to imagine that they owned four mugs from a set of (depending on condition) five, six, seven, eight, nine or ten mugs (see Figure 2.3). They

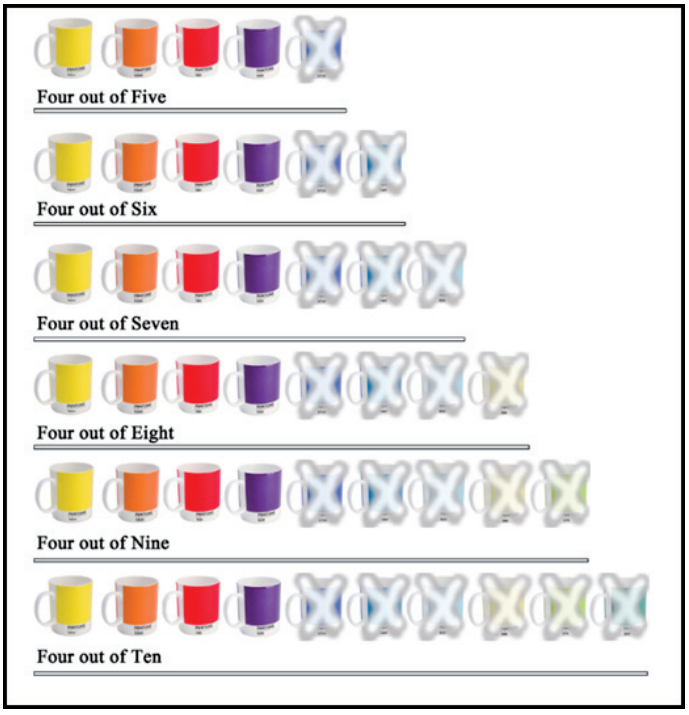


Figure 2.3: Illustration of the sets of mugs used in the six condition in Experiment 2.2.

4 We aimed for 60 participants per cell after exclusions and therefore requested 70 participants per condition. Eleven Amazon mTurk workers did not submit their HIT in time resulting in slightly more participants than requested.

were asked to indicate how much they wanted the fifth mug of the set (on a 100-point scale anchored on the left with *"not at all"* and the right with *"very much"*), and how much they were willing to pay for it (open ended). After this, we asked two attention-check questions (*How many mugs did you own* and *How many mugs did the set consist of?*). After removing the participants who failed the attention check, a total of 338 participants remained (130 females, M_{age} = 29.1, SD = 8.8).[5]

Results and Discussion

All means and standard deviations are reported in Table 2.1. As can be seen there are no strong differences between the conditions where participants imagined they owned four mugs of a set of six, seven, eight, nine or ten. Only when participants imaged owning four of five mugs (80% of the full set) did wanting and willingness to pay increase substantially.

To investigate the pattern further, we ran a similar but more extensive study in which we not only varied the size of the full set, but also the size of the (imagined) endowment. This study is reported in the appendix as Experiment 7.3. Those results match the findings in this experiment, revealing an increase in wanting and willingness to pay when one is missing the last item of the set. However, in that study, small increases in wanting and willingness to pay were also found for participants owning 10/13 mugs, suggesting that the effect does not *only* occur when one is missing the last item of the set but likely emerges once people own a large proportion (70-80%) of a set.

Table 2.1: Wanting and Willingness to pay for the 5th blue mug in Experiment 2.2. Planned contrasts between the set-completion condition (4 out of 5) and the other experimental conditions are provided.

	Wanting		Wanting vs. 4 out of 5		Willingness to Pay		Willingness to Pay vs. 4 out of 5	
Condition	*M*	*SD*	*t*(332)	*p*	*M*	*SD*	*t*(327)	*p*
4 out of 5	62.9	29.1	—	—	$7.14	5.50	—	—
4 out of 6	53.3	30.8	1.7	.09	$5.71	3.05	1.9	.06
4 out of 7	47.5	31.1	2.7	< .01	$6.05	4.33	1.4	.16
4 out of 8	47.8	32.0	2.6	.01	$5.40	4.17	2.2	.03
4 out of 9	46.3	31.3	2.6	.01	$5.43	3.98	1.9	.05
4 out of 10	46.1	30.8	2.9	< .01	$4.97	3.47	2.8	< .01

Note: For Wanting, none of the contrasts between the non-completion conditions were significant ("4 out of 6" vs. "4 out of 10" p = .19, all other p's > .25). For Willingness to Pay none of the contrasts between the non-completion were significant ("4 out of 7" vs. "4 out of 10" p = .16, all other p's > .32)

5 In all conditions around 10-12 participants failed the attention check. In the condition where participants owned four out of a set of nine, 34 participants failed the attention check—most erroneously reported owning 10 mugs.

Experiment 2.3

In Experiments 2.1 and 2.2 we found clear evidence that people want an item more, and are willing to pay more for it, when it completes a set than when it does not. In the current experiment, we aimed to replicate this effect, but also to test the next part of our proposed account—that those who possess a large proportion of a set are more likely to see these items as part of a set than are those who possess a smaller proportion. Furthermore, we expected that seeing the items one owns as part of a set would: 1) correlate with wanting and willingness to pay for an additional item *within* conditions; and 2) mediate the differences in wanting and willingness to pay for an additional item *between* conditions.

Method

We used the same stimuli as in Experiment 2.2, but ran only two conditions. Participants were asked to imagine owning four of a set of ten mugs (control condition) or four of a set of five (set-completion condition). As in Experiment 2.2, participants first indicated how much they wanted the fifth mug (on a 100-point scale anchored on the left with *"not at all"* and the right with *"very much"*) and how much they would be willing to pay for it (open ended). We then asked participants how they would describe the current situation. This question was also answered on a hundred-point scale anchored on the left with "*I own 4 individual mugs*," and on the right with "*I own 4 mugs from a set*."

Results and Discussion

In total, 127 workers on Amazon.com's Mechanical Turk participated (44 females, M_{age} = 28.8, SD = 9.3); nine failed the attention check and were removed from subsequent analysis.[6] As expected, participants in the set-completion condition wanted the fifth mug more than participants in the control condition did; $M_{set\text{-}completion}$ = 59.56, SD = 30.64 vs. $M_{control}$ = 46.49, SD = 31.93, $F(1, 116) = 5.15$, $p = .02$. They were also willing to pay (marginally) more; $M_{set\text{-}completion}$ = 6.93, SD = 5.62, $M_{control}$ = 5.22, SD = 3.91, $F(1,116) = 3.60$, $p = .06$.

Consistent with the set-completion account, participants described their situation more as owning "four mugs from a set" when they owned four of five than when they owned four of ten; $M_{set\text{-}completion}$ = 64.34, SD = 37.98, $M_{control}$ = 45.28, SD = 42.55, $F(1, 116) = 6.61$, p = .01. Within conditions, seeing one's mugs as part of a set correlated significantly wanting ($r_{completion}$ = .46, $r_{control}$ = .38) and willingness to pay ($r_{completion}$ = .39, $r_{control}$ = .47) for a fifth mug (all p's < .001). Finally, seeing one's mugs as part of a set mediated the effect of condition on wanting and willingness to pay (see Figures 2.4a & 2.4b).

---

6 We aimed for, and requested 120 participants but 7 mTurk workers did not return their HITs in time, resulting in 127 participants in total. Not removing the participants who failed the attention check leads to similar results; none of the significant differences become non-significant or the other way around.

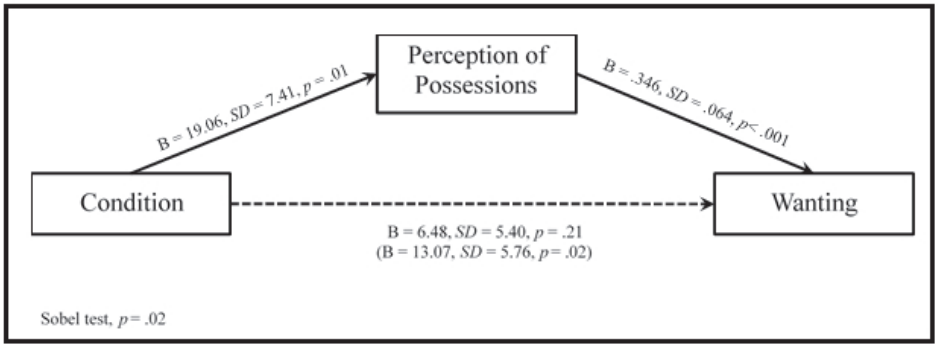


Figure 2.4a: Perception of possessions mediates the effect of condition on wanting.

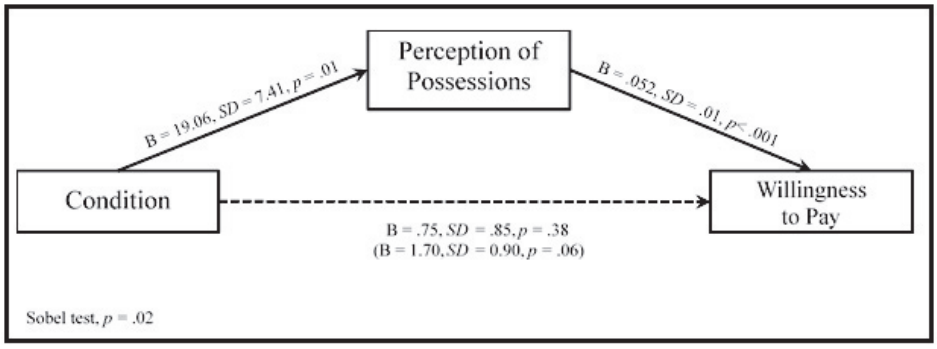


Figure 2.4b: Perception of possessions mediates the effect of condition on Willingness to Pay.

Experiment 2.4

Since we believe that, when owning a large proportion of a set, people are motivated to make it complete by acquiring the last this means that the set-completion premium should not only emerge for generally positive goods, such as the ones used in the previous experiments, but also for aversive stimuli. Finding a set completion premium for disliked items is a strong test of our theory. This was tested in Experiment 2.4.

Method

Because it seems strange to ask participants to pay for aversive goods, we used an adjusted version of the BDM-method (see Experiment 2.1) in this experiment (Becker et al., 1964). Participants (118 students from Tilburg University, 78 females, M_{age} = 20.6, SD = 2.4) were told that they would take part in an experiment where they could uncover hidden pictures.

They read that every time they saw a grey square with a question mark in it, they could click on it to reveal the picture. They were told that for each grey square, the computer would randomly select a number between 1 and 100. If the number of times the participant clicked on the picture was higher than the randomly selected number, the picture would be uncovered. If the number of clicks was lower, the picture would not be revealed. Then three practice trials followed using neutral pictures after which participants in the control condition were presented with a set of ten squares of which five consisted of a negative picture, four were empty squares, and one was a grey square with a question mark that could thus be uncovered. In the completion-condition, participants were presented with 6 squares of which 5 contained the same negative picture, and the sixth was a grey square containing a question mark. In this condition, uncovering the picture behind the grey square would thus mean that the set would be complete (see Figure 2.5).

We expected that, even though the set consisted of clearly negative images, that are unpleasant to look at, participants in the completion condition would still be more motivated to complete the set, and therefore click more often on the grey square to uncover the picture.

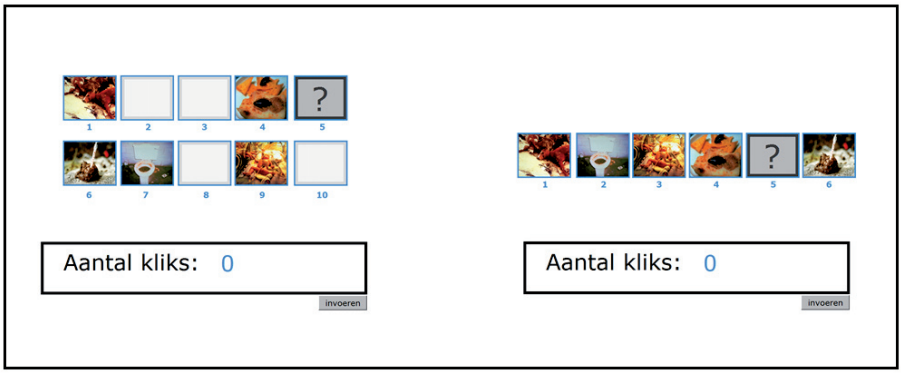


Figure 2.5: Stimuli used in Experiment 2.4 with the control condition on the left and the completion condition on the right. The box on the bottom kept track of the times a participant had clicked on the grey square.

Results and Discussion

As expected, participants in the completion condition clicked more on the grey square than those in the control condition; $M_{control} = 38.9$, $SD = 25.5$; $M_{completion} = 51.1$, $SD = 28.3$; $t(116) = 2.5$, $p = .01$, $d = .46$.[7] This shows that the set-completion premium is not limited to items people are positive about but also emerges for aversive stimuli and as such indicates

7 When planning this experiment, we anticipated that the distribution of clicks would not be close to normally distributed but rather show a bimodal distribution with most people not clicking at all or

that the increased motivation is really a motivation to increase the set and not a general increased liking of the items in question.

General Discussion

In four experiments, we found evidence for a *set-completion premium:* People are more motivated to obtain an item when it completes a set than when it does not. Furthermore, we found that this increased motivation is due to the psychological properties of sets of items. When people own nearly-complete sets, they are apt to see what they own as part of a set rather than as a collection of individual items. Consequently, nearly-complete sets prompt a focus on what is missing (relative to the reference point of a complete set) rather than on what one already owns. This focus on the missing items is what motivates the set-completion premium.

As far as we know, these experiments are the first demonstrations of consumers' increased motivation to obtain set-completing items. Across four studies, the set-completion premium was statistically robust and large in real terms. In Experiment 2.1, willingness to pay increased from 16.1 to 29 cents—an 80% increase. For more expensive items, the set-completion premium was relatively lower, but still meaningful. In Experiment 2.3, willingness the pay for the same mug increased from $5.23 to $6.93—a 32.5% increase. These findings illustrate that even when the target product or experience is exactly the same, merely framing it as completing a set can lead to large changes in behavior. Finally, Experiment 2.4 shows that this effect can even motivate people to acquire goods they normally would not even want to have.

Moderators

As we have shown, perceiving the items one owns as a set is necessary for the set-completion premium to emerge, so anything that makes this more likely—for example, by making the relationship between items more salient—should also increase the likelihood of observing a set-completion premium. This implies that many of the visual gestalt grouping principles should affect the set-completion premium. For example, grouping items closer together (principle of proximity); or forming a "closed group" by separating them on a single shelf in a store (principle of closure); or increasing their similarity (principle of similarity) should all increase the tendency for the items to be seen as a nearly-complete set, and thus raise the set-completion premium (Wagemans et al., 2012). Interestingly, there is some preliminary evidence that this is indeed the case. An independent research

clicking 100 times, but this turned out not to be the case with the distribution being surprisingly similar to a normal distribution.

group has adapted the stimuli we used in the replication of Experiment 2.1 (7.1 in the appendix) adding two conditions in which the items in the set were separated. They replicated the set-completion premium we have described here when the set's items were grouped together, but found that the premium disappeared when the items were physically separated (Paolacci & Van de Bergh, 2014).

The importance of contextual factors (such as grouping) also suggests that there may be "set"-completion premiums for arbitrary sub-sets, as long as people adopt the complete subset as a reference point. Imagine, for example, a box set of 8 different DVDs. If a person owned DVDs #2, #4, and #8 (i.e., all even-numbered items); we would expect an increased motivation to acquire DVD #6, but not #3. On the other hand, if a person has #1, #2, and #4 (DVDs 1-4, except 3); we would expect an increase in motivation to acquire #3, but not #6.

In our studies, we have only examined sets of physical or virtual items. Would we observe a set-completion premium for experiential "sets," like films seen from a top 100 list, or books read from author X? Since the logic of set gestalt outlined above likely also applies to such situations, we believe it will. Anecdotally, the second author's brother recently told him, "I have been to every state in the US except for three. I know that I will probably enjoy going back to a state I have already visited more, but somehow I feel that I should go to one of those three for my next holiday, just to have been in every single state at least once." Even though there are no obviously no tangible items in these cases, the fact that one can easily think of one's experiences either as individual or as part of a set suggests that set-completion premiums are likely to occur (although possibly less strongly, due to the intangibility of the members of the set).

Similarities and Differences between Set-Completion, Curiosity, and Need for Closure

The set-completion account we have proposed here is inspired in part by Loewenstein's, (1994) account of curiosity (Van Dijk & Zeelenberg, 2007). Even though at first glance curiosity and an increased motivation to acquire items may seem very different, the underlying mechanisms show substantial similarities. The model of state-curiosity explains an increased motivation to acquire new motivation using a reference-dependent model. When a person only knows a little bit of information, this is compared to a reference-point of knowing nothing, and extra information only marginally increases ones knowledge. However, when a person knows a large proportion of the information from a certain information-set, this knowledge is compared to a reference-point of knowing everything, because in the second case, the person is on the loss-side of the reference-point, this person feels frustrated and has an increased motivation to acquire the missing information. These strong similarities suggest that it is possible that both state-curiosity, as well as set-completion are specific expressions of a more general underlying tendency for people to want to complete finite groups of which they reached a large proportion regardless of

whether these groups contain material goods, knowledge, or maybe even other things like goals or achievements.

At first glance, it may seem that our effects are also related to the need for closure (Kruglanski, Webster, & Klem, 1993). We initially thought so as well since the literature on the need for closure states that people like things to be complete and "closed". However, when one looks more closely at the operationalization of this need for closure and the way it is usually measured, it appears to measure something different. The need for closure mainly addresses how open-minded people are, and how much ambiguity in information they tolerate. This is clearly reflected in the items the scale consists of; "I hate to change my plans at the last minute" and "I prefer interacting with people whose opinions are different from my own" [reverse]. Even so, we still examined whether need for closure would be related to the set-completion effect. To understand the effect better we administered a selection of items from the need for structure-scale in a conceptual replication of Experiment 2.4 (reported in this dissertation as Experiment 3.1). We explored whether the scores on these scales moderated the effect in such a way that the higher one scores on these scales, the stronger the set-completion effect is (mediated moderation). We found no effect of need for closure. Unpacking the scores into the different subscales only revealed a very small, marginally significant effect for the preference for predictability subscale ($p = .07$). Since with 6 tests, the chance of accidently finding one marginally significant result, even when there is no relation is almost 50%, we think it is very likely there is no, or a negligible relation between the set-completion effect and the need for closure.

Collecting

The current findings are also very relevant for our understanding of collecting. For example, a study of collecting in children found that children who possessed several collectibles preferred to have the final collectible that completed their set over two other toys. Children who did not have any other collectibles did not show such a preference (McAlister, Cornwell, & Cornain, 2011). This is clearly consistent with our model and findings, and underscores that consumers' preferences are influenced by the existence of sets. Similarly, Carey's (2008) formal model of set-completion—which assumes that a collector derives utility not only from the good itself but also from its proportional contribution to the set—predicts that collectors should value set-completing items more highly. However, this model assumes that *only* collectors are motivated to complete sets. The current findings not only provide the first empirical evidence of a set-completion premium, they also explain why this would be observed for non-collectors, a much larger proportion of the population of consumers. Moreover, it can be questioned whether collectors actually try to complete their sets. Research using a large representative sample of the Dutch population (Evers, Lindenberg, & Linksens, 2014) shows that most collectors do not have a very strong idea of their collection with clearly

delimitated borders and organizations, making completion near impossible. Furthermore, for those who do have structured collections, set-completion may not be the goal these collectors strive for, since completing the set is detrimental to the identity of being a collector (Belk et al., 1988). When the set is complete, there is nothing to collect anymore and the collector transforms into merely an owner. From structured interviews with collectors (Evers & Lindenberg, 2014), it appears that collectors actively try to prevent "global" set-completion by expanding the theme of their set right before completion would normally happen. This means that a collector who has almost acquired all Spiderman comics may expand the theme of his collection to also include the other Marvel-comics. This makes sense because it allows the collector the pleasure of completing a set while still having new items to collect. The relationship between collecting and set-completion is thus a quite complex relationship that would be an interesting avenue for future research.

Practical Implications

The most obvious application of our findings is that marketers can (and do) use sets to sell more goods. By producing products that can be perceived as parts of a set, marketers can entice consumers to buy more of them (or pay more for them) than they otherwise would. Of course, in order for set-completion effects to occur, consumers need to already own a large part of the set. One way that marketers have solved this problem is by giving away "free" items for every X dollars spent. The hope is that once the consumer owns enough of these items, he or she will be motivated to complete the set—which, of course, requires spending more money. Manufacturers can also produce larger series of products in different subsets, thereby creating multiple set-completion opportunities. A good example of a company using this strategy is Hallmark (Slater, 2001). In the production of their "keepsakes ornaments," Hallmark carefully releases sets of ornaments such that consumers are constantly adding items to their collection in order to complete the series of ornaments "belonging together." By manufacturing a large set that can be split up in several sub-sets, multiple opportunities for set completion arise—and, consequently, motivation to obtain more ornaments remains high.

Even though companies already seem to understand that giving away freebies can help increase sales, based on these findings some types of giveaways should be more successful than others. For example, a set of 100 collectibles should motivate consumers less than a set of 100 collectibles that can be split up over ten subsets of ten items each. In the first situation, there is only one point in time for set-completion to arise, in the second situation there are eleven time points in which a (sub)set is almost completed. Furthermore, a store using such collectible freebies would be better of giving smaller gifts for free than requiring consumers to save stamps or stickers to trade in for larger gifts. In the first situation, the bar for receiving the gift is very low so more consumers will end up with an incomplete set

they are motivated to complete. In the second situation, only the consumers who actually like the gifts enough to start collecting stamps will acquire the first gifts in the set. Finally, some stores let consumers choose the gifts they receive but other stores use collectible gifts in non-see-through bags. Since the motivation is especially high when one has an almost complete set, preventing consumers from choosing their gift will "drag out" the stage of almost completion and as such lead to a longer time in which the consumer is motivated to acquire the final item.

Our results may also be relevant to many other motivational programs. Firms and other institutions often attempt to motivate behavior by giving consumers small rewards such as stamps, virtual trophies or stickers for doing specific tasks. This ranges from pre-schools giving stickers to children who do their homework well to online communities where consumers can earn virtual trophies and badges for spending time at the gym (e.g., www.fitocracy.com). Our results can be used to design these rewards so that they are maximally motivating (for example using several subsets to collect, so there are many moments in time where one is close to completing a set).

There are, however, cautions we should note for marketers using the set-completion premium as a sales or motivational technique. As we have shown, part of the process underlying the set-completion premium is a focus on what one is missing (compared to the reference point of a complete set) and it could be that this focus on what one is missing evokes negative affect. Evidence for this possibility comes from a study in which participants completed subtasks from a larger set. Some participants saw their progress as "tasks completed" (so compared to a reference-point of zero) whereas others saw "tasks to go." Even though participants in the "to go" conditions were more motivated to complete the tasks, participants who saw how many tasks they already completed were much happier with the task itself (Koo & Fishbach, 2008). If a focus on missing items similarly elicits negative feelings, these feelings could reflect negatively on the products, the set, and the brand. Our experiments were not designed to test this question, but we nonetheless feel that some caution is warranted.

Conclusion

We find in four experiments that people are more motivated to obtain an item when it completes a set than when it does not. This set-completion premium arises because when people own nearly-complete sets, they are apt to see what they own as part of a set rather than as a collection of individual items, and thus are more likely to focus the items they are missing (relative to the reference point of a complete set) rather than on what those they already own. The set-completion premium is statistically robust and large. This suggests that the current results do not merely increase our theoretical understanding of consumer behavior, but could also be used to effect meaningful behavioral changes in the field.

CHAPTER 3

Set-Completion to Increase Motivation

In three Studies we find that people are more motivated to continue playing an (educational) game when they have the opportunity to complete (sub)sets of achievements. The first two Experiments show that participants who were close to possessing a complete set of achievements were more likely to continue playing as compared to when only possessing a small proportion of the total set of achievements. Finally, in Experiment 3, participants played the same game with 25 achievements in total, only this time for some of them these achievements were structured in clearly organized subgroups around a similar theme while for others no such structure was provided. Participants in the first conditions turned out to play the game for longer to earn more achievements. Implications for game-design and social engineering are discussed.

In Chapter 2 we found that people are willing to pay more for products when they complete a set as compared to the same products when they do not complete a set. These effects likely emerge because people automatically perceive an almost complete set, as a complete set with some things missing rather than a collection of individual items. If this is truly the case, set-completion effects should not be limited to the material domain of products, but also emerge in other situations where a set can be completed. Furthermore, organizing a large set into multiple subsets should increase the motivation to acquire missing items even more since in this case there are multiple opportunities for (sub)set-completion. This is what was tested in Chapter 3.

Introduction

Just like in the boy-scouts, where doing certain tasks and reaching set goals is rewarded with a badge, a lot of computer-games also use similar reward-systems. This is a recent development. Whereas 15 years ago virtually no computer-game used such reward-systems, all large gaming platforms (reaching an audience of approximately 200 million gamers) currently provide their members with these types of reward-systems where a gamer earns a badge (also known as trophies or achievements) for each completed goal. This popularization of reward-systems using badges has also generalized to real-world environments, a process known as "gamification" (Deterding, Khaled, Nacke, & Dixon, 2011; Huotari & Hamari, 2011). Popular examples of such "gamified" environments are Foursquare, where people earn achievements by visiting certain stores, and Fitocracy, an app that allows people to log their workouts and provides achievements from reaching certain work-out goals. Even the NSA uses sets of achievements to motivate and train their employees (Poitras, Rosenbach, & Stark, 2013). Seeing the immense popularity of these set-based reward systems, it is surprising that research on the workings and effects of these programs is virtually nonexistent. In this chapter we investigate whether and how sets of virtual badges increase motivation. More specifically, we investigate whether and how people can be motivated to complete sets of virtual badges.

How do Achievement Systems Work?

Most achievement-systems in computer games are designed in such a way that they provide meta-goals outside of regular gameplay. For example, in the on-line game World of Warcraft, the main goal is to increase the strength of ones avatar, which is mostly done by fighting different enemies in the game. Outside of this primary goal of the game, it is currently possible to complete a total of 3790 achievements, the requirements for those ranging from exploring certain areas, catching different types of fish and beating enemies

in specific ways. Because the requirements for completing achievements are frequently separated from the overarching goal of the game, achievement systems have often been defined as secondary to the game (Montola, Nummenmaa, Lucerano, Boberg, & Korhonen, 2009). However, others have argued against this, noting that the collection of achievements can become the primary goal of a game (Hamari & Eranti, 2011; Jakobbson, 2011). In light of our current interests it is noteworthy that these achievement-systems are designed differently across games. In most games, for example all games on Valves Steam network, the set of achievements is merely one long list that a gamer can complete. In others, for example World of Warcraft, the set of achievements is split up into many different subsets of achievements.[8] We have reasons to expect that the organization of these achievements into sub-sets may lead to greater motivation and longer game-play. We will explain these reasons below.

Current Use of Achievement Systems

Since achievements have been largely adopted by the multibillion game-industry, and since achievements are more and more used outside of gaming-domains, insight in whether they work and how they can be designed to work even better is important. Most achievement-systems seem to be intended to either increase the attractiveness of the game and/or to increase time spent on a game. The emphasis mainly seems to be on the latter, which is a logical consequence of the revenue-system of computer games shifting from one-time sales of computer games to subscription and microtransaction based sales (see for example; Totilo, 2012; Evers, van de Ven, & Weeda, 2014). For this reason, it is remarkable that there is virtually no empirical evidence on whether achievements actually increase time spent in the game, and whether the design and organization of achievements moderates these effects.

