

RUNNING HEAD: INDIVIDUAL DIFFERENCE FACTORS IN PATTERNING

Individual difference factors in the learning and transfer of patterning discriminations

Elisa Maes¹, Elias Vanderoost¹, Rudi D’Hooge², Jan De Houwer³ and Tom Beckers¹

¹ Centre for the Psychology of Learning and Experimental Psychopathology, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

² Laboratory for Biological Psychology, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

³ Learning and Implicit Processes Lab, Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent Belgium

Correspondence:

Tom Beckers, Centre for the Psychology of Learning and Experimental Psychopathology, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, box 3712, 3000 Leuven, Belgium; tom.beckers@kuleuven.be

In press, *Frontiers in Psychology*

Abstract

In an associative patterning task, some people seem to focus more on learning an overarching rule, whereas others seem to focus on acquiring specific relations between the stimuli and outcomes involved. Building on earlier work, we further investigated which cognitive factors are involved in feature- versus rule-based learning and generalization. To this end, we measured participants' tendency to generalize according to the rule of opposites after training on negative and positive patterning problems (i.e., A+/B+/AB- and C-/D-/CD+), their tendency to attend to global aspects or local details of stimuli, their systemizing disposition and their score on the Raven intelligence test. Our results suggest that while intelligence might have some influence on patterning learning and generalization, visual processing style and systemizing disposition do not. We discuss our findings in the light of previous observations on patterning.

Keywords: associative learning, patterning, rule-based generalization, feature-based generalization, visual processing style, mental representations

Qualitative differences exist between individuals in what they learn during associative learning tasks, even if their performance appears superficially similar. For example, in a concurrent associative patterning task, people are trained on negative and positive patterning problems simultaneously. In the negative patterning problems, stimuli (e.g., A and B) predict a certain outcome when presented alone, but not when presented in compound (A+/ B+/ AB-). In the positive patterning problems, presentation of a compound of two stimuli predicts the outcome, whereas presentation of the individual stimuli does not (e.g., C-/ D-/ CD+). To solve these simultaneous discrimination problems, people can either learn to associate each specific stimulus configuration presented to them with the absence or presence of the outcome or they can abstract the underlying rule of opposites ('compounds have the opposite outcome as the individual stimuli they are composed of'). Both strategies will result in the same response pattern. However, differences will emerge when participants are subsequently confronted with novel stimulus configurations. That is, when additionally trained on an incomplete pattern (e.g., E+/ F+), they will respond in different ways to novel configurations derived from the incomplete training stimuli (e.g., EF). Participants who abstracted a rule of opposites during training can apply this rule to the novel configuration and may as a consequence predict no outcome on EF trials after E+/ F+ training. In contrast, participants who simply learned to relate each stimulus configuration with its respective outcome without extracting an underlying rule can base their response to the novel configuration only on the degree of featural similarity between the novel configuration and the trained stimuli. Consequently, those participants will tend to predict the outcome on EF trials after E+/ F+ training (Shanks & Darby, 1998). In previous studies, about half of participants exhibited rule-based rather than feature-based generalization after intermixed training on positive and negative patterning problems (Maes et al., 2015; Wills, Graham, Koh, McLaren, & Rolland, 2011; see further analysis reported in Wills, 2014).

The question is now which factors might correlate with those individual differences in performance during negative and positive patterning discrimination and generalization. Performance at the end of training seems to have an important influence on generalization strategy; participants achieving a high level of terminal performance during training showed a stronger tendency to respond in line with rule-based generalization during subsequent generalization testing (Shanks & Darby, 1998). Yet other results suggest that high training performance is not the only factor that determines generalization performance (Little & McDaniel, 2015; McDaniel et al., 2014; Wills, Barrasin, & McLaren, 2011; Wills, Graham, et al., 2011; Winman, Wennerholm, Juslin & Shanks, 2005). In Experiment 2B of Maes and

colleagues (2015), only half of the participants showed rule-based responding in a patterning task, despite the fact that all participants were trained to a performance criterion of at least 75%. Further, Wills and colleagues (Wills, Graham, et al., 2011) observed that participants who received a working memory load during patterning training showed feature-based generalization despite reaching high levels of performance on the training items. Those results indicate that training accuracy is not the only determining factor in generalization style and that qualitatively different forms of learning might underlie similar levels of training performance; availability of working memory resources may be one important factor determining learning and subsequent generalization. Several other studies have also indicated that cognitive resources influence generalization strategy. For example, participants with a high working memory capacity displayed generalization consistent with the opposites rule, while participants with a low working memory displayed generalization consistent with featural overlap (Wills, Barrasin, & McLaren, 2011). In a recent study of Cobos and colleagues (Cobos, Gutiérrez-Cobo, Morís, & Luque, 2017) individuals employed a feature-based generalization strategy when under strict time constraints, whereas individuals not under this strict time constraint demonstrated rule-based generalization. Further, results from the function-learning and categorization literature suggest that fluid intelligence might be another factor that determines generalization style (Little & McDaniel, 2015; McDaniel et al., 2014). However, in the category learning tasks involved, participants needed to pay attention to multiple dimensional features (e.g., color and shape) of a stimulus at the same time in order to classify stimuli correctly. It could be argued that this might have increased the role of intelligence in performance. The same argument can be made about the function learning tasks, because participants needed to map continuous inputs to continuous outputs. Such issues are not at play in patterning tasks that use dichotomous outcomes and involve stimuli that are not dimensional (e.g., food items). As such, the role of intelligence in determining generalization style might be overestimated from the experiments described above. Still, involvement of intelligence in generalization of patterning might be expected, based on at least three observations: (1) A correlation has been reported between the performance in concurrent negative and positive patterning and scores on the Raven's Advanced Progressive Matrices (Winman, Wennerholm, Juslin & Shanks, 2005). From the report by Winman et al. (2005), however, it is unclear whether the patterning performance refers to performance during training, performance on generalization test trials or to rule awareness. (2) As mentioned above, working memory capacity has been shown to modulate feature- versus rule-based generalization, as do time constraints (Cobos et al., 2017; Wills, Barrasin & McLaren, 2011; Wills, Graham, et al., 2011). Working memory capacity and

general intelligence are correlated (for a review, see Conway, Kane, & Engle, 2003). (3) It has been found that rule-based generalizers had greater middle frontal cortex activity than feature-based generalizers in a patterning task (Milton et al., 2017), which might suggest stronger activation of higher cognitive functions in participants demonstrating rule-based generalization in a concurrent patterning task (Fuster, 2001; Siddiqui, Chatterjee, Kumar, Siddiqui, & Goyal, 2008). A first goal of the research reported here was therefore to investigate whether the inclination towards feature- or rule-based generalization after concurrent negative and positive patterning training would indeed be associated with intelligence.

Other factors might also be at play during patterning tasks. Byrom (2013) has argued that negative patterning discrimination requires learning about a configuration independently from learning about its constituent stimuli. Therefore, it was hypothesized that a visual tendency to perceive groups of stimuli as a unitary configuration rather than a cluster of co-occurring stimuli should improve negative patterning performance. As previous work showed that individuals differ in their tendency to focus on global information versus local details (Navon, 1977), it was argued that this variation might be related to variation in learning non-linear discriminations. In line with this idea, Byrom and Murphy (2014) found that people who have a tendency to focus on global aspects rather than local stimulus details (as measured with a Navon task, see below) discriminated better between BC and ABC in a modified negative patterning task (A+/ BC+/ ABC-). Based on this finding, in the current experiment, we investigated whether people with a more global processing style are also better in discriminating negative and positive patterning problems than people with a more local processing style. Importantly, the advantage of configural processing over elemental processing should if anything be stronger for simultaneous negative and positive patterning than for negative patterning alone. Exploratorily, we also assessed whether visual processing style was related to individual differences in feature- versus rule-based generalization.

