

Learning to Decide with and without Reasoning: How Task Experience Affects Attribute Weighting and Preference Stability

TUOMAS LEISTI*  and JUKKA HÄKKINEN

Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland

ABSTRACT

Certain experiments have shown that reasoning may weaken the stability of people's preferences, especially with regard to well-learned perceptual judgment and decision-making tasks, while learning has an opposite, consistency-enhancing effect on preferences. We examined the effects of these factors in a visual multi-attribute decision-making task where reasoning, in contrast, has been found to *benefit* judgments by making them more stable. The initial assumption in this study was that this benefit would be typical for novel tasks, like the one employed here, and that it would decrease when the task is thoroughly learned. This assumption was examined in three experiments by contrasting it with an alternative assumption that this previously obtained beneficial effect is caused solely by learning, not by reasoning. It was found that learning indeed makes preferences more stable by consolidating the weights of the attributes. Reasoning, however, does not benefit this task when it is completely novel but facilitates learning and stability of the preferences long run, therefore increasing the consistency of the participants in the macrolevel. © 2017 The Authors Journal of Behavioral Decision Making Published by John Wiley & Sons Ltd.

KEY WORDS judgment and decision making; reasoning; learning; information integration; attribute weighting; preference stability; routine

Relatively little is known about the interaction between reasoning and learning in judgment and decision making (JDM). In this context, reasoning refers to the conscious thought process leading to a justified resolution to a certain decision problem. It therefore contrasts with intuition, which lacks this deliberate thought process (Kahneman, 2003). In some studies, reasoning seems to benefit relatively novel or unusual JDM tasks, and this advantage disappears with increasing routine (e.g., Berry & Broadbent, 1984, 1987; Hagafors & Brehmer, 1983; Sieck & Yates, 1997). These findings appear quite plausible: Without appropriate experience with a task, an individual can use conscious reasoning to take advantage of the explicit knowledge that has been acquired during the task or that has been acquired from other people, from the surrounding culture, or from other contexts (Baumeister & Masicampo, 2010). More familiar tasks can be executed successfully without reasoning, using existing routines or intuition, which results from learning when individual gathers experience with the task (Glöckner & Wittman, 2010; Hogarth, 2001). The themes of reasoning and learning are therefore closely associated with the current discussion about the dual process models and how deliberative cognitions gradually become intuitions.

The advantage of conscious reasoning in JDM has been suggested to be in *information acquisition* (Dijksterhuis & Nordgren, 2006; Glöckner & Betsch, 2008). The reasoning that is evoked when there is a requirement to provide justifications for decisions may induce a more thorough information search, as the decision maker does not want to appear foolish by not considering all the available information (Leisti, Radun, Virtanen, Nyman, & Häkkinen, 2014; Lerner

& Tetlock, 1999). This may decrease, for instance, the tendency towards biases evoked by the context because participants who are consciously reasoning focus more attention on less salient aspects of the problem (Hamilton, Hong, & Chernev, 2007; Sieck & Yates, 1997). In their results from a dynamic decision task, Berry and Broadbent (1984, 1987) attributed the facilitative effect of verbalization to its ability to focus attention on critical information at the right times. Thus, it may be not only the social pressure of justification, which benefits task performance, but also verbal processing in general, which strategically guides attention in a sequential manner (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003).

Also, learning has been found to influence information acquisition by gradually shifting the learner's focus to more relevant information, usually leading to greater accuracy and speed (Haider & Frensch, 1999), which seems to apply to JDM tasks, too (Orquin, Bagger, & Mueller Loose, 2013). Learning furthermore enhances *information integration*, the process where acquired information is weighted and judgments are formed. Personal preferences can be quite arbitrary and unstable when there is little experience with a task (Ariely, Loewenstein, & Prelec, 2003). Preferences, in these cases, often appear to be dependent on information that is easily available for consideration (e.g., Kahneman, 2003). After making repeated choices with trade-offs, people gradually become more certain of the subjective importance of different attributes and learn to weight them accordingly, consequently making their preferences more stable (Hoeffler & Ariely, 1999). Preferences should therefore not be dependent on how easily the information is accessed.

In familiar, thoroughly learned tasks that have become automatic, reasoning can even be distractive: Several studies have shown that conscious reasoning may sometimes weaken the stability of preferences and therefore lead to regret and reduced satisfaction with choices (Dijksterhuis &

*Correspondence to: Tuomas Leisti, Department of Psychology and Logopedics, University of Helsinki, PO Box 9 (Siltavuorenpenger 1 A), FIN-00014, Helsinki, Finland. E-mail: tuomas.leisti@helsinki.fi

van Olden, 2006; Lee, Amir, & Ariely, 2009; Pham, Cohen, Pracejus, & Hughes, 2001; Rusou, Zakay, & Usher, 2013; Wilson et al., 1993). The cause of this maladaptive effect of reasoning appears to be its interference with automatic information integration processes, which are based on experiences accumulated with similar stimuli (i.e., intuition; Glöckner & Wittman, 2010) or on attribute weights that have been consolidated when attention is directed elsewhere between information acquisition and the actual decision (the unconscious thought effect; Dijksterhuis & Nordgren, 2006). As the role of conscious reasoning is widely assumed to be to evaluate automatic responses and to provide corrections when needed (Kahneman, 2003), it may evoke re-evaluations of initial responses to the stimuli (Betsch, 2011; Simonson & Nowlis, 2000). This re-evaluation may be distracting because it forces individuals to create justifications for preferences, the reasons for which they cannot completely access (Nisbett & Wilson, 1977; Wilson, Dunn, Kraft, & Lisle, 1989). As a consequence, individuals are assumed to concentrate on seeking sound justifications for their choices and to alienate themselves from preferences that they cannot explain. Additionally, automatic judgments are not restricted by working memory capacity and serial processing in the same way as conscious reasoning (Evans & Stanovich, 2013); burden caused by complex decisions may lead to inconsistent weighting of information and unstable preferences, because the decision maker is able to focus consciously only on a subset of all relevant attributes (Dijksterhuis & Nordgren, 2006; Levine, Halberstadt, & Goldstone, 1996).

That excessive experience with a task is an advantage is not the whole truth either; when the choice of a certain option becomes habitual, the individual starts to omit the search for comprehensive information and is therefore unable to adapt to changes in the environment (Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001; Verplanken, Aarts, & van Knippenberg, 1997). The effect of learning on information acquisition is therefore largely context dependent and can potentially lead to maladaptive consequences (Bröder, Glöckner, Betsch, Link, & Ettl, 2013). Individuals may also learn biasing heuristics, when the presentation of the choice problem repeatedly makes certain attributes salient (Amir & Levav, 2008).

Reasoning, as well, does not always benefit unfamiliar choice tasks: In cases where participants are uncertain about their preferences, where there is no default response, or where none of the alternatives dominate others, the most salient feature can provide good justification for reasoning, amplifying certain biases such as the attraction effect (Simonson, 1989). These biases, however, do not concern within-individual consistency, on which this article is focused, but are found when participants are compared with each other (Shafir, Simonson, & Tversky, 1993). In within-subject designs (e.g., Hamilton et al., 2007), participants can always compare their current choices with their previous ones, and they therefore can become more certain about their preferences and deliberately seek out information that had determined their earlier choices.