Set-Completion in Achievement Systems

As explained above, achievement systems provide the gamer (or consumer in a gamified system) with clear goals to reach and virtual pictures (achievements or badges) as a reward for each goal reached. We know that people are usually more motivated when they are close to reaching their goal (Kivetz et al., 2006), so for gamers that try to collect all achievements, motivation should increase when they get closer to having all achievements. However, since achievements are designed as secondary to the games goals, it's not likely that gamers start playing these games with the goal of collecting all achievements.

---

8 Another difference in the design of achievement systems is whether the achievements are public for all to see or private only for the gamer to see. For a discussion on the possible consequences of the way achievements are shared with others, see Antin and Churchill, (2011).

Why would gamers then be motivated to complete sets of these Achievements? Recently, Evers, Zeelenberg, and Inbar (2014, see also Chapter 2) found that consumers are willing to pay more for an item when this item completes a set, as compared to the same item when it does not complete a set. For example, people who own 5 mugs from a set that they believe consists of 6, are willing to pay more for the 6th mug, than people who also own 5 mugs but believe the full set consists of 10 mugs. This "set-completion premium" is likely the result of people being motivated to complete sets. If a completion-goal gets triggered automatically for people who possess a large proportion of a set of products, the same could happen for sets of virtual badges.

Of course achievements differ from products on several dimensions, most notably products are material goods whereas achievements are just virtual images. However, the theoretical model used by Evers et al. (2014) to explain the "set-completion premium" can also be applied to sets of non-material "goods" or experiences. The set-completion premium is supposed to emerge because whereas people who only possess a small proportion of asset focus on the individual items, people who own a large proportion stop focusing on the individual items themselves, but rather focus on the full set, and the items lacking therein. This focus on what is missing then motivates people to acquire the missing items. A similar process could happen for achievements. A person that only completed a small proportion of the achievements will focus on the individual accomplishments, however when a person has completed the majority of achievements all of a sudden the missing ones "pop out" and the person gets extra motivated to complete them. As such, we would predict a similar set-completion premium to emerge for sets of virtual badges. However, instead of an increase in willingness to pay, we would expect this increased motivation to emerge in the form of increased perseverance at the game.

If a set-completion premium also emerges for sets of virtual badges, the organization of these badges should affect motivation as well. Often, games and gamified environments present the achievements as one big unorganized group that can be completed. The same goes for most real-life situations that have been gamified, like Foursquare and Untapped (where one earns badges for drinking beer). A small subset of games does organize achievements into subsets. For example in the on-line game Diablo 3, achievements are organized into subgroups by type (for example class-achievements, crafting achievements and campaign achievements). Within those subgroups, even smaller subgroups of achievements are created (for example the campaign achievements are further subdivided into achievements for Act 1, Act 2, etc.).

The division of the large group of badges into smaller subsets of badges results in multiple opportunities to complete a (sub)set. Thus, if the set-completion premium also emerges for virtual rewards, the latter form of organization (into subsets) should lead to even greater motivation and longer game-play than merely presenting the rewards as one

big set. To illustrate, in the game Halo 2 a person can complete 41 achievements in total, but all of these are portrayed in one big group. Therefore only one moment of increased motivation due to near completion can emerge, for example when the gamer has completed around 38 achievements. On the other hand, Diablo 3 has subdivided its achievements into eight subcategories which are further subdivided into several smaller subcategories. In this case, for each subgroup there will be a moment of near-completion.

In this chapter, we first examined in two experiments whether increased motivation akin to the set-completion premium emerges for sets of non-material goods, specifically virtual achievements. Finally, we tested whether organizing achievements in several subsets increases motivation even more by allowing multiple opportunities for set-completion.

Experiment 3.1

In Experiment 3.1, we tested whether achievements, like material products, also show signs of a set-completion premium. More specifically, we investigated whether people are more likely to continue solving math problems when this would result in a set of achievements being completed as compared to a situation in which continuing to solve the problems would still result in an increase of achievements but not complete a full set.

Method

Eighty-three students from Tilburg University (57 females, $M_{age} = 20.5$, $SD = 2.2$) participated in a weekly testing session in return for course credit. As a cover story, they were told that they would pre-test a children's computer game that involved solving simple math problems. Each time they solved a problem, they were rewarded with a "trophy". They were also told that, because there was only limited time and every part of the game would have to be tested, they would only be able to experience a section of the game. Therefore, some people would start the game with some trophies already earned. Furthermore, we told them that it was likely that time would run out before they could finish the entire game.

Participants were randomly assigned to either the *control* or the *set-completion* condition. In the control condition, participants started the game with zero of ten trophies. After they had earned five trophies, participants were told that the time was up and that they would continue with the next experiment. In the set-completion condition, participants started the game with four out of ten trophies and, like participants in the control condition, were told that the time was up after they had earned five trophies. Thus, in both conditions, participants earned five trophies themselves, but in the control condition the game ended when participants had five trophies of ten, and in the set-completion condition the game ended when participants had nine trophies of ten.

When the game ended, participants first answered some filler questions about the game and we then measured focus on the missing items by asking them to describe their current trophy situation on a (continuous) slider scale anchored on the left with "I am missing 5[1]" (coded as -5) and on the right with "I own 5[9]" (coded as +5). Finally, we informed the participants that they could still continue testing the game, with the option to quit after every problem. However, since the time was up, this would come out of participants' own time (i.e., they would stay at the testing session longer than they would otherwise have to). Thus, participants could decide to spend their own free time to continue playing the game.

We expected that participants in the set-completion condition—who ended the game with nine of ten trophies—would be more likely to spend their free time to continue playing the game as compared to participants in the control condition, who only "owned" five of the ten trophies. Furthermore, we expected this effect to be mediated by the participants' focus on the missing trophies (vs. those they had). For exploratory reasons, we also measured need for closure (Kruglanski et al., 1993) in between several other experiments run that session.

Results

In the control condition (where participants ended the game with five of ten trophies), only six out of 40 (15%) participants decided to continue playing the game. In the set-completion condition this proportion increased significantly to 17 out of 43 (39.5%), $\chi(1, N = 83) = 6.23, p = .01, \phi = .27$. An ANOVA on the focus measure showed that participants in the control condition focused more on the trophies they had ($M = 2.37$, $SD = 3.06$) than did those in the set-completion condition ($M = -0.59$, $SD = 3.73$), $F(1,81) = 15.48, p < .001$, $\eta^2 = .16$. A binary logistic regression regressing the choice to continue playing on focus showed a significant effect of focus, $\chi^2(1, N = 83) = 36.28, p < .001, B = -.485$, meaning that participants who focused more on the trophies they did not have were more likely to use their own time to continue the game. An adjusted Sobel test for logistic regression mediation confirmed that focus mediated the effects of condition on the decision to continue playing, Sobel-$Z = 2.87, p < .01$.

Experiment 3.2

Experiment 3.2 is a conceptual replication of Experiment 3.1. Instead of using a rather tedious task of completing math-problems, a more fun task of finding the differences between two pictures was employed. Furthermore, we kept invested effort the same in Experiment 3.1 by providing one group with 5 already earned badges in the beginning. In Experiment 3.2 we kept invested effort the same by changing the size of the set of achieve-

ments. This prevent the possibility that even though participants did not have to work for the five badges, they may still interpret the existence of those five as some sort of sunk cost (Arkes & Blumer, 1985).

Method

One-hundred-and-fifty-four students from Tilburg University (101 females, $M_{age} = 20.9$, $SD = 2.5$) participated in a weekly testing session in return for course credit and were assigned to one of two conditions. All participants were told that they would be partaking in a task lasting about 4-5 minutes in which they would be presented several sets of two pictures. Both pictures within a set were presented side-by-side and participants were instructed to find the five differences between the two pictures and to indicate the differences by clicking on the locations in the pictures in which the differences were present. In the control condition, 12 black circles were presented under the pictures, in the completion condition only five black circles were presented (see Figure 3.1). For each correctly solved puzzle, participants were told they earned a badge" which appeared in one of the black circles. After correctly completing 4 sets of pictures, participants were told that the time for this task was up and that they were finished with the experimental session and could go home if they wished to do so. However, they were told they also had the opportunity to stay and keep on playing the game.

Figure 3.1: Experiment 3.2, screen presented at the end of the game in the control condition (left) and the completion condition (right).

After choosing to [not] continue, participants answered the following three questions (1 = *not at all*; 7 = *very much*): *"How much fun did you think the game was?; How satisfied are you with the badges you earned?; How much do you want to earn another badge?"* The last question served as a control for our dependent measure of choosing to continue.

The other questions were added for exploratory reasons.[9] We predicted that participants in the completion condition would be more likely to choose to play another round of find the differences and indicate they desired to earn another badge more.

Results

In the control condition (where participants ended the game with four of twelve badges), only nine out of 76 (11.8%) participants decided to continue playing the game. In the set-completion condition this proportion increased significantly to 29 out of 78 (37.2%), $\chi(1, N = 154) = 13.30, p < .001, \phi = .29$.

The data for the three questions showed that participants in the completion condition indeed indicated being more motivated to continue playing the game ($M = 4.36, SD = 1.86$) than those in the control condition ($M = 3.36, SD = 1.73$), $t(1,152) = 3.47, p = .001, d = .56$. Participants in the completion condition also indicated being more satisfied with the badges they earned ($M = 5.31, SD = 1.00$) compared to those in the control condition ($M = 4.70, SD = 1.48$), $t(1,152) = 3.01, p = .003, d = .49$. No significant difference was found for enjoyment of the game with participants in the completion condition liking the game slightly more ($M = 5.09, SD = 1.30$) than those in the control condition ($M = 4.80, SD = 1.70$), $t(1,152) = 1.18, p = .24, d = .19$. This suggests that the completion of sets does not lead to a more positive gaming experience and that the expectation of enjoyment from completing is not likely to be the reason why these participants continued playing, but rather that they continued purely for completion sake.

Experiment 3.3

Since people are more motivated when they can complete a set of trophies, as found in Experiment 3.1 and 3.2, this means that achievements designed in such a way that multiple opportunities for completion arise should be even more effective at increasing motivation to continue playing. This was tested in Experiment 3.3.

---

9 Based on theory, two opposite predictions could be made for these questions. On one hand, having a large proportion of rewards and almost having a complete set could have a positive effect on enjoyment and satisfaction with what one has, since one is on the right track and making good progress. On the other hand, having an almost complete set causes people to focus on the items they are still missing (Evers, Zeelenberg, & Inbar, 2014) which may result in negative feelings such as frustration that can impede enjoyment and cause one to be less satisfied.

Method

A total of 215 Tilburg University students (169 females, M_{age} = 20.0, SD = 2.1)[10] who participated in return for €8,- were randomly assigned to the unstructured or structured condition. All participants read that they would be participating in a topography-game. They read that in this game they would be presented with a screen filled with black squares and that whenever a square lighted up they should click on it and answer the question that appeared. They were told that for each correctly answered question they would receive a lottery ticket with which they could enter a lottery for four movie-vouchers of 15€ each. When participants started the game, they saw that for every correctly answered question, the black square was replaced by the picture of an animal. A total of 25 black squares were presented on the screen, so 25 pictures of animals could be revealed. In the unstructured (control) condition, the 25 black squares were spread seemingly randomly across the screen. In the structured condition the 25 squares were portrayed in a 5´5 block (see Figure 3.2).

The 25 pictures of animals that could be revealed by answering the questions correctly were subdivided into five pictures of five different types of animals. Furthermore, for each type of animal, each picture had a star depicted in the bottom corner in one of five colors. As such, the set of 25 pictures could be subdivided in 10 different subsets (e.g., all zebras, all lions, all red stars, all purple stars, etc., see Figure 3.2). In the unstructured condition, these pictures were spread randomly across the screen making it difficult to identify the subsets. In the structured condition, each row contained 5 pictures of one type of animal and each column contained pictures with one type of colored star. Therefore, the subsets were very easy to recognize in the structured condition.

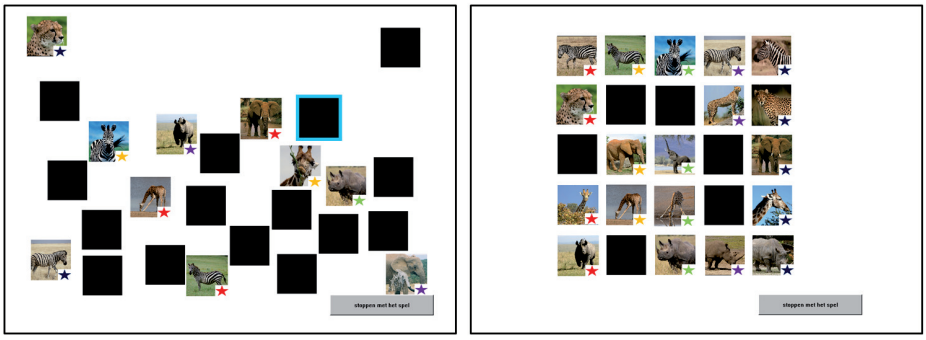


Figure 3.2: Screenshot of the game used in Experiment 3.3 with on the left the unstructured (control) condition and on the right the structured condition.

10 Two participants did not enter their age.

All participants initially answered 9 questions, and thus revealed 9 pictures. After these a pop-up message appeared in which participants were told they completed the compulsory part of the study and could pack up their things and leave since the experiment was officially over. However, if they wished to do so, they had the opportunity to continue answering questions and earning more lottery tickets. Different from the first 9 questions, this time they could choose which question to answer next. They were also told they could quit the experiment at any time. We expected participants to be more likely to continue in the structured condition, because for them almost complete subsets would be more salient.

Results

Consistent with our expectation, participants completed more questions in the structured condition (M= 23.88, SD = 3.6) than in the unstructured condition (M = 22.26, SD = 5.4, $t(1, 214) = 2.63$, $p = .01$, $d = 0.36$). Participants in the structured condition were also more likely to complete the entire field (95/107, 88.8%) than participants in the unstructured condition (80/108, 74.1%, χ^2 $(1, N = 215) = 7.681$, $p < .01$).

General Discussion

In 3 experiments we examined the role of achievements in motivation. We found that the possibility to complete a set of achievements increases motivation and perseverance. As a consequence, designing the achievements in such a way that allows for multiple opportunities of (sub)set completion increases motivation more than unstructured sets of achievements do.

These findings are consistent with the intuition expressed by the game-industry that providing achievements can lead to gamers playing longer. For example companies that wish to develop games for Microsoft systems are obliged to add achievements to their games (see, Jakobbson, 2011), and more and more companies in and outside of gaming seem to add achievements as secondary goal-systems (see for example; Gartner, 2011). However, as mentioned in the introduction, most games provide their achievements in an unstructured way. The results from Experiment 3.3 suggest that it would be better for them to organize the achievements in several subsets.

Even though we found that these achievement-systems increase motivation, it is important to note that we only tested short-term effects. It may be the case that this increased motivation disappears quickly, or even worse, decreases motivation in the long run. There are reasons to believe it may, since extrinsic motivation can be detrimental for intrinsic motivation because it undermines perceived autonomy (Deci, Koestner, & Ryan, 2001). Extrinsic motivation is the increased motivation to reach a target set from outside of the

self, to reach a goal set by others while intrinsic motivation is motivation that results from a personal desire to achieve a goal set by the person him or herself. In the case of achievement systems, it is unclear whether they function as intrinsic or extrinsic forms of motivation. On one hand it can be argued that it is clearly a situation in which an outside force (the requirements for the achievements) determine what the gamer should do. On the other hand, gamers have freedom in choosing whether to pursuit achievements or not. If gamers experience achievement-systems as a form of extrinsic motivation, the consequences of implementing such a system could be that gamers are actually less motivated in the long run, they may feel like the game is forcing them to collect the achievements and lose the intrinsic motivation to play and enjoy the game. In this way, a company implementing achievement-systems to maximize motivation may actually find gamers burned out and less likely to continue to play or purchase new expansions. Understanding these long-term consequences of achievements is thus vital for industries that want to optimally implement them.

Besides effects of set-completion in product domains, there is also evidence that other aspects of sets can influence decision making. Since set-completion effects emerge both for material products as well as virtual badges, it is likely that other processes related to sets of material products also play a role in the effects virtual achievements have on people. For example, it has been found that people generally prefer sets of products that are either completely similar or completely different on salient attributes over those that are partially similar and partially different (Evers, Inbar, et al., 2014). Based on this, we would expect that sets of achievements that follow similar characteristics are more motivating than sets which do not. To illustrate, it could be expected that gamers are more motivated to collect achievements from the left in Figure 3.3 as compared to the set on the right because the set on the left has a clear all-similar characteristic on color, and is all different on design while the set on the right does not have such an underlying structure.

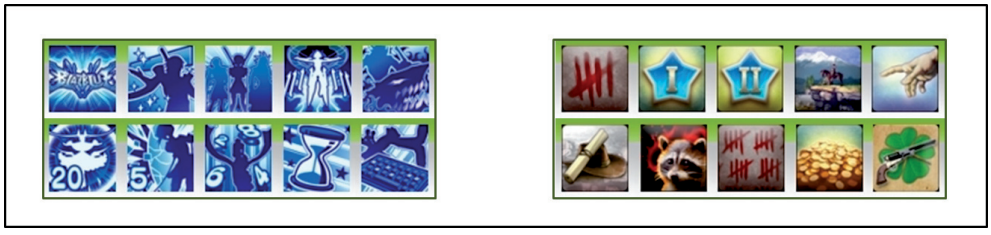


Figure 3.3: Two sets of achievements, on the left an all-similar set from "Blaz-blue calamity trigger", on the right a partially-same partially-different set from "Six-guns".

Since people are motivated to complete a set because the complete set becomes the salient reference-point, it can be expected that a set in which a few achievements are missing is experienced as especially frustrating. The drive to remove this frustration could increase motivation to complete the set, but can also backfire when the last achievement is particularly difficult or even impossible to acquire. If gamers realize that a set of achievements cannot be completed, it may even result in them quitting the game, or at least the hunt for achievements altogether. A recent experience by a colleague indicates that this may indeed be the case. He recently downloaded the Vivino wine-scanner app in which a person can keep track of different wines consumed and acquires badges for reaching certain milestones. One of these badges is the "newbie" badge which one loses as soon as one has consumed more than 5 different wines. As soon as this colleague drank the 6th wine, he realized he lost the opportunity of ever completing all badges and become very demotivated to continue using the app.

Even though we have mainly focused on the role of achievement-systems in computer games in this chapter, similar effects should emerge in gamified real-life situations. This means that achievement-systems can be used as a nudge, a form of social engineering, guiding and helping people to commit to and perform desirable behavior. For example, an app that gamifies working-out (as Fitocracy does) but also implements clear subsets of badges to earn and complete could drastically increase the amount people work out getting healthier in the process, while still respecting autonomy. As such, achievement-systems can be used as a cheap, easy, and non-coercive way of increasing wellbeing.

CHAPTER 4

Set-Fit Effects in Choice[11]

In four experiments, we investigate how the "fit" of an item with a set of similar items affects choice. We find that people have a notion of a set that "fits" together—one where all items are the same, or all items differ, on salient attributes. One consequence of this notion is that in addition to preferences over the set's individual items, choice reflects set-fit. This leads to predictable shifts in preferences, sometimes even resulting in people choosing normatively inferior options over superior ones.

11 This chapter is published as: Evers, E. R. K., Inbar, Y., & Zeelenberg, M. (2014) Set-fit effects in choice. *Journal of Experimental Psychology: General 143*(20), 504-509.

As mentioned in the introduction (see page 8), a lot of sets seem to either all be the same color and design or consist of items that are all different colors with the same design. The set-completion effect in previous chapters is expected to emerge because people automatically group the items in a set and perceive the set as one coherent entity, similar to the Gestalt principle of closure (Wertheimer, 1923). If people automatically group products which for a gestalt of a set, then some organizations of those items should form a better gestalt than others. This was tested in Chapter 4.

Introduction

Imagine choosing between two pens, one of which is clearly superior to the other. In all likelihood, you would choose the superior pen. Now imagine making the same choice, with one difference. This time, regardless of what you choose, you also get three extra pens similar to the inferior pen. Would you now choose differently? If you are like the authors, your intuition is "no." However, as studies reported here will show, this intuition is inaccurate. In fact, people's choices between groups of items take into account how well each set "fits" together, leading to predictable shifts in preferences when choosing between sets rather than individual items.

There is some prior work supporting the intuition that certain sets are seen as better-fitting (and therefore more pleasing) than others. Research in Gestalt psychology has uncovered a number of principles that determine how individual items are perceptually grouped into sets (Koffka, 1935; Wagemans et al., 2012; Wertheimer, 1923; for a recent review article see Wagemans, Elder, Kubovy, Palmer, Peterson, Singh, & von der Heydt, 2012). For example, the Gestalt principle of similarity states that all else being equal, similar stimuli are more likely to be grouped together than dissimilar stimuli. Gestalt perception happens rapidly and automatically, and this feeling of perceptual fluency may make sets that form a good gestalt more pleasing—this would be a specific case of the general rule that easy-to-process stimuli are liked more (Winkielman & Cacioppo, 2001).

In addition to the Gestalt principles, are there any other general criteria of set "goodness"? Philosophers have argued that stimuli that can be described by simpler rules are more aesthetically pleasing (Kintsch, 2012; Schmidhuber, 1997).[12] Similarly, research on the aesthetics of patterns has found that patterns that can be described by simple rules are preferred to those for which no simple descriptive rules exist (Garner, 1970; Garner & Clement, 1963; Glanzer &

12 "Simplicity" is defined in terms of Kolmogorov complexity, a term used in algorithmic information theory to denote the minimal amount of resources needed to describe an object. In other words, the Kolmogorov complexity of an object is the shortest possible description of it.

Clark, 1964). Taken together, this evidence suggests that, all else being equal, sets that follow simple organizing principles will be liked better than sets that do not.

One common organizing principle is the similarity or dissimilarity of a set's members. Sets whose members are all similar or all different can be described by simpler rules than sets where some items are similar and others are different. The latter can only be described in less parsimonious terms—for example, "two items of one kind, three of another, and three more of three different kinds." Of course, "all-similar" or "all-different" are not the only organizing principles that might apply to sets. For example, a set of numbers following the rule "sequentially increasing," as in a straight hand in poker, follows a simple organizing principle, whereas a set of randomly chosen numbers does not. Here, however, given their ubiquity in everyday life, here we focus on all-similar and all-different sets.

We examined the perceived quality of all-similar and all-different sets in a pretest in which we asked participants to rate several sets of pictures of dinosaurs. Some of these sets followed "all similar" or "all different" rules (all dinosaurs were either all of the same type, or all of a different type, and all were either exactly the same color, or all of a different color), and some did not. Consistent with our hypothesis, participants rated sets as better when they followed one of these rules for both shape and color. We also asked the participants how the sets could be improved and found that a large majority (89.6%) of participants suggested changes that would make the set's items either all-similar or all-different on important attributes. For more information about this pretest, see Experiment 7.4 in the appendix.

Based on these results, in the current research we operationalize good sets as those where all items "fit" together according to these simple rules—that is, good sets are those where all items are the same, or all differ, on all salient features. Bad sets are those where one or more items do not "fit"—that is, they violate the all-the-same or all-different principle. We hypothesized that people would choose sets in which items fit together over ones where they did not, even when the bad sets were superior on an item-by-item basis; and that people would be reluctant to choose items that would fit badly with a set of items already in their possession.

Of course, there are cases in which items in a set are complementary, so that the combination of the items adds extra utility. For example, one might prefer a tennis racket over a bowling ball, but a bowling ball and 9 pins over a tennis racket and 9 pins. However, in the cases we will examine, there are no obvious complementarities between goods, so a "good set" is no more useful than a "bad set." Nor, as we will see, is the set-fit effect explained by existing phenomena known to affect choice, such as attraction and compromise effects (Simonson, 1989; Simonson & Tversky, 1992), differences in evaluability (Hsee, 1996; Hsee & Zhang, 2010), or variety-seeking (Ratner, Kahn, & Kahneman, 1999; Simonson, 1990). Below, we test the set-fit effect in four experiments.[13]

---

13 Termination of data collection was decided in advance, based either on a set length of time or a set total amount of participants. All conditions we ran are reported. Finally, only choice was measured

Experiment 4.1

One hundred and four participants recruited via Amazon.com's Mechanical Turk (66 females, M_{age} = 34.79) were randomly assigned to one of two conditions. In the *individual-choice* condition, participants chose between a superior metal-accented pen and an inferior all-plastic pen. In the *set-choice* condition, participants also chose between these two pens,

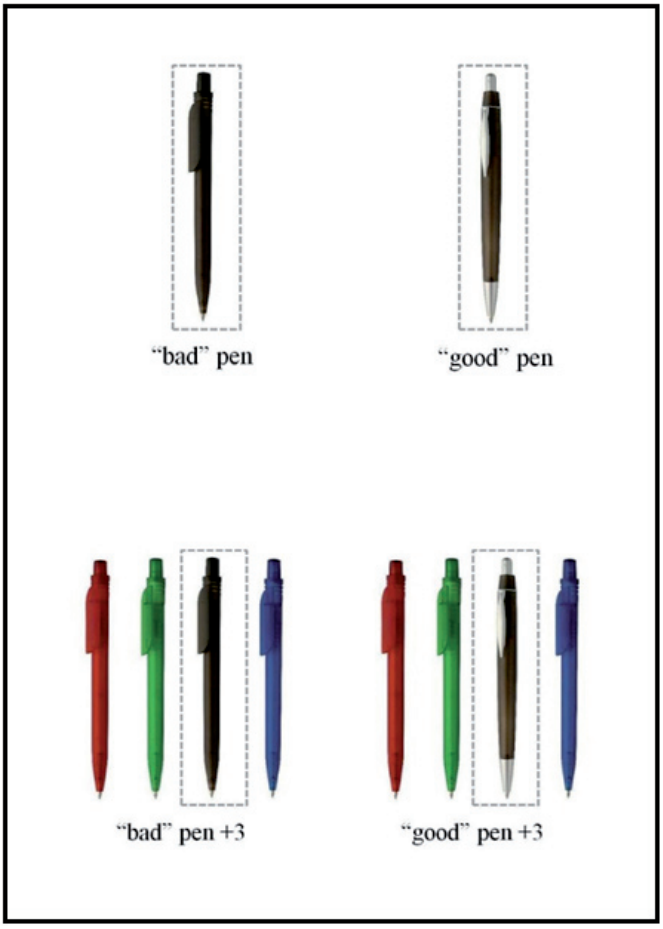


Figure 4.1. Stimuli in the individual-choice condition (top) and set-choice condition (bottom) of Experiment 4.1.

as dependent variable. For exploratory reasons we also asked the participants to explain their choices in Experiment 4.2 and 4.3 (the two experiments in which the participants interacted directly with experimenters). We did not ask for these reasons with any statistical analysis in mind and therefore these "results" will only be addressed in the General Discussion.

but with the same three all-plastic pens added to each option (see Figure 4.1). We expected that participants would prefer the superior over the inferior pen in the *individual-choice* condition, and indeed 78.8% (41/52) chose the superior pen. In the *set-choice* condition, we expected more participants to choose the set including the inferior pen (which fit with the other pens) over the set including the superior pen (which did not). Indeed, 50.0% (26/52) chose the set including the inferior pen, a significant difference from the individual-choice condition, $\chi^2(1, N = 104) = 9.44, p < .01, \phi = .30$.

Experiment 4.2

In Experiment 4.2, we extend the first study in three ways: First, we attempted to conceptually replicate the set-fit effect using different stimuli. Second, all choices were between sets of items, in order to rule out alternative explanations involving extraneous differences in how people choose between sets and single items (Kimchi & Palmer, 1982) .Finally, participants made real choices between goods (Experiment 4.1's were hypothetical).