Finally, for exploratory reasons, we investigated whether systemizing is associated with rule-based generalization. Systemizing is the drive to analyze a system and derive the underlying rules that govern the behavior of the system (Baron-Cohen, 2002). The term system is used in a very broad sense here, and includes technical systems (e.g., a computer, a musical instrument), natural systems (e.g., the weather), and abstract systems (e.g., mathematics, syntax). Systemizing is thought to be an inductive processes that starts with data gathering and results in a rule about how the system works (Baron-Cohen, 2002; Baron-Cohen & Wheelwright, 2003). Therefore, a correlation between systemizing and generalization strategy might be expected.

In the current experiment, participants were first trained on complete and incomplete negative and positive patterning problems simultaneously (see Table 1). Thereafter, they were tested for generalization. Participants then performed a Navon task (to measure global versus local visual processing style), followed by the Systemizing Quotient-Revised (SQ-R) questionnaire to measure systemizing (Wheelwright et al., 2006). Finally, they completed a computerized version of the Raven Standard Progressive Matrices test (RSPM; Raven, 1958) to measure non-verbal intelligence.

Method

Participants and apparatus

Participants were 60 healthy volunteers (15 male, $M_{\text{age}} = 21$ years) who received either 12 euros or course credits for an undergraduate psychology course for their participation. This sample size ensured an a-priori power of .82 to obtain a similar effect in size as the correlation between visual processing style and patterning discrimination reported by Byrom and Murphy ($R^2 = 0.13$) at $\alpha = .05$ (the size of the other potential effects of interest is difficult to gauge from existing research, given the very different tasks used there). All tasks were presented on a computer running Affect software (Spruyt et al., 2010).

Stimuli and Materials

Patterning task

For half of the participants, Stimuli A-P were pictures of fruits, respectively, pineapple, cherry, strawberry, apple, banana, kiwi, raspberry, orange, passion fruit, grape, prune, gooseberry, melon, blackberry, lemon, and mango depicted in a can. For the remaining participants, the foods assigned to A and B were swapped with those assigned to C and D, and similarly for the other sets. Participants could respond by clicking on the “happy” or “sad” button presented on the screen. Further, participants could see the total number of points they gained.

Navon task

The stimuli used in this task closely resembled the stimuli used by Byrom and Murphy (2014). All stimuli consisted of large letters (S or H) composed of smaller letters (S or H), yielding four different stimuli (see Figure 1). Stimuli were presented in a black square spanning 55 mm x 55 mm ($12.6^\circ \times 12.6^\circ$), with the small letters spanning approximately 5 mm x 5 mm ($1.2^\circ \times 1.2^\circ$).

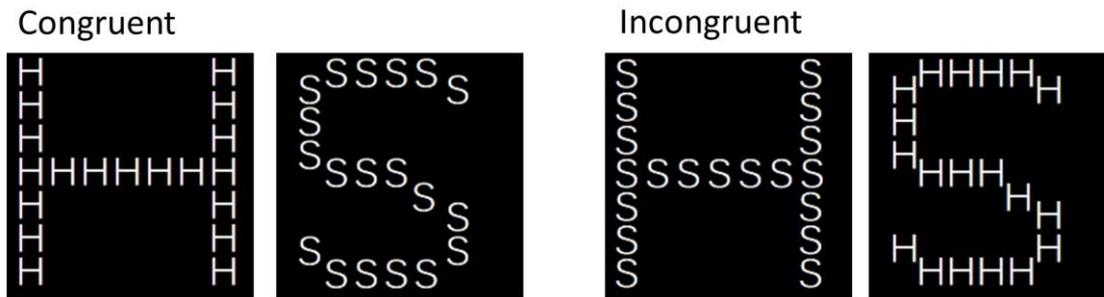


Figure 1. Stimuli used in the Navon task.

SQ-R questionnaire

A purpose-made Dutch translation of the SQ-R questionnaire (Wheelwright et al., 2006) was used to measure systemizing. The SQ-R consists of 75 questions (e.g. “I find it easy to use train timetables, even if this involves several connections”, “When I learn a language, I become intrigued by its grammatical rules.”) answered on a 4-point scale ranging from “strongly agree” to “strongly disagree” and scored from 0 to 2. Total scores on the test range from 0 (low interest in systemizing) to 150 (maximum score, extremely high systemizing).

RSPM

A computerized version (Mindware, Geldermalsen, The Netherlands) of the Raven Standard Progressive Matrices (RSPM) test (Raven, 1958) was used to measure intelligence.

Procedure

All participants gave written informed consent for their participation and the procedure was approved by the social and societal ethics committee of the KU Leuven. All experiments were embedded in a cover story describing a pharmaceutical company that developed new kinds of beverages that could change the mood of a consumer to “happy” or “sad”.

Patterning task

The participants were informed that they needed to learn which beverages (indicated by different tastes of fruit) would lead to a happy mood after consumption and which ones would lead to a sad mood; they would gain 20 points for each correct answer (see Figure 2 for a visual representation of the stimulus display in a single training trial). Participants received up to ten blocks of training, each block comprising two presentations of the 18 training trial types shown in Table 1 in a random order. No time limit for responding was imposed. During training trials

a feedback message was displayed during at least 1500 ms, after which participants could press enter to move on to the next trial. Participants moved on to the test phase when they reached a criterion of at least 32 correct responses in a given 36-trial block (89% correct). The transition to the test phase was accompanied by instructions that feedback would no longer be provided, but that the computer would still keep track of participants' scores. The test phase consisted of two blocks, each comprising one presentation of the 24 test trial types shown in Table 1, in a random order.

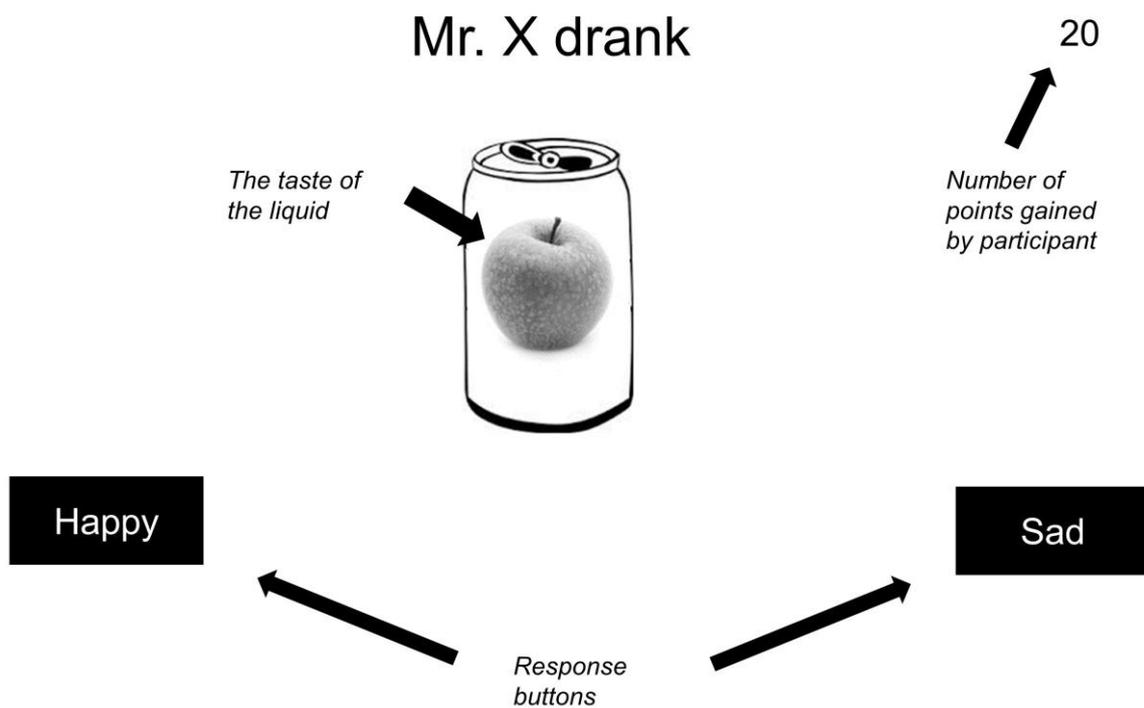


Figure 2. Stimuli as presented on the computer screen during a pattering trial (translated to English). Information in italics is added. Note that colors were removed and that the stimulus dimensions are increased to improve the readability, and are not proportional to the actual dimensions used.