Expertise, preferences, and conscious reasoning

Experience with a certain JDM context and the subsequent learning leads to expertise in the long run (Hogarth, 2001). Experience with many different types of beers, for instance, leads to enhanced ability to evaluate the quality and other aspects of beer (Hoeffler, Ariely, West, & Duclos, 2013). Therefore, it seems plausible that any interaction between reasoning and learning in JDM could be understood by comparing the effect of reasoning on judgments within novice and expert populations. Certain findings so far, however, suggest that reasoning does not benefit novices: Excessive reasoning appears to lead to deviation from expert ratings or to regret, at least when asked to make judgments on art. (McGlone, Kobrynawics, & Alexander, 2005; Wilson et al., 1993; Yamada, 2009). This phenomenon has been associated with novices' inability to describe their experiences (McGlone et al., 2005). Verbalization shifts their attention to aspects that are easier to verbalize, which can then change their stated preferences. The JDM of experts, unlike that of novices, does not suffer from reasoning, because experts are assumed to have explicit knowledge about the reasons for their preferences.

The aforementioned perspective, stemming from the social psychological tradition, appears to be in contradiction not only with the findings discussed in the previous section but also with research on expert JDM, which emphasizes the role of implicit knowledge in expertise (Kahneman & Klein, 2009). Therefore, Dijkstra, Pligt, and Kleef (2013) have suggested that the effect of reasoning is dependent on *both* explicit *and* implicit knowledge: Experts have a large amount of both types of knowledge, and those who are distracted by conscious reasoning are, in fact, intermediates who already have gathered enough experience to have implicit, subsymbolic knowledge but have not yet created explicit, conceptual knowledge about the task (also Cleeremans & Jiménez, 2002). Such intermediates in art evaluation could be, for example, participants recruited from universities, as experience with art is part of the cultural capital of middle-class and upper-class citizens (Bourdieu, 1984). Real novices have neither implicit nor explicit knowledge and perform equally poorly in both conditions.

This model suggests another class of intermediates, with explicit theoretical knowledge but little practical experience (Table 1). This group should benefit from conscious reasoning in novel tasks. As examples of such cases, Dijkstra et al. (2013) present medical doctors (Schmidt & Rikers, 2007) and legal judges in the early stages of their careers. Experimental support for such a pattern has also been found from, for example, dynamic decision-making tasks (Berry & Broadbent, 1984, 1987) and from perceptual identification tasks (Melcher & Schooler, 2004). Reasoning also enables the individual to create explicit knowledge as hypotheses, which can be tested in subsequent decisions (Hagafors & Brehmer, 1983): this suggests that conscious reasoning not only benefits performance when the task is unfamiliar but that it also facilitates learning.

The aforementioned examples are JDM tasks that have objectively correct answers. The question is, then, whether

Table 1. Effects of verbalization within expert, intermediate, and novice populations; adapted from Dijkstra et al. (2013)

	High explicit knowledge	Low explicit knowledge
High implicit knowledge	Experts: Verbalization has no effect on judgments. Possess the verbal skills to describe their intuition.	Intermediates: Verbalization distracts from judgments. Do not have sufficient verbal ability to describe their intuition. Possess gathered experience to make intuitive judgments but not formal (analytic) training.
Low implicit knowledge	Intermediates: Verbalization benefits judgments. Possess adequate conceptual knowledge but lack experience.	Novices: Verbalization has no effect on judgments. Little explicit and implicit knowledge.

these results apply to novel preferential choices with multiple attributes. We will examine this question in three experiments by employing a multi-attribute visual choice task, which has been shown to benefit from reasoning (Leisti & Häkkinen, 2016), contrary to earlier results with visual stimuli (e.g., Levine et al., 1996; Wilson et al., 1993). We aimed to test whether this unexpected effect can be explained by the facilitative effect of reasoning in tasks that are unfamiliar to the participants, who are nevertheless able to describe reasons for their judgments (Radun et al., 2010). Here, conscious deliberation would take advantage of the more organized use of information, using existing explicit knowledge of relevant attributes. Our focus is on preference stability and on weighting of the attributes, while controlling for the effect of participants' reliability in information acquisition.

The current experiments

The basis of our experiments is visual judgments, or more precisely the evaluation of the image quality of photographs. Although almost every mobile phone is equipped with a digital camera, people may not be very experienced in making trade-offs between small differences in image quality, which should mean that there is room for learning. We assume that in an everyday context, people use a simple satisfying heuristic: If the photo fails to reach a certain level of quality with all attributes, it is deleted and a new one is taken (Radun, Nuutinen, Leisti, & Häkkinen, 2016).

Our previous research has shown that with this specific JDM task, conscious reasoning leads to more comprehensive information acquisition by guiding attention to the less salient features of the alternatives (Leisti et al., 2014). This is advantageous in the task not only because of its relative novelty for the participants but also because the differences between the choice alternatives are few and non-salient but visually separable and easily verbalized. This contrasts with the complex and salient but often inef-fable differences between stimuli such as faces or art that have been used in studies that have shown a distractive effect of reasoning.

We focus on the issue of preference stability. A paired comparison setting enables the use of the number of intransitive choices as a measure of stability. Here, intransitivity occurs with three stimuli, A, B, and C, when the participants' paired choices follow the pattern $A \succ B$, $B \succ C$, and $C \succ A$, where \succ indicates preference. Intransitivity reflects

the participants' inability to put the alternatives in order as well as intertrial fluctuations in preferences or noise in the JDM process (Tversky, 1969). In the task employed in this series of experiments, the stability of the preferences is assumed to be a result of the consistent weighting of the attributes, which is estimated from the trials in which differences exist regarding the attribute in question. We also control for the effect of the participants' reliability, which is estimated by examining the participants' choices in trials in which there are differences in only one attribute. Low reliability caused by, for example, carelessness may lead to lower preference stability, which is not related to the low consistency in weighting of information.

Conscious reasoning in these experiments is evoked by asking participants to explain their choices. In other words, they are required to create verbal explanations that link their preferences and visual experiences together, shifting the original process into a more conscious one (Baumeister, Masicampo, & Vohs, 2011; Fox, Ericsson, & Best, 2011). The explanation of perceptual preferences is thus an interpretative process in that participants are not just verbalizing inner speech or task-related thoughts that spontaneously come to their attention.

Purpose of the research

We aim to study the influence of learning and reasoning on preference stability in a task that has earlier been shown to benefit from the eliciting of explanations. Our earlier experiments had not controlled for the effect of learning. Our first aim is to test whether our earlier results can be attributed to learning or whether explaining still has an effect when learning is controlled for. If reasoning influences the preference stability in the task, our second aim is to clarify how it benefits decision making: Does it facilitate learning, or does it have an independent effect on performance? Experiment 1 in this study controls for the role of learning by using a cross-over design whereby the explaining condition is performed both before and after the silent condition in different participant groups. Experiments 2A and 2B then examine the development of preference stability with and without explaining.

Hypotheses

As the first hypothesis, we propose that learning will increase the stability of preferences in this relatively novel task before