Participants (199 Tilburg University students; 137 female, M_{age} = 21.03) chose one of two sets of paperclips to take home. In the *color* condition participants chose between a set of four plastic-covered colored paperclips (red, green, yellow, and blue) and a set of three colored clips (the red, green, and blue clips from the first set) plus one ordinary metal clip. In the *metal* condition, participants chose between a set of the ordinary metal clip plus three metal animal-shaped clips, and a set of the same three metal animal-shaped clips plus

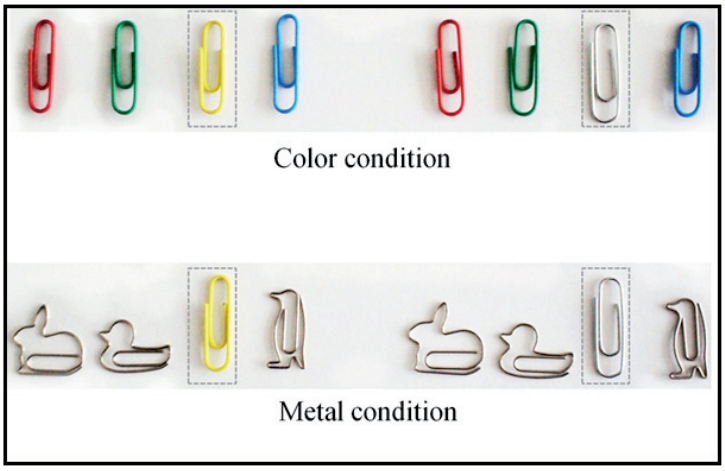


Figure 4.2. Choice sets in the *color* condition (top) and *metal* condition (bottom) of Experiment 2. These sets were physically shown to participants. Note that the yellow and metal clips appear once per set in each condition.

a yellow clip (see Figure 4.2). Thus, the ordinary metal clip and the yellow clip appeared in each condition, but as part of different sets. As predicted by the set-fit account, the set containing the metal clip was preferred by 63% of participants (63/100) in the *metal* condition (i.e., when the metal clip fit with the rest of the set); but by only 18% of participants (18/99) in the *color* condition (i.e., when the yellow clip fit with the rest of the set), χ^2 (1, N = 199) = 41.40, $p < .001$, $\phi = .46$.

Experiment 4.3

Experiment 4.2 demonstrated the set-fit effect in real choices, but perhaps not particularly consequential choices. We therefore conducted Experiment 4.3 with young children using (sets of) marbles. Marbles are of considerable value to Dutch schoolchildren, especially during the spring "marble season" when this experiment was run.

Children from the first through third grades of the elementary school "De Bracken" in Rijen, Netherlands (N = 140; 75 females; M_{age} = 7.51) volunteered and participated individually. The design paralleled that of Experiment 4.1. In the *individual-choice* condition, participants chose between a blue marble with white stripes and a monochrome red marble (pre-testing revealed that children saw the former as more desirable). In the *set-choice* condition, participants also chose between these two marbles, but with the same three monochrome marbles added to each option (see Figure 4.3).

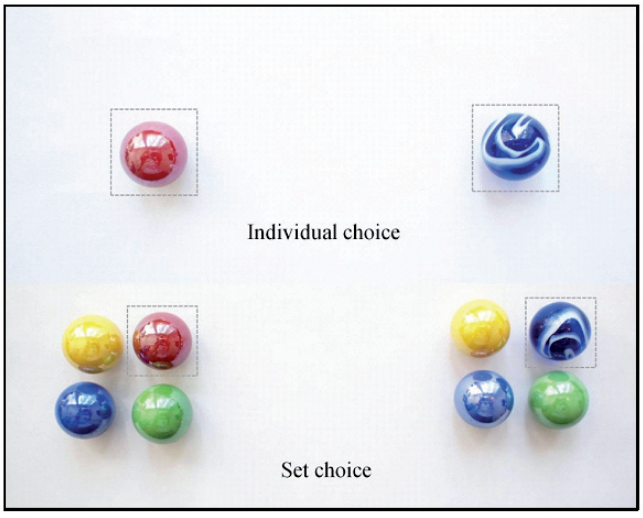


Figure 4.3. Choice options in the individual-choice condition (top) and set-choice condition (bottom) of Experiment 4.3. Choice options were physically shown to participants.

As predicted by the set-fit account, 89.6% of participants (43/48) preferred the superior striped marble over the monochrome red marble in the *individual-choice* condition, but the set including this marble was preferred by only 27.3% (12/44) in the *set-choice* condition, $\chi^2(1, N = 92) = 37.1, p < .001, \phi = .63$.[14]

Experiment 4.4

Experiments 4.1-4.3 showed that people are attracted to sets that fit well together, so much so that they sometimes choose a set of inferior items over one including a superior item. Experiment 4 tests whether making an item fit better or worse with a set of already endowed goods can increase and decrease its choice share, respectively.

Participants (207 Fontys University Tilburg students; 154 female; M_{age} = 19.7) imagined buying specialty beer for a party. Participants were told that they had already chosen three

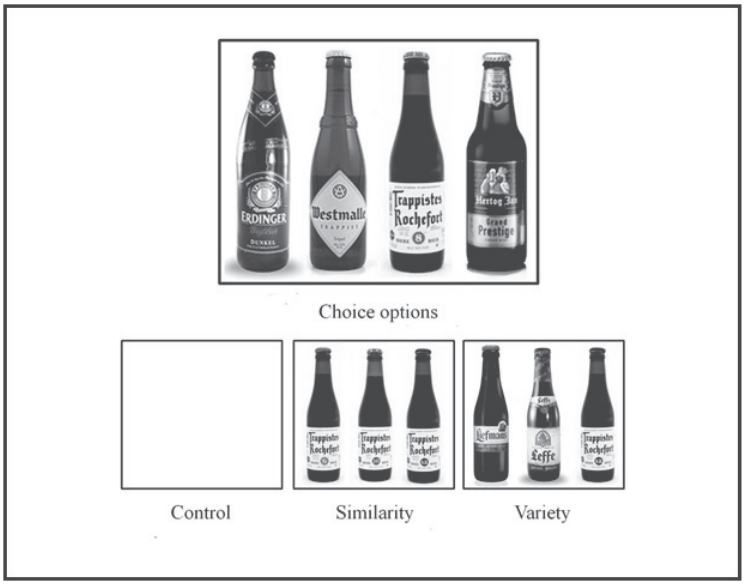


Figure 4.4. Choice options and endowed sets from Experiment 4. The top panel shows the choice options. The bottom panels show the beers participants in each condition were said to have already selected.

14 We also ran one other control condition (= 48) where children chose between a set of four blue/white striped marbles and a set of three blue/white striped marbles and one monochrome red marble. As expected, a large majority of participants in this condition (75%) preferred the set of four blue/white striped marbles.

beers and that they could choose one more from an assortment of four, which included a beer from Rochefort Abbey (see Figure 4.4).

In the *control* condition, participants were given no further information and simply chose which beer they would purchase. In the remaining two conditions, participants saw the three beers they had (allegedly) already chosen: In the *all-similar* condition, these were three beers from Rochefort Abbey; in the *all-different* condition, these were beers from three different brewers including Rochefort. We expected that compared to those in the control condition, participants in the *all-similar* condition would be more likely to choose the Rochefort beer (because it would fit best with the "all-similar" set), and those in the *all-different* condition would be *less* likely to choose this beer (because it would not fit with the "all -different" set). Indeed, compared to the control condition, in which 11/74 (14.9%) chose the Rochefort beer, participants in the *all-similar* condition were more likely to choose the Rochefort (55/69; 79.7%), $\chi^2(1, n= 143) = 60.41, p < .001$, and those in the *all-different* condition were less likely to (2/64; 2.9%), $\chi^2(1, n = 138) = 5.54, p < .001$. Note that although there are alternative explanations for why Rochefort's choice share would increase in the *all-similar* condition (for example, telling participants that they had already chosen three Rocheforts may have been seen as an implicit recommendation or as informative about their or their host's preferences) none of these explanations would explain why Rochefort's choice share would *decrease* (relative to control) in the *all-different* condition.

General Discussion

Four experiments revealed that when choosing between (or adding to) sets of items, people's choices reflect set-fit in addition to preferences over the set's individual items. This can cause predictable preference reversals when an individually-preferred item is bundled with other items with which it fits poorly.

The notion of set-fit can explain behavior that is un- or mis-predicted by well-known accounts of non-normative influences on choice, such as attraction and compromise effects (Simonson, 1989), or contrast or assimilation (Wänke, Bless, & Schwarz, 1999). Attraction or compromise effects occur when choosers must make trade-offs between domains (e.g., price and quality) when choosing between roughly equally-attractive options. These effects do not apply to cases where one choice option is clearly inferior, whereas the set-fit account predicts that inferior options can be preferred when they fit well with other bundled items. Nor can contrast or assimilation explain our results: If participants in Experiments 4.1 and 4.3 contrasted the superior items with the inferior items in the set, this should have increased the choice share of this set by making the superior item seem even more attractive. Similarly, an assimilation account would not predict that people would prefer a well-fitting

set of inferior items (as was the case in Experiment 4.3), and would make no predictions regarding Experiments 4.2 and 4.4, where no item or set was of obviously higher quality.

Even though our approach builds on Gestalt theory's "similarity principle," according to which items similar in shape or hue are perceptually grouped (Wertheimer, 1923), set-fit is not just an application of Gestalt theory. First, the similarity principle explains a perceptual phenomenon, not preferences or choices; second, of course, it would not predict (as we do) that good sets can also follow an "all-different" principle. That said, we do believe our findings fit with the larger "Gestaltian" notion that people naturally structure groups of stimuli in order to reduce complexity. Furthermore, we believe that the similarity and difference principles we investigated here are not the only ones that determine perceptions of set-fit (and thus affect choices). Theoretically, any organizing principle that reduces complexity should affect perceptions of set-fit. Some possibilities include other Gestalt principles, such as common fate (moving together; Wertheimer, 1923), numbering (for example, a set consisting of 1,2,3,4,5,6,7 should be preferred over one consisting of 1,4,8,2,13,6) and perceptual symmetry (Garner, 1970). Testing these possibilities goes beyond the scope of the current research, but we expect that any organizing principle that reduces set complexity should function in the same way as the principles we have focused on here.

One challenge for any alternative to the set-fit account would be to explain why people prefer similarity under some circumstances but variety under others. As we showed in Experiment 4.4, fit (or lack thereof) with either an all-similar or all-different set can affect the attractiveness of a choice option. We ran an additional test of this idea, including all hypothesized effects within one experiment. We ran four conditions in which participants chose between two sets of mugs (see Figure 4.5).

These sets were designed in such a way that in each condition the only difference between the two choices was a green vs. an orange mug. Adding three green mugs to both options (*all-similar green*) we found that the majority (60.4%) chose the set with the fourth green mug. Adding three orange mugs to both options (*all-similar orange*) reversed preferences—now the majority (62.3%) chose the set of four orange mugs ($p = .03$). These two conditions reveal *the all-similar* effect. In the other two conditions we found the *all-different* effect. In the first condition we added a blue, pink and orange mug to both options. In this case the majority (84.3%) chose the set with the green mug. When we added a blue, pink and green mug to the two original options, the majority (94.1%) chose the set with the orange mug ($p < .001$, see also Experiment 5.3 for a conceptual replication of this study involving real choices).

Moderators and Boundary Conditions

One obvious precondition for set-fit effects is that a group of items is actually perceived as one set. Based on the Gestalt principle of proximity (Wertheimer, 1923), we would

Figure 4.5. Participants chose between two sets of mugs (see supporting materials for details). Note that the only difference between the two choice options is always the green vs. orange mug. Boxes are added to the pictures for clarification but were not part of the original stimuli.

expect the shapes in the left panel of Figure 4.6 to be perceived as one poorly-fitting set of three triangles and a circle, whereas the same shapes in the right panel should be more likely to look like one well-fitting set of three triangles plus an extra circle. Thus, subtle perceptual features may have large effects on the influence of set-fit and, therefore, on the attractiveness of sets.

The set-fit effect seems to rely more on intuition than deliberation. Even though investigating how intuitive the effect is was not the primary goal of this research, we did ask participants in Experiments 4.2 and 4.3 why they chose as they did. Participants in Experiment 4.2 often looked confused when they were asked to explain their choices.

Figure 4.6. Left panel: A "bad set" of four three triangles and a circle. Right panel: a "good set" of three triangles, and an additional circle.

Some literally said "I don't know" others said "I guess because these seem to belong together" or "the yellow [clip] just really doesn't seem to belong together with the other ones." Coming up with a reason for the choice also generally took much longer than making the choice itself. The children in Experiment 4.3 reacted similarly, mentioning that they felt some marbles "belonged" together or – pointing at the set they did not choose – that the non-fitting marble "just felt weird".

These reactions are consistent with the idea that set-fit effects are the result of basic perceptual grouping principles (Wertheimer, 1923) that are applied relatively automatically and uncontrollably. Following Gestalt psychologists, we believe that positive evaluations of good sets do indeed occur rapidly and effortlessly. However, this does not mean that better-fitting sets will always be preferred in choice. Whereas the evaluation of the set and its configuration occurs automatically, its effect on choices may be overridden by more deliberate processing of the options or other judgment processes. For example, four boys are a "better set" than three boys and one girl, but we do not believe a pregnant mother of three boys would wish for a fourth in order to have a more pleasing set of children. In this case, other, more deliberative influences—e.g., the anticipated pleasure to be gained from another boy versus a girl, or a commitment to loving a child of either gender equally—are likely to greatly outweigh the influence of set-fit. Set-fit effects on evaluations of set quality *per se* are thus always expected, but set-fit effects on actual choices are likely to be most pronounced when people rely on their immediate intuitive responses, and weaker when decisions are based on deliberative reasoning. This also means that when investigating set-fit effects it is important to keep such other considerations stable.

To what extent does variety-seeking conflict with maximizing set-fit? Note that the processes underlying set-fit and variety-seeking are very different, so they are likely to exert independent effects on choices. For goods or experiences that are chosen and consumed sequentially over time (e.g., Ratner et al., 1999), we would not expect set-fit effects, as it is unlikely that these choices would spontaneously be construed as a set. When choices are bracketed together (Read & Loewenstein, 1995) set-fit effects may *amplify* variety-seeking. If people desire at least some diversification and therefore start by choosing different items, a concern with set-fit should lead them to favor greater diversity (i.e., an all-different set a partially different set). Of course, in cases where the set is all-similar, set-fit and variety-seeking effects operate in opposite directions: Set-fit increases the chance of a similar item being chosen whereas variety-seeking increases the chance of a different item being chosen. In these situations the outcome depends on the relative strength set-fit and variety-seeking. In the experiments described in this paper, considerations that amplify variety-seeking—such as satiation (L. McAlister, 1982), overestimations of satiation due to time contraction (Galak, Kruger, & Loewenstein, 2011), or spreading of risk (Simonson, 1990)—were minor or non-existent, and set-fit predominated, leading people to choose the

all-similar set. In choices where these considerations play more of a role—for example, when physical satiation creates a strong desire for variety—we would expect variety-seeking to predominate.

Conclusion

The results reported here add to a large and growing body of work showing that—counter to the assumptions of classical economics—people's preferences are often unstable and subject to minor contextual factors (Hsee, 1996; Lichtenstein & Slovic, 1971; McGraw, Shafir, & Todorov, 2010). In the case of set-fit, this may have a number of important consequences, including the fact that improvements to products that are seen as part of a set may unexpectedly decrease their appeal when the improvements make the overall set worse. One real-life cautionary example is the case of a Dutch publisher who attempted to improve the design of a popular series of books by changing the covers, paper, and typography; these changes were the result of much careful research regarding consumers' preferences. However, the publisher did not anticipate that many existing customers would be upset because the new, redesigned books would not "fit" with those already purchased. Finally—after many angry letters—the publisher decided to release the books in both old and new formats.

CHAPTER 5

Preference for Order[15]

In 5 experiments we find that choices between bundles of consumption goods exhibit a preference for 'order' that cannot be explained on the basis of utility for consumption itself. The first 3 experiments reveal that this order-preference is strong and produces robust violations of normative properties of decision making; most strikingly dominance. The two final experiments provide evidence that the effect derives specifically from a preference for choice sets that can be encoded in simple terms.

15 This chapter is based on: Evers, E. R. K., Inbar, Y., Loewenstein, G., & Zeelenberg, M. Order Preference. *Working paper available on SSRN: http://ssrn.com/abstract=2466991*

The findings described in the previous chapter show that people generally prefer sets that are either all-similar or all-different on salient characteristics. It is expected that this is due to such sets being less complex than sets that do not adhere to these characteristics. Furthermore, the results described in previous chapter seem difficult, but not impossible, to explain in a way consistent with theories of rational choice. In this chapter we tested whether set-fit effects violate normative choice theories, and tested a preference for simplicity as the underlying process.

Introduction

As highlighted by Herbert Simon's seminal work on 'bounded rationality', the human brain is subject to serious constraints, both in information processing and memory (Simon, 1957). One of the ways humans and other animals cope with these limitations is by encoding the world in ways that minimize storage but maximize information. Research on the perception and processing of groups of visual stimuli has shown repeatedly that humans automatically process and store information in the simplest, or most ordered, form possible (Chater, 1996; Wertheimer, 1923).

In this paper we argue, and provide experimental results which demonstrate, that the brain's automatic propensity to simplify manifests itself not only in low-level processing, such as visual perception and language, but also in preferences. Specifically, we show that choice sets which can be characterized in simple terms are preferred over sets that are not simply described, even when the simpler sets are dominated by the more complicated ones in their constituent items. By descriptive simplicity we mean that all items in a bundle can be summarized by simple rules (Chater & Vitanyi, 2003; Glanzer & Clark, 1964; Kintsch, 2012). "Order," "harmony," "integrity," and "structure" are alternative words that convey a similar meaning. The wifi password "0123456789," for example, would, by our definition, have greater descriptive simplicity than the password "0a7," despite its greater length.

In a series of experiments, all involving real choices between sets of goods, we provide evidence that people have a strong preference for bundles of goods that can be described in simple terms. In the first three experiments we show that people have a strong preference for bundles consisting of sets of items that are either all the same or all different, and show that this preference leads to various choice anomalies, specifically violations of dominance, independence (in a situation in which it would be difficult to argue that there is complementarity between components of the set), and "partition-dependence" effects (Fox & Rottenstreich, 2003) whereby dividing a set up into different subsets of items can have dramatic effects on preference. In the fourth experiment, we show that preferences between two unchanging choice sets can be altered by drawing attention to (and away from) specific

attributes that highlight one or the other set's order. Finally, the fifth experiment shows that this preference for all same or all different results from a more general preference for choice sets that can be described in simple terms.

Earlier research in cognitive psychology found that the brain naturally and automatically tends to make sense of disparate stimuli, perceiving and processing such stimuli at a superordinate level (Wertheimer, 1923). Subsequent research on perception has shown that groups of stimuli are automatically processed in their simplest possible form, using the least amount of informational statements needed (Chater & Vitanyi, 2003). For example, Falk and Konold (1997) found that when subjects were asked to memorize strings of symbols, they organized these strings in the least chunks needed (for example, xxxoxo was memorized as xx, 2*xo rather than xxx, o, x, o) In a similar vein, Ariely (2001) showed that when memorizing sets of objects, people accurately remember the average characteristics of stimuli within that set, but do not remember the characteristics of the individual stimuli in those sets correctly. Because cognitive recourses are scarce, and since for any group of objects the underlying structure is by definition equally simple or simpler than the individual characteristics, it is clearly functional to process a group of stimuli by its overarching structure rather than by its individual contents. Here we show that order-effects go beyond mere processing, and influence how much people like a set of stimuli and as such influence choice between and evaluation of such sets.

The idea that ordered sets may be liked better is consistent with early research on pattern-perception which has consistently revealed that people like patterns of dots more when these patterns have some sort of underlying structure (Garner, 1970; Glanzer & Clark, 1964). More specifically, these researchers found that patterns which are ordered in an easy to describe fashion, and exhibit a more parsimonious structure, are liked better. If people generally like patterns better when they have order, this suggests that the liking of bundles of goods may also be influenced by such structural relationships between goods.

Evers, Inbar, and Zeelenberg (2014) applied insights from this early research to understanding esthetic preferences. In a series of studies they demonstrated that people like groups of objects more when they are either completely similar or completely dissimilar on salient attributes. For example, in one study, participants were asked to indicate how much they liked a set of drawn dinosaurs. Participants rated those sets higher when all the dinosaurs were either completely the same color or all differently colored as compared to sets in which some shared a color but others did not. Similarly, participants liked the sets more when all dinosaurs were either of the same type or completely different types as compared to sets in which they were not completely similar or completely different. Since the preference for order is believed to be the result of a fairly intuitive and automatic process (Evers, Inbar, et al., 2014), one could expect that order only affects preferences for goods consumed for their aesthetic qualities. However, as we will show, the preference for

order affects choices in situations in which one would not expect consumption utility to be affected by aesthetics, such as pens or beers.

In what follows, we show that preferences for groups of consumption goods exhibit a preference for descriptive order. The first three experiments show that a preference for sets of items for which all items are the same or all items are different (which is easy to encode and describe), and that this preference leads to deviations from normative properties of choice. Experiment 5.4 and Experiment 5.5 provide evidence that the preference for sets in which all items are the same or all are different is driven by a more basic preference for sets that can be characterized in simple terms.

Experiment 5.1

In Experiment 5.1, we document the order preference effect and show that it leads to violations of dominance. Specifically, we show that if adding a desirable item to a bundle detracts from the descriptive simplicity of the bundle, it can decrease preference for the bundle. Since a direct choice between a dominant and dominating bundle would be instantly obvious to subjects, we adopt a paradigm previously used to demonstrate dominance violations (Gneezy, List, & Wu, 2006) in which decision makers make choices between the dominating or dominated item and a third item.

Method

Three hundred-and-two students at Tilburg and Fontys University (201 females, M_{age} = 20.2, SD = 3.3) were randomly assigned to one of three conditions. Subjects in all three conditions were given choices between a set of pens and a black notebook. They were told that 10% of them would receive the item(s) of their choice. In one condition (three-same) the set of pens did not exhibit obvious order, it consisted of 9 differently colored pens with

Figure 5.1: Display of the bundle of pens in the three conditions in Experiment 5.1.

3 more brown pens. The sets in the other two conditions did exhibit an obvious strong order (see Figure 5.1).

In the all-different condition, the set consisted of 10 pens, each pen a different color. Finally in the "bonus" condition, the set consisted of the same twelve pens as in the three-same condition, but in this case the bundle of ten different pens was visually separated from the two brown pens, with the intention of creating order in the unordered group of twelve pens. Since people are naturally inclined to group items that are proximal to one another into a separate category from those that are further away (Wertheimer, 1923), it was expected that subjects would perceive the 12 pens as two separate bundles with greater descriptive simplicity; one consisting of 10 all-different pens, and one consisting of 2 pens with the same color.

Similar effects of categorization have been found in work on partition dependence (see for example; Fox, Ratner, & Lieb, 2005) in which subjects were found to have a heuristic of allocating resources approximately evenly over categories (e.g., different categories of charities or different income-groups receiving financial aid), which meant that changing the partitions between categories had a substantial, and normatively unjustifiable, effect on allocations. The second and third conditions in the current study, which differ only in terms of how items are partitioned, also examine a kind of partition-dependence, albeit caused by a different underlying preference pattern from that studied by Fox and colleagues.

After making a choice between the pens and the notebook, participants indicated how much they were willing to pay for each, though these valuations were not incentivized. We expected the differences in WTP to mirror those observed in choice, with items more easily described in simple terms being associated with higher WTPs. We did not predict any changes in WTP for the notebook across conditions, although we would not have been surprised if such differences were observed – for example if subjects contrasted their valuation of the notebook with that of the pens.

Results

Figure 5.2 presents the results, with choices indicated by the left of the paired bars and valuations represented in the right hand bars.

As predicted, subjects in the all-different condition were more likely to choose the pens over the notebook (46%; 47/103) than were subjects in the three-same condition (31%; 31/101, $\chi^2(n = 204) = 4.81, p = .03$), despite the fact that the latter dominated the former. However, when the 12 pens were partitioned in a way that made it transparent that they could be grouped into one all-different and another smaller both-the-same subset, there was a slight but not significant preference for the sets of pens as compared to the all-different condition (55%; 54/98, $\chi^2(n = 201) = 1.8, p = .18$) and a significant preference as compared

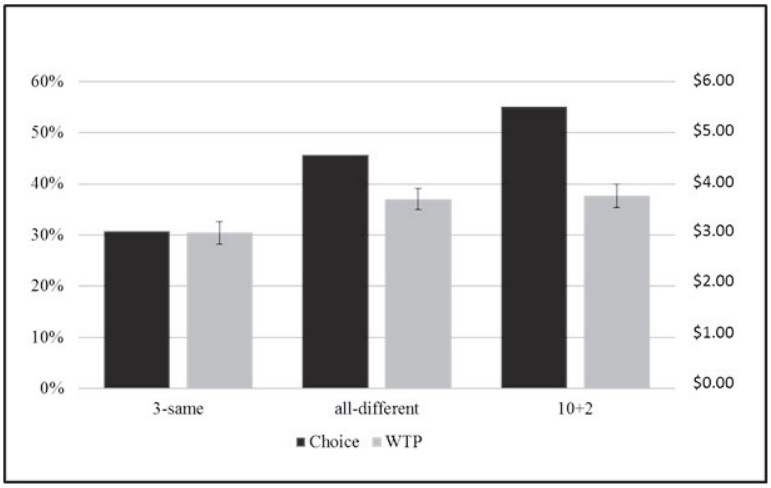


Figure 5.2: Results of Experiment 5.1. For each condition, the bar on the left represents choice proportion for the bundle(s) of pens over the notebook as indicated on the left y-axis. The right bar represents willingness-to-pay as indicated on the right y-axis.

to the three-same condition ($\chi^2(n = 199) = 12.10, p = .001$) even though in both cases they would get exactly the same 12 pens.

Willingness to pay revealed a similar pattern. Participants were willing to pay slightly (but not significantly) more for the set with two bonus pens added separately ($M = €3.77$, $SD = 2.24$) as compared to the condition in which they merely received the good set of ten pens ($M = €3.70$, $SD = 2.12$, $t(202) = .23$, $p = .82$). This is not remarkable because the bonus-set included two more markers than the set in the control condition. However, participants in the three-same condition, who indicated their willingness to pay for exactly the same items as participants in the bonus condition, were willing to pay less for the bundle of 12 pens ($M = €3.04$, $SD = 2.19$, $t(197) = 2.32$, $p = .02$), even compared to participants in the all-different condition in which the bundle only included ten pens, $t(199) = 2.17$, $p = .03$.

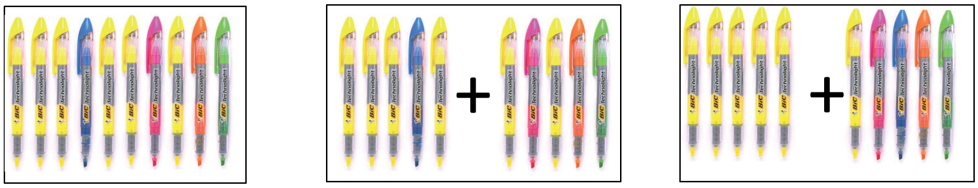
Experiment 5.2

In Experiment 5.1 we demonstrated, via the bonus condition, that regrouping the items could lead to a restoration of standard properties of choice. However, it is possible that this effect arose from the fact that subjects were making a choice between one item (the notebook) and two sets (the two subsets of pens) as opposed to the fact that each bundle of pens was presented in a fashion that highlighted their orderly properties. To avoid this

problem, in Experiment 5.2 we keep the items in the choice-sets completely identical but, as in Experiment 5.1, manipulate partitioning of the constituent items. In one condition the two subsets created by the partition were very easy to describe parsimoniously: one subset is all different, the other subset is all-similar. In the other condition neither subset is easy to describe; neither is all-different or all-similar.