Navon task

The procedure used in this task resembled the procedure used by Byrom and Murphy (2014) as closely as possible. Participants were informed that they would be presented with large letters (S or H) composed out of smaller letters (S or H) and that, in successive trial blocks, they would have to identify either the large (global block) or the small letter (local block) by pressing the S or H key as quickly as possible. Half of the participants started out identifying the large letter, whereas the other half started out identifying the small letter. Global and local blocks alternated until participants completed four blocks of each type. Each 16-trial block

contained four trials of each stimulus type (congruent S, congruent H, incongruent S and incongruent H). Each trial began with the presentation of a fixation cross for 500 ms, followed by the presentation of a single stimulus in the center of the screen. After 40 ms the stimulus disappeared and a masking stimulus was presented until the participant made a response. Trial order within each block was semi-random with no more than two consecutive trials of the same stimulus type.

After completing the patterning task and the Navon task, participants were asked to fill in the SQ-R. Finally, they completed the RSPM task. No time limit was set to complete the different parts of the experiment.

Table 1

Training and test trials for the patterning task; A-P represent different flavours of beverages; + indicates that drinking the beverage results in a happy mood; - indicates that drinking the beverage results in a sad mood. Trials in bold are the crucial generalization test trials

Training			Test		
A+	B+	AB-	A	B	AB
C-	D-	CD+	C	D	CD
E+	F+	EF-	E	F	EF
G-	H-	GH+	G	H	GH
I+	J+		I	J	IJ
		KL-	K	L	KL
M-	N-		M	N	MN
		OP+	O	P	OP

Results

Data and analysis scripts are available at Open Science Framework (https://osf.io/48tc3/?view_only=e3e2433f3c0544e68de3dced6d36952e). Statistical analyses were conducted using both frequentist statistical techniques and Bayesian hypothesis testing. A Bayes Factor (BF) quantifies the strength of the relative statistical evidence for two rivaling hypotheses. It expresses the relative probability of the data under, e.g., the null versus the alternative hypothesis (Dienes, 2011; Gallistel, 2009; Morey, 2015; Rouder, Speckman, Sun, Morey, & Iverson, 2009). If a BF of about 1 is obtained, there is no evidence in favor of either one of the hypotheses, BFs above three can be regarded to provide substantial evidence in favor

of the hypothesis that is in the numerator (or, conversely, values below one third provide substantial evidence for the hypothesis in the denominator). We calculated BFs using JAPS 0.8.1.2 (JASP Team, 2017) and assuming a default prior distribution (Ly, Verhagen, & Wagenmakers, 2016).

Patterning task

None of the participants failed to meet the training criterion within the ten blocks available; on average participants took 3.95 blocks ($SD = 1.72$). We further aimed at characterizing the learning process in more detail. More precisely, we were interested in investigating whether people tended to learn patterning problems in a stimulus-by-stimulus manner or an “once-for-all” manner. If the process of acquiring correct responding is a stimulus-by-stimulus process, performance is expected to increase gradually. In contrast, if the process is a one-for-all process, a sudden jump in performance is to be expected. This jump might indicate that the participant suddenly applied the rule of opposites, because rule application will boost performance (cfr. Mutter, Strain, & Plumlee, 2007). In order to exploratorily investigate the acquisition pattern, we analyzed the difference in percentage correct responses between the penultimate ($M = 76.57$, $SD = 9.17$) and last training block ($M = 93.24$, $SD = 3.17$). A paired t-test confirmed that this difference was significant, $t(59) = -13.94$, $p < .01$, 95% $CI [-19.06, -14.27]$, $d = -1.80$, $BF_{10} > 100$. Although further research is required to confirm this exploratory result, it suggests that acquiring the correct responses in patterning learning may be an “once-for-all” process.

For each participant, the percentage of rule-based responses on the critical generalization trials was calculated as the sum of the number of rule-based responses to the elements (K, L, O, P) divided by 8 (the number of element presentations) and the number of rule based responses to the compounds (IJ, MN) divided by 4 (the number of compound presentations). This sum was divided by 2 and multiplied by 100, such that 100% indicates rule-based responding on all trials and 0% indicates feature-based responding on all trials. At group level, participants exhibited a rule-based generalization strategy ($M = 61.77\%$, $SD = 31.85\%$), one-sample t-test comparing mean with 50%: $t(59) = 2.86$, $p < .01$, 95% $CI [53.54, 70.00]$, $d = 0.37$, $BF_{10} = 5.62$. It is clear from Figure 3, which depicts the distribution of percentage of rule-based responses, that generalization strategy was not simply bimodal (i.e., participants did not apply either a completely feature-based or a completely rule-based strategy). Rather a lot of variability exists in the percentage of rule-based generalization responses among participants with only 6.67% employing a complete feature-based strategy and 13.33% employing a

complete rule-based strategy. Thirty-seven participants could be categorized as rule-based generalizers ($> 50\%$ rule-based responses during generalization test), while 23 participants could be categorized as non-rule-based generalizers ($\leq 50\%$ rule-based responses during generalization test).

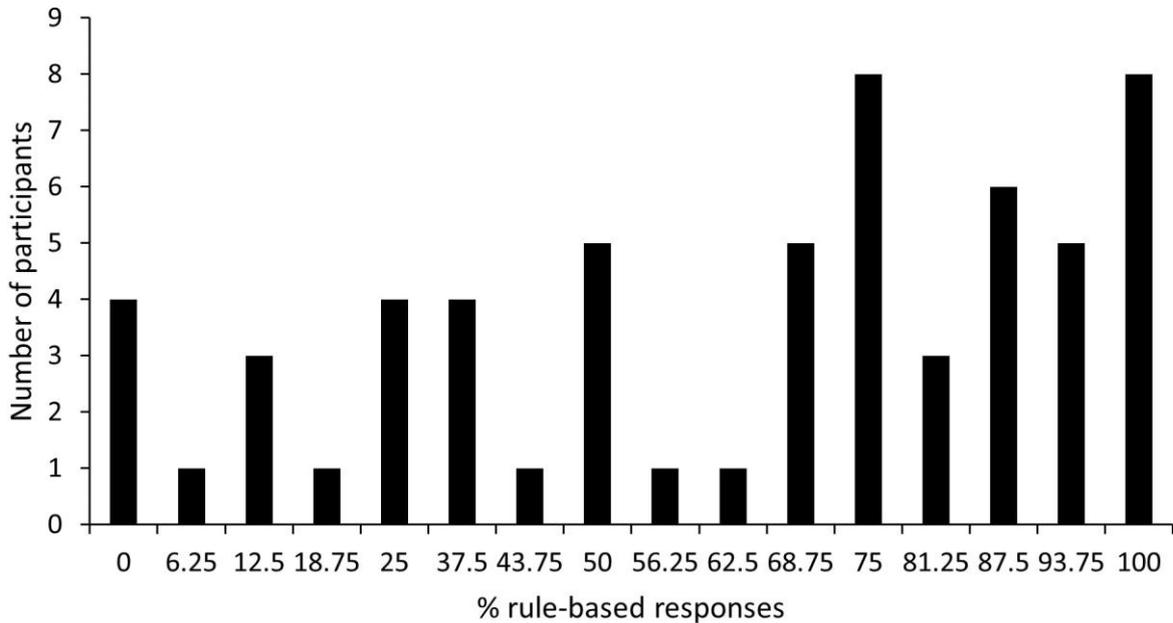


Figure 3. Distribution of percentage of rule-based generalization responses.