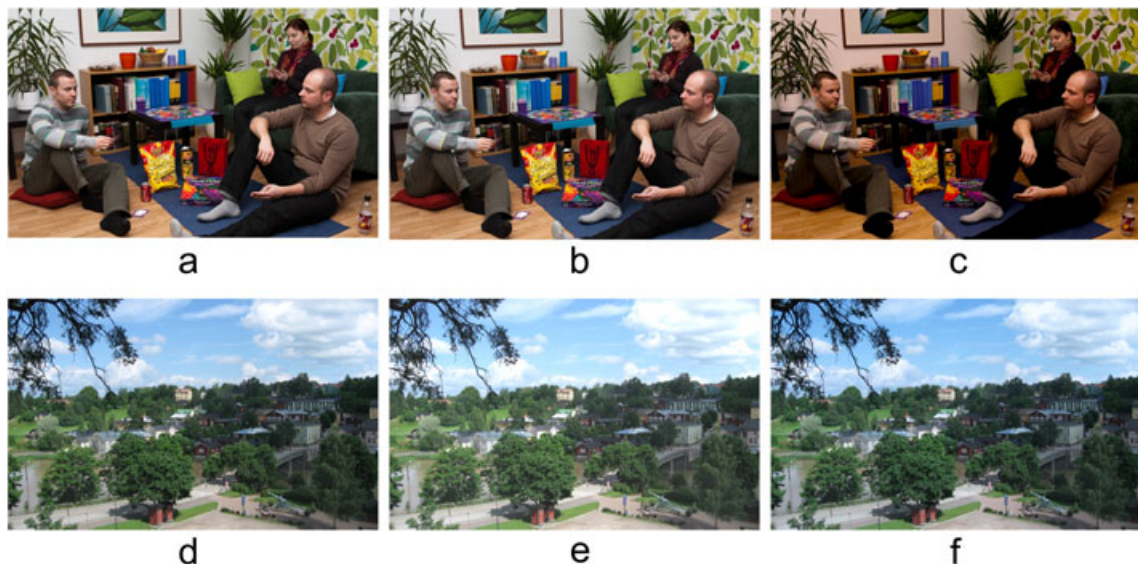


Figure 1. Examples of stimulus images. (a) and (d) represent the original images from Town and Party contents. Blur and noise have been added to (b), whereas (c) has been manipulated by changing the white point and increasing the lightness level. (e) represents the version with added blur and decreased lightness, and (f) represents the version with added noise and a changed white point [Colour figure can be viewed at wileyonlinelibrary.com]

the participants become acquainted with the attributes of the alternatives. This effect is based on the development of consistent information weighting scheme. According to the second hypothesis, reasoning will increase preference stability before the task becomes familiar because it increases the attention paid to less salient attributes. The third hypothesis is that reasoning facilitates learning, as it makes the participants aware of the trade-offs between attributes and requires effortful thinking about the importance of the attributes.

EXPERIMENT 1

Method

Participants

A total of 102 university students (23 men and 79 women) were recruited from university e-mail lists and through Facebook and were given one movie ticket for participation. Participants reported having normal vision. Their color vision was screened with a color vision test (Farnsworth D-15). Two participants failed to pass the test and were excluded from the analyses. The mean age of the accepted participants was 26.1 years (standard deviation (SD) = 6.3 years).

Design

We used a cross-over design: The experiment included two conditions, silent (no explanations) and explanations, which were placed in two separate blocks. All participants thus completed one block with explanations and one without explanations. The order of these two conditions in the two blocks was varied between the participants to examine the learning process during the experiment. Additionally, we included a timing-of-the-verbalization condition as a

between-participant variable because we wanted to test whether the beneficial effect of explaining is restricted to concurrent explaining (Fox et al., 2011).¹

Stimuli

We used stimulus material from our previous experiments (Leisti & Häkkinen, 2016), although we excluded one image and some of the samples to shorten the experiment. We used two image content types (*Town* and *Party*), which were degraded with small amounts of Gaussian blur or noise, or by changing the color balance or luminance level (see Figure 1 for examples). These are considered typical defects in home photography. The amount of blur added to the degraded versions of the images was .45 SD s. Noise was added with a variance of .001 to the Town content and .0006 for the Party content. Concerning the color balance and luminance level, the manipulated samples in the Town content were made slightly more bluish or bright by changing the color temperature from 5600 to 6500 K and by increasing luminance in the L^* channel in the L^* ch color space by a value of 8. The samples in the Party content, however, were made more reddish or darker by changing the color temperature from 3400 to 2700 K and by decreasing the brightness in the L^* ch luminance channel by a value of 12. Altogether, 10

¹In Leisti et al. (2014), we assumed that the benefit of explaining was related to concurrent verbalization. The timing of the verbalization was therefore included as an independent variable. Experiments 1 and 2A show, however, that no difference exists between the concurrent and retrospective conditions. One unpublished experiment first suggested that the effect size would be smaller in the retrospective condition, but this result seems to be related to the amount of accidental incorrect responses because the participants were asked to indicate their choices by pressing keys on a keyboard. To avoid similar incorrect responses due to such lapses, the participants made their choices using a mouse and a graphical user interface in all experiments reported in this study.

degraded versions of the two content types were created by manipulating one or two image-quality parameters. Including the original, unmanipulated image, the total number of stimulus images per content type was 11. Image pairs were presented on two color-calibrated 24.1-in. Eizo ColorEdge CG241W displays with a resolution of 1920 * 1200 pixels.

Procedure

The experiment consisted of two blocks with different image content types, one with explanations and one silent. We randomized the participants into eight experimental groups such that the order of the image content types, the order of the conditions (silent vs. explanations), and the timing of the explanation (concurrent vs. retrospective) were counter-balanced. As there were 11 versions of one content type and all the versions were compared with each other pairwise, the completion of the block required 55 choice trials. The order of the pairs was randomized within both blocks.

Before the experiment proper, the participants were asked for informed consent and were told that they would be evaluating the quality of photographs. They were instructed to always choose the better image from the image pairs on the basis of personal and subjective preferences. The participants were urged to consider which of the photographs they would include in their own photo album.

Within each trial, the participants were presented two images on two displays and were asked to indicate their choices on a third generic display using a graphical user interface (GUI). In the concurrent explanation condition, the GUI consisted of one text field for participants to explain why the chosen image was better and a button for indicating the choice. Between each trial, a blank medium-gray screen replaced the stimulus images for 500 milliseconds. In the retrospective explanation condition, the participants were first presented with the stimulus images and the GUI with a button to indicate the answer. After the answer was given, the stimulus images were removed and replaced with medium-gray blank screens, and a text field appeared in the GUI for participants to explain their choice. There was a short pause between the blocks so that the participants could recover from the first block.

Dependent variables

Intransitive preferences. The stability of the participants' preferences was operationalized as the number of intransitive preferences. For each participant, the number of intransitive choices was calculated in each block of the experiment. One intransitive choice consisted of the pattern $A \succ B$, $B \succ C$, and $C \succ A$, where A, B, and C were different stimuli and \succ indicates preference. A logarithmic transformation was used in the analyses because the number of all possible intransitive patterns grows exponentially when the participant makes more intransitive choices (that is, one choice can be present in several intransitive patterns). In order to avoid missing values for zeros, one was added to the all values before transformation ($\ln(0)$ is not defined, but $\ln(1)$ equals 0).

Weighting of attributes. Calculation of the attribute weights was based on the probability of choosing the alternative with a certain objective parameter manipulation from all pairs with a difference for that parameter. The probabilities were then linearized by using a logit transformation

$$\text{logit}(p) = \ln[p/(1-p)]$$

where p is the probability of choice. After this calculation was performed, the attributes were reordered according to their importance for further analysis. Therefore, the analyses were not made for specific attributes (such as blur or noise) but for the most important attribute, the second most important attribute, and so on. These were determined block-wise, so a participant could have a different order of importance for the attributes in different blocks.

Reliability of choices. Reliability, in this case, refers to the participant's consistency in trials in which there was a difference in only one attribute. Owing to the factorial design of the stimulus manipulation, there were four such cases for each attribute, content type, and participant. For example, there were four pairs in which the only difference was in the image blur. Reliability was based on the logit transform of probability of choosing the sample with the same attribute value in all four cases. When the probability of the choices was 1, the value of .937 was used instead to avoid division by zero, as David (1969) suggests, using the formula $1 - 1/2n$, where n is the number of pairs, to replace the value of 1. The sum of the logits of probabilities of all four attributes then served as a measure of reliability for each condition for each participant.

Results

Stability of the preferences

A mixed analysis of variance (ANOVA) was performed on the data of the log-transformed number of intransitive choices, with the condition (explanation vs. silent) as a within-subject factor, and with the timing of the explanations, the order of the conditions, and the image content as the between-subject factors. The Greenhouse–Geisser correction was used when needed in all reported results in this study.