Method

One hundred and forty three students from Tilburg University (35 females, M_{age} = 20.4, SD = 2.1) were assigned to one of three conditions. In all conditions, subjects made a choice between a bundle of ten markers and a notebook, and were informed, as in the first study, that at the end of the experimental session 10% would receive their choice. The three conditions differed only in the way the bundle of ten markers was presented (see Figure 5.3).



Combined Disorder Order

Figure 5.3: Display of the bundle of markers in the three conditions in Experiment 5.2.

In the two ordered subsets condition, the bundle of markers was split up in two subsets, both of which exhibited an obvious order: One subset of markers was all similar, the other was all different. In the two unordered subsets condition, the same ten markers were divided into two subsets, neither of which exhibited any coherent order. Finally, in the single unordered condition, the markers were presented as a single unordered bundle. After choosing between the markers and the notebook, participants indicated their WTP for the ten pens as well as the notebook; however, as in Experiment 5.1, these valuations were not incentivized.

We expected that participants would be more likely to choose the bundle of markers when the markers were split into two easily described subsets as compared to them being split into two not easily described subsets or into one not easily described set. WTP for the pens was expected to follow the same pattern.

Results

As can be seen in Figure 5.4, participants were more likely (67%; 32/48) to pick the markers over the notebook when the bundle was split up into one all-same and one all-different subset as compared to the condition in which the bundle was split up into two arbitrary subsets (44%; 21/48) or in which the markers were portrayed as one bundle (40%; 19/47); two ordered vs. one bundle, $\chi^2(n = 95) = 6.58$, $p = .01$; two ordered versus two unordered, $\chi^2(n = 96) = 5.10$, $p = .02$. The difference between the large bundle and the low-order subsets condition was not significant; $\chi^2(n = 95) = 0.11$, $p = .75$. A similar pattern was found for willingness to pay, although, due to the high variance in WTP, these differences were only marginally significant. On average participants were willing to pay €3.66 (*SD* = 2.01) for the high-order subsets of markers, but only €3.00 (*SD* = 1.73) for the low-order subsets of markers, $p = .09$. Willingness to pay for the one-bundle condition was slightly higher than the low-order condition but lower than the high-order condition €3.19 (*SD* = 2.19).

While the first and second experiments demonstrated violations of dominance as a result, we claim, of a preference for descriptive simplicity, the third experiment examines its consequences for a weaker property of choice: independence. Ordinarily we would not expect independence to apply to the constituent items composing bundles of goods, since it is perfectly natural that there could be consumption complementarities between items. However, we examine a situation in which order preference produces a strong violation of

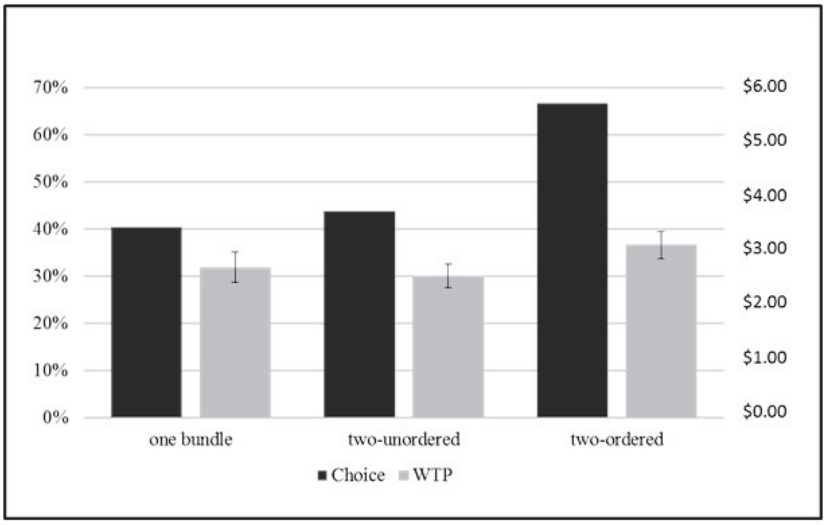


Figure 5.4: Results of Experiment 5.2. For each condition, the bar on the left represents choice proportion for the bundle(s) of pens over the notebook as indicated on the left y-axis. The right bar represents willingness-to-pay as indicated on the right y-axis.

independence in a context in which one would expect consumption complementarities to be small or non-existent because items are consumed one-at-a-time and the consumption of one item is unlikely to affect the marginal utility of consuming other items.

Experiment 5.3

Experiment 5.3 tests the prediction that people prefer bundles that are descriptively simple because they are all completely similar or completely different over bundles that are descriptively more complex. In this study, all decisions were implemented – subjects took home the bundle they chose — unlike in the previous two in which 10% of choices were chosen by chance to count. We also used choice items – bundles of beers — that were familiar to most subjects and that they cared about.

Method

Students from Tilburg University (94 females, 66 Males, M_{age} = 20.1, SD = 2.5) were directed to a cubicle where two bundles of beer were presented (see Figure 5.5). Subjects were told they could choose one bundle of beers which would be theirs to keep and take home.[16]

Subjects were randomly assigned to one of four conditions in each of which they chose between two bundles consisting of four different beers. In all conditions, both bundles were identical to each other for three of the beers; the only difference between the two bundles was that one bundle included a *"La Trappe"* beer (LT from now) as a fourth beer, whereas the other bundle included a *"Scheldebrouwerij"* beer (SB from now) as the fourth beer. In the two "similar" conditions, the three beers that were identical across both bundles were either all beers from the LT brewery (similar LT) or all three from the SB brewery (similar SB). In the two "different" conditions, the three identical beers were one beer by the St. Bernardus brewery (BE), one by the Lefebvre brewery (AB) and one by the LT brewery (different LT condition) or by the SB brewery (different SB condition). All choices are logically equivalent since in all conditions the first, second and fourth beer in the bundles are identical, the third item in the bundle is always LT in one bundle and SB in the other. Therefore, if preferences for bundles are determined only by the combined preference for the individual items in the bundle, then choice proportions should be identical across conditions. However, if subjects have a preference for order – for all-similar and all-different

16 Participants subsequently filled in a Need for Closure scale (Kruglanski et al., 1993) and a Need for Uniqueness scale (Snyder & Fromkin, 1977) on the computer in the cubicle. Results from these scales did not predict patterns of choice (p >.25 for all p's from 4 relevant tests) and are not further discussed.

bundles, choices for the bundle including the LT beer should increase in the similar-LT and different-SB conditions as compared to the similar-SB and different-LT conditions.

Beyond the idea that it would be a remarkable coincidence if consumption complementarities corresponded to order-preferences, the design was intended to produce results that strongly weigh against a consumption complementarity account. One could possibly argue that people enjoy consuming a wide range of beers, which could potentially justify the greater preference for the all-different options in the two Different conditions (the right-hand option from the Different LT condition and the left-hand option from the Different SB condition, in Figure 5.5). However, if people have a preference for variety, then they should prefer the somewhat different over all-same options from the two similar conditions, but this is not what we predicted or observed.

Figure 5.5: The two bundles of beers participants chose between in the four conditions in Experiment 5.3. Note that for each condition, the two bundles only differed on the third beer which was either of the "La Trappe" or "Scheldebrouwerij" brand.

Results

As predicted, subjects in the similar LT condition were more likely to choose the bundle including LT (83%; 33/40) as compared to participants in the similar SB condition (35%; 14/40, χ^2 (1, N = 80) = 18.6, p < .001. In the different LT condition, participants were less likely to choose the bundle including LT (23%, 9/40) as compared to participants in the different SB condition (92.5%, 37/40; χ^2 (1, N = 80) = 40.1, p < .001.

Experiment 5.3 shows that, when making real and consequential choices between bundles, subjects' choices are influenced by the ordering of the bundle. If subjects evaluated the bundles on the basis of the characteristics of the individual items in the bundles, we should have found equal proportions choosing the bundle including LT over the bundle including SB. We found instead that participants were more likely to choose a bundle that had high order, one that was completely-similar or completely-different over a bundle with lower order (partially-similar, partially different).

Experiment 5.4

If the difference in choices observed in Experiments 5.1-5.3 is the result of people preferring bundles with high order over those with lower order, we should be able to affect choices between two identical bundles by varying the salience of characteristics of the bundle. To illustrate, the two "bundles" in Figure 5.6 could either be described by their color or their shape. For color, the bundle on the left has higher order (its contents are all blue), as compared to the bundle on the right (three blue, one grey). For shape, the bundle on the right consists of all-circles and thus has higher order than the bundle on the left with three circles, one square. In this case, drawing attention to color should therefore increase preference for the bundle on the left, since the contents of that bundle are all-similar. Drawing attention to shape on the other hard should increase preference for the bundle on the right, since that bundle is all-similar on shape.

Note that this design is similar to that employed by Fox et al. (2005) who presented subjects with a three choices from a menu of six bottles of wine which were in one condition partitioned by grape-type (two bottles of Chardonnay, two Pinot Grigio, and two Sauvignon Blanc) or in the other by country of origin (two from Australia, two from Italy and two from California). In both conditions subjects split their choices over the three categories resulting in different choices depending on how the bottles were organized. However, whereas Fox et al. created different categories on the basis of these attributes, we keep the bundles the same and merely draw attention to one attribute or the other.

Figure 5.6: Illustration of perceived order being dependent on salience. When color is salient, the left bundle has higher order. However when shape is salient, the right bundle has higher order.

Method

Two hundred students from Tilburg University (130 females. M_{age} = 20.64, SD = 2.3) were assigned to one of two conditions in which they made a choice between two bundles of marbles created for this study by an artisan.[17] Similar to the example in Figure 5.6, the marbles were designed in such a way that on the attribute of color, bundle 1 was all-similar (all aqua) whereas bundle 2 consisted of 3 marbles that were red and 1 that was aqua. For design, bundle 2 was all-similar (all flowers), but bundle 1 consisted of 3 spirals and one flower (see Figure 5.7).

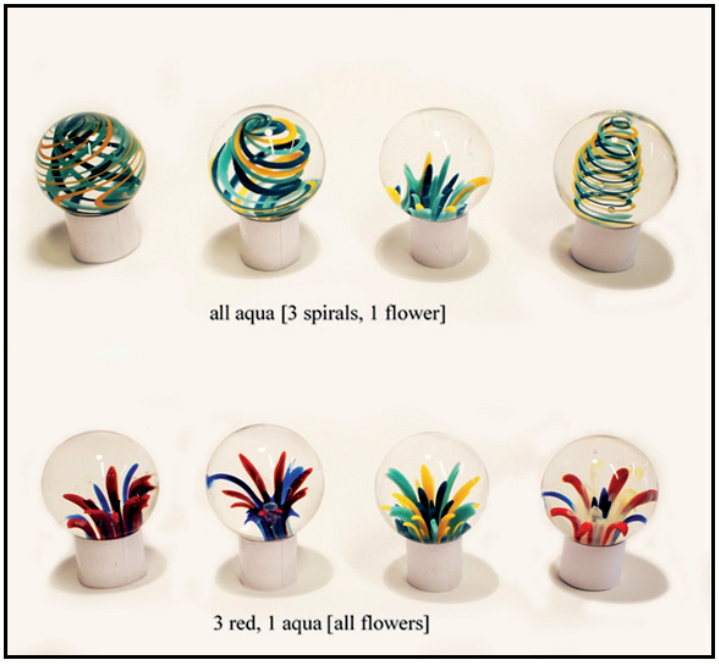


Figure 5.7: The two sets of marbles used in Experiment 4. In condition 1, the sets were described on color, in condition 2 on design.

17 Melissa Fitzgerald of the Pittsburgh Glass Center

Subjects in both conditions were instructed to choose one of these bundles, and were informed that at the end of the experiment one of them would be randomly selected to receive their chosen bundle. In the color-condition, bundle one was described as "all-aqua" and bundle two as "one aqua, three black." In the design-condition, bundle one was described as "three spirals, one flower" whereas bundle two was described as "all-flowers." If people prefer bundles with high simplicity, subjects should be more likely to pick bundle one when described as "all-aqua" as compared to the same bundles described as "three spirals, one flower."

Results

As expected, subjects were more likely to choose the bundle consisting of four marbles that were all aqua over the bundle consisting of three red and one aqua colored marble, when the description of the bundles was based on color (70%; 70/100). However, when design was emphasized, subjects were more likely to choose the bundle consisting of three red and one aqua marble ("all flowers") over the bundle consisting of four aqua marbles ("three spirals and one flower"; 61%; 61/100, $\chi(1, N = 200) = 19.38$, $p < .001$, $\phi = .31$). These results are not only consistent with the idea that the subjective perception of order is what is causing the increased preference, but also imply that the experience of order can be manipulated by emphasizing the (ordered) attributes of goods in a bundle.

Experiment 5.5

Finally, to provide more direct evidence of the idea that a preference for descriptive simplicity underlies the effects demonstrated in Experiments 5.1-5.4, we conducted a study in which, similar to Experiment 4, the bundles were identical across conditions. The only difference between conditions was that subjects were presented with a simple, but arbitrary, description of one bundle or the other. Given our explanation for the observed preferences, based on descriptive simplicity, we naturally predicted that choice sets accompanied by a simple description would be preferred.

It was key to our experimental design that the description we provided would not provide information that would cause the elements of the described choice set to seem more desirable. In the description of the methods, we explain how we designed the study to rule out such an alternative account of the results.

Method

Subjects were 147 students of Tilburg University (93 females, $M_{age} = 21.5$, $SD = 3.0$) who participated in a series of studies in the lab. When leaving the lab, these students were ap-

proached and asked whether they were willing to join a short additional study. All students agreed and were presented two CD box sets. Both of the box sets consisted of a deliberately random set of BB King CD's. In both conditions, boxed-set A consisted of BB King's first, second, fourth, sixth and seventh album, and boxed-set B of BB King's second, third, fifth, sixth and seventh album. The numbers were clearly displayed on the album-cases, and it was expected that subjects would generally perceive these bundles as unordered because no clear structure seemed to exist in the choice of albums. In the first condition, boxed-set A was described as a set of all the albums BB king has ever recorded in Tuscaloosa, to create a simple description that drew connections between the disparate albums. In the second condition, a similar description was provided, but for boxed-set B – that all were recorded in Biloxi. Participants were told that at the end of the experiment one of the boxed-sets would be given to one of the participants and were asked which box they would like to receive if they were the student selected to receive a boxed-set. If the previous effects on choice are the result of a preference for things that can be described simply, then subjects should be more likely to choose whichever box set that was accompanied by a unifying description.

People could possibly have lay theories about the music originating in some cities being better than others. Therefore we conducted a pretest in which a different group of students was asked to indicate on a 100-point slider-scale (1 = expect it to be worse, 100 = expect it to be better; slider originating at 50) the degree to which they expected a music album to be better or worse in quality if they knew it was recorded in six cities; Biloxi, Memphis, Greenville, Tupelo, Tuscaloosa, and Natchez. All cities were rated significantly different from neutral ($N = 210$, all p's $< .001$) with Memphis having the highest expectations ($M = 62$, $SD = 13$) and Greenville expected to be slightly better than average ($M = 55$, $SD = 12$). The others cities, including the two towns selected on this basis for the descriptions in the experiment (Biloxi and Tuscaloosa) were expected to be slightly worse than neutral ($M = 44$, $SD = 11$ for Biloxi; $M = 45$, $SD = 15$ for Tuscaloosa).

Results

As expected, participants were more likely to choose box-a when a unifying description for its contents were provided (all recorded in Tuscaloosa; 69.9%; 51/73) compared to the condition where the unifying description was provided for box-b (all recorded in Biloxi; 48.6%; 36/74,). $\chi^2(1, N = 147) = 6.85$, $p = .01$, $\phi = .22$.

Conclusion

In 5 experimental studies, we found that people have a strong preference for bundle of goods that can be described simply. The first three experiments were designed to provide evidence for the existence of such a preference for descriptive simplicity and to show that it can lead to patterns of choice that violate basic standard axioms of choice. Experiment 5.1 showed that easily described bundles are valued higher than unordered bundles, even when the unordered bundles consist of more desirable goods. Experiment 5.2 showed that organizing an unordered bundle into two ordered bundles which could each be described in simple terms increased subjects' preference for the bundles. Experiment 5.3 showed that choices between bundles are strongly influenced by whether constituent choice objects are organized in a fashion that makes salient a simple description of the bundles.

The final two studies were conducted to provide more direct tests of descriptive simplicity as the causal factor underlying the observed patterns of choice. In Experiment 5.4, drawing attention to an attribute increased choice for the bundle that displayed high descriptive simplicity on that attribute. Finally, in Experiment 5.5, providing a simple description for an otherwise seemingly unordered set of items increased preference for the described set.

These findings show a kind of complementarity in preferences that does not correspond to complementarity in consumption. This could reflect one of two underlying causes; it could be that people mispredict utilities (Kahneman & Thaler, 2006), or that they derive real utility from parsimoniously described bundles that has nothing to do with their consumption. Because the preference for descriptive simplicity is very likely the result of an intuitive and positive response to ordered groups of stimuli (Evers, Inbar, et al., 2014; Garner, 1970; Wertheimer, 1923), choices based on these intuitions can be the correct choice or mispredictions depending on the situation. For some goods, it is perfectly reasonable to expect intuitive responses to be a good indicator for expected utility, for example works of art and other goods mainly acquired for aesthetic appeal. However, in other cases, such as choices between bundles of goods which are mainly consumed in isolation and for their functional characteristics (as in our experiments), an initial affective response to the existence or non-existence of a simple description is likely to result in suboptimal choices.

If such an account is correct, it could be generalized to predict phenomena beyond the situations investigated here. Not only would it imply that similar effects of order should emerge for other stimuli besides goods, but also that other forms of creating or emphasizing order should result in increased liking and evaluation These could be other visual ways of creating ordered categories, for example the gestalt principles of perception are likely to cause similar effects (Wagemans et al., 2012; Wertheimer, 1923). But these properties could also be purely cognitive in nature, like the mere-exposure effect; an increase in preference for objects after repeated exposure (R. Zajonc, B., 1968; Zajonc, 2001), which

is likely the result of an increased ease of processing. These ideas clearly resonate with theories of metacognitive ease (or processing fluency) which posit that things that are easier to process are liked better (Reber, Winkielman, & Schwarz, 1998). However, these theories do not predict *what* is easy to process. Our results strongly suggest that descriptive simplicity plays a causal role.

The fact that people prefer not only all-different bundles over random bundles, but also all-same bundles over random bundles strongly suggests that order preference is not derivative of the human desire for diversification (Ariely & Levav, 2000; Read & Loewenstein, 1995; Simonson, 1990). Instead, the reverse seems possible – i.e., that the preference for descriptive simplicity may contribute to the preference for diversification. Indeed, the fact that people make different choices with respect to diversification when they make all choice simultaneously, as opposed to sequentially (Read & Loewenstein, 1995) fits very nicely with our result that order preference can produce changes in preference that are clearly unrelated to utilities arising from consumption. Similarly, the 1/N heuristic which people are known to employ when spreading investments over multiple portfolios (Benartzi & Thaler, 2001) but also when making choices from multiple menus (Fox, Ratner, et al., 2005) has been interpreted as a form of diversification that people employ to reduce risk (Fox, Bardolet, & Lieb, 2005). Again, it could be that, besides a deliberate strategy to reduce risk, 1/N diversification could also just be a consequence of a general tendency to prefer ordered organizations.

Economists usually view the source of preferences as lying outside the purview of their discipline. Lionel Robbins (1932, p.86), for example, in a treatise that was generally hostile to the introduction of psychological insight into economics argued that: "Why the human animal attaches particular values in this sense to particular things, is a question which we do not discuss. That is quite properly a question for psychologists or perhaps even physiologists." In this paper we seek to show that an investigation of the underlying cause of preferences can be useful, in terms of illuminating phenomena of interest to economists and revealing some of the limitations of conventionally assumed properties of choice.

CHAPTER 6

Discussion

In the previous four empirical chapters, I have presented evidence that sets influence decision making. In Chapter 2 and 3 I find that people are motivated to complete almost complete sets of physical objects (Chapter 2), and virtual items (Chapter 3). Specifically, I found that when people have a large proportion of a set, they are motivated to complete it. This motivation to complete the set can result in an increased willingness to pay for products that complete the set (Chapter 2) or in increased effort and persistence to earn set-completing objects (Chapter 3). Furthermore, in Chapter 4 and Chapter 5, I found that people prefer sets that are either all-similar or all-different on salient attributes (Chapter 4) which is likely the result of a preference for simplicity (Chapter 5). These set-based effects appear to be so strong that they lead to violations of normative choice theories (Chapter 5). As already mentioned in the introduction, previous research on sets in properly incentivized surroundings is limited, if not non-existent. The chapters in this dissertation serve foremost as demonstrations that sets do strongly influence behavior in predictable ways. Because research on this topic is so novel, there are more questions and opportunities for future research. Below I will discuss the most salient questions that I hope to be able to provide answers for in the future. Where applicable I will also provide data from experiments we ran on these questions. Generally, these open questions and future directions can be spread over 4 different broad questions:

- What are the causes of these set-effects?
- What are the consequences of set-effects?
- How do these set-effects relate to known psychological phenomena?
- What practical recommendations can be given based on this work?

Causes

As mentioned in the previous chapters, I believe that Kolmogorov complexity (Chater & Vitanyi, 2003; Kintsch, 2012; Kolmogorov, 1965; Vitányi & Li, 2009) plays a crucial role in the effects described in this dissertation. Below, I will explain why Kolmogorov

complexity is a logical precursor for these effects, discuss literature consistent with this idea, present some data that substantiates this idea, and discuss alternative explanations.

Kolmogorov Complexity

Kolmogorov complexity is a term originating in information theory and mostly discussed in computer-science. This theory was developed by Andrey Kolmogorov (1965) who was interested in finding a way to measure "information" (see; Grünwald & Vitányi, 2014). The Kolmogorov complexity of a string of information is defined as the shortest possible description of this string. The string k429hgw984tmg therefore has high complexity; it cannot be described in simpler terms than the string itself. The string xoxoxoxoxoxoxoxoxoxo on the other hand is very low in complexity; "10 times xo". Similarly, the string 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20 has low complexity "1 to 20". The Kolmogorov complexity of a string therefore does not necessarily reflect the absolute amount of information (or letters) since the xoxoxoxoxoxoxoxoxoxo consists of more letters than k429hgw984tmg but still has lower complexity. Kolmogorov complexity also does not necessarily reflect the amount of different pieces of information since the string 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20 consists of more unique symbols than k429hgw984tmg but still is much lower in complexity. The Kolmogorov complexity of a set of information can thus be seen as the length of the most compressed form of the information. This theory has mainly been picked up in the computer sciences and artificial intelligence domains, because one way to increase the speed at which computers work is by decreasing the computational resources needed to process information.

Simplicity Facilitates Processing

Just like computers, human processing power is limited. We are confronted with large amounts of information that we need to process and store to make sense of the world, but our short term memory can only contain a maximum of three pieces of information[18] to work with. To maximize the contents of the short-term memory, people "chunk" information (De Groot, 1946; Lindley, 1966; Miller, 1956). The chunking of information is the compressing of larger numbers of data-points into a smaller number of "descriptions" that contain the same data-points. For example, the string 1225, consisting of four data-points, can be divided into two chunks: 12 and 25, or one chunk: "Christmas". Because we can only have about 3 chunks of information active in memory at one time, the more information we can contain in one chunk, the more information we can process. As such, it is clearly beneficial to automatically transform information in the fewest chunks possible (Gobet et al., 2001).

---

18 it was believed for a long time that this was 7 +/- 2 popularized by Miller (1956), but recently this belief has been updated (e.g., Mathy & Feldman, 2012).

Since the Kolmogorov complexity of an information-set is by definition the shortest description of the contents of this set, it seems functional for humans to compress data in ways that result in the lowest Kolmogorov complexity. A similar sentiment has been argued more substantively by Chater and Vitányi (2003) who claim that simplicity, defined as the inverse of the Kolmogorov complexity, could possibly be the unifying principle for all models in cognitive science.

Preference for Simplicity

So far, I have provided the logical argument for why I believe people automatically process stimuli in the form of the lowest complexity possible, but I have not yet explained why I believe people would *like* stimuli low in complexity more than stimuli high in complexity. Besides it being clearly advantageous for survival to be able to process data in its simplest form, an argument can also be made for the functionality of liking simple data. Data for which a simple ordering exists means not only that the environment is easy to process and understand but also implies that the environment can easily be predicted and controlled.

The idea that people prefer sets of stimuli that are simple to process is not new. Already in the 1800's, the philosopher Ernst Mach claimed that the cognitive system should and does, prefer patterns that provide simple descriptions of the data (Mach, cited in Chater & Vitanyi, 2003). Similar ideas can be found in work by the gestalt psychologists, on the gestalt principles of perception (Koffka, 1935; Wertheimer, 1923). These works focus mainly on the ways on which people automatically group (visual) stimuli together, but the observation that groups of stimuli that are easier to group are also liked better is heavily implied in these texts, by calling such groups "gute gestalten" (good gestalts) or "pragranz" (good figure). In the original German work by Wertheimer (1923), similar sentiments can be found in the footnotes that are not included in the translated version. In these footnotes Wertheimer (1923) describes how patterns that cannot be easily grouped according to the Gestalt principles can feel bad; when perceiving one of those patterns "a pause occurs, possible confusion, a disorder in the structure" (Wertheimer, 1923; p. 316, translated from German). But, the gestalt psychologists only discuss these gestalt principles as observations that describe how people generally group stimuli. They do not provide a reason for why this may be so, nor provide a model or theory that would allow one to predict whether a pattern will form a good or a bad gestalt.

Later, Attneave (1954) reasoned that perceiving sets of stimuli in the simplest form possible would be the most economical behavior of people since it saves cognitive resources. For this reason the simplest form of perception should be preferred. He also notes that the idea that people prefer simple organizations captures all previous gestalt principles for each one of them imposes some fort of structure on the stimuli. In the 60's, other researchers did empirically investigate affective responses to patterns. For example, Garner and Clement

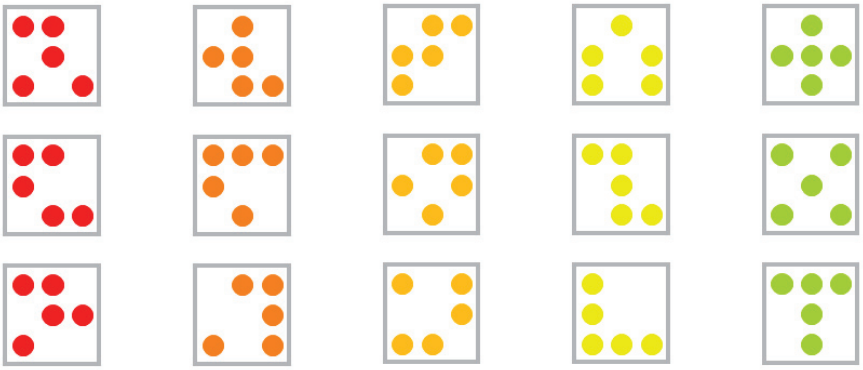


Figure 6.1: Examples of patterns taken from Garner & Clement (1963) that were rated on their "goodness". Colors reflect average rating with the patterns on the left rated lowest on pattern goodness and those on the right highest.

(1963) found that people indicate a much higher liking for some patterns than others (see Figure 6.1). Trying to explain why this was the case, Garner (1962) used the notion of redundancy coming from information theory (Shannon, 1948), in which the redundancy of a (set of) stimuli is a measure of the stimulus in relation to alternative possibilities. The word uncertainty has also been used to mean the opposite of redundancy, and when discussing patterns is seen as the number of alternative patterns that could be formed with these stimuli.