We observed a positive correlation between terminal accuracy (i.e., the percentage of correct responses in the last training block; $M = 93.24\%$, $SD = 3.17\%$) and percentage of rule-based responses on generalization test trials, $r = .34$, $p < .01$, 95% $CI [0.10, 0.55]$, $BF_{10} = 5.53$ (Figure 4). Learning speed (i.e., the number of training blocks needed to reach criterion; $M = 3.95$; $SD = 1.72$) and percentage of rule-based generalization responses were not significantly correlated, $r = -.16$, $p = .22$, 95% $CI [-0.40, 0.10]$, $BF_{10} = 0.34$.

It is conceivable that rule-based and non-rule-based generalizers differ in their acquisition pattern with rule-based generalizers showing a more “once-for-all” pattern and non-rule-based generalizers a stimulus-by-stimulus pattern. In order to investigate this, a repeated-measurements ANOVA with block as within-subjects variable and generalization type as between-subjects variable was conducted (Figure 5). An effect of Block was observed, $F(1, 58)$

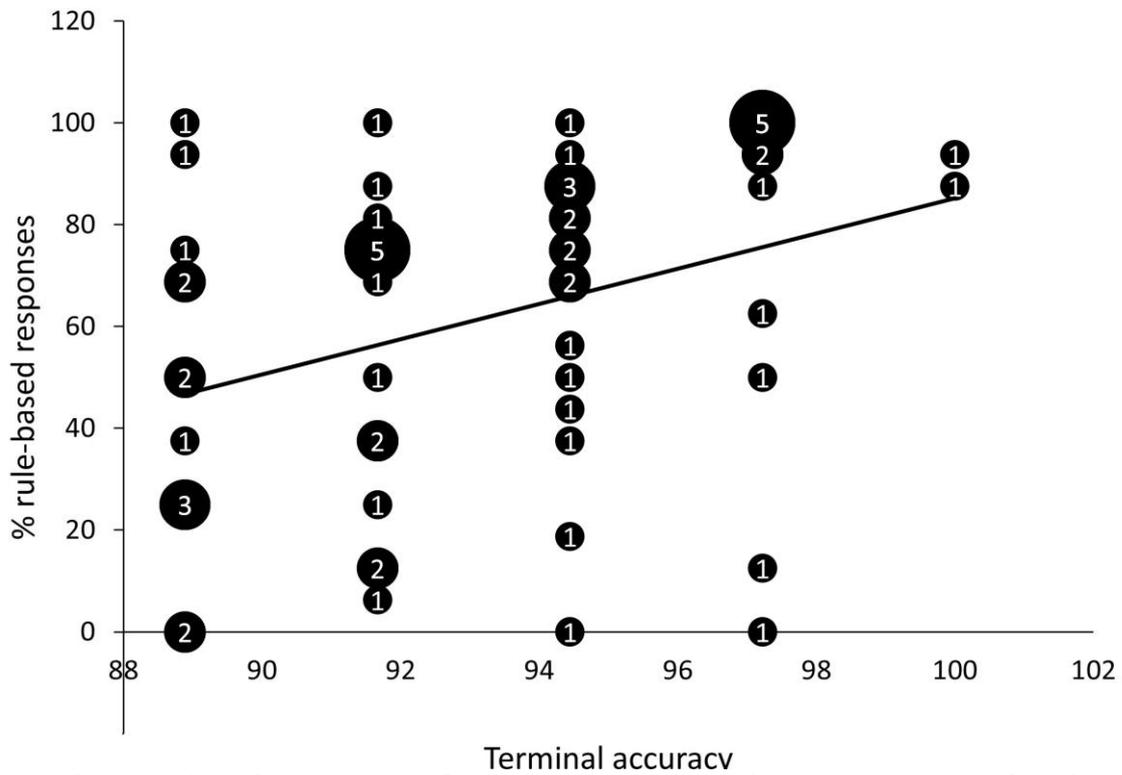


Figure 4. Scatter plot of percentage of rule-based generalization responses as a function of terminal accuracy (percentage of correct responses in last training block). The numbers presented in the bubbles reflect the number of participants with those particular scores.

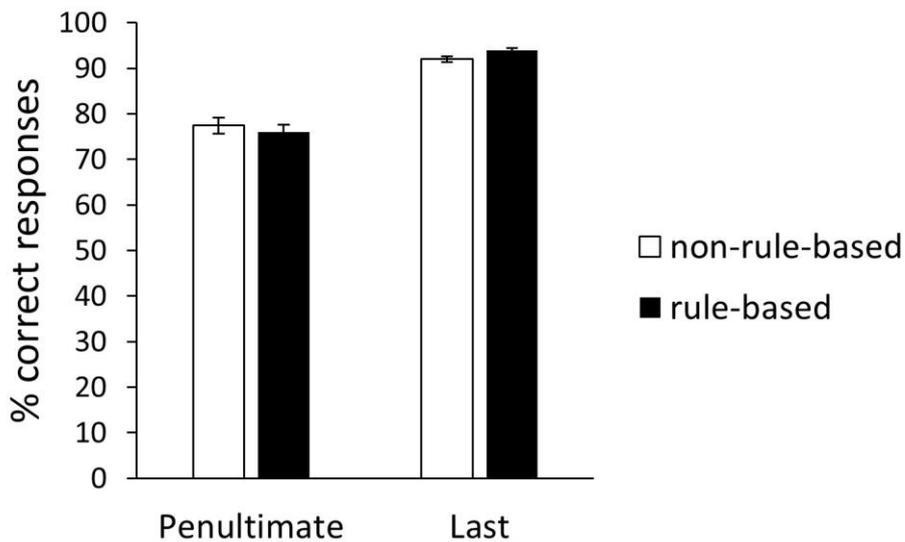


Figure 5. Percentage of correct training responses on penultimate and last training block for rule-based and non-rule-based generalizers. Error bars represent standard error of the mean.

= 177.78, $p < .01$, $\eta_{\text{partial}}^2 = 0.75$. Neither an effect of Generalization type, $F(1, 58) = 0.05$, $p = .83$, $\eta_{\text{partial}}^2 < 0.01$, nor an interaction effect, $F(1, 58) = 1.86$, $p < 0.18$, $\eta_{\text{partial}}^2 = 0.03$, were observed. For the Bayesian analysis, we utilized the Bayesian model comparison technique for factorial designs as introduced by Rouder, Wagenmakers and colleagues (Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017; Wagenmakers et al., in press). In this case, there are five models to consider: (1) the ‘Null model’ that contains only the grand mean, (2) the ‘Generalization type’ model that contains the effect of generalization type, (3) the ‘Block’ model that contains the effect of block, (4) the ‘Generalization type + Block’ model that contains both main effects, and (5) the ‘Generalization type + Block + Generalization type * Block’ model that includes both main effects and the interaction. All models, except the model containing generalization type, receive overwhelming evidence in comparison to the Null model (Table 2). However, the evidence against the two main effects model compared to the ‘Block’ model is roughly a factor 4 (BF_{10} of ‘Block’ model divided by BF_{10} of two main effects model). Further, the evidence against including the interaction compared to ‘Block’ model was roughly a factor 6 (BF_{10} of ‘Block’ model compared with BF_{10} of interaction model). In conclusion, the Bayesian analysis is in line with the standard NHST analysis and further supports the absence of an effect of generalization type. Additional comparisons revealed that the difference between the penultimate and the last block was significant for both the rule-based generalizers, $t(36) = -11.38$, $p < .01$, 95% *CI* [-21.4, -14.75], $d = -1.87$, $BF_{10} > 100$, and the non-rule-based generalizers, $t(22) = -8.24$, $p < .01$, 95% *CI* [-18.29, -10.93], $d = -1.77$, $BF_{10} > 100$ (Figure 5).

Table 2

JASP output table for the Bayesian repeated-measurements ANOVA with block as within-subjects variable and generalization type as between-subjects variable

Models	BF ₁₀
Null model (incl. subject)	1.000
Generalization strategy	0.218
Block	2.081e+24
Generalization strategy + Block	5.636e+23
Generalization strategy + Block + Generalization strategy * Block	3.403e+23

Note. All models include subject.