The main effect of the condition (explanation vs. silent) was not significant ($F(1, 92) = 3.11$; $p = .081$). However, the interaction between the condition and the order of the different conditions in the blocks was significant ($F(1, 92) = 7.97$; $p = .006$; partial $\eta^2 = .08$; Figure 2): The effect of explaining is therefore dependent on whether it precedes or follows the silent block. No other significant effects were found. To clarify the result, a mixed ANOVA was then performed separately for two groups of participants with different orders of experimental conditions. Only those who completed the silent block first performed better in the block that required explanations ($F(1, 46) = 12.02$; $p = .001$; partial $\eta^2 = .21$), replicating our earlier result. The result for the other group was non-significant ($F(1, 46) = .50$; $p = .483$). It appears that the former group benefitted both from conscious reasoning and from learning in the second block,

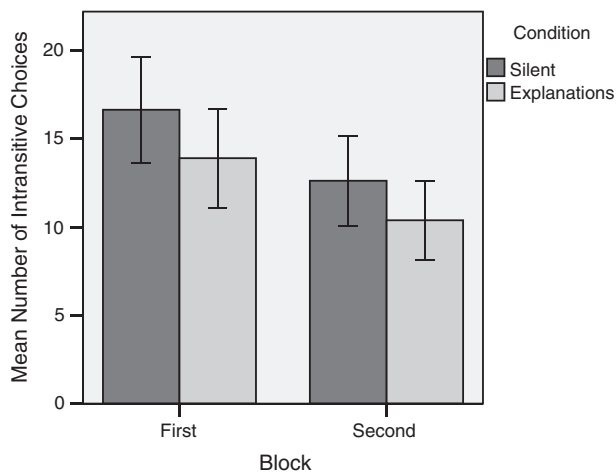


Figure 2. The mean number of intransitive choices in different conditions and blocks. Note the cross-over design: One group performed the silent task in the first block and the task requiring explanations in the second block. The order of the conditions in the blocks was reversed for the other group. The difference between the conditions existed only for the former group

whereas in the latter group, these effects neutralized each other, as they were present in different blocks.

When a mixed ANOVA with the same independent variables was applied for the participants' reliability, no significant main effect of condition was found ($F(1, 92) = 1.66$; $p = .20$). This time, the interaction between condition and the order of the different conditions in the blocks was also non-significant ($F(1, 92) = 3.20$; $p = .077$). All the variation in transitivity cannot be attributed to low reliability in task performance.²

Weighting of the attributes

The only significant result concerned the effect of importance order of the attributes on the attribute weights ($F(3, 276) = 701.55$; $p < .001$; partial $\eta^2 = .88$). Table 2 shows how the participants ordered the attributes.

Possible mediators

Correlational analysis supports our assumption that preference stability is associated with consistent attribute weighting and with choice reliability (Table 3). Also the number of intransitive choices is larger in trials involving less important attributes ($F(3, 276) = 28.53$; $p < .001$; partial $\eta^2 = .24$; Table 2).³ We were interested in whether differences in these factors between the silent and explanation conditions could explain the differences in preference stability in these conditions when the silent condition came first. Mediation analysis (Judd, Kenny, & McClelland, 2001), however,

²Following our earlier findings (Leisti et al., 2014), the effect of the condition still reached statistical significance when the silent condition was located in the first block ($F(1, 46) = 5.04$; $p = .030$; partial $\eta^2 = .01$).

³Repeated-measures ANOVA with the condition (silent vs. verbalization) and the importance of the attribute as within-subject factors, and the timing of the explanations as well as the orders of the condition and the image content as the between-subject factors.

Table 2. Participants' ordering of the attributes according to their importance (i.e., weight), including the mean number of the intransitive choice patterns associated with each rank

Content	Rank	Blur	Noise	Lightness	Color	Intransitive choices
Town	1st	3	66	24	10	14.29
	2nd	23	12	32	33	16.57
	3rd	32	9	29	30	17.06
	4th	42	13	17	27	17.30
Party	1st	1	18	49	35	11.86
	2nd	24	28	17	33	14.16
	3rd	32	29	22	15	14.65
	4th	43	25	13	18	14.02

did not support this pattern. Although there was a significant difference in choice reliability between the conditions (Table 4), it did not predict differences in intransitive choices (Table 5). With attribute weights the result was the opposite: Although no significant difference existed between the conditions, the difference in the weights of the most important attribute was significant.

Although the weights of the most important attribute differed only marginally between the conditions ($p = .061$), we performed a similar mediation analysis for the weight of the most important attribute. This time, the coefficient for the weight of the most important attribute differed from zero ($B = -12.19$; $\beta = -.91$; $t(48) = -16.50$; $p < .001$). Attribute weighting therefore appears to be a plausible candidate for a mediator between learning and more stable preferences.

Learning within blocks

A block with 55 trials is a rather coarse unit with which to examine learning. In addition, different blocks consisted of different image content types with different image manipulations, so the within-block and between-block learning should be slightly different. Therefore, we divided both blocks into five phases consisting of 11 trials—for a total of 55 trials—and calculated how the three choices in each intransitive pattern were divided between these phases. Figure 3 shows the division of these choices in five phases, suggesting that intransitivity decreases with learning within blocks. This was confirmed by the mixed ANOVA ($F(4, 96) = 8.73$; $p < .001$; partial $\eta^2 = .28$), with the log-transformed number of intransitive choices as a dependent variable, and with condition (explanation vs. silent), phase, and the order of the conditions and contents as independent variables.

We then performed mixed ANOVA with the attribute weight as the dependent variable and the condition, the importance of the attribute, and the phase as the within-subject factors, and with the order of the blocks and image content types as between-subject factors. The results show that the attribute weights were dependent on the importance of the attribute ($F(3, 276) = 1111.05$; $p < .001$; partial $\eta^2 = .92$) and on the phase ($F(4, 368) = 2.75$; $p = .03$; partial $\eta^2 = .03$), but a significant interaction existed between attribute importance and the phase ($F(12, 1104) = 2.74$; $p = .002$; partial $\eta^2 = .03$). The condition influenced weights through a rather

Table 3. Correlations of the log-transformed number of intransitive choices with reliability and the weights of the attributes with different levels of importance

	All conditions		Silent first		Explanations first	
	Explanation	Silent	Explanation	Silent	Explanation	Silent
Most important attribute	-.64**	-.65**	-.61**	-.67**	-.65**	-.63**
2nd most important attribute	-.36**	-.25*	-.32*	-.42**	-.40**	-.09
3rd most important attribute	-.41**	-.20*	-.45**	-.12	-.40**	-.29*
4th most important attribute	-.29**	-.19	-.45**	-.18	-.09	-.14
Reliability	-.30**	-.41**	-.10	-.31*	-.43**	-.50**

Note:
 * $p < .05$;
 ** $p < .01$.

Table 4. Significant differences between the first and second blocks of the experiment obtained via *t*-test

	Silent first		Explanations first	
	<i>t</i> (49)	<i>p</i>	<i>t</i> (49)	<i>p</i>
Intransitivity	-3.018	.004	1.067	.291
Reliability	-2.252	.029	.353	.726
Most important attribute	-1.919	.061	.350	.728
Second most important attribute	.403	.689	.488	.628
Third most important attribute	.457	.649	.047	.963
Fourth most important attribute	-1.561	.125	1.225	.226

Table 5. Results of the regression analysis predicting the change in the log-transformed number of intransitive choices between blocks

Variable	Reliability			Weight of the most important attribute		
	<i>B</i>	<i>SE</i> (<i>B</i>)	β	<i>B</i>	<i>SE</i> (<i>B</i>)	β
Difference	-.06	.07	-.13	-.43***	.09	-.59
Centered sum	.09	.07	-.18	.02	.08	.03
(Constant)	.44**	.15		.32*	.12	
<i>R</i> ²	.06			.34		

Note:
 * $p < .05$;
 ** $p < .01$;
 *** $p < .001$.

complicated interaction with the order of the conditions and the image content types ($F(1, 98) = 5.19$; $p = .025$; partial $\eta^2 = .05$). Upon closer examination, only the weight of the most important attribute differed between phases ($F(4, 368) = 11.89$; $p < .001$; partial $\eta^2 = .25$; Figure 4); the weights of the other attributes remained constant. Other effects were not found.