What is interesting, is that Glanzer and Clark (1963) found similar results measuring recall accuracy, but explained these results as being consistent with a verbal loop hypothesis; the shorter the verbal description of the stimuli, the better it will be remembered. This explanation is remarkably similar to the idea behind Kolmogorov complexity. But, complexity as a possible cause of these judgments was not recognized and the verbal loop hypothesis as explanation for these results was rejected since it was nonsensical to assume that people would actively describe things they were perceiving and base their judgments on those descriptions (Garner & Clement, 1963). Only recently, the two ideas have been tested against each other, with (Kolmogorov) complexity[19] predicting human memory for random strings better than Shannon uncertainty / redundancy (Falk & Konold, 1997).

Further support for the idea that things that are easier to process will be liked better comes from research on metacognitive ease and fluency (Reber, Schwarz, & Winkielman, 2004; Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990; Winkielman, Schwarz, Reber,

19 There are several, slightly different, models incorporating the idea of Kolmogorov complexity, such as Minimum Description Length (Risannen, 1978), Minimum Message Length (Wallace & Freeman, 1987), Stochastic Complexity (Risannen, 1987).

& Fazendeiro, 2003). Meta-cognitive ease, or fluency, details the ease that people experience when they perceive or process things. For example, the activity of reading a classical Russian novel will be experienced by most as more difficult than reading a popular magazine. The difficulty with which people process a stimulus affects how much they like it. For example, people like things better when they have seen them before and which are therefore, likely, easier to process (mere exposure effect; Whittlesea, 1993; Zajonc, 1968; Zajonc, 2001). Similarly, it has been found that people seem to smile more when looking at clear pictures than when looking at pictures that have been distorted by adding noise (Winkielman & Cacioppo, 2001). Finally, it has been found that there is a positive correlation between processing speed and judgments of liking (Checkosky & Whitlock, 1973). It must be noted though that most research on fluency and metacognitive ease does find that ease of processing affects (amongst others) liking judgments, but that these theories do not discuss *when* stimuli should be expected to be easy or difficult to process. But, combining knowledge about *how* people process information (creating chunks that maximally decrease the complexity of the information) with work on the *consequences* of this way of processing (metacognitive ease increases liking) leads directly to the hypothesis that what is easier to chunk, and thus has lower complexity, should be liked better.

Further Support for the Complexity-Liking Link

It is impossible to directly test the unique role of complexity in the judgment of sets, since the complexity of a set cannot be changed without simultaneously changing something else (the contents in Chapter 4 and Experiment 5.3, or the organization or description in Chapter 5). It is therefore impossible to prove that (subjective) complexity is what causes these effects. But, besides the evidence presented in previous chapters, where different ways of reducing or increasing complexity lead to similar results, I also conducted other experiments that give support for the link between complexity and liking.

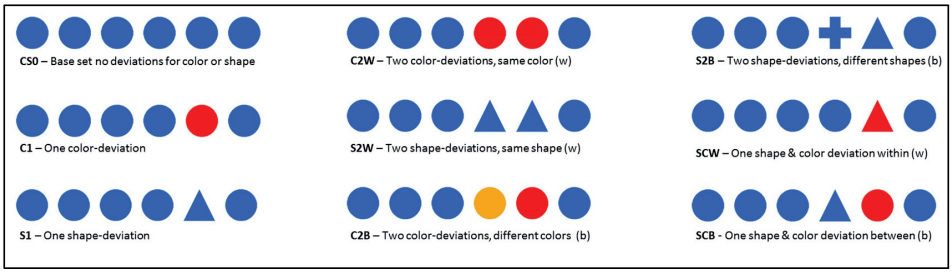


Figure 6.2: Example of one set of trials used in Experiment 7.6b.

The first supporting evidence comes from a set of two experiments in which participants rated many sets of geometrical shapes on how much they liked the sets (see Figure 6.2 for an example; see Experiment 7.6a and 7.6b in the **appendix** for more information).

In both experiments, judgments of liking were very consistent and generally showed that when complexity increases, liking decreases (see for example Figure 6.3).

The results of these experiments also suggest that the relationship between complexity and liking is not a linear one, but rather a concave function, like in Figure 6.3. It is interesting to note that, whereas defining simplicity as Kolmogorov complexity leads to a high proportion of variance explained (81%), defining it as the absolute amount of different stimuli can only very weakly explain the variance between trials (27%).

Further supporting evidence for the link between complexity and liking comes from another experiment in which I manipulated whether participants made deliberate or more intuitive choices. Since the processing of simpler sets should automatically lead to a more positive response, the effect of simplicity should generally be an intuitive effect. Responses from the participants in Experiment 4.2 and 4.3 (page 45 and 46) in which

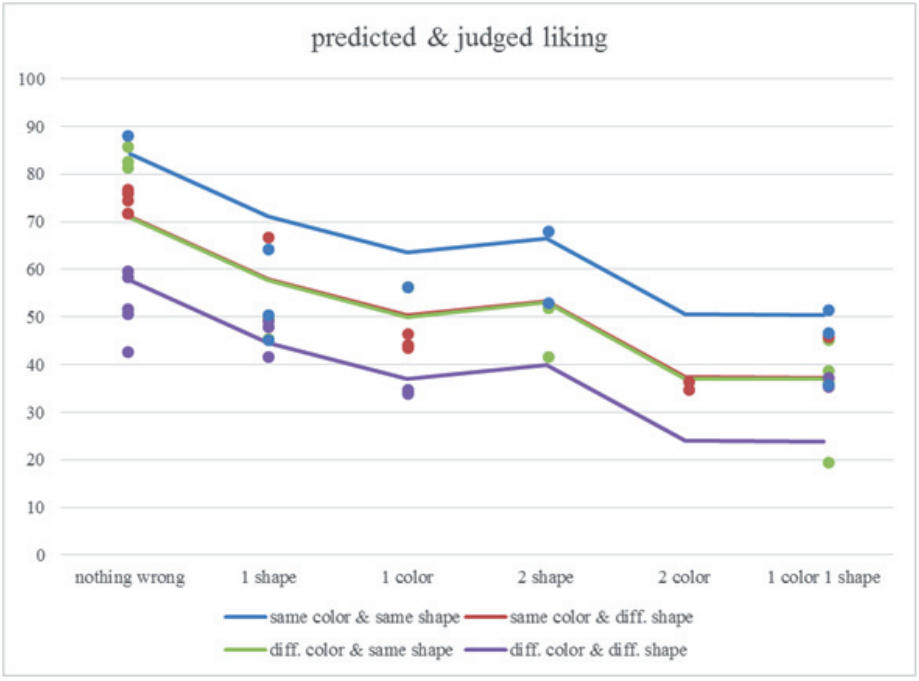


Figure 6.3: Results of Experiment 7.6a, depicted on the y-axis is the average rating of the set, on the x-axis the number of deviating colors and shapes. The solid line represents the predicted rating based on the regression while the dots represent actual ratings of each set.

we asked participants why they chose the set they selected are very consistent with this idea, since participants usually quickly made their choice but took a long time formulating reasons for their choice and often said they made their choice based on intuition. To test this more explicitly, I manipulated whether participants would make the choice more based on intuition or more based on deliberation (this experiment is described more thoroughly in the appendix as Experiment 7.8). In this experiment, participants chose between two sets of pens. One with low-complexity (all-different), the other more complex but also, arguably, more functional (consisting of several blue, black and red pens; see Figure 6.4).

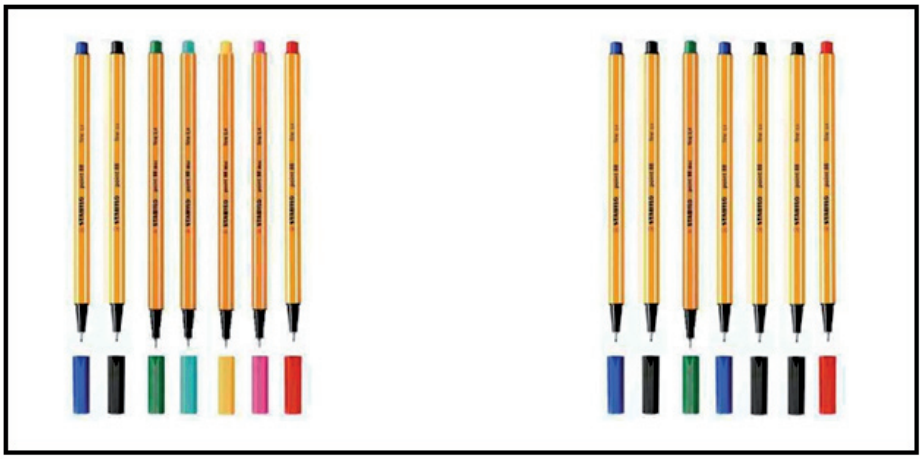


Figure 6.4: The two sets of pens participants chose between in Experiment 7.7

When participants were just asked to choose between these sets, a majority preferred the set low in complexity consisting of all-different pens. However, when participants were first asked to describe common situations in which they would use the pens in the future, choices for the more complex set of black and blue pens increased suggesting that choice for the all-different set was the result of a basic affective response to that set.

An alternative explanation for the effects we found, particularly in the cases where one item in the set seems "odd", could be that people formulate an idea of the structure underlying the set or pattern and that the subsequent violation of this expectation causes frustration and/or anxiety akin to meaning violations (Heine, Proulx, & Vohs, 2006; Kay, Whitson, Gaucher, & Galinsky, 2009). If anxiety due to violations of expectations plays a large factor in the effects we found, then presenting the "odd one out" in the start of a sequentially appearing set should decrease liking less (as compared to a good fitting set) than presenting the odd-one-out at the end of a set. This is exactly what I tested in Experiment 7.8 (described in the appendix). I presented different patterns of geometric shapes (see Figure 6.5)

Figure 6.5: Examples of three of the patterns used to test violation-of-expectations as cause for the difference in liking for different sets.

that appeared sequentially. Each set was presented a total of four times; either the original pattern or the mirror-image, and either starting from the left or from the right. As a result, sometimes the set had bad-fit virtually from the start, which means participants would not be able to form an expectation of the structure of the set. Other times the bad-fit (or the violation of the structure) only appeared near the end of the set, in these cases participants would be able to form an expectation of structure that was subsequently violated. However, we found no differences between the liking-judgments of these sets. This suggests that in general, disconfirmation of expectations and the accompanying anxiety do not play a large role in the effects of sets on preferences that we have detailed in this dissertation.

Exceptions to the Complexity-Liking Link

All results reported in this dissertation are consistent with the idea that when people perceive a collection of individual items as a set, they prefer sets that are low in complexity over those that are high in complexity. However, even though I believe this to be generally true, I do not expect this to be the complete picture. According to the complexity-hypothesis, increases in complexity should *always* lead to lower liking. There are two cases in which I expect this pattern not to hold. The first case is a large set or pattern that seems completely random. There seems to be a point where one does not try to make sense of chaos anymore, but just accepts a pattern or set as "random chaos". Another case in which I believe judgments of liking may deviate from predictions made based on complexity is in case where there is a large coherent structure and only one element out (for an illustration of these hypotheses see Figure 6.6). In such cases, the deviation from complete structure may be es-

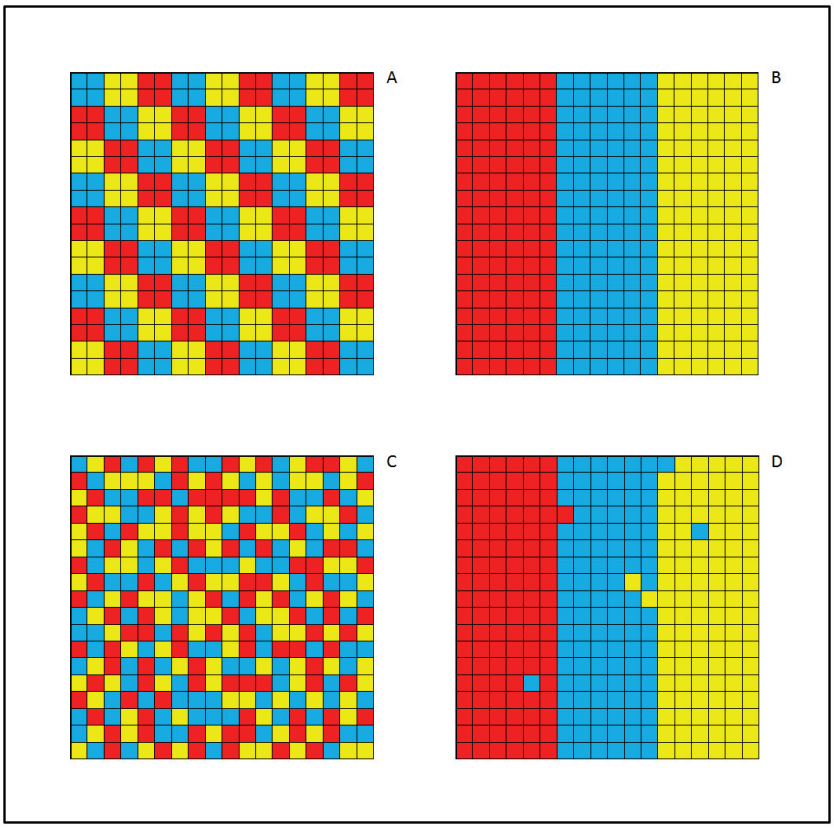


Figure 6.6: Patterns with different amounts of complexity. On top (A&B) two patterns with very low complexity, expected to be liked much. On the bottom-left (C) a highly complex pattern that I do not expect to be disliked very much since it is practically "random". On the bottom-right (D) a not very complex pattern that I still expect to be disliked a lot because it appears to be very structured but then has a few violations to this salient structure.

pecially frustrating. I have not yet empirically investigated these two possible exceptions, but the results presented in Chapter 2 on the set-completion effect seem consistent with the expectations for "almost-simple" structures, showing the highest motivation when only one or a few things are missing from a complete, and thus simple, set. However, it must be noted that in the patterns consisting of 5 or 6 items, I have not found greater dislike for the sets in which only one element was "odd" as compared to sets in which more than one item was "odd". But, those sets only consisted of a very small number of items, it could be the case that with a large set of stimuli, the underlying structure is so salient that any deviation feels especially frustrating. Just like Wertheimer (1923), I would expect that these patterns, that almost form a Gestalt but not quite yet, will feel especially frustrating.

All in all, the effects presented in this dissertation, together with published work in different domains all seems very consistent with the idea that; 1) people process patterns and sets of stimuli in a way that results in the lowest Kolmogorov complexity, and 2) people generally like patterns and sets with lower complexity more than those with higher complexity. This has not only implications for research looking at the relationship between simplicity and liking/preferences, but also holds promising implications for research in other terrains. Essentially, if the Kolmogorov complexity of a set of stimuli predicts how people process information (both the number and contents of the chunks they create, as well as the metacognitive experience of doing so) then it should play a large role in many different cognitive domains dependent on the processing of information, like judgments of likelihood, familiarity, similarity, and others (for an overview, see Chater & Vitányi, 2005).

Before continuing to the consequences of this complexity-liking link, it is important to point out that the concept of Kolmogorov complexity as used in computer-based information processing should not directly be applied to human cognition without recognizing two important differences between humans and computers. Generally, computers are able to process faster and have a larger memory than humans. However, computers process in a sequential way using a single strategy, whereas humans are flexible and can use many different strategies for the same problem. For example, the number "7625597484987" is not very complex for a computer, it is 3^3^3, but for most humans it would be a very complex string. For this same reason, complexity may also differ between people. This is nicely illustrated by the work of De Groot (1946) comparing the memory of chess-players with novices. When these people were asked to remember the way chess-pieces were placed on a chessboard, the chess players greatly outperformed the non-players, since instead of remembering each individual piece they could simplify the set-up by merely remembering game-positions. However, when the board was set up randomly instead of adhering to the rules of the game, chess-players performed on exactly the same level as the non-players. As such, the complexity a human experiences in a set of stimuli is clearly subjective (see also; Strannegård, Nizamani, Sjöberg, & Engström, 2013).

Consequences

In this dissertation I have only looked at the direct consequences of a set-context on judgments and decisions. However, based on the theory presented, other consequences would be expected. First of all, the way in which the set is processed should, besides liking, also affect other forms of cognition, most notably memory. Furthermore, basing a decision on the affective response to complexity can also have several consequences, people may for example regret their choices after some time of deliberation.

Consequences of the Process

Good-fitting and complete sets are processed more easily because they can be described in simple terms: “all-same”, “all-different”, “all Harry Potter books”, “all existing pokemon”, etcetera. Bad fitting sets and incomplete sets need more complicated descriptions such as: “three blue, one red”, “the first three, and fifth Harry Potter book”, and “Pikachu, Bulbasaur, Squirtle, and Charmander”. As a consequence, this implies that “good sets” are processed on a more global level, only the structure between the items in the set is processed. Bad-sets on the other hand have to be processed on a more concrete level since they lack the organizing structure. Therefore, memory for the contents of a set should be better for the items in bad sets as compared to good ones. I have tested this idea in one experiment in which participants were confronted with 6 different sets for which they were asked to indicate their liking (this experiment is described in more detail as Experiment 6.8 in the appendix). Some of these sets had good fit (they consisted of all different items) whereas others had bad fit (some items were present in the set 2 times; see Figure 6.7 left-panel). After that, I confronted participants with items of the same types as the sets before. Sometimes these items were part of one of those sets (always as a single item), sometimes they were similar but a different version of the items they saw before (see Figure 6.7 right panel). I found that, when participants saw a bad-set rather than a good one, they were more likely to correctly indicate seeing the item before when the item was a part of the sets they saw, and were also less likely to indicate having seen an item that was not part of a set before. It must be noted though that this difference was quite small.

Figure 6.7: On the left side is an example of the sets in the three conditions (2 good sets on the top, one bad set at the bottom). On the right two of the memory probes for which participants indicated whether they saw the item before or not.

A stronger effect on memory can be predicted for items that "fix" a set. As we saw, attention is focused on items that complete an incomplete set. I would expect the same salience for other items that somehow break up the simplicity of a set. Thus besides an increased focus on the individual items in bad sets, through increased attention I would expect even better memory for the missing/ badly fitting item akin to pop-out effects (Adler & Orprecio, 2006) and the Zeigarnik-effect (Zeigarnik, 1967). I have tested this idea together with Diogo Ferreira by conducting a study in which participants either gathered an almost complete set of 19 out of 20 badges, or an incomplete set of 9 out of 20 badges (for a more detailed description of this study, see Experiment 7.9 in the appendix). After this, we tested how well the participants in both conditions remembered the 20th badge. We found a strong effect of near-completion on memory with twice as many participants (65%) picking the right badge out of a group of 4 in the almost complete condition as compared to the control condition (32%).

Consequences of the Decision

Decisions based on a greater liking of sets low in complexity may lead to choices that are objectively inferior (see Experiments 4.1, 4.3, & 5.1), result in people overpaying for items (see Chapter 2) and may even cause people to pursuit the acquisition of items they normally would not want (Experiment 2.4). When the items in these sets are subsequently used for their functional characteristics, people may realize that, if they would have made a different choice, they would have been better off. These counterfactual thoughts can result in regret and self-blame (Zeelenberg & Pieters, 2007), but also a dislike of the set which may carry over (through a negative halo-effect; (Thorndike, 1920) into negative attitudes towards other things related to the set (such as the brand or the company).

It is also important to realize that I have only investigated the *direct* liking of low-complex sets. We know from research on art and aesthetics that, people generally prefer slightly complex works of art over simpler works (see; Berlyne, 1971, however evidence for this idea is lacking, see; Martindale, Moore, & Borkum, 1990). For example, Nicki & Moss (1975) found that people generally rated more complex works as more interesting and pleasing, though it must be noted that complexity was measured as the "variation in color and shape" which does not necessarily align with Kolmogorov complexity (for example, one stimulus used was the artwork presented in Figure 6.8, which has a lot of colors and shapes and would therefore be categorized as extremely complex, while it still has a very clear underlying structure).

But more importantly, humans judge art differently than mere patterns. Art is not supposed to be just likeable, it has to be interesting, it has to stay interesting and preferably have deeper meaning as well (Beardsly, 1958/1981). This is where the relationship between complexity and liking becomes very interesting, since it implies that the most aesthetically

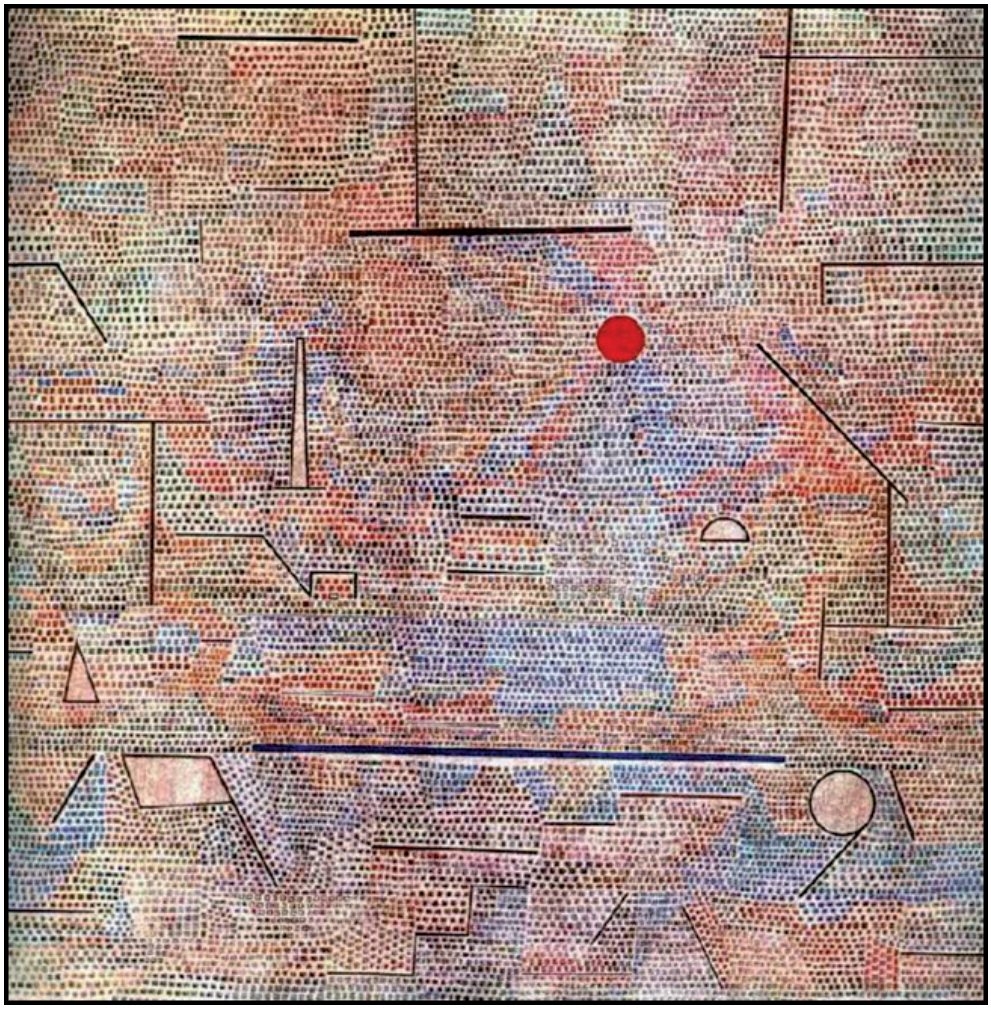


Figure 6.8: Cacodemonic by Klee, used as a complex artwork in Nicki and Moss (1975).

pleasing art should on one hand not be complex, but on the other be complex enough to stay interesting. Hence, the most appreciated art should consist of several low complex patterns that interact in a structured way, because it leads to the piece of art being immediately likeable, but stay interesting because one can alternate ones focus between these different structures (Kintsch, 2012). Based on this reasoning, we could expect that sets that are not primarily bought for their aesthetical appeal but for their functionality (for example a set of mugs) will probably be liked more if they are low-complex sets even in the long run. However, a set bought for its aesthetical value which is looked at often, such as a set of small statues, may become uninteresting in the long run and loose its appeal.

Finally, the motivation to make sets less complex (by completing an incomplete set, Chapter 2; or by choosing from different items to add to a set, Experiment 4.4) can be experienced as an extrinsic drive the person has no control over. In these cases, this desire may feel very unwanted, and this feeling of being trapped or sucked in may be experienced as frustrating (Prentice, 1944).

Relationship with Other Psychological Theories, Traits and Behavior

In this dissertation, I have mainly discussed these set-related effects as if they appear in a vacuum, but there are other psychological theories and traits for which one can expect them to interact with these set-effects. Furthermore, on a global level, set-effects show some remarkable similarities to other psychological effects, suggesting that both the set-effects as well as those other effects may be the result of a more general core-mechanism.

Moderating Traits and Personality Characteristics

In all the studies reported in this manuscript, we have investigated how the presence of a certain set for one group of participants affects behavior compared to another group of participants who were confronted with another set, but it is possible that the presence of a set is moderated by certain psychological characteristics. I have tested some of the moderators for which either I thought it could be possible that they have an effect on set-effects, and some moderators others suggested may play a role but for which I thought it was not likely for them to do so.

One question that has often been asked is whether there are differences in these effects between men and women, but, I have not found consistent differences between the two genders. Besides gender, there are also several personality traits that (partially) explain how a person deals with (a lack of) structure that seem likely moderators for the set-completion and set-fit effects. The need-for-structure (NFS; Neuberg & Newsom, 1993) and the need-for-closure (NFC; Kruglanski et al., 1993) scales are two psychological constructs aimed at measuring preference for structure and aversion of ambiguity. As such, it seems likely that these traits correlate with set-effects such that people scoring high on need for structure show stronger set-effects (liking of simple sets as well as an aversion to complex sets). However when one looks at what these scales really measure, it seems that the "structure" targeted by these scales is something different than the structure that reduces complexity in our experiments. To illustrate, both NFS and NFC include questions such as *"I like having friends who are unpredictable"* and *"I find that a well ordered life with regular hours suits my temperament"*. Even though it seems unlikely that these traits correlate with set-effects, I still measured them in 3 experiments (2 on set-completion, 1 on set-fit), but they never showed a strong relationship (I found one positively correlated subscale once but that is

likely due to chance). Two other traits I measured were trait-autism and trait-OCD, but neither of these correlated significantly with set-effects either.

I did find a correlation for another trait I measured; Type-D curiosity (Litman & Jimerson, 2004). Generally, there seem to be two types of curiosity. Type-I, or curiosity out of interest, is the curiosity one experiences when one is out to discover new things and which involves the expected pleasure experienced from finding out new things. Type-D, or curiosity out of deprivation, is the curiosity one experiences when trying to reduce uncertainty and eliminate undesirable states of ignorance. Whereas type-I is mainly related to an intrinsic motivation to learn more, type-D is mostly an extrinsically induced need state. In one set-completion experiment I also measured state curiosity, and found that type-I curiosity does not correlate with the set-completion effect, but that for people scoring high on type-D curiosity, the motivation to complete the set was even bigger in the set-completion condition. Even though it may initially sound surprising that this measure of trait-curiosity moderated the set-completion effect, inspecting the items in the scale suggest that there may be a more global trait that underlies both the reaction to incomplete sets and incomplete information. For example, two items in this curiosity scale are; "*It really gets on my nerves when I am close to solving a puzzle, but still cannot figure it out"* and *"I can spend hours on a single problem, I just cannot rest without knowing the answer"*. If curiosity and set-completion (and possibly even goal-gradient effects) are all expressions of some general underlying desire for completion, this would not only change our theoretical understanding of these concepts, but also hold practical implications. For example, companies could use this insight to selected and motivate employees.