Navon task

Response times on the Navon task were recorded for four different trial types: global congruent, local congruent, global incongruent and local incongruent, where global and local refer to the level the participants were asked to respond to. We copied the data processing procedure of Byrom and Murphy (2014), analyzing response times from correct trials only and removing response times deviating one or more standard deviations from the mean for that trial type¹. A visual processing score was calculated for each participant by subtracting the mean reaction time for global incongruent trials from the mean reaction times for local incongruent trials, higher scores reflecting faster identification of targets on global trials (Byrom & Murphy, 2014). The average visual processing score ($M = -116.52$ ms; $SD = 209.52$ ms) was significantly below zero, $t(59) = -4.31$, $p < .01$, 95% $CI [-170.65, -62.40]$, $d = -0.56$, $BF_{10} = 331.6$, suggesting that participants in our sample showed a bias towards the local level. While it is often assumed that humans have a global processing advantage (Navon, 1977), this advantage has not always been observed (e. g., Billington et al., 2008; Hoffman, 1980) and various factors have been reported to influence this tendency (for an overview see Kimchi, 1992).

To estimate the reliability of the Navon task, we calculated a visual processing score based on the first half of global incongruent and local incongruent trials and a visual processing score based on the second half of global incongruent and local incongruent trials. A strong correlation was observed between those split-halve scores, $r = .75$, $p < .01$, 95% $CI [0.62, 0.85]$, $BF_{10} > 100$, indicating that our Navon task had a good internal reliability.

SQ-R and RSPM

The average score on the SQ-R was 46.63 ($SD = 13.86$), which is somewhat lower than the mean score in previous work (Wheelwright et al., 2006: $M = 55.56$, $t(59) = -4.99$, $p < .01$, 95% $CI [-43.05, 50.21]$, $d = -0.64$, $BF_{10} > 100$). All participants answered “slightly disagree” or “strongly disagree” on question 43 of the SQ-R (“If there was a problem with the electrical wiring in my home, I’d be able to fix it myself”) which resulted in a score of 0 for all participants on this question. Excluding this question resulted in a Cronbach’s alpha of 0.84.

¹ Because the criterion for determining outliers employed by Byrom and Murphy was rather strict, we also analyzed our data using a less strict criterion, by which only values that deviated more than 2.5 SD from the mean were removed. The same conclusions were reached when using this criterion.

The average score on the RSPM was 52.55 ($SD = 5.19$). To calculate the Cronbach's alpha of the RSPM ten items had to be removed because all participants replied correctly to those items and the variance was thus 0. After removing those items a Cronbach's alpha of 0.83 was obtained.

Correlations between tasks

First, we investigated whether there was a relationship between RSPM and SQ-R on the one hand and patterning discrimination and generalization on the other hand. Three different parameters were used to assess patterning discrimination: learning speed or the number of training blocks needed to reach criterion ($M = 3.95$; $SD = 1.72$), overall accuracy during training ($M = 76.12$; $SD = 3.81$), and a discrimination difference score, calculated as the difference between the percentage of correct responses after the first block and the percentage of correct responses after the second block ($M = 19.77$; $SD = 11.58$; all participants took minimally two blocks to reach criterion)². First, intelligence seemed not to correlate significantly with speed of learning, $r = -.18$, $p = .18$, 95% $CI [-0.41, 0.08]$, $BF_{10} = 0.39$. However, it is conceivable that intelligence is especially important for rule learning, but not so much for simply learning specific relationships. Therefore, we computed correlations separately for the rule-based and non-rule-based generalizers. For rule-based generalizers (>50% rule-based responses during generalization test), but not for non-rule-based generalizers, the number of training blocks

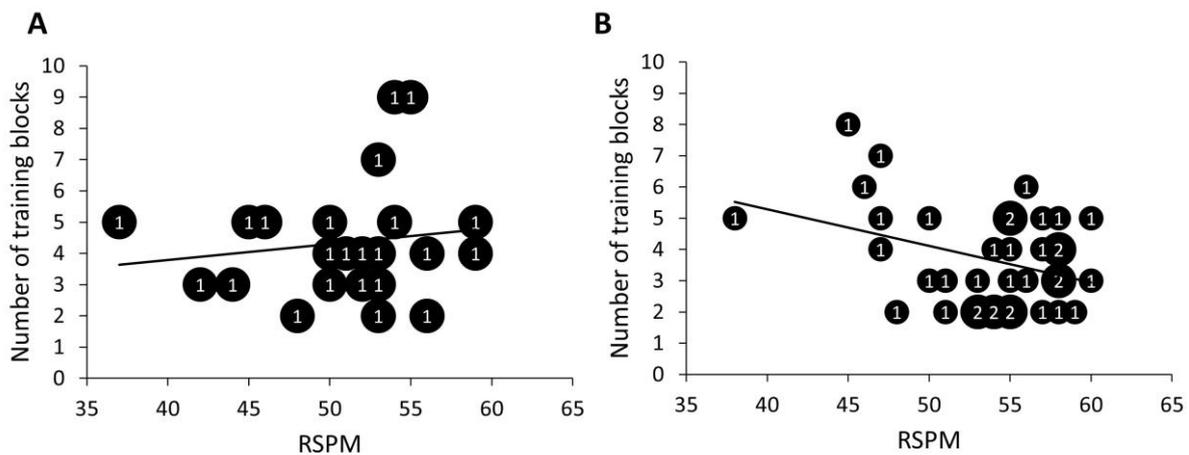


Figure 6. Scatter plots of scores on Raven Standard Progressive Matrices (RSPM) as a function of terminal accuracy (percentage of correct responses in last training block) for non-rule-based generalizers (A) and for rule-based generalizers (B). The numbers presented in the bubbles reflect the number of participants with those particular scores.

² Byrom and Murphy also calculated a discrimination difference score, but over four blocks of training.

needed to reach criterion decreased with increasing RSPM scores according to the standard null-hypothesis methodology, $r_{\text{rule}} = -.36$, $p_{\text{rule}} = .03$, 95% $CI [-0.62, -0.04]$, $BF_{10} = 2.14$; $r_{\text{non-rule}} = 0.15$, $p_{\text{non-rule}} = .51$, 95% $CI [-0.28, 0.53]$, $BF_{10} = 0.323$ (Figure 6). None of the other discrimination parameters correlated significantly with either RSPM or SQ-R. Appendix A provides detailed statistics of the correlations between the patterning discrimination parameters on the one hand and RSPM and SQ-R on the other hand. Further, participants with higher scores on the RSPM were more likely to show rule-based generalization, $r = .25$, $p = .05$, 95% $CI [0.00, 0.48]$, although Bayesian analysis was indifferent between the presence or absence of a correlation, $BF_{10} = 1.05$ (Figure 7). SQ-R was not associated with rule-based generalization, $r = .06$, $p = .67$, 95% $CI [-0.20, 0.31]$, $BF_{10} = 0.18$.