EXPERIMENT 2A

A significant interaction between the conditions and condition order in Experiment 1 supports our first and second hypotheses that both learning and explaining influence preference stability in the particular task. The within-block

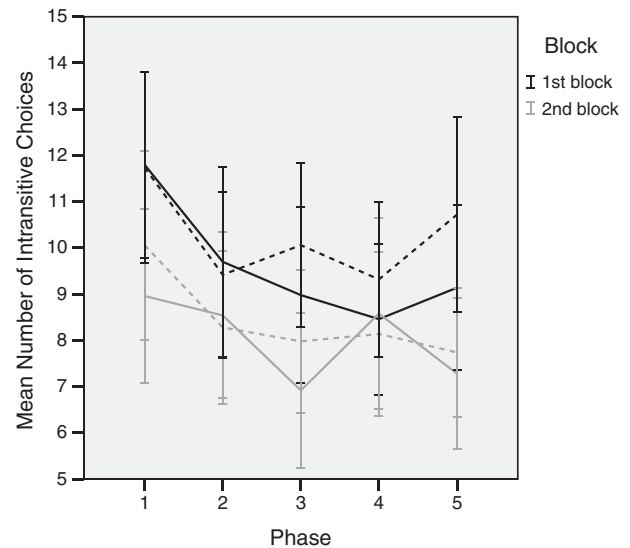


Figure 3. Distribution of choices present in intransitive preference patterns in phases 1–5. Solid lines indicate reasoning conditions and dashed lines silent conditions

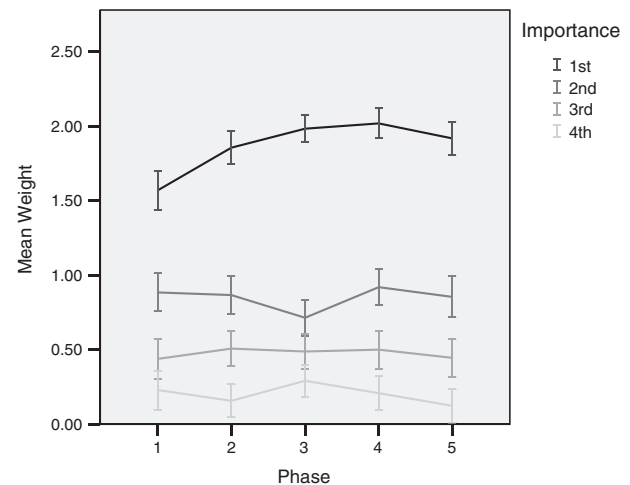


Figure 4. Development of the weights of the attributes in phases 1–5 of the experiment block

examination of intransitivity and of attribute weights further supported the role of learning in the task. The results also suggest that we overestimated the effect of conscious

reasoning in our previous studies. The result concerning learning, however, does not answer the question of why there was no difference between the silent and reasoning blocks when the reasoning block was presented/completed first. Did reasoning facilitate learning in this group; that is, did participants learn a better strategy when they were asked to explain their choices first and then apply this strategy in the subsequent, silent condition? Or do explaining and learning simply have an independent, additive effect on preference stability? In order to explore this, we conducted two additional experiments to isolate the effect of learning. Experiment 2A included explaining in both blocks; in Experiment 2B, both blocks were silent.

Methods

Participants

The participants were 63 university students, of whom 54 were female and 8 were male. They were recruited from university e-mail lists and were given one movie ticket for participation. The mean age was 25.5 years ($SD = 6.6$ years). Screening tests for near visual acuity, near contrast vision (near F.A.C.T.), and color vision (Farnsworth D-15) were administered to the participants before the experiment to ensure normal vision. All participants passed the tests, but the data from one participant were lost owing to a programming error and were therefore excluded.

Stimuli and procedure

The same stimulus images used in Experiment 1 were used in this experiment. The procedure was the same as in Experiment 1, except that the participants gave explanations in both blocks.

Results

We performed a mixed ANOVA with the block as a within-subject factor, and with the timing of the explanations and the order of the image contents as the between-subject factors. We found a significant learning effect; the log-transformed number of intransitive choices decreased from the first block ($M = 2.70$, $SD = .76$) to the second block ($M = 2.34$, $SD = .90$; $F(1, 60) = 6.71$; $p = .012$; partial $\eta^2 = .10$). Concerning the attribute weights, a significant main effect existed for both the block ($F(1, 58) = 8.69$; $p = .005$; partial $\eta^2 = .13$) and attribute importance ($F(3, 174) = 350.94$; $p < .001$; partial $\eta^2 = .86$). No difference was found in the participants' reliability between the blocks ($F(1, 58) = .42$; $p = .521$).

This time, the only factor that differed between blocks (in addition to image content) was learning within the experiment. We wanted to know how learning enhances participant performance; therefore, we performed a similar mediation examination as in Experiment 1. Table 6 shows the t -test results between blocks. The only candidate for the mediator was the weight of the most important attribute. We performed a regression analysis with a difference of log-transformed numbers of intransitive choices between blocks

Table 6. Significant differences between the first and second blocks of the experiment obtained via t -test

	$t(61)$	p
Intransitivity	2.60	.012
Reliability	-.661	.511
Most important attribute	-2.298	.025
Second most important attribute	-1.714	.092
Third most important attribute	-1.830	.072
Fourth most important attribute	-1.553	.126

as the dependent variable, and difference and the centered sum of the weights of the most important attribute between blocks as the predictors. According to the regression model ($R^2 = .62$), the coefficient for the difference of the weights significantly differed from zero ($B = -.57$; $t(61) = -9.68$; $p < .001$), but this was not the case for the intercept ($B = -.114$; $t(61) = -1.24$; $p = .221$) or the coefficient of the centered sum of the weights ($B = -.06$; $t(61) = -1.02$; $p = .314$). It appears, therefore, that learning affects the consistency of decision making in this task mainly through more consistent weighting of the attributes.

We also analyzed the within-block learning (Figure 5a). Following the method employed in Experiment 1, we divided both blocks into five phases. The log-transformed number of intransitive choices decreased within blocks ($F(4, 240) = 7.35$; $p < .001$; partial $\eta^2 = .11$). The attribute weights in the phases were dependent on attribute importance ($F(3, 174) = 518.705$; $p < .001$; partial $\eta^2 = .90$), on block ($F(1, 58) = 7.339$; $p = .009$; partial $\eta^2 = .11$), and on phase ($F(4, 232) = 6.013$; $p < .001$; partial $\eta^2 = .09$; Figure 5a).

EXPERIMENT 2B

Experiment 2A does not rule out the possibility that learning is different when the participants are asked to explain their choices. Therefore, one additional experiment was required. Experiment 2B was similar to Experiment 2A in all other respects except that all photo choices were made without having the participants provide an explanation.

Methods

Participants

The participants in this experiment were 63 university students (15 men and 48 women). They were recruited from university e-mail lists and were given one movie ticket for participation. The mean age was 26.1 years ($SD = 4.9$). The same vision tests were administered to the participants as in Experiment 2A. All participants passed the tests.

Stimuli and procedure

The experiment used the same stimuli and procedure as in Experiment 2A, except that all choices were made without explanation.