Assortments & Consumer Decision Making

The findings described in this dissertation also seem related to research on assortments. One salient difference between that work and the work presented in this dissertation is that research on assortments usually either looks at judgments of the entire assortment, or choice from assortments, whereas I mainly looked at the choices between bundles of goods or the addition of goods to an already owned set. Still, it seems like similar processes play a role in both cases, most notable the complexity of a set and the judgments of such sets.

A problem with comparing work on assortments with itself and with the findings presented herein is that different definitions of simplicity or structure are used. Most research on assortments either looks at the absolute number of distinct items, or the redundancy of the assortment when investigating the effects of structure (see for an overview; Chernev, 2012). This is problematic because what is simple under one definition may be complex under the other and the other way around. For example, a string consisting of 10 different numbers (from 1-10) is more complex than a string consisting of 5 different numbers (1-5) using absolute number of stimuli or redundancy as a benchmark for simplicity. However,

under Kolmogorov complexity the 10 number string may be much simpler than a 5 number string depending on the organization of the numbers within the string (for example 12345678910 is simpler than 15423). As a consequence, conclusions from this research that often seem to contradict each other may just be an artifact of the definition of simplicity used. Because of that, it would be a fruitful endeavor to collect these studies on the effects of the structure of assortments and reanalyze their results using all three definitions of structure to see whether inconsistencies in results can be explained by the difference in definitions and to see which definition is most accurate and consistent in predicting the effects of assortment-structure.

Partition-Effects, Choice Bracketing and 1/N.

Even though the research on how people choose between sets of goods is very limited, there is a line of research looking at how choices distributed over bundles (or sets) is affected by the organization of those bundles; partition dependence (Fox, Ratner, & Lieb, 2005; Fox & Rottenstreich, 2003). Most of the experiments revealing partition dependence effects use a set-up where participants can choose products from several bundles of different sizes. The organization of these bundles is varied between conditions, but in total exactly the same items are presented. For example, in Study 3 in Fox, Ratner, and Lieb (2005), participants were told they would receive 3 free lunches and were asked to indicate on a calendar when they would like to receive these free lunches. All these participants were Duke-students whose academic year consists of the fall semester, which itself consists of term I and term II, and the spring semester consisting of term III and term IV. For some of the participants the calendar-year was presented as "term I, term II and spring". For others it was presented as "fall, term III, and term IV".[20] Because participants were likely to divide their choices equally over categories (one in each), participants were more like to plan the lunches earlier in the "term I, term II, and spring" condition as compared to the "fall, term III, term IV" condition. This tendency to divide choices over categories has also been investigated in work on choice-bracketing (see for example; Read & Loewenstein, 1995), and is often called a 1/N strategy or heuristic (see for example; Benartzi & Thaler, 2001).

Different explanations have been proposed for these effect. In some cases, the categorization of items could be seen as a way of conveying information by the person who created the categories. However, these effects are also found when such an interpretation is very unlikely (Fox, Ratner, & Lieb, 2005). Another possibility that has been proposed to explain these strategies is that they are used to reduce risk, especially when the quality of the options is uncertain (Kahn & Lehman, 1991; Simonson, 1990), but this cannot fully explain the effects either (Read & Loewenstein, 1995). It could very well be that people

20 Besides these two conditions, the experiment also included a control condition in which the year was presented only in a fall and spring semester.

show this diversification bias purely because it is the least complex, and as such the most attractive, structure of a choice-set.

Goal Gradient & Round Numbers

As already discussed in Chapter 2 and Chapter 3, the set-completion effect seems very similar to a goal-gradient effect, but the set-completion effect differs in a crucial way from regular goal gradients. The goal gradient effect is the increased motivation people and other animals show when nearing a goal. We find evidence for an increase in motivation and willingness to pay when people are nearing a complete set. But, this increase in motivation emerges even when people did not have a goal of completing the whole set and even emerges for aversive stimuli. As such, the goal gradient cannot explain the effects I present in Chapter 2 and 3, but may definitely play a role in the increased motivation when, and only when, the proportion of items one owns from a set is high enough to trigger a completion goal.

Another effect that shows some similarities with the set-completion effect is the increased motivation found when people's achievements near round numbers (Pope & Simonsohn, 2011). For example, it was been found that when students that had a SAT-score just below a round number (e.g., 1000, 1100, 1200 etc.) are more likely to retake the test to try and improve to reach the round number, resulting in a clustering of the distribution of scores near these round numbers. The authors explain these findings as people consciously setting goals for themselves, and because goals are often set at round numbers, they also set their goals at a round number. A very consistent but slightly alternative explanation could be that, similar to the set-effects described in this dissertation, round numbers automatically form natural, gestalt-like, reference-points that people are motivated to achieve as soon as they become salient. If this is the case, that would mean we should also find clustering around round numbers in situations where people do not explicitly set goals for themselves.

Collecting & Collectors

When thinking about sets and collections of products, one sub-group of the population comes to mind immediately; collectors. It is claimed that a large proportion of the general population are collectors (33% in O'Brien, 1981; 62% in Schiffer, Downing, & McCarthy, 1981), and the organization and acquisition of products that are a part of a set is an essential part of collecting behavior.

Very little research has actually been conducted on collecting and collectors. Regardless, collecting has been claimed to be very structured and organized, all about completing clearly bounded sets of goods (e.g., McIntosh & Schmeichel, 2004). This idea is also reflected in Carey's (2008) model of set-completion in which it is assumed that collectors, unlike non-collectors, are willing to pay more for items that complete sets. There are two

reasons I believed this may not be the case. First of all, to be a collector one has to have something to collect. Completing a set completes the collection and as a consequence, the person stops being a collector (see also; Belk et al., 1988). Assuming that collectors like collecting, if anything they should be more hesitant to complete their collection. Another reason is that, even though it is assumed collectors are very structured in their collecting (McIntosh & Smeichel, 2004; Belk, 1995a; but see Danet & Katriel, 1989), it appears that a lot of people that collect do so without a strongly defined plan.

Together with Siegward Lindenberg, I examined collecting behavior. Two studies, described in more detail as Experiments 7.9 and 7.10 in the appendix, provide evidence for the ideas expressed above. In one study I visited the national collectors fair, together with two research assistants, and conducted structured interviews with 61 self-proclaimed collectors. When we asked these collectors to describe their entire collection and indicate if their collection had any subsets, a large group of these collectors expressed that they started with a very small collection around a very specific theme. However, when their collection grew and they were about to complete their collection, they expanded the theme of their collection to prevent completion of the collection allowing them to continue being a collector. A good example of this is the middle-aged man who explained that he started collecting graphic novels starring Olivier B. Bommel[21] when he was still a student and he received a free novel when he bought a certain product at the supermarket. After he received a second novel, he actively started to collect them all. But then, when he almost completed his set, he expanded his collection now also collecting any products with the image of the main character. When that came to a slow halt, he also included any other work by the original artist (Marten Toonder). Now, even further down the road, he actually collects anything ever made by anyone ever working with Marten Toonder and his collection contains more than 20.000 unique items. This example does not only show that many collectors do not aim at completing their collection (if anything they try to prevent it), but also that for most collectors there is not a clear "complete set" nor a structured strategy of what to add to their collection next.

To corroborate this intuition, I asked a representative sample of the population (5069 Liss-panel members) whether they collected something, and if so, how structured their collecting behavior was (see Experiment 7.12). Opposite to most claims in the literature (with an exception of Belk et al., who mention in their 1988 paper that collections seldom begin purposefully), most people indicated that their collecting was not structured or planned at all. Taken together, I do not believe collectors are more likely to want to complete sets or more strongly affected by this motivation when they have an almost complete set. Similarly, there are currently no reasons to expect that set-fit effects affect collectors in a different way than non-collectors. However, it should be noted that even

21 The main character in a comic book series written by Marten Toonder

though approximately 30% of the population collect something, there has been virtually no empirical research conducted on this behavior with the exception of Russel Belk who has done extensive work on the more "extreme" collectors (see for example; Belk, 1995a; Belk, 1995b; Belk et al., 1991). Since claims about how collectors approach sets are so prevalent while empirical evidence for those claims is lacking, I believe it is crucial for our understanding of collecting to empirically investigate these core dimensions of collecting.

Practical Implications

As mentioned in the introduction to this dissertation, much of the inspiration for this work came from observations on how sets are used by stores and manufacturers. As a consequence, most of the findings have direct practical implications. Some of these implications follow directly from the papers presented in this dissertation, others implications follow logically from the theory and findings presented there. Where possible, I will illustrate these implications with real-life examples of companies using the strategy, or companies not using the strategy correctly and how it can be approved. Because I am from the Netherlands, most of these examples will use Dutch companies.

First of all, the findings in presented in Chapter 2 and 3 clearly suggest that people are motivated to complete (sub)sets when they possess a large proportion of that set. Two direct implications to take form this work is that to reach the largest audience possible, companies can start with giving away freebies to reach the largest audience possible. Furthermore, because the increased motivation happens predominantly when the set is almost complete, a company could try to prevent the set from being completed to easily. This can be done by either making some parts of the set more difficult to get than others (like the "rare" cards in collectible cardgames), by providing random items from the set (like the Panini stickers that come in non-see-through bags) or a combination of both. Furthermore, since set-completion also emerges for almost complete sub-sets (see Chapter 3) dividing larger sets into smaller subsets should increase motivation even more. This last implication also brings up a vital characteristic for set-completion effects to emerge, which is that the set actually should be experienced by the consumer as a structured and coherent set. If the collection of items is not experienced as a set but merely as several individual items, none of these effects would emerge. Finally, these effects seem to be very fleeting, similar to state-curiosity (Loewenstein, 1994), I would expect these effects only to emerge when both the set and the missing items therein are very salient. As such, it is important for the company to make sure the consumer realizes what is missing, preventing the consumer to re-order the set in a way that feels simple and complete.

These implications may seem obvious and something that companies already do, but looking at how companies using sets design them, I doubt they realize these effects since they often do not seem to implement sets optimally. As already mentioned in Chapter 3, some games and gamified social media implement sets that are impossible to complete, for example the wine-app Vivino has a "newbie" achievement one can only possess as long as one hasn't drunk more than 5 wines. This will not increase motivation and may even decrease motivation turning customers away from the product. Furthermore, some companies actively hide the items missing in the set. For example Hearthstone, an online collectible card-game, does not automatically show the gamer which cards he or she is still

Figure 6.9: Illustration of showing vs. not showing missing cards for a person owning a large or small proportion. In the first situation (on top) showing the missing cards is expected to increase motivation to acquire them all. However, in the second situation, showing all missing cards may backfire by over whelming the gamer.

missing. Showing the missing cards, instead of keeping them hidden, will make the gamer more motivated to acquire them, and thus spend money. However, it is possible that when just starting out, the gamer may feel overwhelmed by all the missing cards. In such a case, the optimal strategy would be to not show the missing cards for gamers just starting out, but to automatically show them as soon as the gamer possesses about 75% or more of the cards (see Figure 6.9).

Similarly, even though it may feel intuitive to organize a larger set into several subsets with their own coherent theme, this is not what Panini did for their Dutch 2012 European Championships album in which instead of organizing the pictures by country and position the pictures were placed seemingly randomly and they did not even include full teams but only a selection from each team. Not having clearly delineated subsets probably makes it less likely for the consumer to actually perceive those individual stickers to be part of a subset, preventing set-completion from emerging.

A company famous in the Netherlands for often using sets as promotions is the large supermarket chain "Albert Heijn". They often implement sets suboptimally. I will give two examples to illustrate this point. In 2012 they released a line of 51 "mini-groceries". For every €15,- spent, the shopper would receive a miniature product that kids could use to play shop with (see Figure 6.10). Furthermore, they released a mini-shop made out of cardboard in which the kids could stall their groceries. However, no clear coherent structure was given to these items. By organizing these groceries in several subsets with

Figure 6.10: Albert Heijn mini-groceries. A total of 51 were released without coherent structure or organization.

a common theme, numbering them, and displaying dedicated spaces to each product on the mini-shop, they could have increased motivation to acquire these miniatures, and thus shop, a lot more.

Another example by the Albert Heijn comes from 2011 when customers could collect several Walt Disney dolls. For each 15 Euros spent, the customer received a sticker. These stickers could be collected and when the customer collected 6 stickers they could be exchanged for one of the 6 Walt Disney figures. As we have seen, endowing people with part of a set makes them more motivated to acquire the remainder of that set. In this case, however, the customer is not endowed with any part of a set, but has to actively and consciously decide to start acquiring them. Furthermore, because the customer can pick and choose which next doll to receive next, there is only a very small window in which the increased motivation to complete a set is present. Handing out the freebies in a random fashion increases, on average, the duration of a near-complete state and as such should increase motivation, and thus sales, even more.

The results presented in Chapter 4 and 5 clearly suggest that when designing a set that people will collect and use together, it is important to keep the structural relationships between items in that set in mind. More specifically, a manufacturer should attempt to make the structural characteristics of the set as simple as possible, while still being interesting. This can be done by creating consistent attributes shared by items in the set, such as design, shape or color, while varying other characteristics across *all* items in the set. If one needs to deviate from these rules, it would be best to do so in a symmetrical way. For example, it would be better to have two mugs of each color in a set, than one of each color, but one double.

If somehow such a clear structure cannot be achieved because of other constraints, the results of experiment 5.4 and 5.5 show that structure can also be imposed on an unstructured set by creating a narrative that focusses the consumer on the implicit structure within that set.

Finally, it is important for manufacturers to realize that there may be caveats with using these strategies, most notably consumers may become frustrated by feeling like the set drives their choices instead of themselves, and by regretting their choices in the long run.

Conclusion

The work described in this dissertation started with the question *if* and, if so, *how* the presence of a set influences decision making. In four empirical chapters combined with extra studies reported in the appendix, we found that sets do indeed affect decision making, often in ways that violate normative theories of choice. More specifically, we found that people

are motivated to complete sets and that people prefer sets in which all items are either completely similar or completely dissimilar on all salient characteristics over sets that are not. Finally, we also found that when expanding a set, people are more likely to choose items that keep the set all-similar or all-different over those that do not. These set-effects are likely the result of a general preference for sets that are low in complexity. Studies reported in the appendix support this interpretation. Even though the data described in this dissertation is consistent with classical theories on the perception and liking of patterns, this work is the first to systematically investigate how people interact with sets of goods. Besides theoretical insights and implications, this work also holds practical implications for manufacturers and designers of products and services that can be perceived as part of a set.

Appendix

Some studies mentioned in the main body of this dissertation are not described in detail, either because they are conceptual replications of the studies included in the main text, or because they represent preliminary tests of the underlying processes and consequences described in the general discussion. These studies are described in more detail in this appendix. Specifically, Experiments 7.1-7.3 are conceptual replications of the set-completion effect described in Chapter 2. Experiment 7.4 is a pre-test of, and Experiment 7.5 a conceptual replication of set-fit effects presented in Chapter 4. Experiment 7.6-7.8 are tests of the complexity-liking link as underlying process and are discussed in the general discussion. Experiment 7.9-7.11 are also discussed in the general discussion and test the effects of sets on memory (Experiment 7.9) and the relationship of collectors with sets (Experiment 7.10-7.11).

Experiment 7.1

Experiment 7.1 is a conceptual replication of Experiment 2.1 (which is described on page 16).

Method

During a weekly testing session, 96 students (75 females, M_{age}= 19.7, SD = 2.7) from Tilburg University were randomly assigned to one of only four conditions with varying degrees of set-completion. They were all asked to imagine that they owned four books of an author they liked. Depending on the condition, they subsequently read that this author had written in total five, six, seven, or eight books. After reading this scenario, participants indicated how much they wanted the fifth book (*1 = not at all, 7 = very much*), and how much they were willing to pay for it in Euros. We expected a higher wanting and willingness to pay when the 5th book was set-completing (i.e., when the author had only written five books in total). We had no explicit expectations of differences between the other conditions, since we do not know exactly when the reference point switches from zero to the full set. Even though we do not have any strong predictions about differences between

the non-completing conditions, they should either have similar means, or slightly increase with an increase in proportion possessed.

Results and Discussion

As expected, participants wanted a book more and were willing to pay more for it when it completes a set as compared to when it does not. Differences between conditions were found on both wanting, F (3, 125) = 5.87, p = .001, η^2 = .12 (one participant failed to complete the wanting-measure), and willingness to pay, F(3, 126) = 7.39, p < .001, η^2 = .15. For means and standard errors, see Figure 7.1.

These results confirm our expectations that consumers want, and are willing to pay more for a product when it completes a set as compared to the same product when it does not complete a set. Interestingly, consumers in the 4/6 condition also seemed to want the prod-

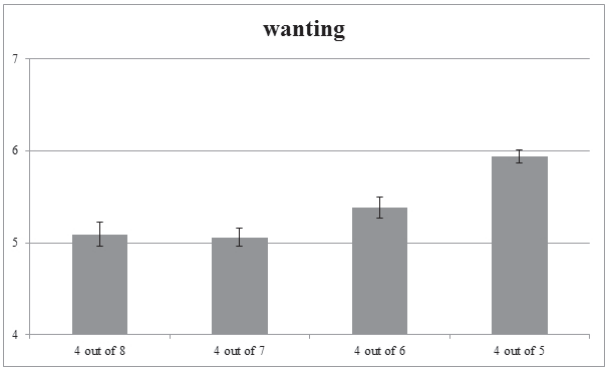


Figure 7.1a: Conceptual Replication Experiment 2.1, wanting for the 5th book.
Note: Wanting was reported on a 7-point scale anchored on the left with "not at all" and on the right with "very much". Error bars represent standard-error of the means.

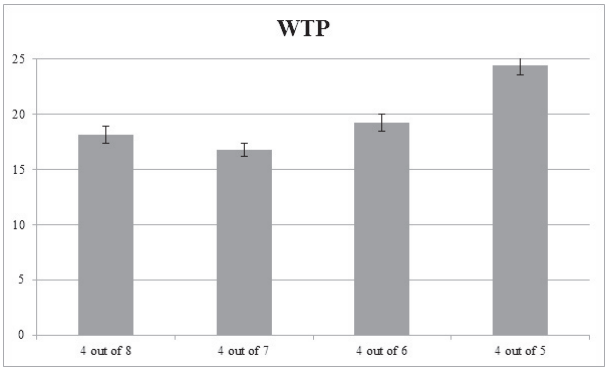


Figure 7.1b: Conceptual replication Experiment 2.1, Willingness to Pay for the 5th book.
Note: Willingness to Pay in dollars. Error bars represent the standard-error of the means.

uct slightly more as compared to consumers in the 4/7 condition (marginally significant for wanting, ns. for willingness to pay). It could be that for several participants, owning 4/6 already shifts the reference-point to the full set.

Experiment 7.2

In this second conceptual replication of Experiment 2.1 (see page 16) we also added an extra condition to this experiment in order to provide further insight into why consumers want to complete the set. In this condition consumers already owned a substitute product with similar functional characteristics as the missing item, but not the same overall fit with the set. This way we were able to distinguish between wanting the product for its individual and functional characteristics versus wanting the product for its "completing" aspect. Because we hypothesize that the completion-premium is the result of feelings of deprivation resulting from the experience of incompleteness and not an increased liking of the product for its individual characteristics, we still expected a set-completion effect for this condition.

Method

Participants were 105 students from Tilburg University (78 females, M_{age}= 21.1, SD = 3.2) who were randomly assigned to one of only four conditions (almost complete set, complete set with substitute, and two control conditions). In all conditions, participants read a scenario in which they were asked to imagine liking a band. In the two control conditions they either read that they owned two albums out of eight made by this artist, or seven albums out of 18. The other two conditions were the completion conditions. In one of these conditions participants read they owned seven out of eight albums made. In the other condition they also read that they owned seven albums out of eight, but crucially, they also owned a burned copy of the eighth album. In other words, they owned the music of all eight albums, but did not own the *original* eighth album. Thus in two conditions (owning seven out of eight, owning seven out of eight and having the eighth as a burned copy) set-completion is possible, whereas the other two conditions (owning two out of eight, owning seven out of 18) are not close to completion. Subsequently, participants indicated their willingness to pay for the eighth (third in control-condition 1) album and their wanting of the album on a 9-point Likert scale (*1 = not at all, 9 = very much*). Finally, participants indicated why they wanted the album by ticking a 13 cm line anchored on the left with "*reducing feelings of frustration for not having it*" and on the right by "*expecting pleasure from owning it*".

Results

An ANOVA over the four conditions (control two out of eight, control seven out of 18, set-completion seven out of eight, and set-completion burned copy) on willingness to pay reveals significant differences between conditions, $F(3,101) = 7.01, p < .001, \eta^2 = .17$, such that willingness to pay for the eighth album is higher for participants in the completion conditions compared to participants in the two control conditions. For all means and contrasts, see table 7.1. An ANOVA on wanting shows significant differences between conditions in the expected direction, $F(3,101) = 8.98, p < .001, \eta^2 = .21$, such that participants in the completion conditions indicate a higher wanting than participants in the control conditions (see table 7.1).

Table 7.1: Results Experiment 7.2, conceptual replication Experiment 2.1; Willingness to Pay, and Wanting for the 8th album, and reason why participants wanted the 8th album.

	Control (2 out of 8)	Control (7 out of 18)	Completion (7 out of 8)	Completion (burned copy)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Willingness to Pay	€19.83 (7.30)[a]	€20.46 (11.47)[a]	€37.33 (21.92)[b]	€34.57 (25.37)[b]
Wanting	6.69 (0.74)[a]	6.35 (1.20)[a]	7.60 (0.77)[b]	7.35 (1.27)[b]
Reason for Wanting	4.44 (3.45)[a]	4.22 (2.97)[a]	8.72 (3.60)[b]	9.73 (3.64)[b]

Note: Wanting and deprivation were reported by ticking an 11 cm scale. Conditions with different subscripts differ with $p < .01$ from each other. Reason for wanting was ticked on 13-cm line with higher scores agreeing more with "*reducing feelings of frustration for not having it*" and lower values agreeing more with "*expecting pleasure from owning it*".

Finally, an ANOVA on the reason for wanting the items (to reduce the feelings of deprivation for not having it vs. the expected pleasure of having it) recoded so that higher numbers reflect higher feelings of deprivation, revealed that participants in the two completion conditions were significantly more likely to indicate wanting the album to reduce the feelings of deprivation associated with not having it compared to participants in the other 2 conditions, $F(3, 101) = 17.78, p < .001, \eta^2 = .35$, (see Table 7.1).

Discussion

This study replicates the set-completion effect, finding that a completion-premium still emerges when consumers already own an item that has a similar functioning as the missing item (the burned copy). This suggests that the completion-premium is not a result of a higher appreciation of a products individual characteristics, but more about resolving the feelings of something missing.

Experiment 7.3

We initially ran a more extended form of Experiment 2.2 (see page 18), however, because the results of this design are quite complicated to describe parsimoniously, we decided to rerun a simpler version for the main body of the article and move this one to the appendix.

529 participants (250 females, M_{age} = 31.5, SD = 11.8) approached via Amazon's mTurk were assigned to one of 15 conditions (we aimed for 750 participants, but only 529 participated during the time the experiment was online). In these conditions participants saw a set of three [five, seven, nine, or eleven] mugs of which they owned two [four, six, eight, or ten]. This is technically a 5´5 design, but since it is impossible to, for example, own six out of three mugs, this design results into 15 conditions in total. We expected that, as soon as one would be close to completing a set (i.e., having four out of five, of eight out of nine mugs), people would want a mug they did not have more and would be willing to pay more for it. Furthermore, we also asked participants whether they felt the situation was more like "having Y of a group of X mugs" or more like missing "(X-Y) of a group of X mugs". So in the 4/11 condition, participants chose between "having 4 out of 11 mugs" or "missing 7 out of 11".

Results

For all means, see Figure 7.2a and 7.2b. The multivariate analysis with the number of mugs missing and the number of mugs possessed as between-subject factors reveals a significant main effect of missing on both wanting and willingness to pay ($F_{want}(4, 512) = 4.42$, $p < .01$, $\eta^2 = .033$; $F_{WTP}(4,512) = 7.30$, $p < .001$, $\eta^2 = .054$. A further marginal main effect was found for the number of mugs possessed on willingness to pay; ($F(4, 512)=2.22$, $p = .07$, $\eta^2 = .017$, but not on wanting; $F(4, 512) = 1.69$, $p = .15$, $\eta^2 = .013$. No interaction effect was revealed (p's > .23). Visual inspection of the plots (see Figure 7.2) suggests that near set-completion, people want a new mug more and are willing to pay more for it. All conditions, except for the ones in which one possesses six out of "x" show this pattern. Since in total there are 15 conditions, we think the effect being absent in the "six out of x" is probably a Type-II error (especially since had less participants than aimed for).

To investigate the differences more thoroughly, we ran contrast analyses between the conditions in which the number of mugs owned was the same but the total amount varied (for example 6 out of 7, 6 out of 9, and 6 out of 11). This showed differences between conditions where one was missing one item, compared to conditions in which one missed more than one on both wanting and willingness to pay for all conditions except for the previously mentioned "possessing 6 out of x". Contrast analyses comparing missing 1 to the other conditions revealed the following effects on wanting: 2/3 vs others, $p = .019$, 4/5 vs. others, $p < .001$, 6/7 vs. others, $p = .91$, 8/9 vs. others, $p = .072$. And for willingness to

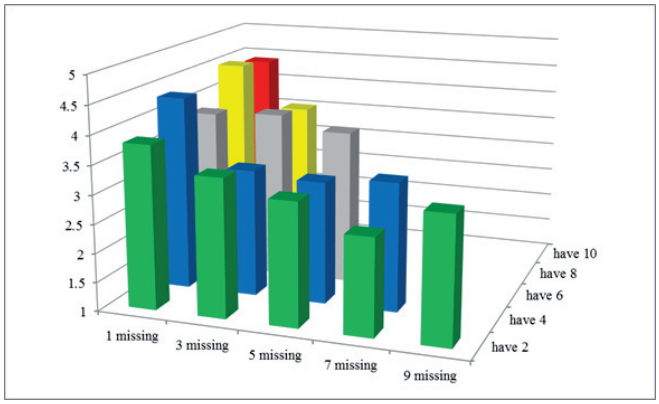


Figure 7.2a: Wanting for the (missing) blue mug on a 7-point scale.

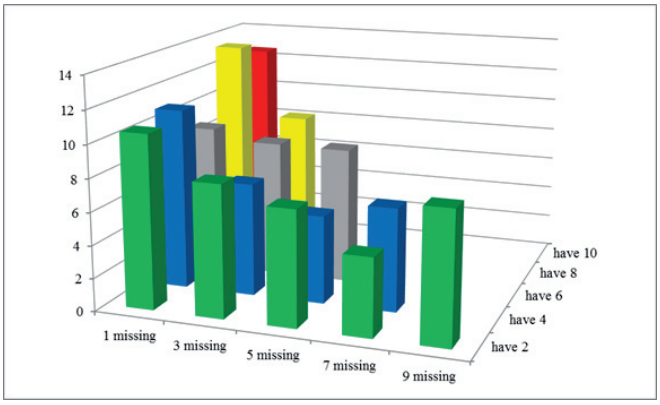


Figure 7.2b: Willingness to Pay for the (missing) blue mug in dollars.

pay: 2/3 vs. others, $p = .001$, 4/5 vs. others, $p < .001$, 6/7 vs. others, $p = .62$, 8/9 vs. others, $p = .013$.