We next investigated whether we could replicate previous observations of a relationship between visual processing style and patterning discrimination, as observed by Byrom and Murphy (2014). Surprisingly, we failed to find a correlation between visual processing score and any of the patterning discrimination parameters. Bayesian analysis in fact provided

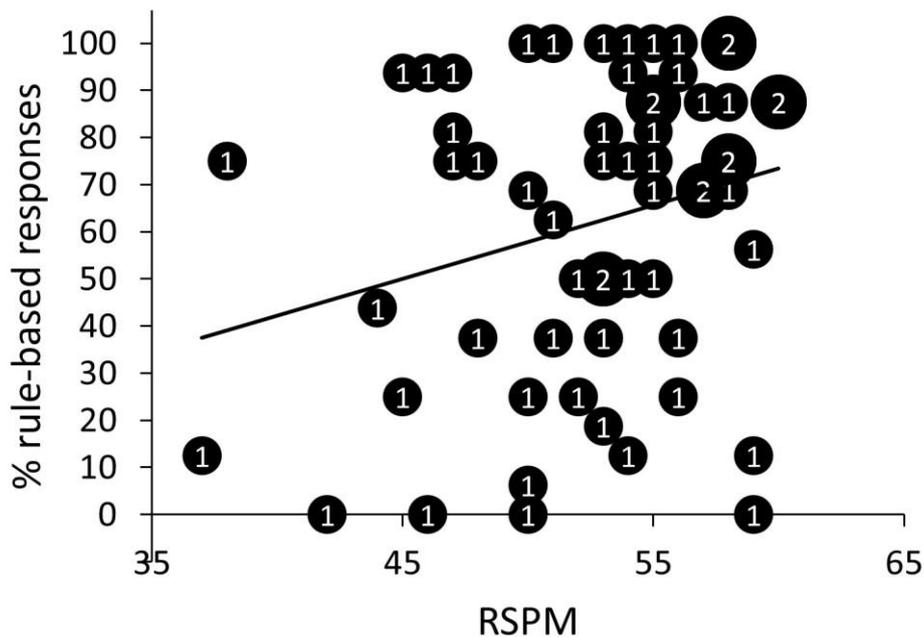


Figure 7. Scatter plot of scores on Raven Standard Progressive Matrices (RSPM) as a function of the percentage of rule-based responses. The numbers presented in the bubbles reflect the number of participants with those particular scores.

³ Note that, according to Bayesian statistics, there is not indicating that the number of training blocks decreased with increasing RSPM scores for rule-based generalizers. It does, however, provide evidence for the absence of a correlation between those variables for non-rule-based generalizers.

evidence in support of the absence of a correlation for all three parameters (see Table 3 for detailed statistics). Byrom and Murphy reported a correlation between visual processing score and patterning discrimination with an R^2 of 0.13, corresponding to a medium effect size (Cohen, 1992). Based on their power, our sample size yields an a-priori power of 0.82 to detect a comparable effect. Moreover, the standard error of the visual processing score ($SEM = 27.05$) was higher than the one reported by Byrom and Murphy (2014) ($SEM = 1.20$), providing ample room for a correlation with visual processing style to emerge. This suggests that our failure to find a correlation was not due to a lack of power. Visual processing style was not correlated with the percentage of rule-based generalization responses either (see Table 3).

Table 3
Correlations between visual processing score and patterning discrimination parameters

		Visual processing score
Number of training blocks	<i>r</i>	.03
	<i>p</i>	.79
	95% <i>CI</i>	[-0.22 , 0.29]
	<i>BF</i> ₁₀	0.17
Overall accuracy	<i>r</i>	.15
	<i>p</i>	.25
	95% <i>CI</i>	[-0.11 , 0.39]
	<i>BF</i> ₁₀	0.31
Discrimination difference score	<i>r</i>	.07
	<i>p</i>	.57
	95% <i>CI</i>	[-0.18 , 0.32]
	<i>BF</i> ₁₀	0.19
Percentage of rule-based generalization responses	<i>r</i>	.10
	<i>p</i>	.43
	95% <i>CI</i>	[-0.15 , 0.35]
	<i>BF</i> ₁₀	0.22

Multiple linear regression

In order to discern which factors most prominently predict rule-based generalization a multiple linear regression was conducted. Given that some previous work suggests an influence of training performance (Shanks & Darby, 1998), terminal discrimination accuracy was entered in the model first. Thereafter, RSPM score, SQ-R score and visual processing score were entered into the model simultaneously.

The models without and with RSPM, SQ-R and visual processing score both predict the percentage of rule-based generalization responses to a statistically significant degree, $R_1^2 = .12$, $F_1(1, 58) = 7.79$, $p_1 < .01$, $R_2^2 = .17$, $F_2(4, 55) = 2.89$, $p_2 = .03$. The change in R^2 which was not significant, $R^2_{change} = .06$, $F(3, 55) = 1.23$, $p = .31$, and only terminal accuracy seemed to contribute significantly to the prediction of generalization performance in either model (see Table 4). Bayesian analysis provides substantial evidence in favor of the alternative model over the null model for both the model that contains terminal accuracy only ($BF_{10} = 6.18$) and the model that contains terminal accuracy and RSPM score ($BF_{10} = 4.67$) (see Table 5). We also calculated the Bayes factor for each model relative to the full model BF_{mf} , which is given by $BF_{mf} = BF_{m0} / BF_{f0}$ (Rouder & Morey, 2012). The model with the greatest evidence relative to the full model is the model with only terminal accuracy as a predictor (see Table 5).

Table 4

Coefficients of the regression model. Model 1 included terminal accuracy as predictor. Model 2 additionally includes RSPM score, SQ-R score and visual processing score as predictors.

Model		b_1	SE_b	t	p	95% CI
1	(Constant)	-260.75	115.65	-2.26	.028	[-492.24 , -29.27]
	Terminal accuracy	3.46	1.24	2.79	.007	[.98 , 5.94]
2	(Constant)	-302.49	119.50	-2.53	.01	[-541.97 , -63.01]
	Terminal accuracy	3.13	1.28	2.45	.02	[.57 , 5.69]
	RSPM score	1.19	.78	1.54	.13	[-.36 , 2.75]
	SQ-R score	.25	.29	.88	.38	[-.32 , .82]
	Visual processing score	.02	.02	.87	.39	[-.02 , .05]

Note. SE_b : Standard error of b_1 . RSPM: Raven Standard Progressive Matrices

Discussion

In accordance with earlier studies (Maes et al., 2015; Wills et al., 2011; see further analysis reported in Wills, 2014), we observed that about half of the participants exhibited rule-based generalization whereas the other half exhibited feature-based generalization after simultaneous negative and positive patterning training. Notwithstanding that exploratory results indicate that the acquisition pattern for the initial training stimuli did not differ between rule-based and non-rule-based generalizers. Further, a statistically-significant correlation between intelligence and rule-based generalization was observed (using a standard null-hypothesis methodology). Most importantly, visual processing was not correlated with patterning discrimination or rule-based generalization.

Table 5

Results of Bayesian linear regression with terminal accuracy, RSPM score, SQ-R score and visual processing score as predictors

Models	BF_{10}	BF_{mf}
Null model	1.000	
Terminal accuracy	6.18	4.65
RSPM	1.37	1.03
Terminal accuracy + RSPM	4.70	3.53
SQ-R	0.28	0.21
Terminal accuracy + SQ-R	2.600	1.95
RSPM + SQ-R	0.53	0.40
Terminal accuracy + RSPM + SQ-R	2.38	1.79
Visual processing	0.34	0.26
Terminal accuracy + Visual processing	2.481	1.87
RSPM + Visualprocessing	0.67	0.50
Terminal accuracy + RSPM + Visual processing	2.37	1.78
SQ-R + Visual processing	0.14	0.10
Terminal accuracy + SQ-R + Visual processing	1.24	0.93
RSPM + SQR + Visualprocessing	0.31	0.23
Terminal accuracy + RSPM + SQ-R + Visual processing	1.33	1

Note. BF_{10} Bayes factor in favor of the model compared to the null model; BF_{mf} Bayes factor for each model relative to the full model; RSPM: Raven Standard Progressive Matrices.

To some extent, individual differences in generalization performance observed in the current experiment may be the simple consequence of differences in the extent to which the various relations presented during training have been learned; humans will exhibit more accurate generalization if they have acquired the relations from which to generalize more firmly. However, even under conditions of high terminal accuracy (at least 89% correct), considerable individual variability was observed, with about half of the participants not generalizing according to the rule of opposites. Moreover, the regression model that included terminal accuracy as a predictor was able to explain only 12% of the variance observed in percentage of rule-based generalization responses. Those results support the idea that although differences in initial patterning performance exist, those differences cannot fully account for the observed differences in generalization performance. The capacity to learn the rule of opposites and the tendency to apply this rule to new stimuli might further contribute to the variation in generalization responses. The absence of a difference in acquisition pattern between rule-based generalizers and non-rule-based generalizers suggests that the later factor (differences in application tendency) plays a role in determining generalization performance, although future research should confirm this.