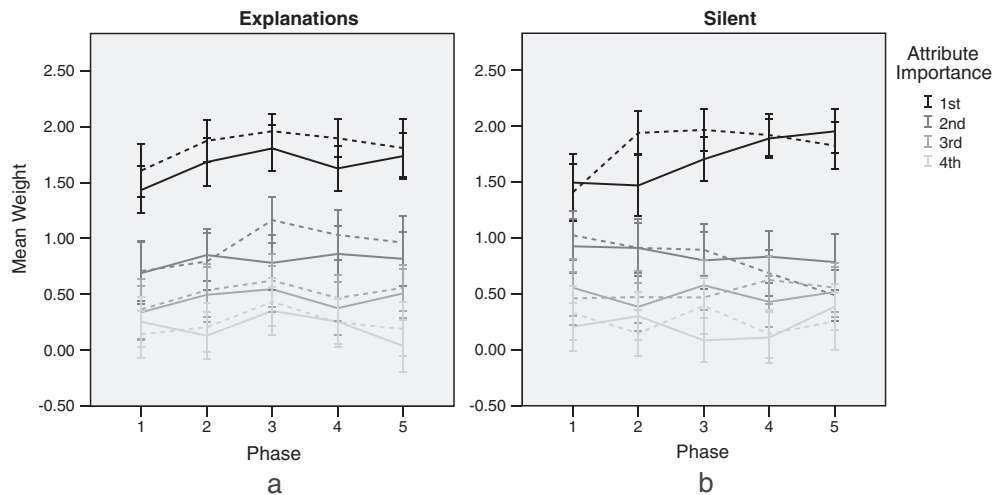


Figure 5. Weights of the attributes according to their importance in Experiment 2A (a) and in Experiment 2B (b). The solid lines indicate weights in the first block, and the dashed lines indicate weights in the second block

Results

The effect of learning in this case did not reach the level of statistical significance when measured as the log-transformed number of intransitive choices between blocks ($F(1, 61) = 2.62$; $p = .11$; first block: $M = 2.72$, $SD = .82$; second block: $M = 2.51$, $SD = .80$). In this experiment, learning did not even have a significant effect on the weights of the attributes across the blocks; the weights depended only on the attribute importance ($F(3, 180) = 345.05$; $p < .001$; partial $\eta^2 = .85$). No statistically significant change was found in participants' reliability either ($F(1, 61) = 2.34$; $p = .13$).

When the blocks were divided into five phases, no significant difference existed even between phases in the log-transformed number of intransitive choices ($F(4, 244) = 2.21$; $p = .073$). Concerning the differences in attribute weights, the result was more complicated: The main effect of phase was absent ($F(4, 240) = .41$; $p = .80$), but there was an interaction between the phase and the importance of the attribute ($F(12, 732) = 2.78$; $p = .001$; partial $\eta^2 = .04$). Figure 5b illustrates this interaction in which the weight of the most important attribute increases at the expense of the second most important attribute. A significant interaction also existed between the order of the image content and the block ($F(1, 61) = 8.39$; $p = .005$; partial $\eta^2 = .12$). When repeated-measures ANOVA was performed only for the weights of the most important attribute, a significant main effect for the phase ($F(4, 244) = 6.67$; $p < .001$; partial $\eta^2 = .10$) was found. This also applied to the second most important attribute ($F(4, 244) = 2.68$; $p = .032$; partial $\eta^2 = .04$) but with a decreasing linear trend ($F(1, 61) = 8.85$; $p = .004$). For the most important attribute, the block and the phase significantly interacted ($F(4, 244) = 3.38$; $p = .01$; partial $\eta^2 = .05$).

DISCUSSION

In our previous research, we found that having participants explain their choices in certain visual JDM tasks made their preferences more stable (Leisti & Häkkinen, 2016). The aim of this study was to examine whether this could be

attributed to learning and to the novelty of the task. In Experiment 1, we found support for our first hypothesis that learning increases participants' preference stability. Reasoning, however, benefitted preference stability as well. As learning has been associated with a more consistent weighting of decision attributes (Hoeffler & Ariely, 1999), we also examined the changes in the weights of the attributes, but the between-block differences in attribute weights in Experiment 1 did not reach statistical significance. Enhanced reliability did not explain the increased preference stability either. Within the blocks, learning was found to increase both preference stability and the weight of the most important attribute. After a certain level, further trials did not appear to yield any advantages. Regarding reasoning, Experiment 1 did not reveal a definite cause for its facilitative effect.

We performed two additional experiments to understand the effect of reasoning by isolating the effect of learning from reasoning. The first experiment, where participants were required to explain their choices in both blocks, showed that learning makes the participants' preferences more stable across the blocks, and this is caused by changes in the weights of the attributes. In the silent version of the experiment, however, the between-block learning concerning the stability of the preferences did not achieve statistical significance. The same also applied to the differences in the attribute weights between the blocks. Reasoning therefore appears to facilitate learning between blocks.

Our initial hypothesis that reasoning would lead to more stable preferences earlier in the experiment, when learning has less influence on decisions, was not supported by our results. Rather, the effect of reasoning was more evident in the latter part of Experiment 2A, in which the weights of the less important attributes remained at their earlier level, in comparison with Experiment 2B, in which the weight of the less important attributes decreased. This was accompanied by significant decreases in intransitive choices in Experiment 2A, both within and between blocks. It might be that reasoning did not have much of an effect in the earlier parts of the experiment because the participants were already using a deliberative strategy by thinking about the reasons for each

choice; the task was novel to them, and they did not have an existing way to approach it (Aarts, Verplanken, & van Knippenberg, 1998).

Reasoning facilitates the transfer of knowledge

The starting point of our experimentation was the adapted model of Dijkstra et al. (2013), which differentiates between explicit and implicit knowledge. The original version of this model suggested that reasoning distracts those individuals who have already gathered implicit knowledge but do not yet have adequate explicit knowledge to verbalize their judgments. The extension of this model concerned situations where the individual had gathered some explicit knowledge of the relevant judgment criteria but had no experience in the particular task.

Contrary to our initial hypothesis, reasoning started to benefit participants when the task was no longer completely novel, after several trials and in the second block. Explicit knowledge about quality attributes is not therefore enough; rather, some experience with the task is required. It appears that the more efficient performance required participants to first generate explicit knowledge and then to apply it deliberately in the subsequent trials. It has been presented that one of the benefits of reasoning or explaining is the transfer between different tasks (Baumeister & Masicampo, 2010; Fox et al., 2011). The transfer is relevant in these experiments also because the blocks were not exactly the same: Both blocks consisted of different image content types and image manipulations. In order to benefit from learning occurring in the first block, participants should have used their knowledge from the previous block to detect the differences between the alternatives and then reconsider their influence on the overall quality. It appears, however, that the participants in the silent experiment were paying less attention to less important attributes (in terms of attribute weights) when they proceeded to the second block, which may have left them uncertain about the effect of these attributes on their preferences. The weights of the attributes suggest that the participants in the second block of the silent Experiment 2B continued the routine they acquired in the first block. This assumption is supported by the interactions between the image content types and blocks on attribute weights, which were present in the silent Experiment 2B but absent in Experiment 2A. In other words, the performances of participants under the reasoning condition were less context dependent than were those of participants under the silent condition. It may also be that when participants under the silent condition acquired a strategy that was less compensatory, they had difficulty adapting to a new context that required attention to different attributes. Those under the reasoning condition, on the other hand, may have paid more attention also to less important attributes, which led them to give those attributes increased weight and made their choices more consistent.