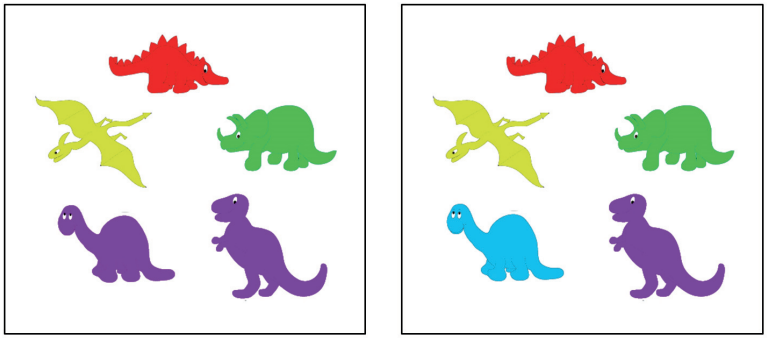
Finally, we looked at the correlations between "being in a situation of near set-completion (missing one) or not", how participants described the situation (using a reference-point of nothing vs. the full set), wanting, and willingness to pay. All these measures correlated with each other in the predicted directions (all p's < .001). Having an almost complete set correlates with a higher willingness to pay ($r = .29$), a higher wanting ($r = .23$) and describing the situation with the full set as reference-point ($r = .32$). Describing the full set as the reference-point also strongly correlated with wanting ($r = .42$) and willingness to pay ($r = .37$).

Experiment 7.4

This pretest (mentioned on page 43) was conducted to test whether sets whose members are completely identical or completely distinct on each salient feature are rated as better than sets for which this is not the case (i.e., some members are the same, but some differ, on a salient feature).

Method

Participants ($N = 81$, 72 females, $M_{age} = 19.28$) were students from Tilburg University who participated in return for course credit. Participants were shown nine pictures depicting sets of five cartoon dinosaurs. They were instructed that "some sets of things feel better than other sets of things. Please indicate how good the set feels and (if applicable) how the set can be improved". The instructions were intentionally left vague because we wanted to test whether people have a notion of what a "good" set is without defining what good was and thereby creating demand effects. For each set we created a "good" version for which all dinosaurs were the same—or all differed—on the attributes of color and type (i.e., shape), and a "bad" version for which this was not the case (i.e., some dinosaurs were of the same color/type, while others differed; see Figure 7.3). Participants were randomly shown either the good or the bad version of each of the nine sets. For each set, they indicated how good it felt on a 7-point scale (1 = not at all good, 7 = very good). They were also asked if, and if so how, the set could be improved.



Bad set Good set

Figure 7.3. "Good" and "bad" versions of Set 1 used in Experiment 7.4. On the left is the bad set: All dinosaurs are of different types, but some share the same color whereas others do not. On the right is the good set: All dinosaurs completely differ both on type and color.

Results

We excluded four participants from analysis. One participant failed to mark down the correct trials, the other three gave responses to the questions on how to improve the set that were classified as "weird" for all 9 trials (for example, one participant suggested for every set giving the dinosaurs camouflage colors, the other suggested making them look more angry so they appeared more dangerous). Including these participants from the analysis did not lead to any meaningful differences in the ratings of the goodness of the sets (all test statistics are still $p < .001$).

Rated goodness of set

Participants indeed judged good sets as better than the bad sets: Wilk's $\Lambda = .23$, $F(9,67) = 25.40$, $p < .001$, $\eta^2 = .77$. For univariate results, see Table 7.2.

Table 7.2. Means and standard deviations of rated goodness for the good and bad versions of each of the 9 sets in Experiment 7.4.

	Good set		Bad set			
	M	*(SD)*	*M*	*(SD)*	*p*	η^2
Set 1	5.26	(1.61)	3.31	(1.34)	< .001	0.31
Set 2	5.74	(1.74)	4.39	(1.70)	< .001	0.14
Set 3	5.77	(1.37)	4.42	(1.31)	< .001	0.21
Set 4	5.97	(1.50)	3.62	(1.35)	< .001	0.41
Set 5	6.38	(1.14)	3.82	(1.54)	< .001	0.48
Set 6	6.58	(0.72)	3.67	(1.53)	< .001	0.60
Set 7	6.26	(0.94)	4.08	(1.57)	< .001	0.42
Set 8	5.59	(1.62)	3.95	(1.27)	< .001	0.25
Set 9	6.58	(0.76)	3.38	(1.51)	< .001	0.64

Suggested Improvements

When coding the responses to the open-ended question on how to improve the set, we ran into an unexpected regularity. Although we intentionally varied the dinosaurs on two aspects (color and type) we unintentionally also varied them on two other aspects. One of the dinosaurs was a pterodactyl (meaning it could fly), and one of the dinosaurs was facing the right where others were all facing the left. Because changing both these aspects was suggested as an improvement several times, we coded the responses in the following categories: 1) no improvement suggested; 2) improvements following the our proposed rules (in the example bad set this would be changing the color of one of the purple dinosaurs to a color not used yet or removing one of them); 3) other improvements following rules (in

good set example flipping the red dinosaur over so all are facing the same direction); 4) expanding the set; and 5) other (e.g., spacing the dinosaurs more evenly).

For bad sets, almost 90% of the participants suggested that the set could be improved in a way that follows the proposed rules for good sets (see Figure 7.4). This is further evidence that there indeed seem to be general principles that determine whether a set feels good or not. These rules were also apparent for good sets. Here a lot of people suggested (more than 20%) that the set should be expanded by adding another dinosaur according to the hypothesized principles (for example for Figure 7.3; adding an orange dinosaur of a different type). Since the dinosaurs unintentionally differed on other aspects as well, we also looked at differences between the conditions for suggestions relating to these differences. Interestingly, for bad sets only 2.9% gave suggestions pertaining to these differences (e.g., in Figure 7.3, removing the yellow flying dinosaur since it is the only one that can fly) whereas for good sets 14.7% gave such suggestions, suggesting that when differences in type or color were present, these were more salient than the other aspects (left/right facing and flying/non-flying).

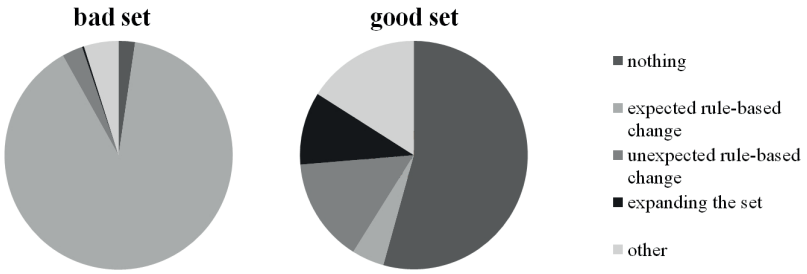


Figure 7.4. Improvements (%) suggested for the good and bad sets in Experiment 7.4.

Experiment 7.5

Experiment 7.5 was an additional Experiment conducted to test set-fit effects and is discussed on page 49 of Chapter 4. An experiment using a similar design but having participants keep the products they chose (a set of beers) is described as Experiment 5.3 reported on page6177.

Method

Participants (N = 204, 50 females, M_{age}= 27.41) recruited on Amazon.com's Mechanical Turk, were randomly assigned to one of four conditions in which they made a choice between two sets of mugs (see Figure 4.5 on page 50). In each condition, the only difference

between the two choice options (the sets between which the participants chose) was a green vs. an orange mug. In the two all-similar conditions the green and orange mugs were accompanied by 3 more green mugs in the all-similar-green condition, and by 3 more orange mugs in the all-different-orange condition. In the two all-different conditions, we added a blue, pink and green mug to both options in one condition, and a blue, pink and orange one in the other condition. Based on set-fit, we expected that adding three green would increase the green mug's choice share. Adding three orange mugs should increase choice for the orange mug. In the all-different conditions set-fit predicts that adding the single green mug (plus the pink and blue one) decreases the green mug's choice share, and adding the single orange one (plus the pink and blue) decreases the orange mug's choice share.

Results

In the all-similar orange condition, 62.3% (33/53) of the participants chose the set including the orange mug. In the all-similar green condition, only 39.6% (19/48) chose the set including the orange mug; $\chi^2(1, n = 101) = 5.19, p = .03, \phi = .23$. In the all-different condition in which a pink, blue and green mug was added, 94.1% (48/51) chose the set including the orange mug. When a pink, blue and orange mug were added, only 15.7% (8/51) chose the set including the orange mug; $\chi^2(1, n = 102) = 63.4, p < .001, \phi = .79$.

Experiment 7.6

Experiment 7.6 consists of two different experiments in which participants rated the goodness of a large set of patterns. These experiments were mainly conducted to gain more insight into set-effects by investigating general tendencies occurring in the judgment of patterns and are discussed on page 74 in the general discussion. Because it is impossible to change the complexity of a pattern (or set) without also changing the contents of the pattern (for example to make a pattern of four blue squares more complex, another color or shape needs to be added) we can of course not be certain that differences in judgments are a consequence of the complexity of the pattern rather than the change in the individually changed shape. However, if we consistently find the same effects, regardless of whether we for example change a red dot into blue, green or purple, the complexity-liking link is the most parsimonious explanation for these results. I have repeated these experiments with different participants but results were very similar so I am only reporting one set of results for each design.

Method

Participants in Experiment 7.6a where 117 students (101 female, M_{age} = 20.1, SD = 3.9) from Tilburg University. Participants in Experiment 7.6b were also students from Tilburg University (N = 246, 164 females, M_{age} = 19.9, SD = 2.3, Experiment 6b was conducted 2 years after experiment 6a so it is not likely that more than a few people participated in both studies). In both experiments, participants were confronted with a large amount patterns, presented one at a time on their computer screen. Participants were instructed to indicate for each pattern how "good" this pattern felt on a 100-point slider scale anchored on the left with "really bad" and right with "really good".

Experiments 7.6a consisted of "good" patterns that were either all-similar or all-different on shape while also being all-similar or all-different on color, and "bad" patterns that deviated from these rules by 1 or more part of the pattern. Experiment 7.6b was similar, however, in this experiment we systematically introduced these deviations using a good-set as base (see Figure 6.2 on page 73). Furthermore, we only used patterns that were all-similar on shape as base-patterns, with some also being all-similar on color while others were all-different on color.

All data were analyzed by calculating the average rating of each pattern across all participants. Then, for each pattern, a variable was created indicating whether the basis of the set was all-similar [0] or all-different [1] for 1) color, and 2) shape. Furthermore, a variable was created indicating whether the pattern deviated [1] or not [0] from this all-similar or all-different rule on 3) color or 4) shape. Finally, a variable was created that represented whether, above and beyond the deviation under 3) and 4), more deviations were present for 5) color or 6) shape. For clarification, see Figure 7.5.

We then conducted a hierarchical linear regression with basis of the set (IV 1 and 2) in step 1, deviations from structure (IV 3 and 4) in step 2, and finally additional deviations (IV 5 and 6) in step 3. The reason we included additional deviations as a separate variable is that it allows us to see whether the negative effect of deviations is linear (the B-weight of IV 3 and IV 5 are the same in the regression) or in- or de-creasing.

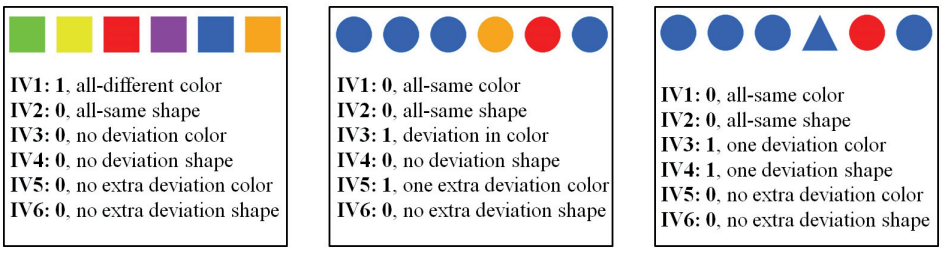


Figure 7.5: Examples of the coding of the independent variables.

Results

In Experiment 7.6a, a hierarchical multiple regression shows that the model does not significantly increase explained variance adding the dummy-variables for whether the set is all-same or all-different on shape and color (IV1 & IV2, $R^2 = .31$, $F = 2.3$ $p = .11$). The model is improved significantly in step 2 when adding whether a violation was detected in shape or color (IV3 & IV4, $R^2 = .76$, $F = 32.7$, $p < .001$), and in step 3 where the amount of violations is included ($R^2 = .81$, $F = 28.1$, $p < .001$).[22]

Looking at the B-weights for each variable we can see the following; sets that are all different rather than all-similar on color are liked slight less; $B = -13.4$, $p < .001$. The same

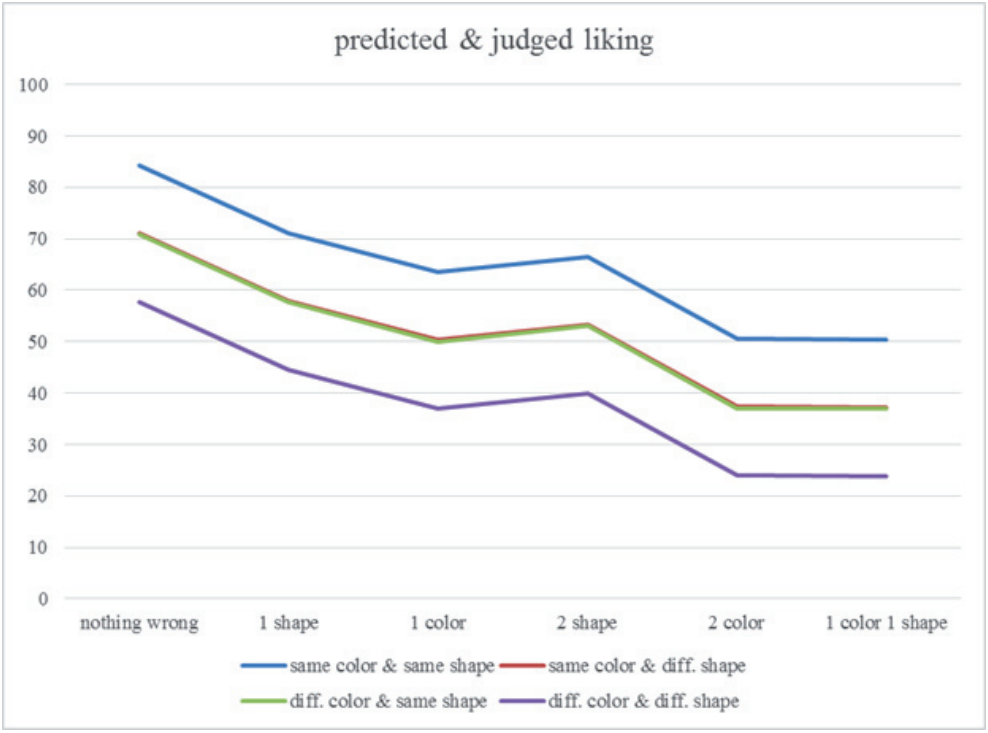


Figure 7.6: On the y-axis the rated goodness of the pattern, on the x-axis the different violations within the patterns. The color indicates whether the base-set is all-similar or all-different on shape and color. Lines represent predicted averages based on the regression coefficients. Dots are the average ratings given by the participants. Patterns with more than 1 deviation on shape *and* color have been grouped together with patterns with only 1 deviation on both shape and color.

22 *F*- and *p*-values represent a test of the model against the previous step, so the *F*- and *p*-value for the model including IV5 and IV6 represent the test of the full model (IV1 – IV6) against the model specified in step 2 (IV1 – IV4).

holds for sets that are all-different shapes rather than all-similar; B = -13.2, $p < .001$. If a set contains a deviation in color, liking is further reduced; B = -20.8, $p < .001$ and the same holds for deviations on shape; B = -13.2, $p < .001$. Finally, additional deviations decrease liking even more, but not as strong as the first deviation; for color; - 13.0, $p = .01$ and not significantly for shape B = -4.6, $p = .24$. The predicted scores (based on the regression coefficients) are plotted as lines, and the average judgments are plotted as dots in Figure 7.6.

In Experiment 7.6b, the three step regression shows that the model does not significantly increase explained variance adding the dummy-variables detailing whether the set is all-same or all-different on shape and color ($R^2 = .29$, $F = 3.1$ $p = .09$), but the model is improved significantly in step 2 when adding whether a violation was detected in shape or color ($R^2 = .78$, $F = 16.9$, $p < .001$) and in step 3 where the amount of violations is included ($R^2 = .85$, $F = 15.2$, $p < .001$).

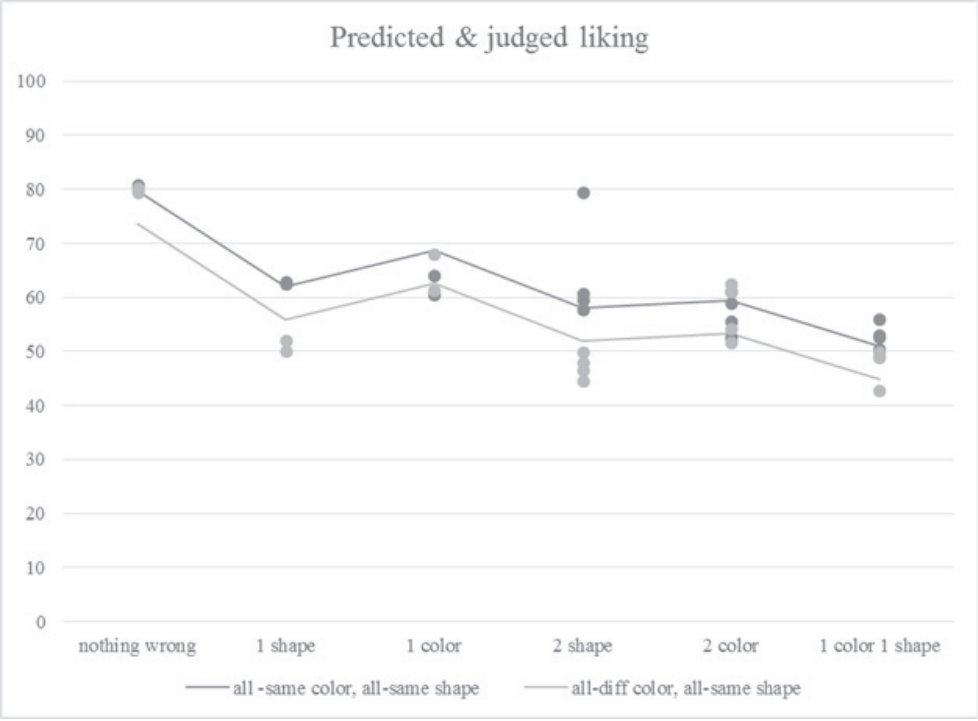


Figure 7.7: On the y-axis the rated goodness of the pattern, on the x-axis the different violations within the patterns. The color indicates whether the base-set is all-similar or all-different on color (all patterns were all-similar on shape in the base-set). Lines represent predicted averages based on the regression coefficients. Dots are the average ratings given by the participants. Patterns with more than 1 deviation on shape *and* color have been grouped together with patterns with only 1 deviation on both shape and color.

Looking at the b-weights for each variable we can see that sets that are all different rather than all-similar on color are liked slightly less; B = -6.1, p = .01. If a set contains a deviation in color, liking is further reduced; B = -11.0, p = .001 and the same holds for deviations on shape; B = -17.7, p < .001. Finally, additional deviations decrease liking even more, but not as strong as the first deviation; for color; B = - 9.3, p = .01 and for shape, B = -4.0, p = .25

The predicted scores (based on the regression coefficients) are plotted as lines, and the average judgments are plotted as dots in Figure 7.7.

Experiment 7.7

Experiment 7.7 was conducted to test whether violations of expectations (for example; Heine, Proulx, & Vohs, 2006) could explain set-fit and set-completion effects. This alternative explanation is discussed in the general discussion on page 75. In this experiment, participants judged multiple patterns for which the individual parts in this pattern appeared sequentially. Sometimes these patterns were all-same or all-different but other times the sets were more complex deviating from the all-same or all-different rule. Furthermore, we varied whether this deviation happened at the beginning of the set or at the end. If sets that are more complex are disliked particularly because they imply some inherent structure that is subsequently violated, we should find a greater dislike of sets showing a deviation at the end as compared to sets showing this deviation in the beginning.

Method

Participants were 116 students from Fontys University of applied science (70 females, M_{age} = 20.0, SD = 2.1) who participated in a series of studies in return for €5.-. Participants were told that in this experiment they would be presented with several different patterns and that we would like them to indicate how good the pattern feels.

Participants were then presented with 60 patterns for which the order was determined randomly. These 60 patterns were based on 20 base-patterns which were either all-similar or all-different, and 20 versions of these patterns in which one shape deviated from the all-similar or all-different rule. Furthermore, this deviation either happened in the beginning of the pattern or at the end. As such, we have 20 patterns that were rated three times by each participant. Once in its good form, and twice the rating of the pattern in a (identical) bad form with the deviation happening at the start or the end of the pattern.[23]

---

23 We counterbalanced direction of appearance between the sets with 3*10 appearing from the left and 3*10 appearing from the right, to exclude the possibility that people having a general preference for one or the other confounding our results.

Results

Figure 7.8 presents the results for all twenty trials. The first two bars represent the ratings of the pattern with the deviation happening in the beginning (left-bar) or at the end (middle-bar). The right bar represents the average judgment for the pattern when it had good form (either all-similar or all-different on color and shape). As can be seen, besides replicating a general higher liking for sets that are all-similar or all-different, no other differences were detected.

Figure 7.8: Ratings of sets for which the contents appeared sequentially with 0 representing *"really dislike"* and 100 *"really like"*. On the left the patterns for which the deviation happened in the beginning of the pattern, in the middle those for which it appeared at the end. The right bar reflects judgments of the same set without any deviation.

Experiment 7.8

In Experiment 7.8 (discussed on page 75), we tested whether effects of complexity are attenuated by deliberation. More specifically, we had participants choose between two sets of pens, of which one had high complexity while the other was low in complexity, either without any other instructions, or by forcing them to deliberate about the choice.

Method

Participants were 304 workers on Amazon.com's Mechanical Turk (126 females, M_{age} = 30.1, SD = 10.0) who were randomly assigned to one of two conditions. In both conditions, participants made a choice between two sets of pens.[24] One of these sets was a set of x pens with good set-fit (all different) whereas the other set consisted of 2 blue, 3 black, 1 red, and 1 green pen (see Figure 7.9).

24 We also measured whether participants would switch their choice and for that reason included a third condition in which participants first chose, then deliberated, and then had the opportunity to switch. However, in none of the three conditions more than 2 participants switched and for that reason those results will not be discussed further. Since for the first choice, participants in the third condition saw identical instructions as those in the control condition, those two conditions were combined. Results without combining the conditions are similar.

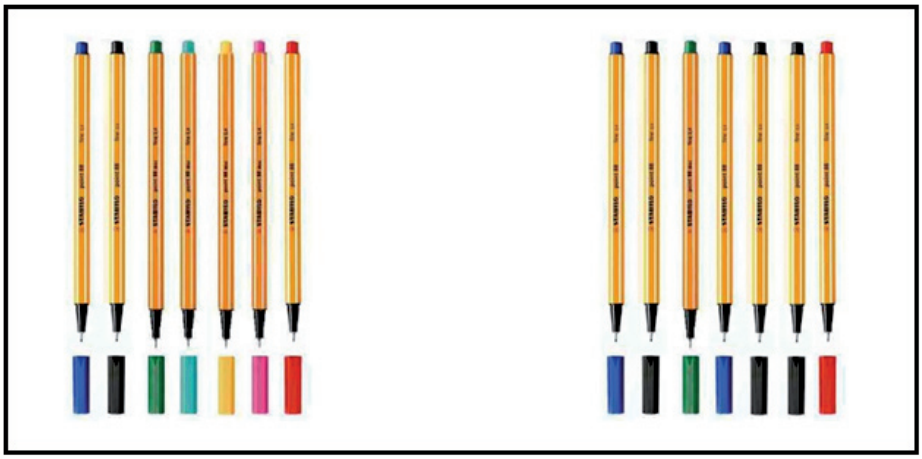


Figure 7.9: The two sets of pens participants in Experiment 7.8 chose between.

In the control condition, participants just made a choice between these two sets. In the deliberation condition, we asked participants to describe three situations in which they would use the first set of pens (type was counterbalanced) and three situations in which they would use the second set of pens. After answering these questions, participants in this condition also made a choice between the two sets. It was expected that participants would be less likely to choose the set of all-different pens after deliberating, because deliberation would give the affective response less relative weight and focus participants more on the functional characteristics of the pens.

Results & Discussion

As expected, participants in the deliberation-condition were more likely to choose the (bad-fitting) set of 3 black, 2 blue, 1 red, and 1 green pen (60.8%, 115/189) as compared to participants in the control-condition who did not deliberate about the functional characteristics of the pens (54.8%, 63/115), $\chi(1, N = 304) = 7.06, p = .01, \phi = .15$). This suggests that the preference for good-fitting sets is a very intuitive automatic process.

Experiment 7.9

Experiment 7.9 was conducted to test whether memory is better for items that complete a set. Reasoning behind the expectation that set-completing items are remembered better is discussed on page 80 in the general discussion.

Method

Students from Tilburg University (N = 158, 101 females, M_{age} = 21.5, SD = 2.9) were assigned to either a control- or a completion-condition. In both conditions participants were told that they would be writing a short Portuguese poem. On the left on the screen, 20 pictures were presented. In the control-condition all these pictures were portrayed semi-transparent, in the completion-condition only 10 of the 20 pictures were presented semi-transparent while the other 10 were presented non-transparent (see Figure 7.10).

Participants were presented with a sentence in Portuguese and asked to retype the sentence in the answer-box. When they typed in the sentence correctly, one of the semi-transparent pictures was replaced by a non-transparent version. In total, nine sentences were presented that had to be entered. As a result, participants in the control condition finished this experiment with 9 out of 20 pictures non-transparent whereas those in the completion-condition ended with 19 out of 20, see Figure 7.11.

Figure 7.10: Start-screen for participants in the control condition (left) or completion condition (right).

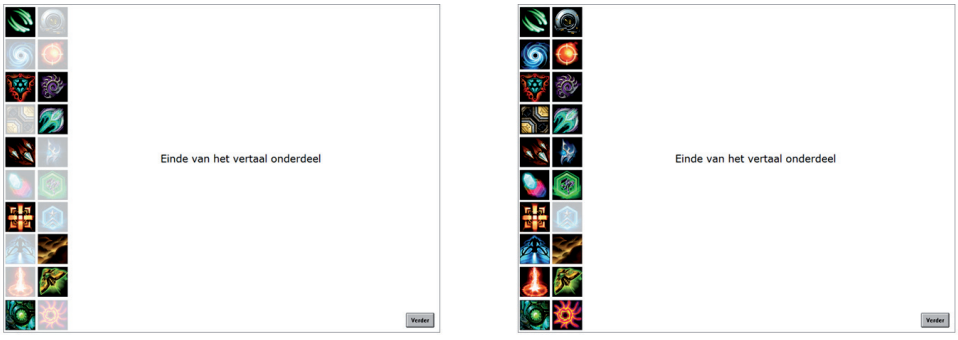


Figure 7.11: End-screen for participants in the control condition (left) or completion condition (right).

After correctly entering the nine sentences, participants continued to a screen in which the pictures were again presented on the left, only this time the 14th picture was missing. In both conditions this was a semi-transparent picture, but crucially, it was the only "missing" picture from the set in the completion-condition. Participants were also presented 4 pictures, and were asked to indicate which of the four pictures was the one in the empty spot (see Figure 7.12). It was expected that participants in the completion-condition would have better memory for the set-completing item and thus would be more likely to give the correct answer.

Figure 7.12: The memory-check.