The question raised here is whether certain cognitive factors correlate with the interindividual variance in generalization style. The first factor taken into consideration was intelligence. A correlation between RSPM scores and speed of learning was observed in the rule-based group (although Bayesian statistics was inconclusive), suggesting that fluid intelligence determines how fast people pick up on patterning rules. However, the results regarding the relation between fluid intelligence and generalization are more equivocal. While the results of null hypothesis significance testing suggest a correlation between RSPM scores and the percentage of rule-based generalization responses, Bayesian analysis is indecisive. Further, RSPM scores did not seem to predict the percentage of rule-based generalization responses above and beyond terminal training accuracy in the multiple linear regression analysis. The lack of a clear relation with intelligence might seem surprising given previous findings (see Introduction). One possible limitation of the current study is that we used the Raven Standard Progressive Matrices test, whereas in previous studies in which a correlation with fluid intelligence was observed, the Raven Advanced Progressive Matrices test was used (Little & McDaniel, 2015; McDaniel et al., 2014; Winman et al., 2005); the latter might be more appropriate in a sample mainly consisting of university students. However, evidence from previous studies does not in fact yield a very clear picture. The relation between intelligence

and generalization style observed in function-learning and categorization tasks might have been inflated due to the complexity of the tasks employed. Even so, McDaniel et al. (2014) observed a correlation between intelligence and generalization style in one of three function-learning tasks only ($N_1 = 62$, $N_2 = 76$, $N_3 = 37$). When the data of the three experiments were pooled, a weak overall correlation ($r = .23$) was obtained. From the work of Winman and colleagues it is unclear whether we should expect a correlation between intelligence and terminal accuracy or between intelligence and generalization style.

Second, we did not observe a correlation between SQ-R and generalization strategy, suggesting that systemizing was not related to rule use in the present patterning task. However, it has been argued that SQ-R is not be a valid instrument to measure systemizing (Ling, Burton, Salt, & Muncer, 2009; Morsanyi, Primi, Handley, Chiesi, & Galli, 2012). Other researchers have even questioned whether systemizing is a meaningful concept to begin with (Ling et al., 2009; Morsanyi et al., 2012). The results reported here further fuel those debates.

Lastly, and most importantly, we investigated whether the tendency to focus on local details or global aspects of stimuli would influence the learning and generalization of patterning rules. Byrom and Murphy (2014) observed a correlation between global versus local processing and accuracy of learning an A+/ BC+/ ABC- discrimination. Based on this observation, they suggested that, because negative patterning requires learning about configurations of stimuli (Byrom, 2013), a global processing advantage facilitates configural learning (Byrom and Murphy, 2014). This boils down to the assumption that how information is visually processed (i.e., whether the separate components rather than the configuration of the different parts are processed) influences the kind of associations that will be formed (i.e., whether associations between individual elements and the outcome will be formed as is the case in an elemental strategy or between configurations of elements and the outcome as is the case in a configural strategy). In other words, their assumption entails that visual processing style corresponds to mental representation style. Learning about positive and negative patterning problems simultaneously, as the participants in the current experiment were required to do, should also be enhanced by a configural strategy (and thus, according to the preceding analysis, by a global processing style). However, despite an almost identical Navon procedure (be it with a somewhat larger visual angle for stimulus presentation) and sufficient power, we did not observe a correlation between visual processing style and discrimination performance. We believe that a crucial difference between both experiments was the kind of patterning task employed. In the experiment of Byrom and Murphy each stimulus consisted out of nine colored shapes of one type (triangle, square, circle...) presented in a black grid of 6 x 6 squares. A compound of two

stimuli consisted of the same grid, but filled with 18 shapes. Empty squares in the grid were filled with darkly colored circles. When focusing on the global aspects of a stimulus the difference between, say, a BC trial and an ABC trials was probably immediately apparent in that particular procedure. However, when focusing on the local aspects of the compounds those differences might not have been so clear, because the compounds shared a lot of local elements. Thus, people with a more global visual processing style might simply have been better in discriminating the different trial types of the negative patterning problem in the task used by Byrom and Murphy. Because there was no visual overlap between the different stimuli used in their linear discrimination control problem (D-/ EF-/ GHI+), visual processing type would not have influenced performance on that discrimination. We therefore argue that the correlation between visual processing style and association-formation performance observed by Byrom and Murphy was due to their specific experimental task and design. Importantly, we believe that there is no compelling evidence to extrapolate from their findings that configurations in a spatial sense reflect configurations in terms of mental representations. Of course, to ascertain that the nature of the stimuli used in the patterning task is indeed of crucial importance, it would be necessary to conduct a follow-up study in which all participants receive the same Navon task, but a patterning task that employs either the stimuli used by Byrom & Murphy (2014) or those used in the current experiment. If our analysis is correct, a correlation of visual processing style should be expected with performance in the former but not the latter patterning task.

In sum, our findings suggest that differences in learning speed and accuracy between participants are not sufficient to explain differences in the acquisition and generalization of associative patterning. While the relation of individual differences in intelligence with patterning performance remains unclear on the basis of the present results, individual differences in systemizing appear to be irrelevant for associative patterning. Finally, and most importantly, our results do not support previous suggestions of a principled association between the learning of patterning discriminations and visual processing style.

Further research will be needed to ascertain the exact source of previous observations of an apparent relation between visual processing style and patterning in associative learning. It would also be interesting to further investigate the acquisition pattern in more detail and whether rule-based and non-rule-based generalizers differ in their acquisition pattern. Our results seem to suggest that there is no difference in acquisition pattern between rule-based and non-rule-based generalizers. However, it is possible that more sensitive experiments would be able to uncover potential differences. Finally, to the best of our knowledge, no research has

been conducted to separate the influence of the ability to abstract the rule of opposites from the tendency to apply this rule to a new set of stimuli on generalization performance.

Acknowledgements

The research reported here was supported by Research Grant G.0766.11N of the Fund for Scientific Research (FWO-Flanders) awarded to Tom Beckers and Jan De Houwer, InterUniversity Attraction Pole Grant P7/33 of the Belgian Science Policy Office awarded to Tom Beckers and Jan De Houwer, and KU Leuven Program Funding Grant PF/10/005 awarded to Tom Beckers. Elisa Maes is supported by a postdoctoral fellowship of the KU Leuven (PDM/16/077). The authors have no financial interest or benefit to disclose.

The authors would like to thank Ruth Van der Hallen for helpful comments on the Navon task.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6, 248–254. doi:10.1016/S1364-6613(02)01904-6
- Baron-Cohen, S., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society London B*, 358, 361–374. doi:10.1023/B:JADD.0000022607.19833.00
- Billington, J., Baron-Cohen, S., & Bor, D. (2008). Systemizing influences attentional processes during the Navon task: an fMRI study. *Neuropsychologia*, 46, 511–20. doi:10.1016/j.neuropsychologia.2007.09.003
- Byrom, N. C. (2013). Accounting for individual differences in human associative learning. *Frontiers in Psychology*, 4, 588. doi:10.3389/fpsyg.2013.00588
- Byrom, N. C., & Murphy, R. A. (2014). Sampling Capacity Underlies Individual Differences in Human Associative Learning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40, 133–143. doi:10.1037/xan0000012
- Cobos, P. L., Gutiérrez-Cobo, M. J., Morís, J., & Luque, D. (2017). Dependent Measure and Time Constraints Modulate the Competition Between Conflicting Feature-Based and Rule-Based Generalization Processes. *Journal of Experimental Psychology: Learning Memory and Cognition*, 43, 515–527. <http://doi.org/10.1037/xlm0000335>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7, 547–552. <http://doi.org/10.1016/j.tics.2003.10.005>
- Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science*, 6(3), 274–290. doi:10.1177/1745691611406920
- Fuster, J. M. (2001). The prefrontal cortex - An update: Time is of the essence. *Neuron*, 30(2), 319–333.
- Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, 116(2), 439–53. doi:10.1037/a0015251
- Hoffman, J. E. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology. Human Perception and Performance*, 6(2), 222–234. doi:10.1037/0096-1523.6.2.222

- JASP Team. (2017). JASP (Version 0.8.1.2)[Computer software]. <https://jasp-stats.org/download/>
- Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: a critical review. *Psychological Bulletin*, *112*(1), 24–38. Retrieved from <http://psycnet.apa.org/journals/bul/112/1/24/>
- Ling, J., Burton, T. C., Salt, J. L., & Muncer, S. J. (2009). Psychometric analysis of the systemizing quotient (SQ) scale. *British Journal of Psychology*, *100*, 539–552. doi:10.1348/000712608X368261
- Little, J. L., & McDaniel, M. A. (2015). Individual differences in category learning: Memorization versus rule abstraction. *Memory & Cognition*, *43*(2), 283–297. doi:10.3758/s13421-014-0475-1
- Ly, A., Verhagen, A. J. & Wagenmakers, E.-J. (2016). Harold Jeffreys's Default Bayes Factor Hypothesis Tests: Explanation, Extension, and Application in Psychology. *Journal of Mathematical Psychology*, *72*, 19-32. <http://dx.doi.org/10.1016/j.jmp.2015.06.004>
- Maes, E., De Filippo, G., Inkster, A. B., Lea, S. E. G., De Houwer, J., D'Hooge, R., ... Wills, A. J. (2015). Feature- versus rule-based generalization in rats, pigeons and humans. *Animal Cognition*. doi:10.1007/s10071-015-0895-8
- McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual Differences in Learning and Transfer: Stable Tendencies for Learning Exemplars Versus Abstracting Rules. *Journal of Experimental Psychology: General*, *143*, 668-693. doi:10.1037/a0032963
- Milton, F., Bealing, P., Carpenter, K. L., Bennattayallah, A., & Wills, A. J. (2017). The neural correlates of similarity- and rule-based generalization. *Journal of Cognitive Neuroscience*, *29*(1), 150–166. <http://doi.org/10.1017/CBO9781107415324.004>
- Morey, R., D. (2015). Jasp and Bayes: A fresh way to do statistics [PowerPoint slides]. Retrieved from <https://static.jasp-stats.org/presentations/August2015/RM/#1>
- Morsanyi, K., Primi, C., Handley, S. J., Chiesi, F., & Galli, S. (2012). Are systemizing and autistic traits related to talent and interest in mathematics and engineering? Testing some of the central claims of the empathizing-systemizing theory. *British Journal of Psychology*, *103*(4), 472–496. doi:10.1111/j.2044-8295.2011.02089.x
- Mutter, S. a, Strain, L. M., & Plumlee, L. F. (2007). The role of age and prior beliefs in contingency judgment. *Memory & Cognition*, *35*(5), 875–84. <http://doi.org/10.3758/BF03193462>

- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. doi:10.1016/0010-0285(77)90012-3
- Raven, J. C. (1958). *Standard Progressive Matrices*. London: H. K. Lewis.
- Rouder, J. N., Morey, R. D. (2012). Default Bayes Factors for Model Selection in Regression. *Multivariate Behavioral Research*, *47* (6), 877-903. doi: 10.1080/00273171.2012.734737
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. doi:10.3758/PBR.16.2.225
- Shanks, D. R., & Darby, R. J. (1998). Feature- and Rule-Based Generalization in Human Associative Learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 405–415. doi: 10.1037/0097-7403.24.4.405
- Siddiqui, S. V., Chatterjee, U., Kumar, D., Siddiqui, A., & Goyal, N. (2008). Neuropsychology of prefrontal cortex. *Indian Journal of Psychiatry*, *50*(3), 202–208. doi: 10.4103/0019-5545.43634
- Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A free software package for implementing psychological and psychophysiological experiments. *Experimental Psychology*, *57*(1), 36–45. doi:10.1027/1618-3169/a000005
- Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, A. J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Smira, M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. N., & Morey, R. D. (in press). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*.
- Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., ... Wakabayashi, A. (2006). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). *Brain Research*, *1079*(1), 47–56. doi:10.1016/j.brainres.2006.01.012
- Wills, A. J. (2014). Data and analysis unit NUS1. <http://www.willslab.co.uk/nus1/>
- Wills, A.J., Barrasin, T.J., & McLaren, I.P.L. (2011). Working memory capacity and generalization in predictive learning. In L. Carlson, C. Holscher, & T. Shipley (Eds.). *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3205-3210). Austin, TX: Cognitive Science Society.

- Wills, A. J., Graham, S., Koh, Z., McLaren, I. P. L., & Rolland, M. D. (2011). Effects of concurrent load on feature- and rule-based generalization in human contingency learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(3), 308–316. doi:10.1037/a0023120
- Winman, A., Wennerholm, P., Juslin, P., & Shanks, D. R. (2005). Evidence for rule-based processes in the inverse base-rate effect. *The Quarterly Journal of Experimental Psychology*, *58A*(5), 789–815. <http://doi.org/10.1080/0272498044300033>

Appendix A

Table A

Correlations between patterning parameters on the one hand and RSPM and SQ-R on the other hand

Patterning parameter	Sub-group		RSPM	SQ-R
Learning speed	Rule-based	<i>r</i>	-.36	-.04
		<i>p</i>	.03	.80
		95% <i>CI</i>	[-0.62 , -0.04]	[-0.36 , .29]
		<i>BF</i> ₁₀	2.14	0.21
Overall accuracy	Rule-based	<i>r</i>	.16	-.02
		<i>p</i>	.33	.93
		95% <i>CI</i>	[-0.17 , 0.46]	[-0.34 , 0.31]
		<i>BF</i> ₁₀	0.32	0.21
Discrimination difference	Rule-based	<i>r</i>	.18	-.08,
		<i>p</i>	.29	.63
		95% <i>CI</i>	[-0.15 , 0.48],	[-0.40 , 0.25]
		<i>BF</i> ₁₀	0.36	0.23
Learning speed	Feature-based	<i>r</i>	<i>r</i> = .15	-.06
		<i>p</i>	.51	.78
		95% <i>CI</i>	[-0.28 , 0.53]	[-0.46 , 0.36]
		<i>BF</i> ₁₀	0.32	0.27
Overall accuracy	Feature-based	<i>r</i>	<i>r</i> = .16	<i>r</i> = -.27
		<i>p</i>	.47	.21
		95% <i>CI</i>	[-0.27 , 0.54]	[-0.62 , 0.16]
		<i>BF</i> ₁₀	0.33	0.55
Discrimination difference	Feature-based	<i>r</i>	-.001	.16
		<i>p</i>	.996	.46
		95% <i>CI</i>	[-0.41 , 0.41]	[-0.27 , 0.54]
		<i>BF</i> ₁₀	0.26	0.34
Learning speed	All	<i>r</i>	- .18	- .07
		<i>p</i>	.18	.61
		95% <i>CI</i>	[-0.41 , 0.08]	[-0.32 , 0.19]
		<i>BF</i> ₁₀	0.39	0.18
Overall accuracy	All	<i>r</i>	.19	-.09

		<i>p</i>	.14	.50
		95% <i>CI</i>	[-0.06 , 0.43]	[-0.34 , 0.17]
		<i>BF</i> ₁₀	0.47	0.20
Discrimination difference	All	<i>r</i>	.14	.01
		<i>p</i>	.30	.93
		95% <i>CI</i>	[-0.12 , 0.38]	[-0.24 , 0.27]
		<i>BF</i> ₁₀	0.28	0.16
Percentage of rule- based generalization responses	All	<i>r</i>	.25	.06
		<i>p</i>	.05	.67
		95% <i>CI</i>	[-0.00 , 0.48]	[-0.20 , 0.31]
		<i>BF</i> ₁₀	1.05	0.18

Note. RSPM: Raven Standard Progressive Matrices