The accumulation of task-specific explicit knowledge therefore appears to make adaptation to new tasks easier. The prerequisite for the positive transfer effect is that individuals can somehow relate this explicit conceptual knowledge to the new task. We showed in our earlier article that

reasoning facilitates preference stability only if the explanations are predictable, that is, if subjective attributes are used consistently with the objective properties of the alternatives when explaining the choices (Leisti & Häkkinen, 2016). This ability may reflect the successful transfer of knowledge across the blocks and the role of conscious reasoning in the context-independent application of learning in JDM tasks.

Much evidence suggests that reasoning may weaken the quality of JDM in nonexpert populations. Previous results, however, have concerned only unique judgments in certain fields, such as art, where participants do not have experience in reasoning. A more useful approach to reasoning in these contexts would take larger scale and learning into account. If reasoning in the course of a single judgment turns out to be a disadvantage, what occurs in the case of dozens of judgments that involve reasoning? Would the benefits of reasoning become visible when individuals have gathered more explicit knowledge in the same context? This definitely appears worth studying; according to the contemporary view (Baumeister & Masicampo, 2010), the function of conscious thinking is not to control single actions, but rather to make an individual's behavior coherent in the long run, and the results of this study support this view.

Reasoning prevents routine effects that decrease preference stability

In addition to transfer, explaining appears to maintain more compensatory strategies when participants acquire a routine with the choice task. Thus, learning without reasoning appears to denote routinization in which an economical search strategy gradually replaces a more thorough information search to preserve effort (Aarts et al., 1998; Betsch et al., 2001). Changes in the decision weights across the separate trials may also reflect attempts by the participant to make decisions as coherent as possible by increasing the weight of the most important attribute in comparison with less important attributes (Glöckner & Betsch, 2008; Simon, Pham, Le, & Holyoak, 2001), which simplifies the choice problem and information search for later trials such that a simple take-the-best heuristic can be used (Gigerenzer & Gaissmaier, 2011). At least in certain contexts, novices differ from experts in their use of compensatory strategies instead of more frugal noncompensatory strategies (Garcia-Retamero & Dhimi, 2009). Incorrect diagnoses of experienced medical doctors have been found to usually be caused by a failure to identify all of the relevant information, whereas the incorrect diagnoses of medical students are caused by the inability to integrate information to form the correct diagnosis (Groves, O'Rourke, & Alexander, 2003). Routine does not mean automatization of a compensatory strategy in our task; rather, the shift is towards a more noncompensatory strategy, which is probably a result of effort caused by the information search. Reasoning appears to prevent or delay this gradual change towards noncompensatory strategies. Reasoning may demand that the participants deliberately focus their attention on all available information and therefore avoid the harmful effects of routinization, such as ignoring relevant information or not

adapting their information search patterns according to changing task needs (Betsch et al., 2001).

Our approach cannot fully separate the roles of information integration and information acquisition. Additionally, we do not know about the causal relation between these processes: Did the decreased weights of the less important attributes in the latter part of the blocks follow from the weighting scheme, which aimed at a more efficient information search, or do the decreased weights simply reflect that these attributes were ignored in the information search and that most choices were thus made using information about the most important attribute?

Conclusions

Learning appears to have a significant effect on preference stability. Its effect is most profound for the weighting of information, which becomes more consistent when the individual acquires experience with trade-offs between attributes. Excessive routinization, however, can lead to ignoring the less important attributes, which may result in decreasing preference stability. Reasoning may help maintain a wider focus on several attributes, leading to a more compensatory-like strategy, and help in transferring the experience to similar tasks that are not exactly the same, thereby leading to an increased stability in preferences in these tasks. The results suggest that the benefits of reasoning manifest themselves as coherence in the macrolevel not in single decisions.

ACKNOWLEDGEMENTS

We are grateful to the anonymous reviewers for their helpful comments. This research was funded by Academy of Finland (Doctoral Program in User-Centered Information Technology) and by grants from the Alfred Kordelin Foundation, the Finnish Cultural Foundation, and the Ella and Georg Ehrnrooth Foundation. We would like to thank Jaakko Tähti and Erno Paunonen for conducting part of the experiments.

REFERENCES

- Aarts, H., Verplanken, B., & van Knippenberg, A. (1998). Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology*, 28(1355–1), 374–374. <https://doi.org/10.1111/j.1559-1816.1998.tb01681.x>.
- Amir, O., & Levav, J. (2008). Choice construction versus preference construction: The instability of preferences learned in context. *Journal of Marketing Research*, 45, 145–158. <https://doi.org/10.1509/jmkr.45.2.145>.
- Ariely, D., Loewenstein, G., & Prelec, D. (2003). “Coherent arbitrariness”: Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118, 73–106. <https://doi.org/10.1162/00335530360535153>.
- Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, 130, 641–657. <https://doi.org/10.1037/0096-3445.130.4.641>.
- Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review*, 117, 945–971. <https://doi.org/10.1037/a0019393>.
- Baumeister, R. F., Masicampo, E. J., & Vohs, K. D. (2011). Do conscious thought cause behavior? *Annual Review of Psychology*, 62, 331–361. <https://doi.org/10.1146/annurev.psych.093008.131126>.
- Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology*, 36, 209–231. <https://doi.org/10.1080/14640748408402156>.
- Berry, D. C., & Broadbent, D. E. (1987). Explanation and verbalization in a computer-assisted search task. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 39, 585–609. <https://doi.org/10.1080/14640748708401804>.
- Betsch, T. (2011). The stability of preferences—A social-cognition view. *Frontiers in Psychology*, 2, 1–3. <https://doi.org/10.3389/fpsyg.2011.00290>.
- Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., & Fiedler, K. (2001). The effects of routine strength on adaptation and information search in recurrent decision making. *Organizational Behavior and Human Decision Processes*, 84, 23–53. <https://doi.org/10.1006/obhd.2000.2916>.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. London, UK: Routledge & Kegan Paul.
- Bröder, A., Glöckner, A., Betsch, T., Link, D., & Ettlin, F. (2013). Do people learn option or strategy routines in multi-attribute decisions? The answer depends on subtle factors. *Acta Psychologica*, 143, 200–209. <https://doi.org/10.1016/j.actpsy.2013.03.005>.
- Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French, & A. Cleeremans (Eds.), *Implicit learning and consciousness* (, pp. 1–40). Hove, UK: Psychology Press.
- David, H. A. (1969). *The method of paired comparisons*. London, UK: Charles Griffin & Company Limited.
- Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, 1, 95–109. <https://doi.org/10.1111/j.1745-6916.2006.00007.x>.
- Dijksterhuis, A., & van Olden, Z. (2006). On the benefits of thinking unconsciously: Unconscious thought can increase post-choice satisfaction. *Journal of Experimental Social Psychology*, 42, 627–631. <https://doi.org/10.1016/j.jesp.2005.10.008>.
- Dijkstra, K. A., Pligt, J., & Kleef, G. A. (2013). Deliberation versus intuition: Decomposing the role of expertise in judgment and decision making. *Journal of Behavioral Decision Making*, 26, 285–294. <https://doi.org/10.1002/bdm.1759>.
- Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual task investigation. *Journal of Memory and Language*, 48, 148–168. [https://doi.org/10.1016/S0749-596X\(02\)00511-9](https://doi.org/10.1016/S0749-596X(02)00511-9).
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223–241. <https://doi.org/10.2139/ssrn.1090866>.
- Fox, C. M., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting practices. *Psychological Bulletin*, 137, 316–344. <https://doi.org/10.1037/a0021663>.
- Garcia-Retamero, R., & Dhami, M. K. (2009). Take-the-best in expert-novice decision strategies for residential burglary. *Psychonomic Bulletin & Review*, 16, 163–169. <https://doi.org/10.3758/PBR.16.1.163>.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>.
- Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1055–1075. <https://doi.org/10.1037/0278-7393.34.5.1055>.
- Glöckner, A., & Witteman, C. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgement and decision making. *Thinking & Reasoning*, 16, 1–25. <https://doi.org/10.1080/13546780903395748>.