Results

As expected, more participants in the completion-condition (52/80, 65%) correctly identified picture number 1 to be the missing picture as compared to the control condition (25/78, 32.1%, χ^2 (1, N = 158) = 17.16, p < .001).[25]

25 As can be seen, the effect-size found in this study is really big. Before running this study, I was hesitant about whether we would find the effects of an almost complete set on memory because, even though

Experiment 7.10

`Experiment 7.10 consisted of structured interviews with 61 Visitors (45 males, 16 females, M_{age} = 46.0, SD = 16.8), of a collectors-fair in Utrecht the Netherlands on November 20th in 2011 and was conducted as part of a project with Siegward Lindenberg on collectors (see, Evers & Lindenberg, 2014). Parts of those results give insight into the relation between set-effects and collectors. This is discussed on page 86 in the general discussion. Visitors of

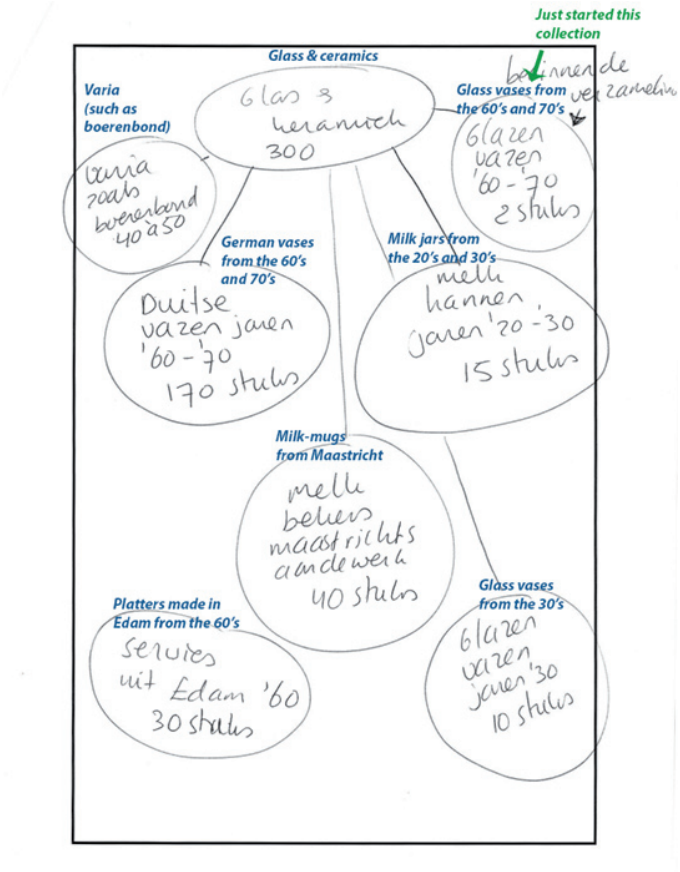


Figure 7.13: Example of one of the collections and the subsets therein.
Note: In blue the English translations of the sub-sets, in green the translated remark made by the collector that she only just expanded her collection to include that subset.

theoretically we would predict an effect, I expected the effect-size to be quite small. Since effect-sizes are quite variable (see for example Lakens & Evers, 2014), it is possible that this effect-size is an overestimation of the true effect-size.

the collectors-fair were approached by one of three investigators[26] and asked if they were willing to be interviewed about their collecting-hobby for about 10 minutes. No monetary compensation was provided, but they were offered a small candy bar. The questionnaire consisted of 6 pages of questions about their collecting behavior, detailing what they collected, how they came to collecting that, collecting-style, the organization of the collection and finally collecting-motivations. Participants were asked how they came to collect what they collected and were further asked to draw the organization of their collection indicating sub-sets when applicable, for an example, see Figure 7.13. Virtually all collectors indicated having several related, but only partially overlapping, sub-collections. This suggests that collectors actually prevent completing their collection by continually expanding it when they reach near-completion. For information about the other measures, see Evers & Lindenberg (2014).

Experiment 7.11

Experiment 7.11 was a large questionnaire on collecting behavior collected using a representative sample of the Dutch population which was conducted for a project on collecting together with Siagward Lindenberg (see; Evers & Lindenberg, 2014). A selection of the results is reproduced below since those findings give insight into the relation between set-effects and collectors. This is discussed on page 86 in the general discussion. The panel data were acquired through the Liss-panel by CentERData (Tilburg University, The Netherlands). In total, 5069 panel-members (2726 females, Mage = 50.93, SD = 17.6) filled in a questionnaire in which they were asked whether they collected something or not. After this question, a follow-up questionnaire was distributed under those people that indicated collecting something. Collectors were asked what they collected, how their collection started, collecting-style and collecting motivations. Furthermore, we had access to all other questionnaires completed by these panel-members including detailed demographic information as well as personality measures such as the big-five.

We found that a substantial proportion of the panel-members indicated collecting something (33.9%). We also asked the panel-members whether they would call themselves a collector (on a 5-point scale anchored on the left with *not at all* and on the right with *very much*). Even though a large proportion indicated collecting something, only a small proportion would call themselves a collector (of the 33.9%, 5.7% indicates a *4*, 2,4% a *5* on the 5-point scale). For the effects of personality on collecting and information about different motives for collecting, see Evers & Lindenberg, 2014 or Evers, Lindenberg & Linskens, 2014.

---

26 Lonne Landsmeer, Lieke van Dorst, & Ellen Evers

We also looked at how structured the collectors indicated their collecting to be with the six questions presented in table 7.3 (all answered on a 7-point scale).[27]

We calculated a structure score by averaging the six items. If collectors believed their collecting style to be neither structured nor unstructured, their score should be around the midpoint of the scale (4). The distribution of scores on the structure-scale are plotted in Figure 7.14 As can be seen, most of those who indicate collecting something do so in a very unstructured way with 83% being on or below the midpoint of the scale. Even when looking only at those who indicated seeing themselves as a true collector, collecting becomes more structured on average, but still 35% of the extreme collectors indicate their collecting to be neither structured nor unstructured. These data suggest that instead of

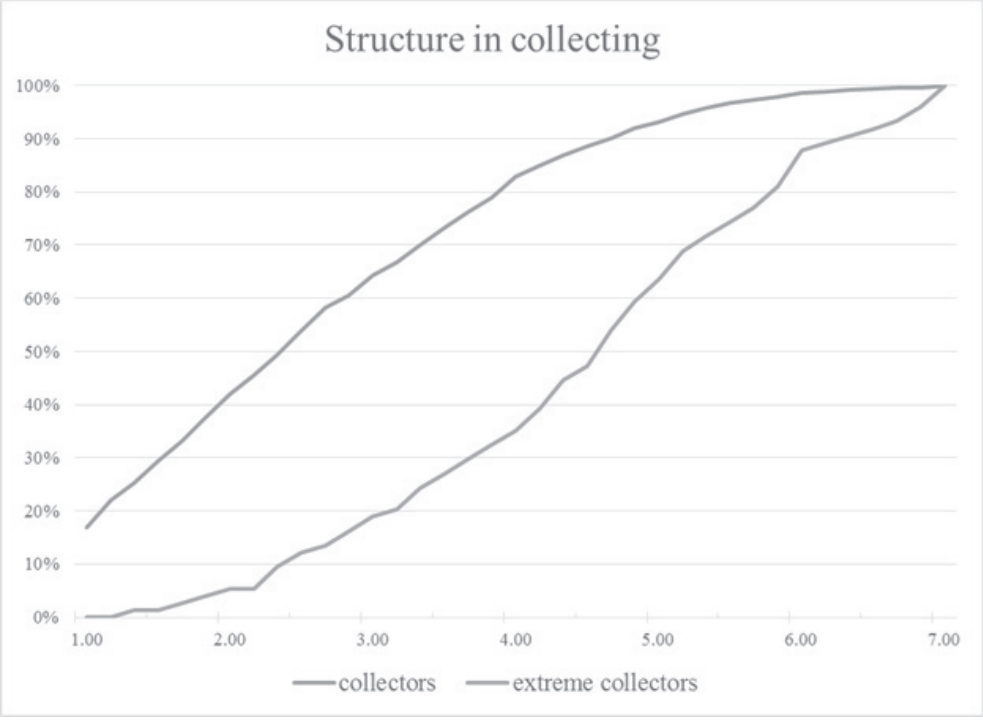


Figure 7.14: Cumulative proportion of collectors on the structure score with lower values meaning less structure. The orange line represents all collectors who scored a 1-4 on the question "*How much would you consider yourself a true collector?*" while the blue line represents those scoring 5 on this question.

27 The original questionnaire contained 12 questions, the six mentioned above and the reverse-stated version of all six. The scale appeared to have higher internal consistency when only regular-coded questions were used. A factor analysis also indicates two different factors, one that loads on all items and is consistent with structure, one that only loads on the reversed questions suggesting that some participants were just more likely to indicate higher (lower) numbers in general. Analyzing all 12 questions does not change the conclusions in any meaningful way.

seeing collecting behavior as one specific kind of behavior, there may be different types of collectors who collect in completely different ways. This is consistent with ideas expressed by Danet and Katriel (1989, see also Belk, 1995b) on type A and type B collectors, but this idea has been neglected in theoretical models on collecting.

Table 7.3: Statements about the structure of collecting, responses were made on a 7-point scale anchored on the left with "completely disagree" and the right with "completely agree".

1. In general I know exactly what I need for my collection.
2. My collection consists of clearly demarcated sub-collections.
3. I have a clearly worked-out plan in my head as to all the items I wish to add to my collection in the future.
4. My collection is highly structured.
5. I grow my collection very purposefully.
6. My collection is structured according to clearly demarcated categories.

References

Adler, S. A., & Orprecio, J. (2006). The eyes have it: Visual pop-out in infants and adults. *Developmental Science, 9*(2), 189-206. doi: 10.1111/j.1467-7687.2006.00479.x

Algemeen Dagblad (2009). Bloedneus gemept om voetbaalplaatjes [Being punched a bloody-nose for soccer-stickers]. Retreived from http://www.ad.nl/ad/nl/1038/Rotterdam/article/detail/2015130/2009/02/11/Bloedneus-gemept-om-voetbalplaatjes.dhtml on 08-28-2014.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157-162. doi: 10.1111/1467-9280.00327

Ariely, D., & Levav, J. (2000). Sequential choice in group settings: Taking the road less traveled and less enjoyed. *Journal of Consumer Research, 27*(3), 279-290. doi: 10.1086/317585

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*(1), 124-140. doi: 10.1016/0749-5978(85)90049-4

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61*(3), 183-193. doi: 10.1037/h0054663

Beardsly, M. C. (1958/1981). *Aesthetics: Problems in the Philosophy of Criticism*. Hackett Publishing Company Inc. Cambridge, MA.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9*(3), 226-232. doi: 10.1002/bs.3830090304

Belk, R. W. (1995a). Collecting as luxory consumption: Effects on individuals and households. *Journal of Economic Psychology, 16*(3), 477-490. doi: 10.1016/0167-4870(95)98956-x

Belk, R. W. (1995b). *Collecting in a consumer society*. New York: Routledge.

Belk, R. W., Wallendorf, M., Sherry, J., Holbrook, M., & Roberts, S. (1988). Collectors and collecting. In M. J. Houston (Ed.), *Advances in Consumer Research* (Vol. 15). Provo, UT: Asociation for Consumer Research

Belk, R. W., Wallendorf, M., Sherry, J. F., & Holbrook, M. B. (1991). Collecting in a consumer culture. In R. W. Belk (Ed.), *Highways and Buyways: Naturalistic Research From the Consumer Behavior Odyssey* (pp. 238). Provo, UT: Association for Consumer Research.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review, 91*(1), 79-98. doi: 10.1257/aer.91.1.79

Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Carey, C. (2008). Modeling collecting behavior: The role of set completion. *Journal of Economic Psychology, 29*(3), 336-347. doi: 10.1016/j.joep.2007.08.002

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review, 103*(3), 566-581. doi: 10.1037/0033-295x.103.3.566

Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *Quarterly Journal of Experimental Psychology, 52*(2), 273-302. doi: 10.1080/027249899391070

Chater, N., & Vitanyi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences, 7*(1), 19-22. doi: 10.1016/s1364-6613(02)00005-0

Checkosky, S. F. & Whitlock, D. (1973). Effects of pattern goodness on recognition time in a memory search task. *Journal of Experimental Psychology, 100*(2), 341-348. doi: 10.1037/h0035692

Chernev, A. (2012). Product assortment and consumer choice: An interdisciplinary review. *Foundations and Trends in Marketing, 6*(1), 1-61.

De Groot, A. D. (1946). *Het denken van den schaker. Een experimenteel-psychologische studie*. Amsterdam, The Netherlands: Noord-Hollandsche Uitgevers Maatschappij.

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71*(1), 1-27. doi: 10.3102/00346543071001001

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011). Gamification: Toward a definition. *Gamification Workshop: CHI2011*.

Evers, E. R. K., Inbar, Y., & Zeelenberg, M. (2014). Set-fit effects in choice. *Journal of Experimental Psychology: General, 143*(2), 504-509. doi: 10.1037/a0033343

Evers, E. R. K., & Lindenberg, S. (2014). How collectors collect. *Working Paper.*

Evers, E. R. K., Lindenberg, S., & Linksens, S. J. (2014). Characteristics of collecting. *In Preparation*.

Evers, E. R. K., van de Ven, N., & Weeda, D. (2014). The hidden cost of microtransactions: How increasing ones power through microtransactions decreases status. *Under Review.*

Evers, E. R. K., Zeelenberg, M., & Inbar, Y. (2014). The set-completion premium. *Under Review.*

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*(2), 301-318. doi: 10.1037/0033-295x.104.2.301

Fox, C. R., Bardolet, B., & Lieb, D. S. (2005). Partition dependence in decision analysis, managerial decision making, and consumer choice. In R. Zwick & A. Rapoport (Eds.), *Experimental business research, volume III*. Dordrecht, The Netherlands: Kluwer.

Fox, C. R., Ratner, R. K., & Lieb, D. S. (2005). How subjective grouping of options influences choice and allocation: Diversification bias and the phenomenon of partition dependence. *Journal of Experimental Psychology: General, 134*(4), 538-551. doi: 10.1037/0096-3445.134.4.538

Fox, C. R., & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science, 14*(3), 195-200. doi: 10.1111/1467-9280.02431

Galak, J., Kruger, J., & Loewenstein, G. (2011). Is variety the spice of life? It all depends on the rate of consumption. *Judgment and Decision Making, 6*(3), 230-238.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: John Wiley and sons inc.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist, 58*(1), 34-42.

Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior, 2*(5-6), 446-452. doi: 10.1016/s0022-5371(63)80046-8

Glanzer, M., & Clark, W. H. (1964). The verbal-loop hypothesis: Conventional figures. *American Journal of Psychology, 77*, 621-626. doi: 10.2307/1420773

Gneezy, U., List, J. A., & Wu, G. (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *Quarterly Journal of Economics, 121*(4), 1283-1309. doi: 10.1162/qjec.121.4.1283

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, L., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236-243. doi: 10.1016/s1364-6613(00)01662-4

Hamari, J., & Eranti, V. (2011). Framework for designing and evaluating game achievements. *Proceedings of the DiGRA 2011 Conference: Think Design Play*.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38*(1), 79-109. doi: 10.1006/cogp.1998.0708

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review, 10*(2), 88-110.

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*(3), 247-257. doi: 10.1006/obhd.1996.0077

Hsee, C. K., & Zhang, J. A. (2010). General evaluability theory. *Perspectives on Psychological Science, 5*(4), 343-355. doi: 10.1177/1745691610374586

Huotari, K., & Hamari, J. (2011). Gamification from the perspective of service marketing. *Gamification Workshop, CHI2011*.

Jakobbson, M. (2011). The achievement machine: Understanding Xbox 360 achievements in gaming practices. *Game Studies, 11*(1).

Kahn, B. E., & Lehman, D. R. (1991). Modelling choice among assortments. *Journal of Retailing, 67*, 274-299.

Kahn, B. E., & Wansink, B. (2004). The influence of assortment structure on perceived variety and consumption quantities. *Journal of Consumer Research, 30*(4), 519-533. doi: 10.1086/380286

Kahneman, D., & Thaler, R. H. (2006). Anomalies: Utility maximization and experienced utility. *Journal of Economic Perspectives, 20*(1), 221-234. doi: 10.1257/089533006776526076

Kay, A. C., Whitson, J. A., Gaucher, D., & Galinsky, A. D. (2009). Compensatory control: Achieving order through the mind, our institutions, and the heavens. *Current Directions in Psychological Science, 18*(5), 264-268.

Kimchi, R., & Palmer, S. E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology-Human Perception and Performance, 8*(4), 521-535. doi: 10.1037/0096-1523.8.4.521

Kintsch, W. (2012). Musings about beauty. *Cognitive Science, 36*(4), 635-654. doi: 10.1111/j.1551-6709.2011.01229.x

Kivetz, R., Urminsky, O., & Zheng, Y. H. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research, 43*(1), 39-58. doi: 10.1509/jmkr.43.1.39

Koffka, K. (1935). *Principles of Gestalt Psychology*. London: Lund Humphries.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems Information Processing, 1*(1), 1-7.

Koo, M., & Fishbach, A. (2008). Dynamics of self-regulation: How (un)accomplished goal actions affect motivation. *Journal of Personality and Social Psychology, 94*(2), 183-195. doi: 10.1037/0022-3514.94.2.183

Kruglanski, A. W., Webster, D. M., & Klem, A. (1993). Motivated resistance and openness to persuasion in the presence or absence of prior information. *Journal of Personality and Social Psychology, 65*(5), 861-876.

Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science, 9*(3), 278-292. doi: 10.1177/1745691614528520

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling situations. *Journal of Experimental Psychology, 89*, 46-55.

Lindley, R. H. (1966). Recoding as a function of chunking and meaningfulness. *Psychonomic Science, 6*(8), 393-394.

Litman, J. A., & Jimerson, T. L. (2004). The measurement of curiosity as a feeling of deprivation. *Journal of Personality Assessment, 82*(2), 147-157.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75-98.

Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological theory. *American Journal of Psychology, 103*(1), 53-80. doi: 10.2307/1423259

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition, 122*(3), 346-362. doi: 10.1016/j.cognition.2011.11.003

McAlister, A. R., Cornwell, T. B., & Cornain, E. K. (2011). Collectible toys and decisions to share: I will gift you one to expand my set. *British Journal of Developmental Psychology, 29*(1), 1-17. doi: 10.1348/026151010x526353

McAlister, L. (1982). A dynamic attribute satiation model of variety-seeking behavior. *Journal of Consumer Research, 9*(2), 141-150. doi: 10.1086/208907

McGraw, A. P., Shafir, E., & Todorov, A. (2010). Valuing money and things: Why a $20 item can be worth more and less than $20. *Management Science, 56*(5), 816-830. doi: 10.1287/mnsc.1100.1147

McIntosh, W. D., & Schmeichel, B. (2004). Collectors and collecting: A social psychological perspective. *Leisure Sciences, 26*(1), 85-97. doi: 10.1080/01490400490272639

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97. doi: 10.1037//0033-295x.101.2.343

Miller, N. E. (1944). Experimental studies of conflict. *Personality and the Behavior Disorders*. Oxford, UK: Ronald Press.

Montola, M., Nummenmaa, T., Lucerano, A., Boberg, M., & Korhonen, H. (2009). Applying game achievement systems to enhance user experience in a photo sharing service. *Ek '09: Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*.

Mourey, J. A., Oyserman, D., & Yoon, C. (2013). One without the other seeing relationships in everyday objects. *Psychological Science, 24*(9), 1615-1622. doi: 10.1177/0956797613475631

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual-differences in the desire for simple structure. *Journal of Personality and Social Psychology, 65*(1), 113-131.

Nicki, R. M., & Moss, V. (1975). Preference for non-representational art as a function of various measures of complexity. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 29*(3), 237-249. doi: 10.1037/h0082029

O'Brien, G. (1981). Living with collections. *New York Times Magazine*, 25-42.

Palmer, S. E. (1991). Goodness, gestalt, groups, and Garner: Local symmetry subgroups as a theory of figural goodness. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The Perception of Structure* (pp. 23-39) Washington DC: American Psychological Association.

Paolacci, G., & van de Berg, B. (2014). Distant is different: Mere physical proximity between products affects choice. *Working Paper*.

Poitras, L., Rosenbach, M., & Stark, H. (2013). Ally and target: US intelligence watches Germany closely. *Der Spiegel, 33*.

Pomerantz, J. R., & Lockhead, G. R. (1991). Perception of structure: An overview. In G. R. Lockhead & J. R. Pomeranz (Eds.), *The Perception of Structure*. Washington DC: American Psychological Association.

Pope, D., & Simonsohn, U. (2011). Round numbers as goals: Evidence from baseball, SAT takers, and the lab. *Psychological Science, 22*(1), 71-79. doi: 10.1177/0956797610391098

Prentice, W. C. H. (1944). The interruption of tasks. *The Psychological Review, 51*(6), 329-340.

Ratner, R. K., Kahn, B. E., & Kahneman, D. (1999). Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research, 26*(1), 1-15. doi: 10.1086/209547

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separate choices. *Journal of Experimental Psychology: Applied, 1*(1), 34-49. doi: 10.1037/1076-898x.1.1.34

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*(4), 364-382. doi: 10.1207/s15327957pspr0804_3

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*(1), 45-48. doi: 10.1111/1467-9280.00008

Rissanen, J. (1978). Modeling by the shortest data description. *Automatica 14*, 465-471.

Rissanen, J. (1987). Stochastic complexity. *Journal of the Royal Statistical Society B*, 49, 223-239.

Robbins, L. (1932). *An essay on the nature and significance of economic science*. London, UK: Macmillan and co. Limited.

Sardy, S., & Velenik, Y. (2010). Paninimania: Sticker rarity and cost-effective strategy. *Swiss Statistical Society Bulletin*(66), 2-6.

Schiffer, M. B., Downing, T. E., & McCarthy, M. (1981). Waste not, want not: An ethnoarchaeological study of refuse in Tuscon, Arizona. In M. Gould & M. B. Schiffer (Eds.), *Modern Material Culture: The Archaeology of Us*. (pp. 67-86). New York: Academic Press.

Schmidhuber, J. (1997). Low-complexity art. *Leonardo, Journal of the International Society for the Arts, Sciences, and Technology., 30*(2), 97-103.

Shannon, C. E. (1948). The mathemetical theory of communication. *Bell Systems Technical Journal, 27*, 379-423.

Simon, H. (1957). A behavioral model of rational choice. *Models of Man, Social and Rational: Mathematical Essays on Behavior in a Social Setting*. New York: Wiley

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16*(2), 158-174. doi: 10.1086/209205

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research, 27*(2), 150-162. doi: 10.2307/3172842

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29*(3), 281-295. doi: 10.2307/3172740

Slater, J. S. (2001). Collecting brand loyalty: A comparative analysis of how Coca-Cola and Hallmark use collecting behavior to enhance brand loyalty. In M. C. Gilly & J. MeyersLevy (Eds.), *Advances in Consumer Research, Vol Xxviii* (Vol. 28, pp. 362-369). Valdosta, GA: Association for Consumer Research.

Snyder, C. R., & Fromkin, H. L. (1977). Abnormality as a positive characteristic: The development and calidation of a scale measuring need for uniqueness. *Journal of Abnormal Psychology., 86*, 518-527.

Strannegård, C., Nizamani, A. R., Sjöberg, A., & Engström, F. (2013). Bounded Kolmogorov complexity based on cognitive models. *General Intelligence, 7999*, 130-139.

Te Pas (2009). Aangfite na gevecht om AH-plaatjes. [Criminal-complaint filed after fight for Albert-Heijn stickers.] Retrieved from http://www.distrifood.nl/Formules/Algemeen/2009/2/Aangifte-na-gevecht-om-AH-plaatjes-DIS132107W/ on 08-28-2014.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25-29.

Totilo, Stephen (2012). The strange, scary, fascinating, exciting future of video games, according to a giant. Retrieved from http://kotaku.com/5919847/the-strange-scary-fascinating-exciting-future-of-video-games-according-to-a-giant on 08-28-2014.

Trouw (2009a). Ruzie om voetbalplaatjes. [Fights for soccer-stickers.] Retrieved from http://www.trouw.nl/tr/nl/4556/Onderwijs/article/detail/1135990/2009/02/08/Ruzie-om-voetbalplaatjes.dhtml# on 08-28-2014.

Trouw (2009b). Vechten voor een sticker. [Fighting for a sticker.] Retrieved from http://www.trouw.nl/tr/nl/4556/Onderwijs/article/detail/1135991/2009/02/08/Vechten-om-een-sticker.dhtml on 08-28-2014.

Van Dijk, E., & Zeelenberg, M. (2007). When curiosity killed regret: Avoiding or seeking the unknown in decision-making under uncertainty. *Journal of Experimental Social Psychology, 43*(4), 656-662. doi: 10.1016/j.jesp.2006.06.004

Vitányi, M. B., & Li, M. (2009). *An introduction to Kolmogorov complexity and its applications*. NY: Springer.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*(6), 1172-1217. doi: 10.1037/a0029333

Wallace, C., 6 Freeman, P. (1987). Estimation and inference by compact coding. *Journal of the Royal Statistical Society Bulletin*, 49, 240-251.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psycologische Forschung, 4*, 301-350.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology-Learning Memory and Cognition, 19*(6), 1235-1253. doi: 10.1037//0278-7393.19.6.1235

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an atributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language, 29*(6), 716-732. doi: 10.1016/0749-596x(90)90045-2

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology, 81*(6), 989-1000. doi: 10.1037//0022-3514.81.6.989

Winkielman, P., Schwarz, N., Reber, R., & Fazendeiro, T., A. (2003). Cognitive and affective consequences of visual fluency: When seeing is easy on the mind. In L. M. Scott & R. Batra (Eds.), *Persuasive Imagery: A Consumer Response Perspective. Advertising and Consumer Psychology.* (pp. 75-89). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Wänke, M., Bless, H., & Schwarz, N. (1999). Assimilation and contrast in brand and product evaluations: Implications for marketing. *Advances in Consumer Research, Vol 26, 26*, 95-98.

Zajonc, R., B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1-27. doi: 10.1037/h0025848

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*(6), 224-228. doi: 10.1111/1467-8721.00154

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology, 17*(1), 3-18. doi: 10.1207/s15327663jcp1701_3

Zeigarnik, B. (1967). On finished and unfinished tasks. In D. Ellis (Ed.), *A Sourcebook of Gestalt Psychology* (pp. 300-304). New York: Humanities press.

Acknowledgements

In the three years as grad-student, I have been blessed to be surrounded by a large group of interesting, inspiring, friendly and fun people all of whom have affected me as a person and a researcher in one way or the other.

The research presented in this dissertation would not be close to what it is without the questions and suggestions by friend and colleagues from Tilburg Social Psychology, Wharton OPIM, CMU SDS, and those of you I have met at conferences.

Friends and family, thank you for being there in times of need, for accepting (or at least pretending to accept) my at times erratic and obnoxious behavior, and for the fun times.

Even though I cannot possibly give everyone the credit they deserve on this page, there are a few people I would like to mention explicitly.

First of all Bart, I am not easy to deal with, I don't know how you do it, but I am happy you do! I am afraid that the chaos I sometimes bring puts even Durban rush hour to shame… It is impossible to express in words how important you are to me and how happy I am having you in my life. Thank you for being there through all uncertainty and stress, you make every part of my life better.

Marcel and Yoel "baas van de club" Inbar, you guys are amazing. I could not have wished for better teachers, colleagues, and friends. I am extremely grateful for you giving me the freedom to try my own things, while always being there to step in when I needed directions. Both of you have made me feel fully supported and like you always had my best interests at heart. Because of the two of you, grad-school has been an extremely pleasant experience.

GIDEON, you are the sababa-est! Thank you for introducing me to doing research. Uri, you have had a huge influence on my understanding of research and I have learned so much during my visits, thank you for providing the opportunity. George, you are amazing, thank you for everything! Finally my two paranymphs, Flip & Koen. Flip, sharing an office (as well as Spanish gansta-rap shows) with you has been amazing and I miss it a lot. I really hope we can somehow, somewhere, share an office again. Koen, you are my favorite brother. I always have a great time hanging out with you. I am really lucky to have a brother like you.

Mama, ik wou dat je erbij kon zijn. Alles dankzij jou.