- Groves, M., O'Rourke, P., & Alexander, H. (2003). Clinical reasoning: The relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. *Medical Teacher*, 25, 621–625. <https://doi.org/10.1080/01421590310001605688>.
- Hagafors, R., & Brehmer, B. (1983). Does having to justify one's judgments change the nature of the judgment process? *Organizational Behavior and Human Decision Processes*, 32, 223–232. [https://doi.org/10.1016/0030-5073\(83\)90122-8](https://doi.org/10.1016/0030-5073(83)90122-8).
- Haider, H., & Frensch, P. A. (1999). Information reduction during skill acquisition: The influence of task instruction. *Journal of Experimental Psychology: Applied*, 5, 129–151. <https://doi.org/10.1037/1076-898X.5.2.129>.
- Hamilton, R., Hong, J., & Chernev, A. (2007). Perceptual focus effects in choice. *Journal of Consumer Research*, 34, 187–199. <https://doi.org/10.1086/519147>.
- Hoeffler, S., & Ariely, D. (1999). Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology*, 8, 113–139. https://doi.org/10.1207/s15327663jcp0802_01.
- Hoeffler, S., Ariely, D., West, P., & Duclos, R. (2013). Preference exploration and learning: The role of intensiveness and extensiveness of experience. *Journal of Consumer Psychology*, 23(3), 330–340. <https://doi.org/10.1016/j.jcps.2012.10.007>.
- Hogarth, R. M. (2001). *Educating intuition*. Chicago, IL, USA: University of Chicago Press.
- Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods*, 6, 115–134. <https://doi.org/10.1037/1082-989X.6.2.115>.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *The American Psychologist*, 58, 697–720. <https://doi.org/10.1037/0003-066X.58.9.697>.
- Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *The American Psychologist*, 64, 515–526. <https://doi.org/10.1037/a0016755>.
- Lee, L., Amir, O., & Ariely, D. (2009). In search of homo economicus: Cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research*, 36, 173–187. <https://doi.org/10.1086/597160>.
- Leisti, T., & Häkkinen, J. (2016). The effect of introspection on judgment and decision making is dependent on the quality of conscious thinking. *Consciousness and Cognition*, 42, 340–351. <https://doi.org/10.1016/j.concog.2016.04.008>.
- Leisti, T., Radun, J., Virtanen, T., Nyman, G., & Häkkinen, J. (2014). Concurrent explanations can enhance visual decision making. *Acta Psychologica*, 145, 65–74. <https://doi.org/10.1016/j.actpsy.2013.11.001>.
- Lerner, S. J., & Tetlock, E. P. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125, 255–275. <https://doi.org/10.1037/0033-2909.125.2.255>.
- Levine, G. M., Halberstadt, J. B., & Goldstone, R. L. (1996). Reasoning and the weighting of attributes in attitude judgments. *Journal of Personality and Social Psychology*, 70, 230–240. <https://doi.org/10.1037/0022-3514.70.2.230>.
- McGlone, M., Kobrynowics, D., & Alexander, R. (2005). A certain je ne sais quoi: Verbalization bias in evaluation. *Human Communication Research*, 31, 241–267. <https://doi.org/10.1111/j.1468-2958.2005.tb00871.x>.
- Melcher, J. M., & Schooler, J. W. (2004). Perceptual and conceptual training mediate the verbal overshadowing effect in an unfamiliar domain. *Memory & Cognition*, 32, 618–631. <https://doi.org/10.3758/BF03195853>.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259. <https://doi.org/10.1037/0033-295X.84.3.231>.
- Orquin, J. L., Bagger, M. P., & Mueller Loose, S. (2013). Learning affects top down and bottom up modulation of eye movements in decision making. *Judgment and Decision making*, 8, 700–716.
- Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research*, 28, 167–188. <https://doi.org/10.1086/322896>.
- Radun, J., Leisti, T., Virtanen, T., Häkkinen, J., Vuori, T., & Nyman, G. (2010). Evaluating the multivariate visual quality performance of image-processing components. *ACM Transactions on Applied Perception*, 7, 3–16. <https://doi.org/10.1145/1773965.1773967>.
- Radun, J., Nuutinen, M., Leisti, T., & Häkkinen, J. H. (2016). Individual differences in image-quality estimations: Estimation rules and viewing strategies. *ACM Transactions on Applied Perception*, 13, 14:1–14:22. <https://doi.org/10.1145/2890504>.
- Rusou, Z., Zakay, D., & Usher, M. (2013). Pitting intuitive and analytical thinking against each other: The case of transitivity. *Psychonomical Bulletin & Review*, 20, 1–7. <https://doi.org/10.3758/s13423-013-0382-7>.
- Schmidt, H. G., & Rikers, R. M. J. P. (2007). How expertise develops in medicine: Knowledge encapsulation and illness script formation. *Medical Education*, 41, 1133–1139. <https://doi.org/10.1111/j.1365-2923.2007.02915.x>.
- Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, 49, 11–36. [https://doi.org/10.1016/0010-0277\(93\)90034-S](https://doi.org/10.1016/0010-0277(93)90034-S).
- Sieck, W., & Yates, F. J. (1997). Exposition effects on decision making: Choice and confidence in choice. *Organizational Behavior and Human Decision Processes*, 70, 207–219. <https://doi.org/10.1006/obhd.1997.2706>.
- Simon, D., Pham, L. B., Le, Q. a., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1250–1260. <https://doi.org/10.1037/0278-7393.27.5.1250>.
- Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16, 158–174. <https://doi.org/10.1086/209205>.
- Simonson, I., & Nowlis, S. M. (2000). The role of explanations and need for uniqueness in consumer decision making: Unconventional choices based on reasons. *Journal of Consumer Research*, 27, 49–68. <https://doi.org/10.1086/314308>.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48. <https://doi.org/10.1037/h0026750>.
- Verplanken, B., Aarts, H., & van Knippenberg, A. (1997). Habit, information acquisition, and the process of making travel mode choices. *European Journal of Social Psychology*, 27, 539–560. [https://doi.org/10.1002/\(SICI\)1099-0992\(199709/10\)27:5%3C539::AID-EJSP831%3E3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-0992(199709/10)27:5%3C539::AID-EJSP831%3E3.0.CO;2-A).
- Wilson, T. D., Dunn, D., Kraft, D., & Lisle, D. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. *Advances in Experimental Social Psychology*, 22, 287–343. [https://doi.org/10.1016/S0065-2601\(08\)60311-1](https://doi.org/10.1016/S0065-2601(08)60311-1).
- Wilson, T. D., Lisle, D. J., Schooler, J. W., Hodges, S. D., Klaaren, K. J., & LaFleur, S. J. (1993). Introspecting about reasons can reduce post-choice satisfaction. *Personality and Social Psychology Bulletin*, 19, 331–339. <https://doi.org/10.1177/0146167293193010>.
- Yamada, A. (2009). Appreciating art verbally: Verbalization can make a work of art be both undeservedly loved and unjustly maligned. *Journal of Experimental Social Psychology*, 45, 1140–1143. <https://doi.org/10.1016/j.jesp.2009.06.016>.

Authors' biographies:

Tuomas Leisti, MA (psychology), is a doctoral student at University of Helsinki. He is currently finishing his doctoral dissertation about the role of conscious thought in subjective preferences. His

research concerns subjective quality perception and accompanying judgment and decision making processes.

Jukka Häkkinen, received PhD in experimental psychology at University of Helsinki, Finland. Currently he is a Principal Investigator at the Department of Psychology and Logopedics at University of Helsinki, where he leads Visual Cognition ResearchGroup.

Authors' addresses:

Tuomas Leisti, Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland.

Jukka Häkkinen, Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland.