Acta Psychologica 218 (2021) 103356

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Extrapolation accuracy underestimates rule learning: Evidence from the function-learning paradigm

Nadia Said [a,b,c,*], Helen Fischer [b,d]

[a] Department of Psychology, University of Tübingen, Schleichstr. 4, 72076 Tübingen, Germany
[b] Institute of Psychology, Heidelberg University, Hauptstr. 47-51, 69117 Heidelberg, Germany
[c] InterdisciplinaryCenter for Scientific Computing (IWR), Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany
[d] Max Planck Institute for Human Development, Center of Adaptive Cognition, Lentzeallee 94, 14195 Berlin, Germany

ARTICLE INFO

Keywords:
Function-learning
Rule-based vs exemplar-based learners
Non-linear processes
Understanding

ABSTRACT

Understanding the development of non-linear processes such as economic or population growth is an important prerequisite for informed decisions in those areas. In the function-learning paradigm, people's understanding of the function rule that underlies the to-be predicted process is typically measured by means of extrapolation accuracy. Here we argue, however, that even though accurate extrapolation necessitates rule-learning, the reverse does not necessarily hold: Inaccurate extrapolation does not exclude rule-learning. Experiment 1 shows that more than one third of participants who would be classified as "exemplar-based learners" based on their extrapolation accuracy were able to identify the correct function shape and slope in a rule-selection paradigm, demonstrating accurate understanding of the function rule. Experiment 2 shows that higher proportions of rule learning than ruleapplication in the function-learning paradigm is not due to (i) higher a priori probabilities to guess the correct rule in the rule-selection paradigm; nor is it due to (ii) a lack of simultaneous access to all function values in the function-learning paradigm. We conclude that rule application is not tantamount to rule-learning, and that assessing rule xlearning via extrapolation accuracy underestimates the proportion of rule learners in function-learning experiments.

1. Introduction

Non-linear processes abound in human life, ranging from small-scale examples such as fuel consumption to large-scale, global processes such as the developments of economies, populations, or greenhouse gas emissions. A long-standing question in the cognitive literature is whether humans acquire an understanding of the underlying function rule when making predictions about the development of such processes. This question is often investigated in the function-learning paradigm, where participants learn about the beginning of a process with input-output pairs sampled from the underlying function, and predict the future development of that process. Typically, extrapolation accuracy, the distance between participants' predictions and the actual function values, is used to infer whether participants acquired an understanding of the function rule: It is argued that when predictions are sufficiently close to the correct function, participants must have learned the correct function rule; when predictions deviate sufficiently from the correct function, for example by showing flat extrapolations of highly non-linear processes, participants did not learn the correct function rule.

Here we argue, however, that even though sufficiently correct extrapolations necessitate previous rule learning, the reverse does not necessarily hold: Incorrect extrapolations do not exclude rule-learning. Rather, incorrect extrapolations can result from other processes, such as implementation failure. Based on this theoretical argument, we investigate in how far accuracy of extrapolations coincides with rule-learning of three different exponential declining[1] processes in two function-learning experiments.

In function-learning experiments, participants learn to predict continuous output (y-values) from continuous input (x-values) variables. To do so, participants are presented with an input value (for example, a time point; Fischer & Holt, 2016), and then predict the corresponding outcome value. During training, participants receive

* Corresponding author at: Department of Psychology, University of Tübingen, Schleichstr. 4, 72076 Tübingen, Germany.
E-mail address: nadia.said@uni-tuebingen.de (N. Said).

[1] Please note that the term "exponential declining" conventionally refers to e^{-x}, however we will use this term throughout the paper to refer to $-e^{x}$ which is the negative of the exponential function.

https://doi.org/10.1016/j.actpsy.2021.103356
Received 24 March 2020; Received in revised form 15 June 2021; Accepted 22 June 2021
Available online 25 June 2021

feedback on their predictions; during test (interpolation or extrapolation), no feedback is given.

Research has shown that there are two fundamentally different types of learning style that participants may employ in function-learning experiments: Rule-based and exemplar-based learning (McDaniel et al., 2014). In exemplar-based models, participants try to memorize the given exemplars, whereas in rule-based models, participants learn the function rule underlying the to-be predicted process. Among the class of exemplar-based models, at least three different accounts exist on what participants do with the stored exemplars during extrapolation. Simple exemplar-based models, first, hold that participants extrapolate using exemplars that are identical (or at least highly similar) to learned exemplars, thereby for example producing flat extrapolations that correspond to the stored exemplars (DeLosh et al., 1997). The Extrapolation-Association Model (EXAM; DeLosh et al., 1997), second, holds that participants retrieve the two best-matching exemplars, and extrapolate linearly through these exemplars. And the Population of Linear Experts model (POLE; Kalish et al., 2004), third, holds that participants store mappings between x-values and matching linear functions that they retrieve for extrapolation. Rule-based models, in contrast, hold that participants use the training information provided to abstract a rule describing the ensemble of x-y pairings (McDaniel & Busemeyer, 2005).

While function-learning studies differ in many aspects, such as the functions used (for example V-shaped, McDaniel et al., 2014, or periodic, Bott & Heit, 2004), the input format (entering a number, MacKinnon & Wearing, 1991, or clicking on a bar, McDaniel et al., 2014), the number of learning trials (for example 200, McDaniel et al., 2014, or 10 Fischer & Holt, 2016), and the experimental design (such as one learning, followed by one extrapolation phase, Lewandowsky et al., 2002, as opposed to several interspersed extrapolation phases, Bott & Heit, 2004), most function-learning studies have in common that extrapolation accuracy is used as a proxy for learning style. Specifically, not only is high extrapolation accuracy interpreted as signaling rule-learning, but also low extrapolation accuracy is interpreted as signaling exemplar-based or simple exemplar-based learning.

In one of the classic function-learning studies (DeLosh et al., 1997), absolute deviations of participants' extrapolations from the correct quadratic function were used to infer learning-type, and the authors concluded that flat extrapolations to a quadratic function were reflective of simple exemplar-based learning. In another experiment using quadratic functions (Lewandowsky et al., 2002), about 20% of participants were classified as being unable to learn the underlying rule based on the low fit of their extrapolations with the correct function. In a study with periodic functions participants were able to extrapolate (surprisingly) accurately compared to the results in other studies. The authors suggested that this difference in results may be due to participants in other experiments being unable to learn the function rule (Bott & Heit, 2004). And in a more recent study explaining individual differences in learning style, participants who showed relatively flat extrapolations to a V-shaped function were categorized as exemplar-, as opposed to rule-based learners (McDaniel et al., 2014).

The reasoning behind these studies is summarized in a theoretical argument of Kwantes and Neal (2006) who argue: “To show that you have really learned the concept, you need to demonstrate two things: You need to perform reasonably well on new items that fall within the bounds set by the training examples (so-called interpolation items), and you need to perform reasonably well on new items that fall outside the bounds set by the training examples (so-called extrapolation items)”. The authors thus argue that extrapolation accuracy separates participants who learned a function rule (or “concept”) from those who did not learn a function rule.

In sum, function-learning studies share the (often implicit) assumption that provided that, participants did acquire an understanding of the correct rule, they also apply it when extrapolating. If this assumption holds, inaccurate extrapolations can indeed be interpreted as signaling the absence of rule-learning. If this assumption does not hold, however, inaccurate extrapolations are also compatible with accurate rule-learning. In other words: while accurate extrapolations are an implication of rule-learning, inaccurate extrapolations indicate learning styles other than rule-learning if, and only if, the assumption of rule application holds.

Here we put the assumption of rule-application given rule-learning to an experimental test. The reasoning behind this is that participants neither need to apply a learned rule per se, nor do they need to apply it correctly. For example, participants may fail to accurately implement a learned rule, potentially because deriving extrapolation points from the abstracted rule requires substantial cognitive resources such as working memory capacity (Fischer & Holt, 2016). Also adjusting each consecutive extrapolation to previous extrapolations may be error-prone. Participants may even deliberately use comparatively simple linear extrapolations despite better knowledge. Indeed, in the category-learning literature, participants were found to employ an exemplar-based categorization style despite being told the correct rule beforehand (Allen & Brooks, 1991; Regehr & Brooks, 1993).

We will use the term (*a*) *function rule* to refer to the general trend (declining), shape (exponential), and slope of a presented process. Depending on whether participants acquire an understanding of only one, two, or all three of these aspects, increasingly stricter conditions of rule-learning are met. We use the term (*b*) *extrapolation style* to refer to participants' extrapolations as either (1) simple exemplar-based, that is, linear extrapolation parallel to the x-axis (DeLosh et al., 1997), (2) exemplar-based, that is, linear extrapolation through the two best-matching learning points (DeLosh et al., 1997) or (3) rule-based, that is, extrapolation according to a function rule.

To investigate in how far classifying participants based on their extrapolation accuracy reflects the extent of rule-learning, we compared (i) performance in a typical number-based input format to performance in alternative formats, and (ii) extrapolation “accuracy” when used typical linear deviation measures (rRMSES) compared to alternative measures based on slopes. We used a common, purely number-based input format (Fischer & Holt, 2016; MacKinnon & Wearing, 1991) that avoids “mixed” formats that also entail graphical elements (e.g. bars with labeled ticks McDaniel et al. (2014)), and compared it two types of graphical input formats (picture choice and drawing the function on a grid). Humans have been shown to be sensitive to subtle differences in the task format (Kalish, 2013), such that the proportion of rule-learners may be result from specifics of the task format used previously.

We employed graphical input formats for the reason that visual displays arguably constitute the typical format of displaying time series (such as climate predictions, economic growth, or disease spread) that participants should be most familiar with. Hence, if familiarity with non-linear time series being displayed graphically affects participants' ability to extrapolate these time series, the input format should affect extrapolation accuracy, and, consequently, the proportion of participants who demonstrate an understanding of the underlying function rule. We compare extrapolation accuracy using simple linear deviation measures with more complex slope-based measures because participants might deviate substantially from the to-be extrapolated function in absolute, or also relative terms (simply because they considerably over-or undershoot, for example), but still possess an understanding that the function is increasingly non-linear.

We report the results of two large function-learning experiments demonstrating that a substantial proportion of participants who would be classified as “exemplar-based learners” based on their extrapolation accuracy actually acquired an understanding the correct function rule. These results shed doubt on the assumption of rule-application given rule-learning. Furthermore, these results also deliver a comprehensive estimate of the extent to which rule-learning is underestimated by means of extrapolation accuracy.

2. Experiment 1

Experiment 1 investigated the extent to which participants who would be classified as "exemplar-based" based on their extrapolation accuracy in a classical function-learning paradigm had acquired an understanding of the function rule. Participants completed two tasks: A standard function-learning task to assess extrapolation accuracy, and a rule-selection task to assess whether participants could identify the correct function shape and slope. In the function-learning task, participants extrapolated the development of three exponential declining processes. The task consisted of one learning phase, and one extrapolation phase per process (Fischer & Holt, 2016). Participants received the instructions to extrapolate the development of different types of bacteria cultures, "Ain", "Bin", and "Cin". After the learning phase, participants completed the rule-selection task. Participants identified the function rule of the process they had just learned by selecting one of a total of six pictures displaying different function shapes and slopes.

2.1. Method

All data and the analysis scripts (R) are stored under https://figshare.com/s/6be582d99deee687c126.

2.1.1. Participants

A total of 520 participants completed the experiment. Participants were recruited over MTurk, and received 1.05$. Data from nine participants were excluded because they already participated in the pretest, thus n = 511 participants were included in the final analysis. Participants were instructed not to use pen, paper or any other help during the study. The sample size was determined by a power analysis based on a small effect size of $r = 0.18$ (Fischer & Holt, 2016), $p = .05$ and $\beta = 0.8$, resulting in a sample size of $n = 240$ per condition.

2.1.2. Materials

(a) Processes

We used three variations of exponential functions based on the equation:

$$y = 1500 - e^{a\cdot(x+50)+2} \tag{1}$$

with a=[0.045,0.040,0.046]. In the following we refer to the function

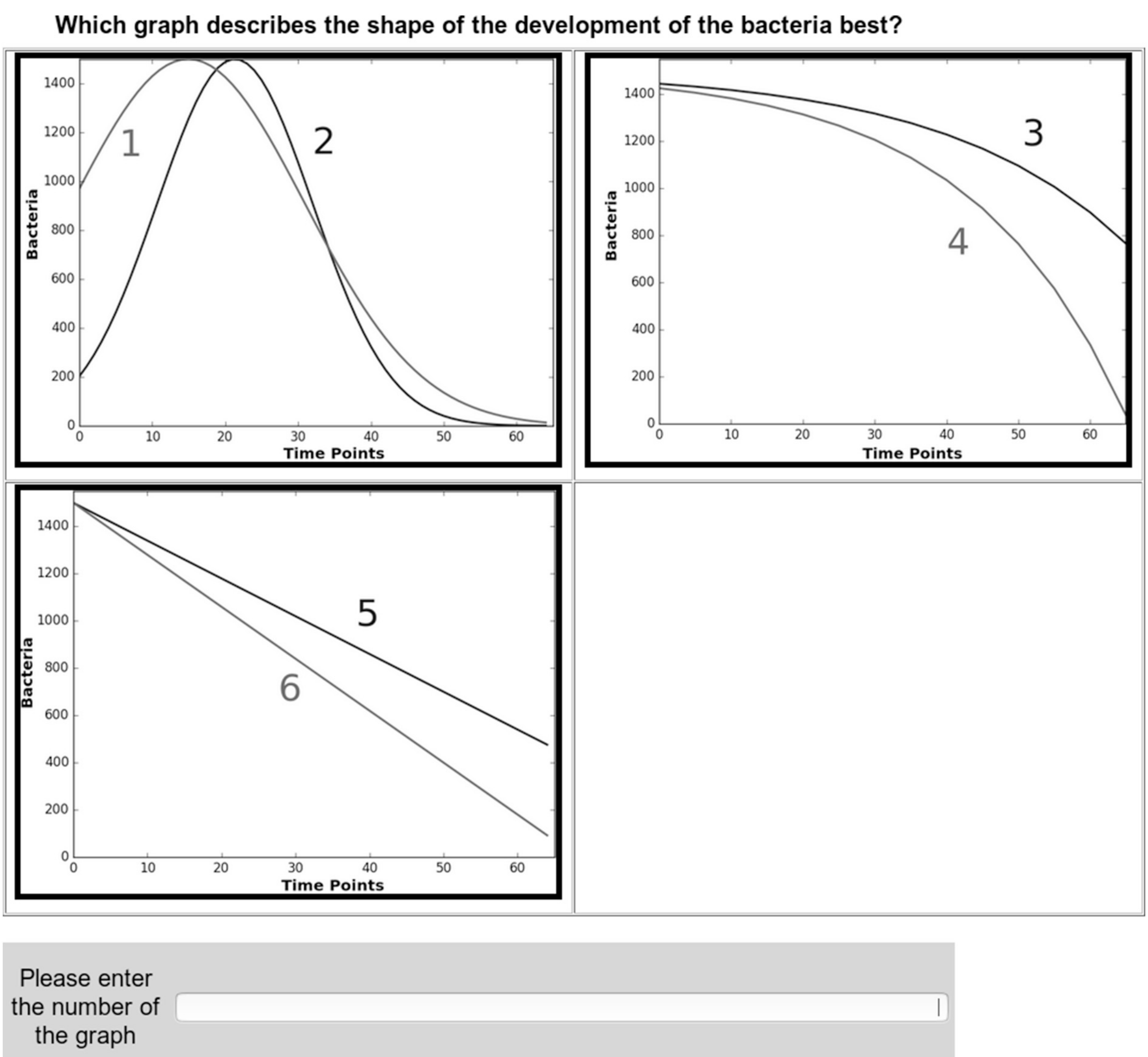

Fig. 1. Example of rule-selection task for process Bin. The figure displays the functions given to participants with the instructions to choose the process they have just learned (Correct: Function no. 3.)

with a_1=0.045 as process Ain, with a_2=0.040 as Bin, and with a_3=0.046 as Cin.

Exponential declining functions were chosen because they represent a particularly difficult function to extrapolate (Busemeyer et al., 1997), and hence because of their strong deviation from the "cognitive default" of positive linearity, a particularly strong case for rule-learning can be made (DeLosh et al., 1997; Kalish et al., 2004; Kwantes & Neal, 2006). Functions were constructed such that they were as similar as possible to each other to allow aggregation across functions, but still sufficiently different from each other to avoid solving the extrapolation task by simply giving function values from the respective previous function.

(b) Rule-selection task

Participants were presented with three graphs, displaying three different function shapes of two different slopes each: two linearly decreasing functions, two exponential declining functions, and two Gaussian functions. Participants were asked to indicate, "Which graph describes the shape of the development of the bacteria best?". Participants entered the number of the graph into a text box (Fig. 1).

Linear functions were chosen as they represent the most basic and frequently found extrapolation style (Busemeyer et al., 1997; Carroll, 1963) that is furthermore employed in exemplar-based, as well as simple exemplar-based extrapolations; Gaussian functions were chosen to assess whether participants believed the process to be non-monotonical; and the exponential declining functions were chosen to assess whether participants could correctly identify the correct function shape, and potentially also slope. The slopes displayed were 0.045_{correct} and 0.040 for Ain, 0.040_{correct} and 0.046 for Bin, and 0.046_{correct} and 0.043 for Cin. Slopes for all functions were chosen in a way such that y-values of functions remained between 0 and 1500.

2.1.3. Procedure

Each participant extrapolated 3 processes. Each process consisted of 13 trials, 8 learning and 5 extrapolation trials. The range of the x-values was x={5,...,40} in increments of 5 points for the learning trials and x={45,...,65} in increments of 5 points for the extrapolation trials. The range of the y-values was between 1445 and 34. Since we were interested in how participants extrapolate processes (that is, time series), the x-y points were shown in chronological order (rather than unordered). At the beginning of each process, participants were given the starting point of that process, that is, the number of bacteria at time point 0 (1430 for Ain, 1445 for Bin, and 1426 for Cin). Processes were shown in a fixed order. Specifically, for all participants the order of the processes was as follows: Ain, Bin, Cin. During each trial, participants were shown the current time point and predicted the number of bacteria for that time point by entering their extrapolation as a number into a text box ("I guess the number of bacteria is …"). During the learning phase, participants received feedback in terms of the correct number of bacteria for each time point, immediately after entering their extrapolation ("You guessed: … Actual number: …").

Interpolation trials were not included since previous research showed that rule and exemplar learners performed quite similar regarding interpolation (DeLosh et al., 1997; McDaniel et al., 2014), suggesting that interpolation trials are not well-suited to differentiate rule-from exemplar learners.

To control for the effect of different function shapes in the rule-selection task on extrapolation accuracy, participants were randomly allocated to complete the rule-selection task immediately before, or immediately after the extrapolation phase.

2.2. Results

2.2.1. Dependent variables

To measure extrapolation accuracy in the function-learning task, the relative root mean square error (rRMSE) was used:

$$\text{rRMSE} = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} \left(\frac{(y_i - z_i)}{z_i}\right)^2}, \tag{2}$$

with y_i : extrapolation and z_i : correct function value.

To measure understanding of the function rule, the rule selection task distinguished between identifying the correct function *shape*, that is, choosing either of the two exponential functions; and additionally identifying the correct function *slope*, that is, choosing the exponential function AND the correct slope.

2.2.2. Outliers

We excluded individual extrapolations more than five standard deviations above or below the mean of each time point (0.51% of the total number of extrapolations). If more than two out of the five extrapolation trials were excluded, the process was treated as missing for this participant (Fischer & Holt, 2016). In total, for 4 participants processes were excluded, resulting in 509 participants for process Ain, and 510 participants for processes Bin and Cin.

2.2.3. Order of rule-selection and extrapolation

In order to assess whether showing participants pictures of the correct function impacted extrapolation accuracy, we compared extrapolation accuracy in the group completing the rule-selection task before (M=3.52, SD=2.07) versus after (M=3.55, SD=2.02) the extrapolation phase. Accuracy was marginally but not significantly higher in the group performing the rule-selection task before the extrapolation phase, $F(3,503)=2.36$, p=.07, Pillais' Trace = 0.014. Thus, in the following, results for both groups are presented together.

2.2.4. Extrapolation accuracy by function slope

We assessed whether extrapolation accuracy varied by function slope. As the assumptions of homogeneity of variances ($F(2,1526)=324.12$, $p<.001$) as well as of normality (W=0.69, $p<.001$) were not met, Kruskal-Wallis rank sum tests were conducted. Results showed that extrapolation accuracy differed between the three processes as a function of slope $\chi^2(2)=1204.3$, $p<.001$. Prediction accuracy was highest for Bin and lowest for the steepest function Cin. That is, prediction accuracy was higher for Bin (M_{Bin}=0.30, SD_{Bin}=1.00) compared to Ain (M_{Ain}=1.50, SD_{Ain}=1.12) z=17.82, $p<.001$, and higher for Ain (M_{Ain}=1.50, SD_{Ain}=1.12) compared to Cin (M_{Cin}=8.86, SD_{Cin}=5.4) $z = -16.86$, $p<.001$. Interestingly, the drop in accuracy was particularly steep from Bin to the steepest function Cin (M_{Cin}=8.86, SD_{Cin}=5.4) $z = -34.70$, $p<.001$. These results are in line with previous findings that participants have a tendency toward linear extrapolation, and hence extrapolation accuracy decreases as function slope increases.

2.2.5. Proportion of participants per extrapolation style

Extrapolation styles were categorized based on their extrapolation accuracy (McDaniel et al., 2014). Specifically, we determined the deviation (rRMSE) of each participant's extrapolation accuracy including a 95% confidence interval from these three cases: (1) $\textbf{rRMSE}_{\textbf{Exp}}$: The deviation from the correct function, (2) $\textbf{rRMSE}_{\textbf{LinSlope0}}$: the deviation from a linear extrapolation with slope 0 through the last learning point, (3) $\textbf{rRMSE}_{\textbf{Lin}}$: the deviation from a linear extrapolation through the last two learning points. Confidence intervals were calculated as follows:

$$\text{CI}_{\pm} = \bar{y}_i \pm 2.776 \cdot \frac{\sigma_i}{\sqrt{n}} \tag{3}$$

with $\bar{y}_i$: rRMSE for each time point i.

Extrapolations were categorized into the different groups based on whether their entire $\text{CI}_{\pm}$ was (a) above $\textbf{rRMSE}_{\textbf{LinSlope0}}$, indicating extrapolation parallel to the x-axis (simple exemplar-based extrapolation), (b) below $\textbf{rRMSE}_{\textbf{LinSlope0}}$ but above $\textbf{rRMSE}_{\textbf{Lin}}$, indicating linear extrapolation through the last two learning time points (exemplar-based extrapolation); or (c) below **rRMSE$_{\mathbf{Lin}}$**, indicating the most accurate extrapolation (rule-based extrapolation). Fig. 2 displays an exemplatory categorization for process Bin.

2.2.6. Relationship between rule-learning and rule-based extrapolation

To investigate the relationship between rule-learning and extrapolation style, we determined the association between rule selection and extrapolation style in logistic regressions, separately for each of the three processes, Ain, Bin, and Cin. We distinguished between (i) choosing the correct function shape, and (ii) choosing the correct function shape AND slope. Choosing the correct function shape, (i), was significantly related to extrapolation style for each process, $\chi^2_{\text{Ain}}(1)=11.13$, $\chi^2_{\text{Bin}}(1)=14.86$, and $\chi^2_{\text{Cin}}(1)=13.73$, each $p<.001$, suggesting that, unsurprisingly, learning of the correct function shape was related to rule-based extrapolations. Choosing the correct function shape AND slope was associated with extrapolation style for processes Ain and Cin,

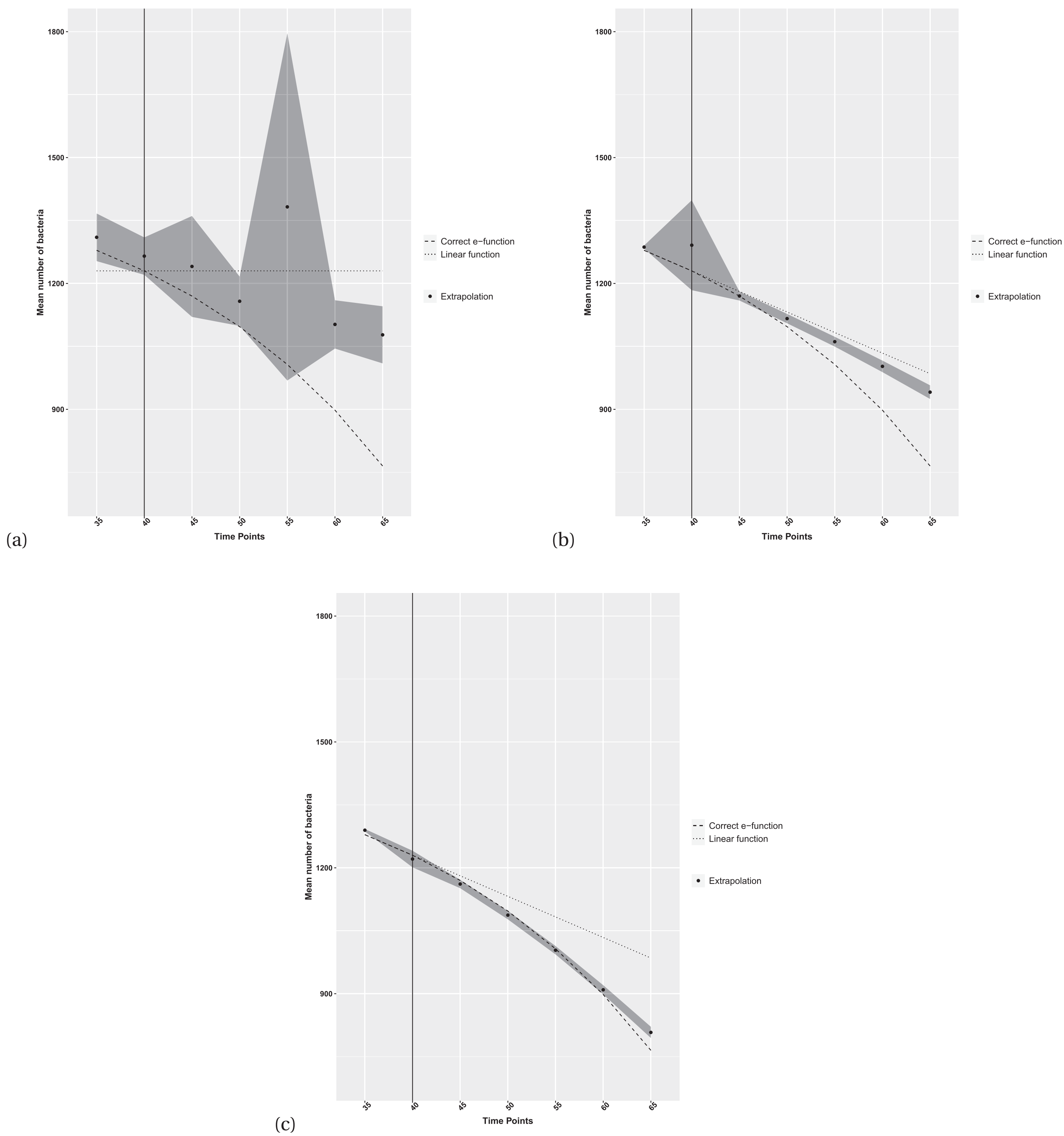

Fig. 2. Example of extrapolations for process Bin. The figure displays the number of bacteria predicted by participants who were classified as (a) simple-exemplar based, (b) exemplar-based and (c) rule-based learners based on their extrapolation accuracy. Dark-grey area: 95% confidence band. The vertical line denotes the last training trial before extrapolation.

$\chi^2_{\text{Ain}}(1)$=36.93, $\chi^2_{\text{Cin}}(1)$=16.83, p<.001, but not for process Bin, $\chi^2_{\text{Bin}}(1)$= 1.11, p=.29, suggesting that learning of the correct function slope was related to rule-based extrapolations, except for extrapolation of the process with the lowest slope.

To test the extent to which exemplar-based, or simple exemplar-based extrapolation styles exclude rule-learning, Table 1 displays the proportion of participants who could correctly identify the correct function shape, separately for each extrapolation style. In the group displaying simple exemplar-based extrapolation, 46% of participants were able to identify the correct function shape, while only 31% of participants estimated the function was actually linear. A similar pattern held for the group displaying exemplar-based extrapolation, where 61% of participants chose the correct function shape, and only 24% estimated the function to be linear. For the group displaying rule-based extrapolation, 65% chose the correct function shape. In sum, the relative majority of participants displaying simple exemplar-based extrapolations, and even the absolute majority of participants displaying exemplar-based extrapolations could identify the correct function shape as exponential declining.

In total, the proportion of participants who had acquired an understanding of the correct function rule in the learning phase (as indicated by the rule-selection task) but did not apply this in the extrapolation phase (as indicated by classifications of their extrapolation style based on extrapolation accuracy) was 47% for process Ain, 45% for Bin, and 36% for Cin (Table 2).

As the stricter criterion of rule-learning, (ii) we determined the number of participants choosing not only the correct function shape but also function slope, per extrapolation style. As Table 3 shows, the proportion of participants choosing the correct function slope increased with extrapolation style, from simple exemplar-based, to exemplar-based, to rule-based. Across all three processes, 24% of participants displaying simple-exemplar-based extrapolations, and 37% of participants displaying exemplar-based extrapolations, were able to identify the correct function slope. These results suggest that even among those participants who had acquired a deep understanding of the function rule in that they could identify the correct shape AND slope, a considerable proportion of participants did not apply this understanding when extrapolating, but rather used exemplar-based, or even simple exemplar-based extrapolation styles.

The last column of Table 3 displays the proportion of participants who could identify the correct slope, out of those who could identify the correct shape. Results show that while for processes Bin and Cin, around half of participants who could identify the correct shape also identified the correct slope, results were different for the steepest process Cin in that the vast majority of participants who identified the correct shape also identified the correct slope.

Table 2
Proportion of participants identifying the correct exponential function shape, and applying exemplar-based and simple exemplar-based extrapolations, per process (Ain, Bin, and Cin).

Style	Process	Exponential
Simple-& Exemplar-based	Ain	237 (47%)
	Bin	229 (45%)
	Cin	185 (36%)

Interestingly, for all three extrapolation styles, the proportion of participants who could correctly identify the correct function shape dropped 15% for the steepest function Cin compared to the proportion of participants who could correctly identify the correct function shape for processes Ain and Bin. This result contrasts results on extrapolation accuracy for Cin (M_{Cin}=8.86) which dropped by 83% compared to Ain (M_{Ain}=1.50), and even 97% compared to Bin (M_{Bin}=0.30).

2.2.7. *Prevalence of rule-learning based on rule-selection task vs. extrapolation accuracy*

Table 4 compares the proportion of participants who would be classified as rule-learners based on accuracy in the rule-selection task as opposed to extrapolation accuracy. Results show that while the minority (<20%,χ^2(2,N=1529)=238,p<.001) of participants would be classified as rule-learners based on extrapolation accuracy, the relative majority of participants could identify the correct function shape (>50%,χ^2(2, N=1529)=308,p<.001), and even slope (>25%, χ^2(3,N=1529)=53.33, p<.001).

2.3. *Summary 1*

Experiment 1 showed that a substantial proportion of participants who had acquired an understanding of the correct function rule in the learning phase of a function-learning experiment (as indicated by the rule-selection task) did not apply their understanding in the extrapolation phase (as indicated by classifications of their extrapolation style based on extrapolation accuracy).

We will focus our discussion on the results of the first process Ain since the order of processes was fixed, such that after the first process participants were more familiar with the task, which might affect prediction strategies used. Thus, results for the first process Ain are the most

Table 1
Proportion of participants selecting one of the three function shapes in the rule-selection task, per extrapolation style.

Style	Process	Gaussian	Linear	Exponential	χ^2-test
Simple exemplar-based	Ain	59 (21%)	73 (27%)	144 (52%)	$\chi^2(2, N = 276) = 45.15^{***}$
	Bin	59 (28%)	45 (22%)	104 (50%)	$\chi^2(2, N = 208) = 27.41^{***}$
	Cin	49 (18%)	127 (47%)	97 (35%)	$\chi^2(2, N = 273) = 34.02^{***}$
	Total	**22%**	**32%**	**46%**	
Exemplar-based	Ain	19 (14%)	26 (19%)	93 (67%)	$\chi^2(2, N = 138) = 72.57^{***}$
	Bin	36 (18%)	41 (20%)	125 (62%)	$\chi^2(2, N = 202) = 74.27^{***}$
	Cin	23 (14%)	56 (33%)	88 (53%)	$\chi^2(2, N = 167) = 37.95^{***}$
	Total	**15%**	**24%**	**61%**	
Rule-based	Ain	16 (17%)	14 (15%)	65 (68%)	$\chi^2(2, N = 95) = 52.70^{***}$
	Bin	16 (16%)	12 (12%)	72 (72%)	$\chi^2(2, N = 100) = 67.52^{***}$
	Cin	15 (22%)	17 (24%)	38 (54%)	$\chi^2(2, N = 70) = 13.91^{***}$
	Total	**18%**	**17%**	**65%**	*** $p \le .001$

Table 3
Proportion of participants selecting the correct function shape AND slope, per extrapolation style.

Style	Process	Correct shape AND slope	**Total**	$\frac{\text{Correct shape AND slope}}{\text{Correct shape}} \cdot 100$
Simple exemplar-based	Ain	44 (16%)	**24%**	31%
	Bin	65 (31%)		63%
	Cin	70 (26%)		72%
Exemplar-based	Ain	51 (37%)	**37%**	55%
	Bin	76 (38%)		61%
	Cin	62 (37%)		70%
Rule-based	Ain	36 (45%)	**44%**	55%
	Bin	50 (36%)		69%
	Cin	35 (50%)		92%

Table 4
Proportion of participants classified as rule-based, exemplar-based or simple exemplar-based learners in the function-learning and rule-selection paradigm, per process (Ain, Bin and Cin).

	Traditional FL paradigm			**Rule-selection paradigm**			
	Simple exemplar-based	Exemplar-based	Rule-based	Gaussian	Linear	Exponential (shape)	Exponential (slope)
Ain	276 (54%)	138 (27%)	95 (19%)	94 (19%)	113 (22%)	302 (59%)	138 (27%)
Bin	208 (41%)	202 (40%)	100 (19%)	111 (22%)	98 (19%)	301 (59%)	177 (35%)
Cin	273 (53%)	167 (33%)	70 (14%)	87 (17%)	200 (39%)	223 (44%)	167 (33%)
Total	50%	33%	17%	19%	27%	54%	32%

readily interpretable. For the first process Ain, out of those participants who would be classified as exemplar-based learners based on their extrapolation accuracy, 67% were able to accurately identify the correct function shape, and 37% were able to identify the correct function shape AND slope. Moreover, only 27% of participants showing simple exemplar-based, and 19% of participants showing exemplar-based extrapolations believed the functions to be actually linear. Subsequent processes showed a similar pattern of results.

These results suggest that (i) extrapolation accuracy underestimates rule-learning in the classical function-learning paradigm, and that (ii) up to certain extent, participants are aware of the non-linearity of the process, even if their extrapolations suggest otherwise.

Interestingly, the reverse case was also evident in our data in that some participants who correctly extrapolated exponentially, nevertheless picked an incorrect rule in the rule-selection task. This phenomenon appeared to be partially driven by the order of the two tasks: Participants were slightly better at selecting the correct shape when they had extrapolated the function beforehand. This suggests that, to some degree, learning of the function shape still took place while completing the tasks. This additional learning was lacking in participants who completed the rule-selection task first.

There are three limitations to Experiment 1. First, as the task for the participants was not only to identify the correct function shape but also slope, the graphs (i) contained the trained x-values and (ii) displayed the whole function (learning and extrapolation phase). This could have enabled participants to solve the task by checking points that they memorized against the graphs. However, performance did not differ between displaying the graphs directly before or after the extrapolation phase. If participants inferred the correct values from seeing the graphs, performance should have been better for participants who saw the graph before the extrapolation phase. Second, even though selection of the exponential shape was clearly above guessing rate (that is 33%) in all three extrapolation styles, a priori probabilities to guess correctly were considerably higher in the rule-selection task compared to extrapolation in the standard function-learning task. To address those limitations, we introduced an additional condition in Experiment 2 that required participants to indicate their understanding of the function shape not by selecting a picture, but by drawing their understanding of the function shape into a grid. Third, it remains unclear why participants fail to apply their rule understanding in the classical function-learning paradigm. One plausible explanation is that participants make implementation errors in the classical function-learning paradigm where x-y pairings are given only consecutively, whereas in the rule-selection task participants have simultaneous access to all function values. To address this third limitation, Experiment 2 introduced another control condition, where all extrapolations were displayed on the same page for a given process, so that current as well as all previous extrapolations were visible to participants.

3. Experiment 2

Experiment 1 showed that a substantial proportion of exemplar-based or even simple exemplar-based extrapolators had acquired an accurate understanding of the function rule, indicating that the number of rule-learning was underestimated previously by using extrapolation accuracy as a proxy for rule-learning. Experiment 2 provided equal *a priori* probabilities between a standard function-learning condition and an alternative paradigm in which participants indicated their understanding of the process by drawing the function into a grid (grid condition). Furthermore, we added a third condition in which extrapolations were displayed on one screen instead of in consecutive order (summary function-learning condition).

3.1. Method

3.1.1. Participants

A total of 918 MTurk participants completed the experiment. Sample size was determined based on a power analysis with $f = 0.1$, p=.05 and $\beta = 0.8$, resulting in a sample size of n=323 per condition. Data from 176 participants were removed because inspection of MTurk IDs revealed participants had taken part in either Experiment 1, or a pretest. Participants were instructed not to use pen, paper or a calculator during the study. We included a statement at the end of the study in which participants had to confirm that they did not do so (“I confirm that I did NOT use a calculator, pen or paper”). A total of 15 participants were excluded because they reported having used aids. Additionally, in order for the grid condition to be fully comparable with the function-learning conditions, we checked for the position (one click per time point) and order (clicks starting from time point T1 followed by T2 and so on) of extrapolations. Out of the 232 participants in the grid condition, 67 were excluded because they violated this requirement. A total of N = 660 participants were included in the final data set.

3.1.2. Materials

Two of the processes from Experiment 1 were used, the development of the bacteria cultures Ain and Bin.

3.1.3. Procedure

Participants were randomly assigned to three groups: (a) A *standard function-learning* condition identical to Experiment 1 that served as a baseline. Participants entered their extrapolations as numbers, and extrapolations were displayed sequentially; (b) A *grid* condition where participants drew the function shape by clicking the respective positions on a grid; and (c) A *summary function-learning* condition where participants entered their extrapolations in numbers and all time points were displayed as value-pairs on one screen, so that participants were able to see all previous function values (Fig. 3). To ensure comparability between the three conditions, the maximum and minimum extrapolations were restricted to values 0–1550, in steps of 1; and the clicks on the grid were restricted to the same number as the entries in both function-learning conditions.

3.2. Results

3.2.1. Dependent variables

Calculation of **extrapolation accuracy (rRMSE)** in the two function-learning conditions was identical to Experiment 1. We applied the same procedure regarding outliers as in Experiment 1 (0.11% of the total number of extrapolations).

To assess rule-learning in the grid condition, we used two types of approaches:

(1) Calculating the first derivatives

Calculating the first derivatives allows us to determine whether participants extrapolated (a) linearly through the last two learning points (exemplar-based extrapolation), or (b) according to the function rule (rule-based extrapolation) (McDaniel et al., 2014). This is because

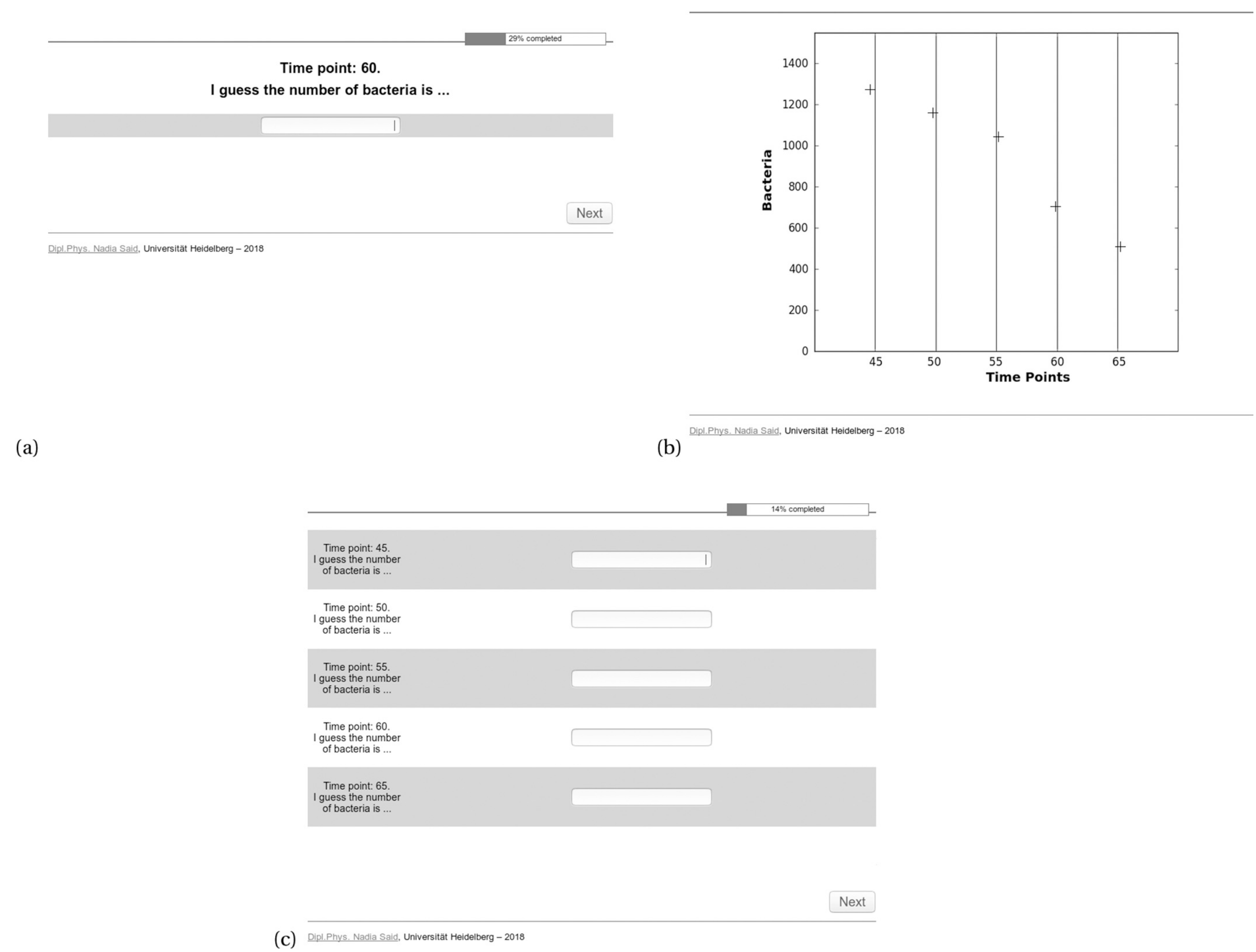

Fig. 3. Screenshots of the three conditions. The figure displays the three conditions: (a) standard function learning, (b) grid, and (c) summary function-learning. Valid numbers participants could enter: 0–1550.

in case of (a), derivatives must be constant, whereas in case of (b), derivatives must be strictly monotonic decreasing as $\frac{d}{dx}\left(-e^x\right) = -e^x$ with the negative sign reflecting the trend of the process. This allows us to evaluate whether participants abstracted a rule about the exponentiality of the process in that it is *increasingly declining*. To do so, we used three different approaches varying in strictness of what counts as strictly decreasing derivatives: (1) slopes of the lines through the first and the second, as well as the first and the last extrapolation point were strictly decreasing; (2) slopes through the first and two other extrapolation points have to be strictly decreasing; and (3) slopes though the first and all other extrapolation points have to be strictly decreasing. Please note that regarding the first criterion (1) other non-linear functions could also meet this condition. However, it is safe to say that in all three cases, participants must have acquired an understanding of the non-linearity of the process.

(2) Least squares approach

Using a least squares approach, we classified the functions drawn in the grid condition as rule-based, exemplar-based, or non-distinguishable based on the deviation (RMSE) of clicks on the grid from three models: If

$$\mathrm{RMSE}_{\mathrm{linear}} > \mathrm{RMSE}_{\mathrm{exponential}} + \mathrm{RMSE}_{\mathrm{linear}}{*}25\%, \tag{4}$$

participants were classified as rule-based learners; if

$$\mathrm{RMSE}_{\mathrm{linear}} < \mathrm{RMSE}_{\mathrm{exponential}} - \mathrm{RMSE}_{\mathrm{linear}}{*}25\%, \tag{5}$$

participants were classified as exemplar-based learners; for all other cases, participants were classified as non-distinguishable. To ensure comparability between the models, only the slope parameter *a* was allowed to vary, and all other parameters were fixed.

3.2.2. Extrapolation accuracy in standard versus summary function-learning

To assess whether comparatively low proportions of (accurate) rule-application in the classical function-learning paradigm were due to a lack of simultaneous access to all previous function values, we compared extrapolation accuracy based on participants' rRMSEs in (a) the standard function-learning, with (b) the summary function-learning condition. Results showed that extrapolation accuracy was not higher in the summary (M_{Ain1}=1.54, SD_{Ain1}=0.80, M_{Bin1}=0.27, SD_{Bin1}=0.19), compared to the standard function-learning condition (M_{Ain2}=1.53, SD_{Ain2}=0.84, M_{Bin2}=0.26, SD_{Bin2}=0.19), $F(2,492)$=0.027, p=.97, Pillais' Trace = 0.0001. This result suggests that having access to all function values did not increase rule application per se, nor did it increase the accuracy of rule-application.

3.2.3. Proportion of rule-based extrapolation in the two function-learning conditions

For conditions (a) standard function-learning and (b) summary function-learning, participants' extrapolation styles were classified based on extrapolation accuracy. Table 5 shows that for both function-learning conditions, between 19% (Ain) and 17% (Bin) were classified as rule-based extrapolators.

3.2.4. Proportion of rule-learning in the grid condition

(1) Calculating the first derivatives

We calculated the first derivatives ($d_{1,\,2}$, $d_{1,\,3}$, $d_{1,\,4}$, $d_{1,\,5}$) by calculating the slopes between the first extrapolation point and the following 4 points. Participants were classified as having understood the function rule as being non-linear if (a) they captured the trend of the process ($d_{1,\,j}$<0, $\forall j \in \{2,3,4,5\}$) and (b) if slopes of the lines through the first and the second, as well as the first and the last extrapolation point were strictly monotonic decreasing ($d_{1,\,5} < d_{1,\,2}$). Following that classification, 48% of participants had abstracted the function rule for process Ain, and 56% for process Bin (Fig. 4). We employed the same method for the next stricter criterion (three out of the four slopes), resulting in 28% of participants having abstracted the function rule for process Ain, and 39% of participants having abstracted the function rule for process Bin. Only when using the strictest possible criterion where the values *for all four slopes* have to decrease strictly monotonically, results were comparable to classifications based on rRMSEs in that 20% of the participants were classified as having abstracted the function rule for process Ain, and 29% for process Bin (Table 6).

(2) Least squares approach

Using the least squares approach outlined above, 52% of participants were classified as having understood the correct function shape for process Ain, and 59% for process Bin (Fig. 5). These results are broadly in line with results using the derivatives approach.

3.2.5. Proportion of rule-based learners in all three conditions

Table 7 shows the proportion of participants classified as having acquired an understanding of the function rule for both processes, per condition. Across the two function-learning conditions, standard and summary, 19% of participants were classified as rule-based learners for process Ain and 17% for process Bin. In contrast, in the grid condition 50% of participants were classified as rule-based learners for Ain and 57% for Bin.

3.2.6. Hierarchical Bayesian analysis

As in the previous analysis, participants were classified based on five extrapolation trials only, reliability of the classification of participants can be low at the individual level, and aggregating over these probabilistic classifications can distort parameter estimates at the group level. We therefore conducted a hierarchical Bayesian analysis to validate results from the individual-level classifications at the group level. Specifically, first, we investigated whether the mean rRMSEs of the three individual-level classifications: rule-based, exemplar-based, and simple exemplar-based were different. If individual-level classifications are reliable, Bayesian analyses should reveal the smallest mean rRMSE for all participants classified as rule-based, followed by those classified as

Table 5
Proportion of participants classified as showing each of the three extrapolation styles based on extrapolation accuracy, per processes (Ain and Bin).

	Simple exemplar-based		Exemplar-based		Rule-based	
	Standard FL	Summary FL	Standard FL	Summary FL	Standard FL	Summary FL
Ain	148 (60%)	139 (56%)	58 (23%)	58 (24%)	42 (17%)	50 (20%)
Total Ain	58%		23%		19%	
Bin	126 (51%)	115 (47%)	77 (31%)	94 (38%)	45 (18%)	38 (15%)
Total Bin	49%		34%		17%	

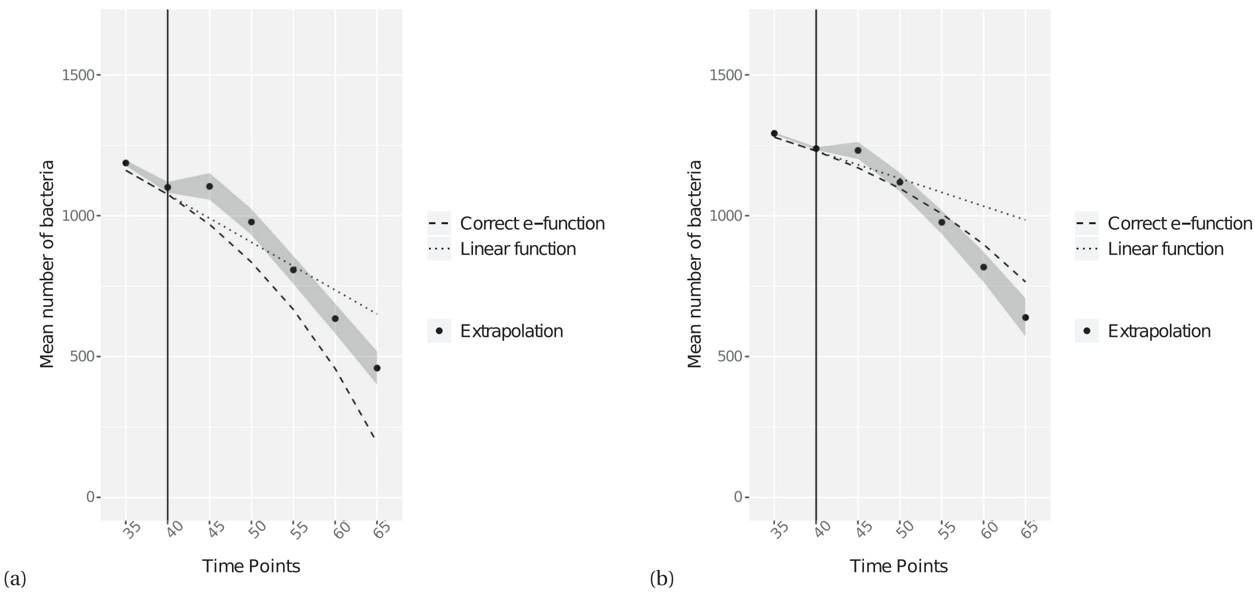

Fig. 4. Mean number of bacteria for Ain and Bin for first derivatives approach. The figure displays the mean number of bacteria estimated by participants in the grid condition who were classified as having abstracted a rule about the underlying process for (a) Ain and (b) Bin by calculating the first derivatives. The vertical line denotes the last training trial before extrapolation.

Table 6
Proportion of participants classified as "rule-based learners" in the derivatives approach.

	$d_{1,5} < d_{1,2}$	$d_{1,5} < d_{1,3} < d_{1,2}$ or $d_{1,5} < d_{1,4} < d_{1,2}$	$d_{1,5} < d_{1,4} < d_{1,3} < d_{1,2}$
Ain	80 (48%)	92 (51/41) (28%)	33 (20%)
Bin	93 (56%)	128 (66/62) (39%)	48 (29%)
and $d_{1,j} < 0, \forall j \in \{2,3,4,5\}$ for all four classification criterions.			
χ^2_{Ain}	$\chi^2(1) = 55.87^{***}$	$\chi^2(1) = 9.33^{**}$	$\chi^2(1) = 0.08, p = .77$
χ^2_{Bin}	$\chi^2(1) = 30.35^{***}$	$\chi^2(1) = 3.83, p = .05$	$\chi^2(1) = 0.34, p = .56$

Comparison of proportions FL (standard & summary) and derivatives approach, $^{**}p \leq .01$, $^{***}p \leq .001$.

exemplar-based, and followed by those classified as simple exemplar-based. And second, we investigated whether the mean rRMSEs of the three conditions: standard FL, summary FL, and grid condition were different. If the classical function-learning paradigm underestimates the proportion of rule-learners compared to the grid condition, Bayesian group-level results should reveal smaller mean rRMSE for the grid, compared to the standard FL paradigm.

We used the R-script provided by Kruschke (2014). The model was implemented in JAGS. For analysis we used Markov chain Monte Carlo (MCMC) sampling with four chains of 11,000 samples, 20 thinning steps and 1000 burn-in steps. To account for heterogeneous variances each group was provided with its own standard-deviation parameter σ_j. For more robustness against outliers t distributed data was assumed. Prior choices were taken from Kruschke (2014), p.573. An overview of the model is given in Table 8.

To compare groups we used a region of practical equivalence (ROPE) of $[-0.1 \cdot \sigma_y, 0.1 \cdot \sigma_y]$ with $\sigma_y = \max(\sigma_1, \sigma_2, \sigma_3)$, with 1–3 referring to the three categories/conditions. The region of practical equivalence corresponds to a "null" hypothesis and is used to test whether a parameter is significant. That is, if the high density intervals (HDI) are completely outside the ROPE results are significantly different (Kruschke, 2014; Makowski et al., 2019).

Figs. 6, 7, and 8 show the mean rRMSEs for the different individual-level classifications in all three conditions and for both functions. Furthermore the figures display the posterior predictive check to assess whether the model fits the observed data. The greater the overlapping between the observed data and the predictive distribution the better the model represents the data (Gelman et al., 2013).

Hierarchical Bayesian results corroborate our individual-level classifications in that, for both processes Ain and Bin, mean rRMSEs were smallest for participants applying rule-based extrapolations, followed by those applying exemplar-based, and simple exemplar-based extrapolations.

To furthermore test for effect sizes and statistical differences of these group-level results, Figs. 9 to 14 show pair-wise comparisons of the mean rRMSES of participants applying rule-based vs exemplar-based vs simple exemplar-based extrapolations. For both functions Ain and Bin, HDIs did not overlap with the respective ROPEs. Calculating the difference between all extrapolation-styles and for both functions showed that all HDIs fell outside the ROPE. That is, rRMSEs for all three groups classified based on their extrapolation trials were significantly different.

To evaluate our individual-level results suggesting that the standard-function learning paradigm underestimates rule knowledge compared to rule application, Fig. 15 compares the mean rRMSEs for the three conditions: standard FL, summary FL, and grid condition and demonstrates that prediction accuracy was indeed higher for the grid paradigm, compared to both FL paradigms. These results suggest that participants were better able to apply their rule knowledge when given graphical as opposed to purely number-based input formats.

Figs. 16 and 17 show the difference and the effect size comparing the standard and the summary function-learning paradigm with the grid paradigm for Ain and Bin. Group-level results revealed that, for both functions Ain and Bin, the standard and the summary FL paradigm did not differ regarding mean rRMSEs. However, for Ain participants

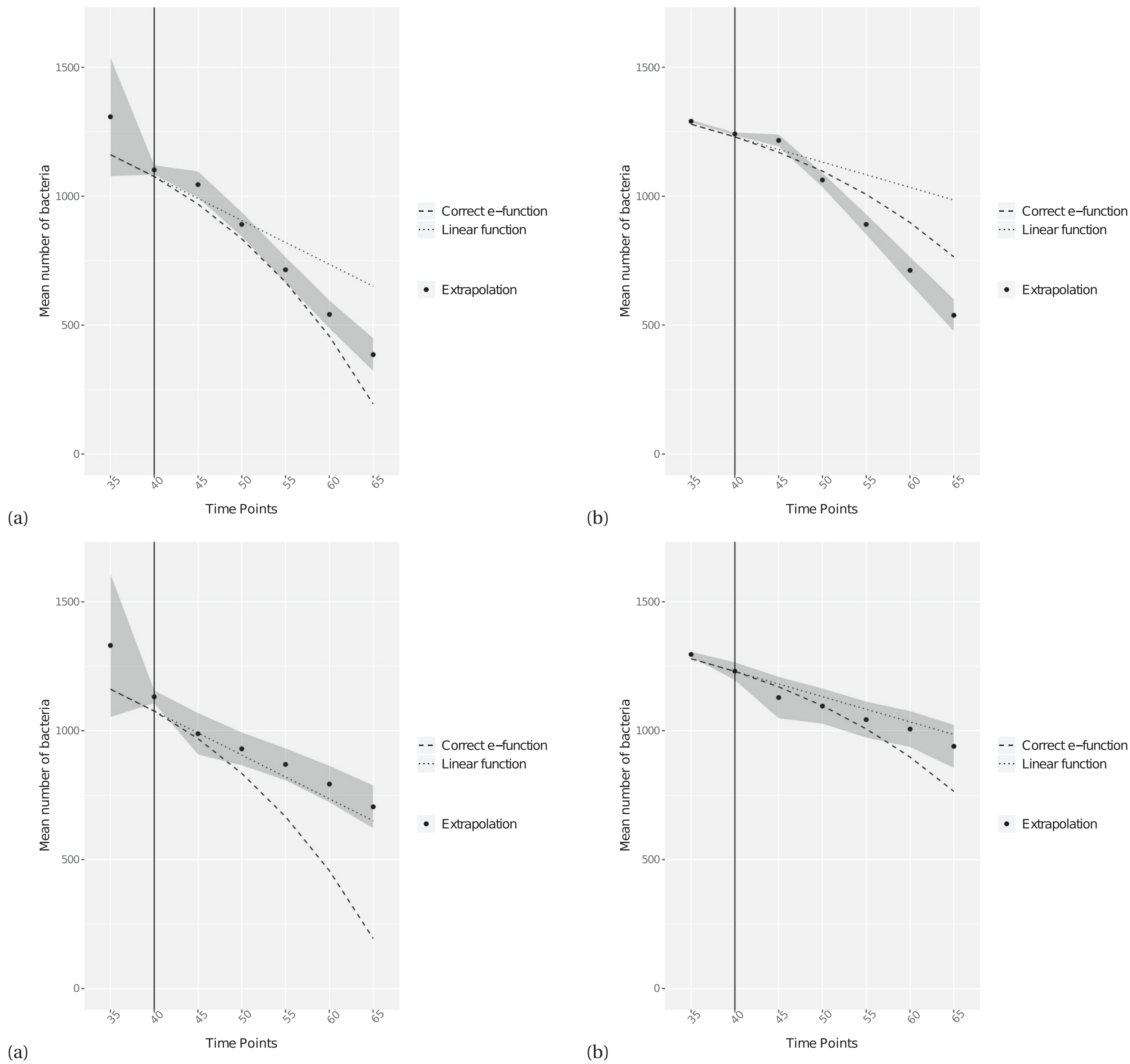

Fig. 5. Mean number of bacteria for Ain and Bin for least squares approach. The figure displays the mean number of bacteria estimated by participants in the grid condition who were classified as having abstracted a rule about the underlying process for (a) Ain and (b) Bin applying a least squares approach. Pictures (c) and (d) display participants who were classified as extrapolating linearly. The vertical line denotes the last training trial before extrapolation.

Table 7
Proportion of participants classified as "rule-based learners" in all three conditions.

	Standard FL	Summary FL	(a) Total: Standard & Summary	Derivatives	Least-Squares	(b) Mean: Grid condition	Comparison of proportions of (a) & (b)
Ain	42 (17%)	50 (20%)	92 (**19**%)	80 (48%)	86 (52%)	166 (**50**%)	$\chi^2(1) = 91.20^{***}$
Bin	45 (18%)	38 (15%)	83 (**17**%)	93 (56%)	97 (59%)	190 (**57**%)	$\chi^2(1) = 147.09^{***}$

Comparison of proportions (a) FL and (b) grid condition, ***$p \leq .001$.

performed considerably better in the grid paradigm (μ_1=1.01) than in the other two conditions (μ_2=1.53, μ_3=1.53). Specifically, Fig. 16 demonstrates that for Ain, the HDI was completely outside the ROPE when comparing the grid condition with the standard FL condition. This was not the case for Bin (Fig. 17), suggesting that differences in the input format did not affect extrapolation accuracy for this process.

Importantly, when classifying participants based on their rRMSEs for the grid condition, the majority were classified as displaying a rule-based extrapolation style (45%), while 24% were classified as exemplar-based and 32% as simple exemplar-based extrapolators for Ain. Specifically, for Ain the number of rule-based extrapolators were broadly the same as when employing alternative quantifications to the rRMSEs (derivative: 48% or least squares: 52%). For Bin percentages for the different extrapolation styles were similar to the other two conditions. That is, the majority of participants were classified as simple exemplar-based extrapolators (52%), 33% were classified as exemplar-

Table 8
Model parameters and priors for hierarchical Bayesian model.

	Parameter	Prior
Data: Student's t-distribution	Normality parameter ν	Exponential distribution
	Scale parameter σ_j	Gamma distribution
	Predicted value μ_j	
Predicted value $\mu_j = \beta_0 + \sum_j \beta_j x_j$	Baseline parameter β_0	Normal distribution
	Group deflection parameter β_j	Normal distribution

based and 16% as rule-based extrapolators. These results suggest that, as expected, both the input format, as well as means to quantify "accuracy" can affect the proportion of participants that count as having acquired rule knowledge.

3.3. Summary 2

Experiment 2 investigated whether unequal a priori probabilities to guess correctly could explain increased rule-learning compared to rule application in the standard function-learning paradigm. To do so, we introduced a condition where participants drew their understanding of the progress of the function into a grid where the numbers of clicks as well as the range of possible values were restricted to the same values as extrapolations in the standard function-learning task. The function shapes were evaluated using two different types of approaches, calculating the first derivatives, and a least squares approach of varying strictness. Both methods produced broadly similar results in that 50% of participants were classified as rule-based learners for process Ain and 57% for processes Bin, compared to 19% of participants showing rule-based extrapolation in both function-learning conditions for process Ain, and 17% for process Bin.

To furthermore investigate whether (accurate) rule-application is reduced in the standard function-learning paradigm because participants lack access to all previous function values, we compared extrapolation accuracy in the standard function-learning, with a summary function-learning condition. Results showed that there was no difference in extrapolation accuracy between the two conditions, suggesting that lacking access to all function values did not affect rule-application.

We additionally conducted a hierarchical Bayesian analysis to validate the individual-level classifications of extrapolation style that might be noisy and yield low reliabilities. Results corroborated the individual-level classifications in three ways: First, by demonstrating that, at the group level, rRMSEs for the three individual-level classifications of extrapolation styles, simple exemplar-based, exemplar-based, and rule-based, were in fact different. Specifically, rule-based extrapolators had the highest extrapolation accuracy followed by exemplar-based and simple exemplar-based extrapolators. Second, Bayesian analyses also corroborated that the input format can exert an effect on extrapolation accuracy. Specifically, for Ain (but not Bin), extrapolations were significantly more accurate in the grid condition, compared to the two function-learning conditions. Third, for Bin, results showed that methods of analysis matter independently from the input format. Applying three different methods of evaluation (rRMSE, first derivatives, and least squares) to the same input format (grid condition) resulted in 16% participants being classified as rule-based extrapolators

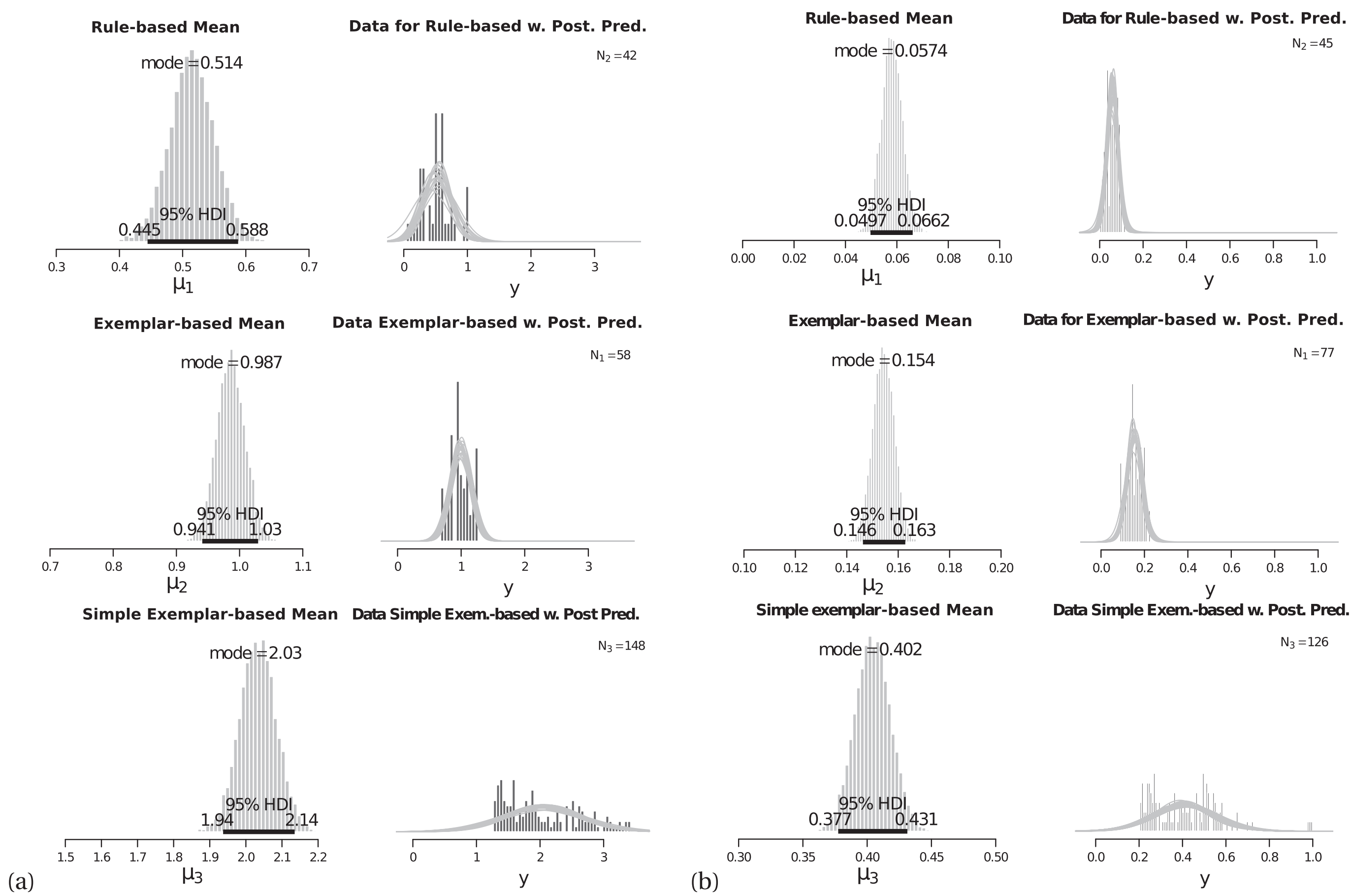

Fig. 6. Mean rRMSEs in the standard function-learning paradigm for the individual-level classifications based on the hierarchical Bayesian procedure. The figure displays the posterior distributions of the mean rRMSEs μ for rule-based, exemplar-based, and simple exemplar-based extrapolators (left) and the corresponding posterior predictive check (right) for the two functions: Ain (a) and Bin (b). Note that for better visibility of the results the x-axis displays a different range of values for each category.

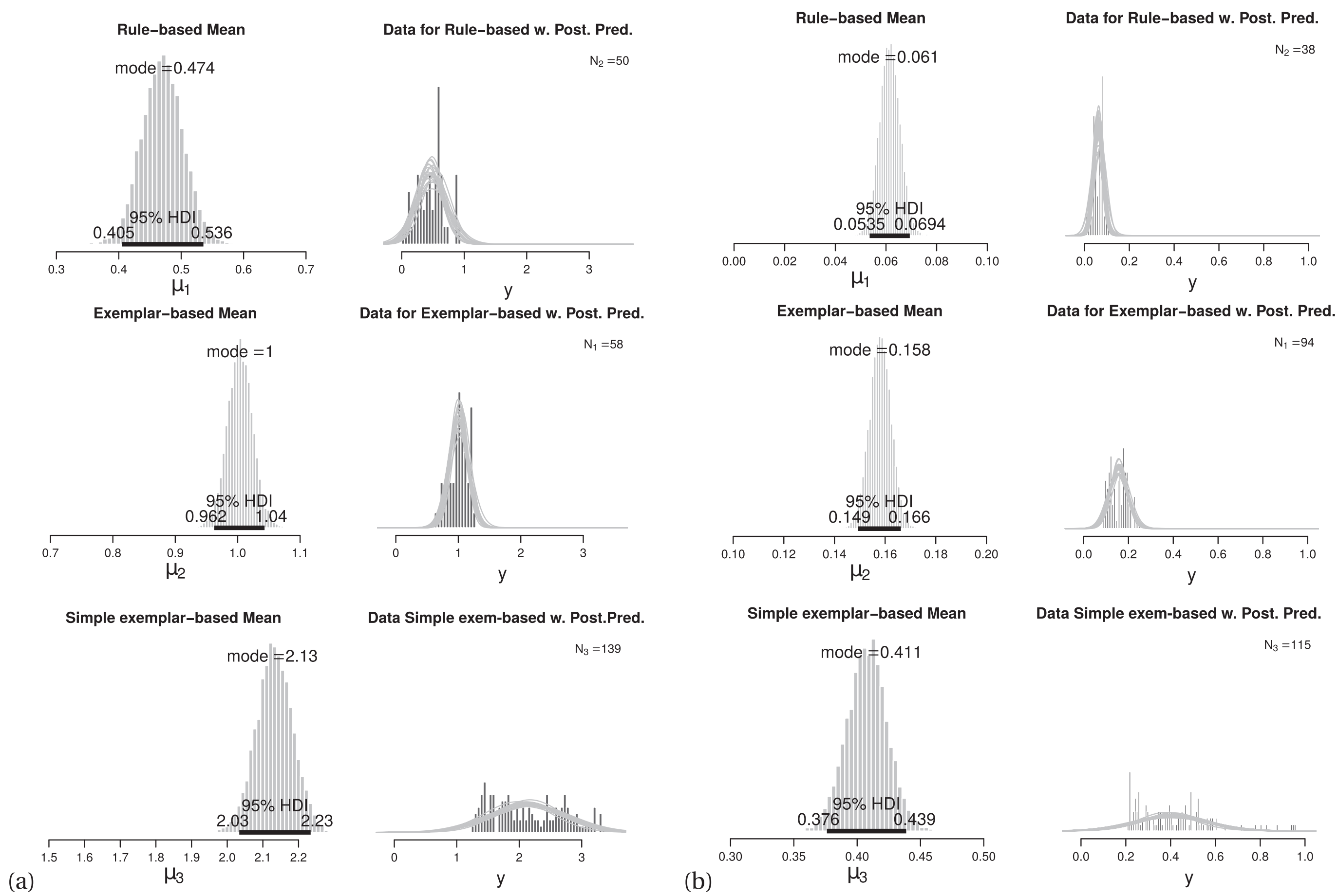

Fig. 7. Mean rRMSEs in the summary function-learning paradigm for the individual-level classifications based on the hierarchical Bayesian procedure. The figure displays the posterior distributions of the mean rRMSEs μ for rule-based, exemplar-based, and simple exemplar-based extrapolators (left) and the corresponding posterior predictive check (right) for the two functions: Ain (a) and Bin (b). Note that for better visibility of the results the x-axis displays a different range of values for each category.

when calculating extrapolation accuracy (rRMSE) compared to 57% being classified as rule-based extrapolators when applying a first derivatives or least squares approach. And fourth, similarly to the individual-level, frequentist results, Bayesian analysis revealed no difference in prediction accuracy between the two function learning paradigms, standard and summary.

4. General discussion

In the function-learning paradigm, understanding of the function rule that underlies the to-be extrapolated process is typically measured by means of extrapolation accuracy (Bott & Heit, 2004; DeLosh et al., 1997; Kwantes & Neal, 2006; Lewandowsky et al., 2002; McDaniel et al., 2014). Here we argue, however, that even though accurate extrapolations necessitate rule-learning, the reverse does not necessarily hold: Inaccurate extrapolations do not exclude rule-learning. Using inaccurate extrapolations to infer learning styles therefore hinges upon the assumption of rule-application given rule-learning. In two function-learning experiments with exponential declining functions, we showed that the proportion of participants who demonstrated an understanding of the correct function rule was almost twice as high as the proportion of participants who would be classified as rule-learners based on extrapolation accuracy in the standard function-learning experiment. We therefore conclude that (i) using extrapolation accuracy as a proxy for rule-learning severely underestimates people's actual ability to abstract the correct function rule; and that (ii) for a substantial proportion of participants, the assumption of rule-application given rule-learning does not hold.

In the following we will, again, focus on discussing the results for the first process Ain as the order of the processes was fixed and thus, even though participants' performance did not increase with subsequent processes, participants were likely to be more familiar with the task format after the first process which might affect strategy-use. For the first process a majority of participants who would be classified as "exemplar-based learners" (67%) or "simple exemplar-based learners" (54%) based on their extrapolation accuracy as measured with rRMSEs, were able to identify the correct function shape in the rule-selection paradigm, and 37% of the "exemplar-based learners" were able to identify the correct function shape AND slope. The grid paradigm that ensured equal a priori probabilities between drawing one's understanding of the rule and extrapolations in the classical function-learning paradigm produced broadly similar results: Half of participants showed an accurate understanding of the function rule, both when analyzing their understanding of the exponentiality of the process via the first derivatives, and via a least squares approach. For both experiments, subsequent processes showed similar response patterns.

These results therefore suggest that a substantial proportion of participants did not apply their rule-understanding when extrapolating. In other words, extrapolation accuracy was considerably lower than what would be expected based on participants' understanding of the function rule. However, these results are limited in that we showed this for only one specific input format, namely for entering numbers. As previous research has shown, differences regarding the input format affects performance (Kalish, 2013). Hence, the generalizability of our findings to

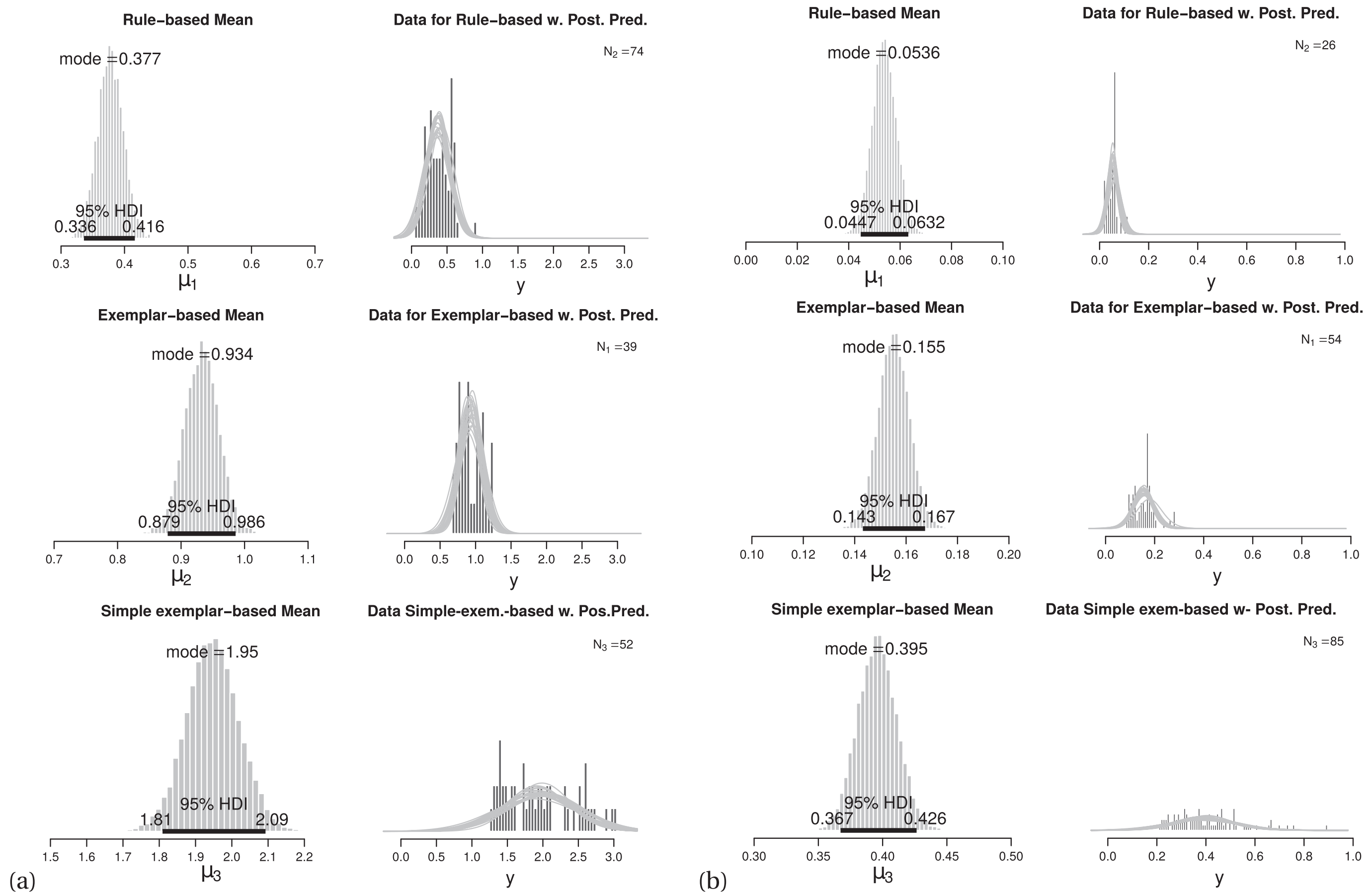

Fig. 8. Mean rRMSEs in the grid paradigm for the individual-level classifications based on the hierarchical Bayesian procedure. The figure displays the posterior distributions of the mean rRMSEs μ for rule-based, exemplar-based, and simple exemplar-based extrapolators (left) and the corresponding posterior predictive check (right) for the two functions: Ain (a) and Bin (b). Note that for better visibility of the results the x-axis displays a different range of values for each category.

other input formats is unknown. Comparing performance in the graph paradigm to other input formats such as bars might affect performance, which can only be clarified by research testing a wider range of input formats.

Furthermore, results showed that for the first process Ain (but not for Bin) in the grid format, (i) extrapolation accuracy was generally higher than in the other two formats and (ii) even based on rRMSEs most participants were classified as rule-based extrapolators. That is, for Ain the number of participants classified as rule-based extrapolators based on their rRMSEs were broadly the same as when employing the first derivatives or the least squares approaches for classification. These results corroborated previous findings that demonstrated that the input format can severely affect the proportion of participants demonstrating rule-application in the function-learning paradigm (Kalish, 2013).

For the second process Bin, the different methods of evaluation resulted in different proportions of rule-based learners for the same paradigm (grid format). Specifically, 16% participants were classified as rule-based extrapolators when calculating extrapolation accuracy (rRMSE) compared to about 57% who were classified as rule-based extrapolators when applying a first derivatives or least squares approach. Hence, these results suggest that both the input format and the method of evaluation affect outcomes when trying to assess whether or not participants had understood the non-linearity of a process.

In the grid paradigm, we employed different criteria varying in strictness of what counts as understanding of the function rule as exponential declining. Specifically, in the derivatives approach we varied the number of slopes that had to decrease strictly monotonically. It is important to note that even though strictly decreasing slopes are a characteristic feature of exponential declining functions, other non-linear functions could also meet this condition.

When the first and last slopes, as well as three out of four slopes were strictly monotonically decreasing, the proportion of rule-learners was higher than the proportion of rule-based extrapolators. Only when using the strictest criterion for rule-understanding, that all four slopes be strictly monotonically decreasing, the proportion of participants who were classified as having understood the rule was broadly in line with the proportion of participants who were classified as rule-based learners based on their extrapolation accuracy. This suggests that a considerable proportion of participants who had acquired an understanding of a characteristic feature of the function rule, namely that later extrapolation points should be steeper than earlier extrapolation points, could not implement this understanding when extrapolating. However, the comparability between performance in the grid task and the standard function-learning task might be limited as differences in the proportion of rule-learners might also arise because of the different modes of presentation (Kalish, 2013).

A hierarchical Bayesian analysis validated the individual-level classifications. If individual-level classifications are reliable, group-level rRMSEs for the three individual-level classifications (simple exemplar-based, exemplar-based, and rule-based) should differ. Indeed, hierarchical Bayesian results corroborated our individual-level results: Individual-level classifications based on extrapolation accuracy yielded three distinct groups differing in accuracy at the group level. Furthermore, results corroborated findings of previous research in that the input format could have an impact on performance. More specifically, the grid paradigm could reduce extrapolation errors at the group level. For the steeper function Ain, providing a more visual (rather than number-based) input format considerably increased rule-application.

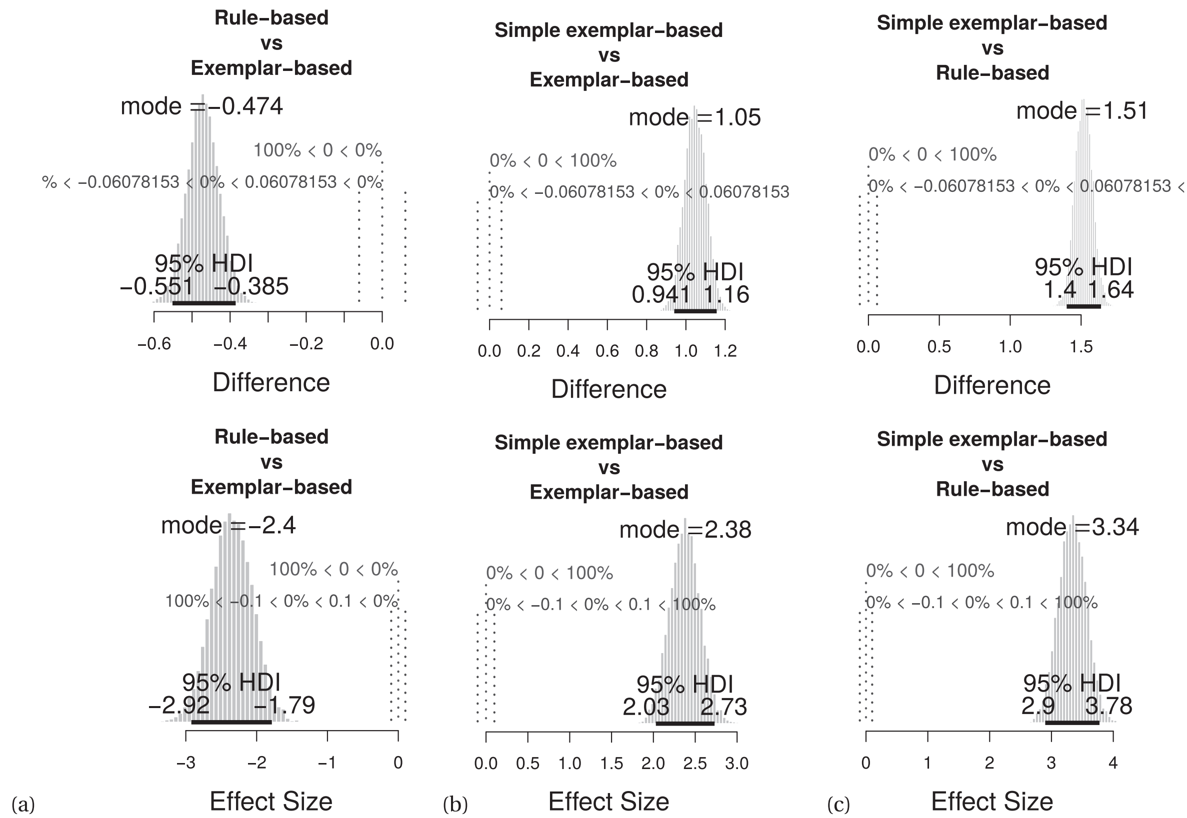

Fig. 9. Comparison between the different categories in the standard function-learning paradigm for Ain. The figure displays the difference and the effect size for (a) rule-based vs. exemplar-based extrapolators and (b) simple exemplar-based vs. exemplar-based extrapolators.

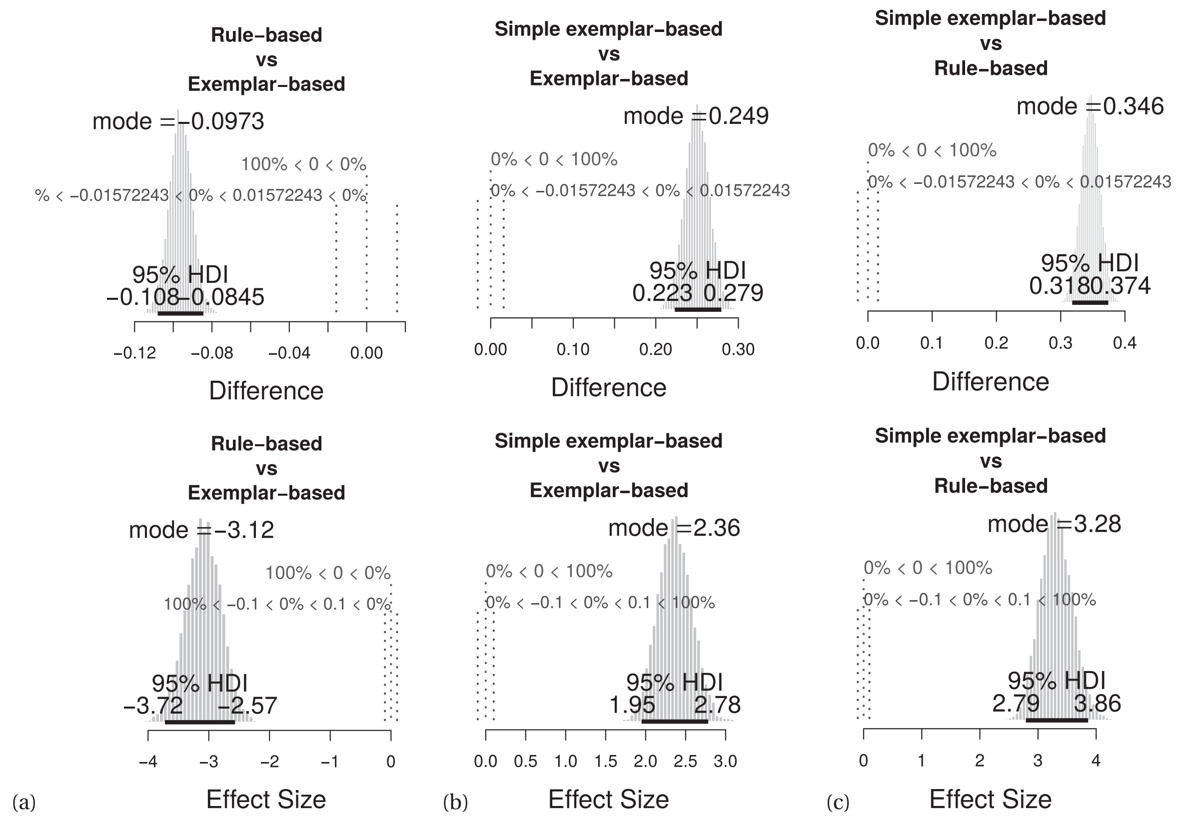

Fig. 10. Comparison between the different categories in the standard function-learning paradigm for Bin. The figure displays the difference and the effect size for (a) rule-based vs. exemplar-based extrapolators and (b) simple exemplar-based vs. exemplar-based extrapolators.

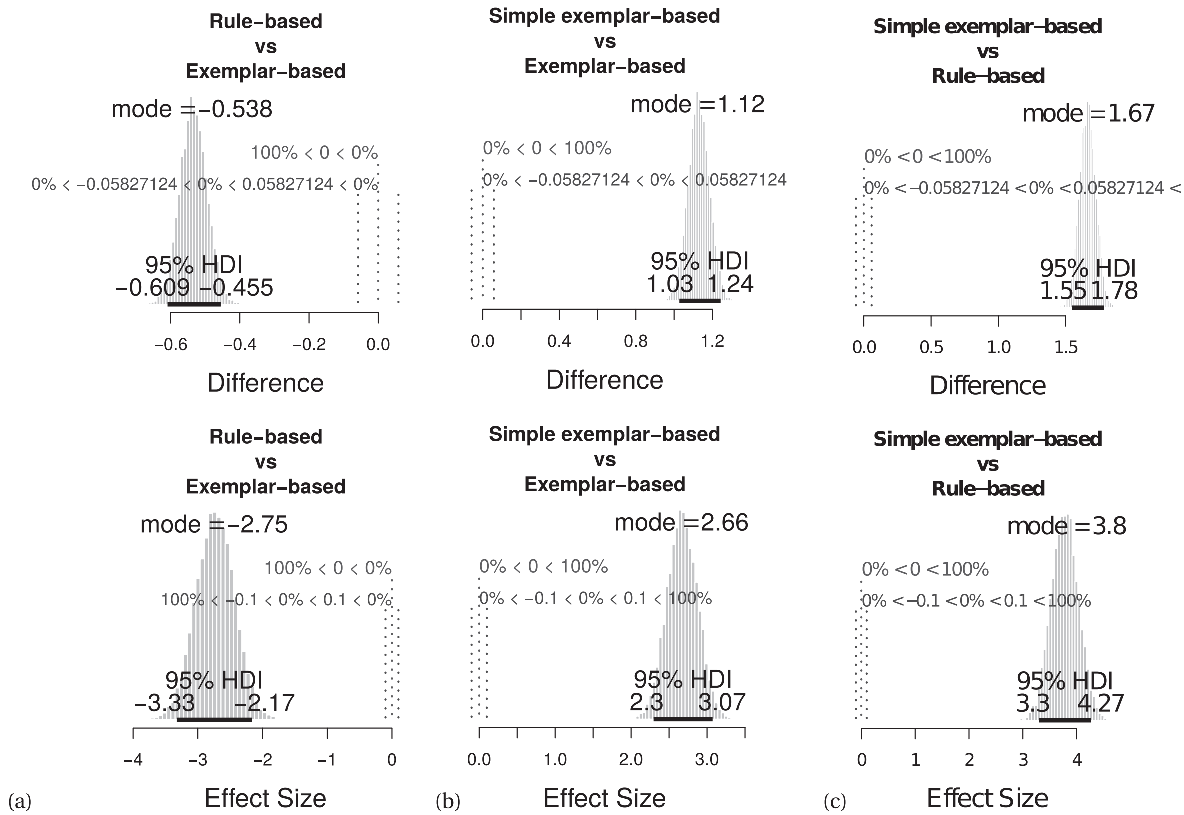

Fig. 11. Comparison between the different categories in the summary function-learning paradigm for Ain. The figure displays the difference and the effect size for (a) rule-based vs. exemplar-based extrapolators and (b) simple exemplar-based vs. exemplar-based extrapolators.

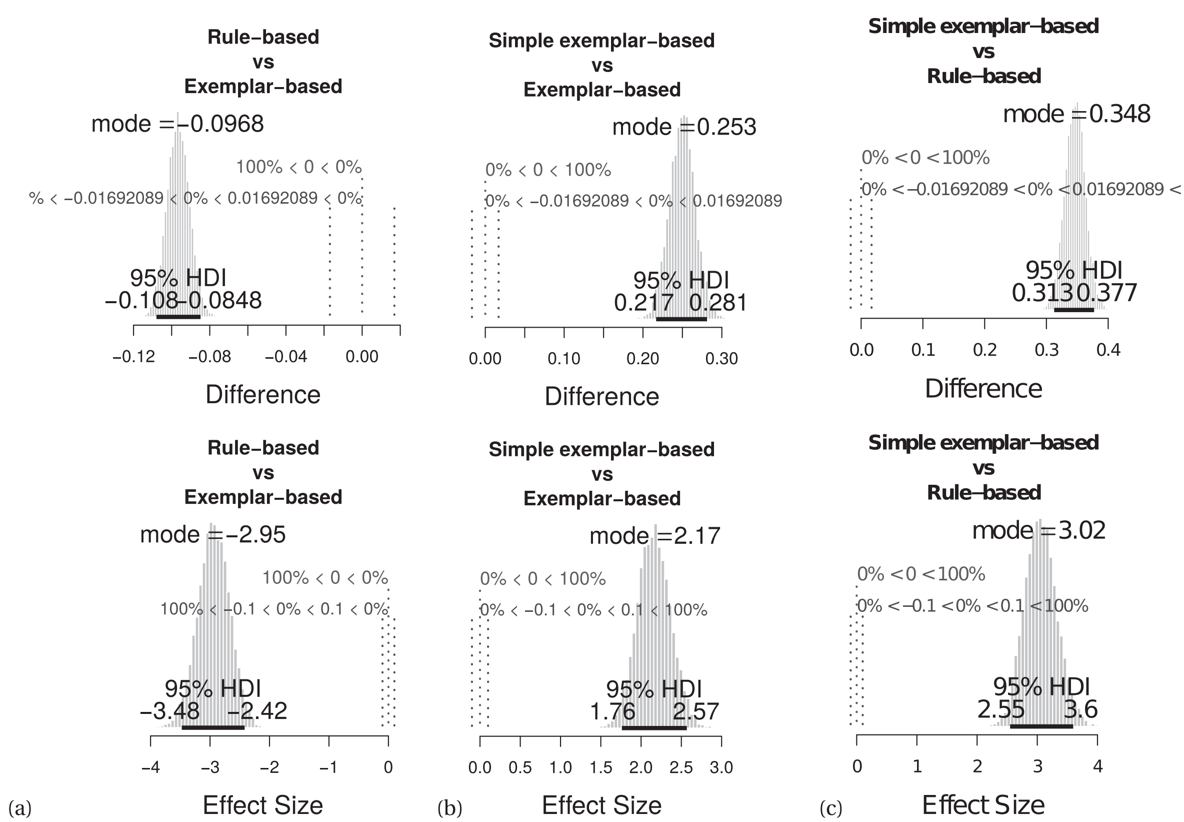

Fig. 12. Comparison between the different categories in the summary function-learning paradigm for Bin. The figure displays the difference and the effect size for (a) rule-based vs. exemplar-based extrapolators and (b) simple exemplar-based vs. exemplar-based extrapolators.

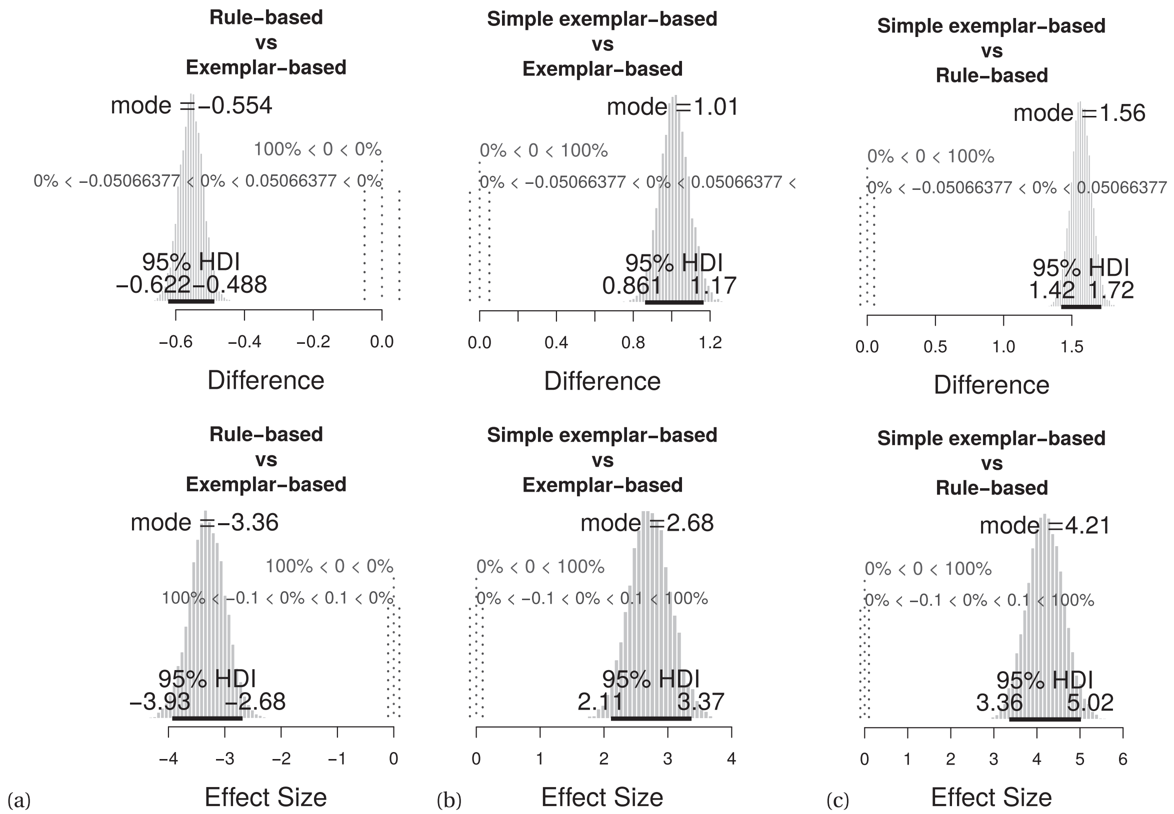

Fig. 13. Comparison between the different categories in the grid function-learning paradigm for Ain. The figure displays the difference and the effect size for (a) rule-based vs. exemplar-based extrapolators and (b) simple exemplar-based vs. exemplar-based extrapolators.

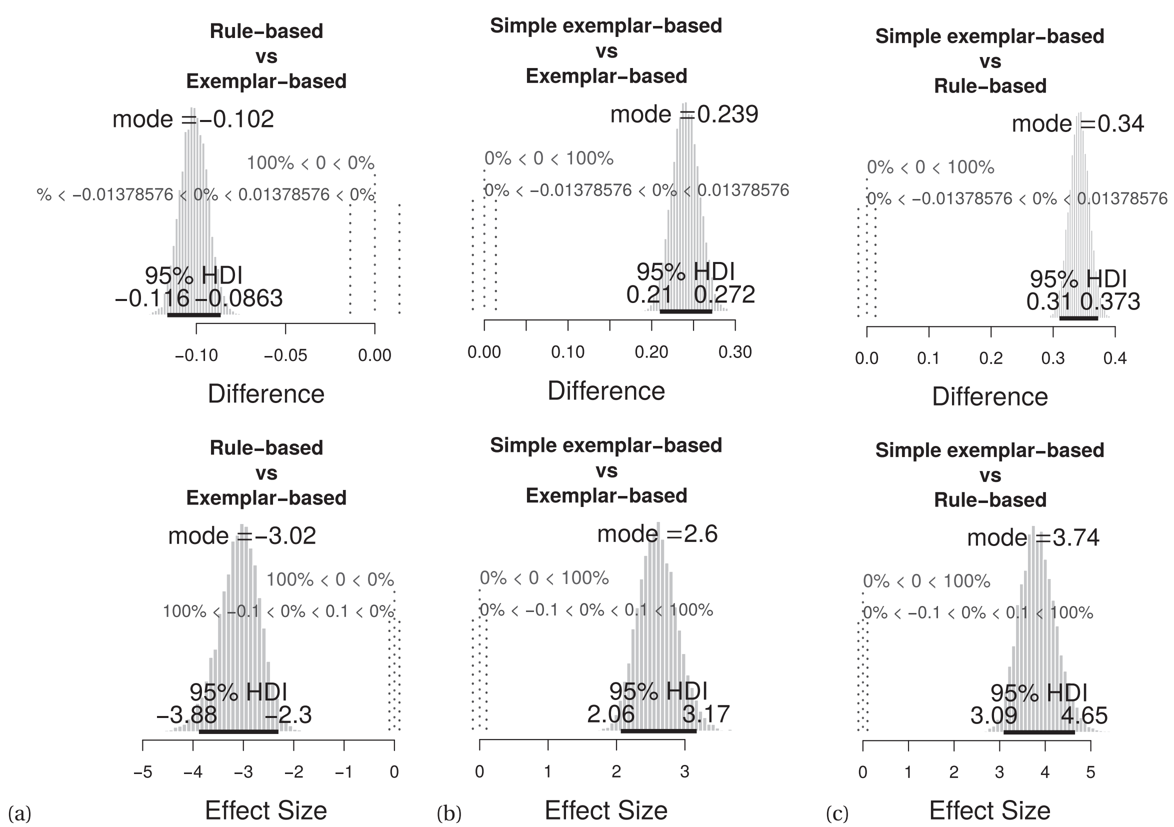

Fig. 14. Comparison between the different categories in the grid function-learning paradigm for Ain. The figure displays the difference and the effect size for (a) rule-based vs. exemplar-based extrapolators and (b) simple exemplar-based vs. exemplar-based extrapolators.

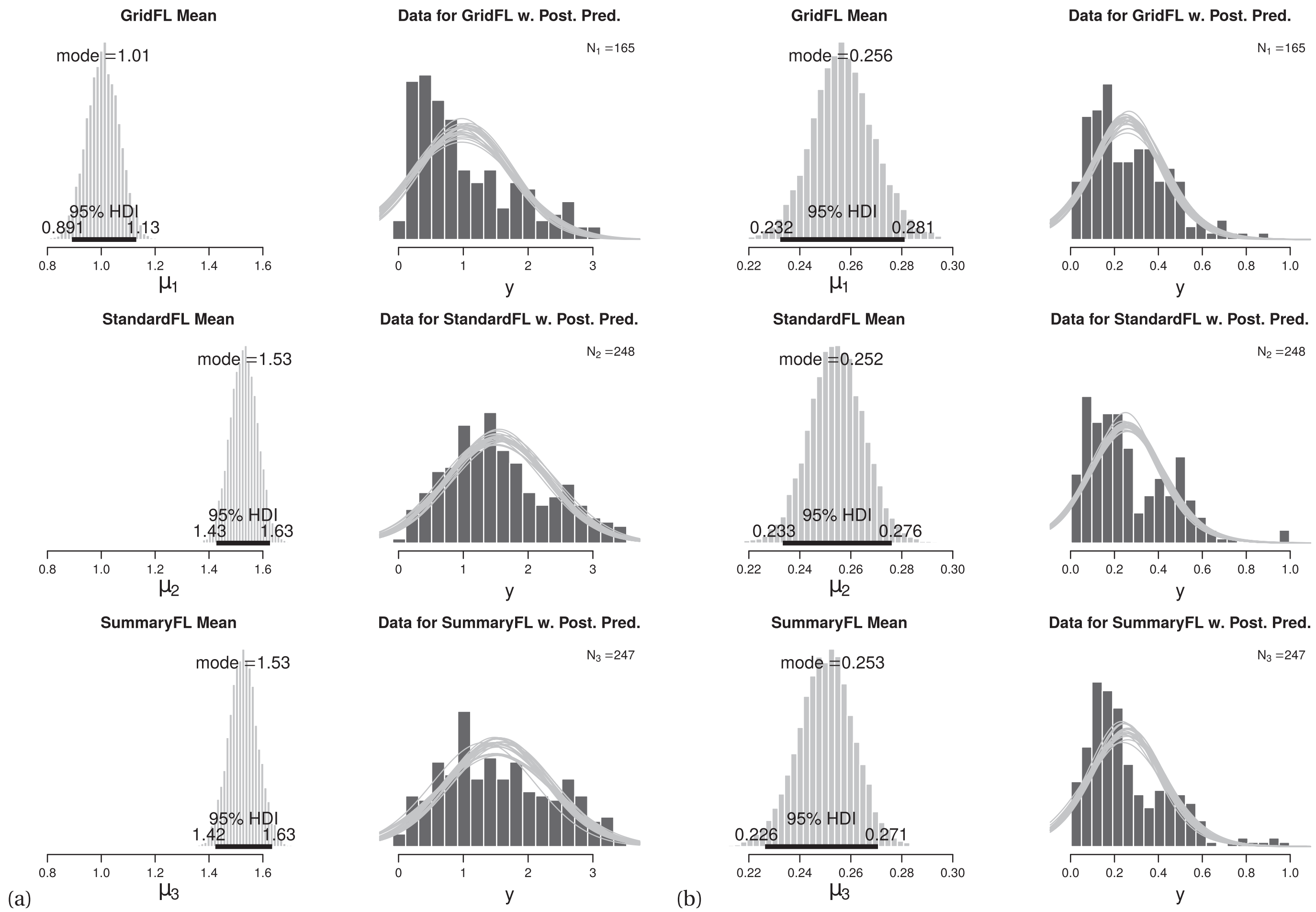

Fig. 15. Mean rRMSEs for the different conditions (grid, standard FL, summary FL) based on the hierarchical Bayesian procedure. The figure displays the posterior distributions of the mean rRMSEs μ for the standard FL, summary FL, and grid condition (left) and the corresponding posterior predictive check (right) for the two functions: Ain (a) and Bin (b).

Taken together, the present results advance our theoretical understanding of the cognitive processes involved in function-learning tasks in two ways. First, a considerable proportion of participants may well possess an understanding of the non-linearity of processes that is captured only when using measures of extrapolation performance that capture the core-feature of the functions displayed (in our case increasingly declining slopes), rather than simple measures of deviation (such as rRMSES). And second, a considerable proportion of participants may furthermore acquire an understanding of the non-linearity of a process that they struggle to implement in number-based input formats, but may well be able to implement in graphical input formats. Hence, in order to capture the full range of participants acquiring a deeper understanding of the to-be extrapolated processes, it may prove fruitful for future function-learning studies to consider several ways of assessing “extrapolation accuracy” as well as not relying on one, but rather several input formats. To conclude, both the input format and the method of evaluation may matter for a reliable measurement of understanding.

While different measures of quantifying extrapolation accuracy produced different results for all processes tested, different input formats produced different results for one of the processes (Ain), but not for the other (Bin). While we can only speculate about the specific reasons for this divergence, a plausible reason might be that Bin was the function with the flattest slope, and thus the easiest to extrapolate. Participants might therefore be able to extrapolate relatively accurately irrespective of the input format. Another potential explanation is that after the first function (Ain), participants were probably more familiar with the more difficult task format, which might have increased their ability to estimate correct number values.

Contrary to our expectation, extrapolation accuracy was not affected by implementation errors caused by a critical feature of classical function-learning experiments, the successive (as opposed to instantaneous) presentation of function values. Extrapolation accuracy was not higher in an alternative presentation format (summary function learning condition) that provided participants instantaneous access to current, as well as previous function values. This result suggests that while cognitive resources (working memory capacity) may be a limiting factor for rule-induction (McDaniel et al., 2014), they seem to be less relevant for rule-application during extrapolation. This result is akin to findings from research in analogical reasoning: usually participants have access to the whole set of problem elements (such as words) necessary to solve the task (e.g. Viskontas et al., 2004), thereby minimizing working memory requirements. However, also presenting the problem elements consecutively did not affect performance (Cho et al., 2007). Hence, when finding rules in analogical reasoning, as well as in function-learning, instantaneous vs. consecutive display of the necessary task elements seems to affect performance surprisingly little.

Interestingly, while performance generally dropped as a function of slope, extrapolation accuracy was more strongly affected by function slope (>80% drop for Cin compared to Ain and Bin) compared to learning of the correct function shape (approx. 15% drop for Cin compared to Ain and Bin). Furthermore, learning of the correct function shape AND slope was not influenced by function slope, suggesting that function slope impairs rule learning to a lesser extent than extrapolation accuracy. This result suggests that the well-established tendency toward

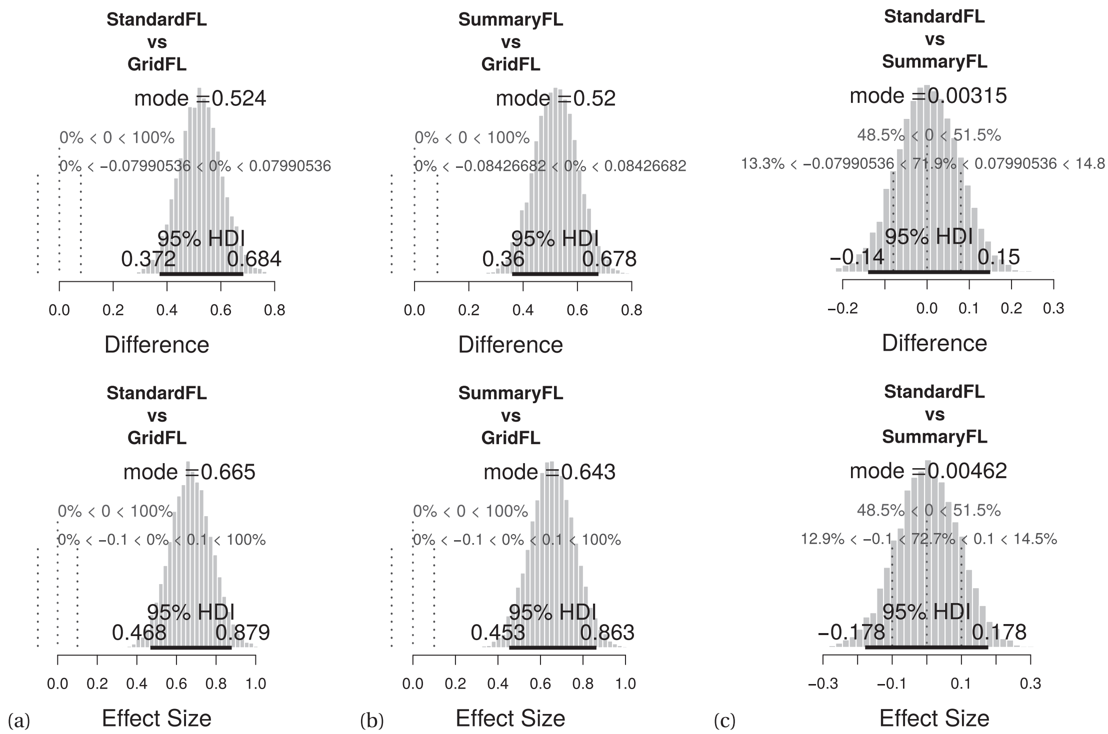

Fig. 16. Pair-wise comparisons between the different conditions (grid, standard FL, summary FL) for Ain. The figure displays tests for statistical significance of the difference as well as effect sizes.

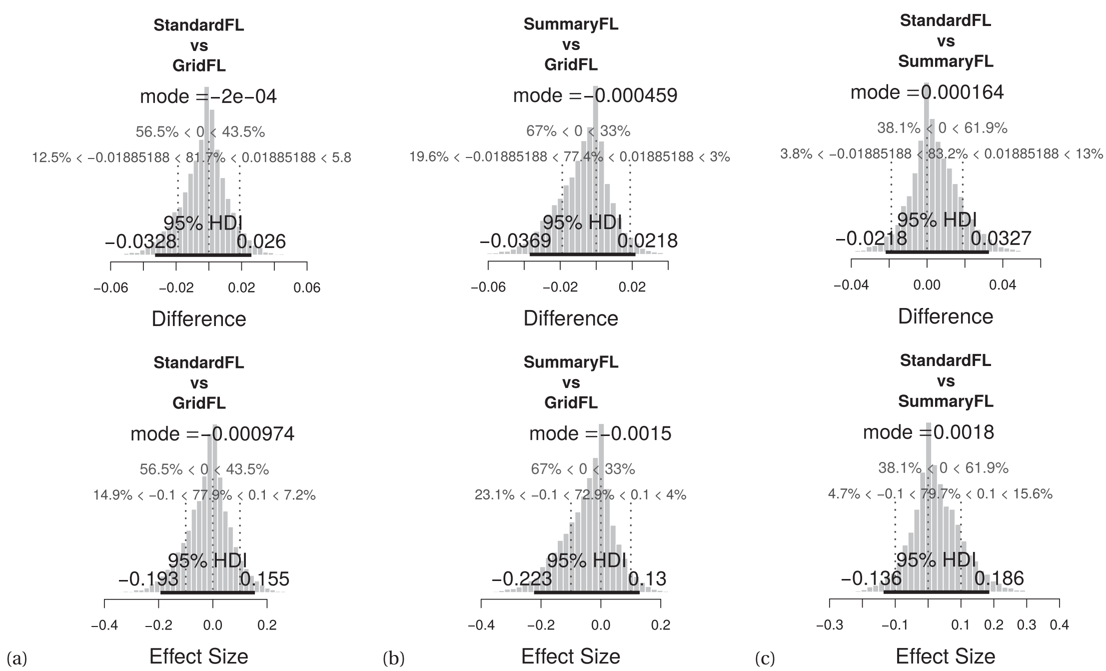

Fig. 17. Pair-wise comparisons between the different conditions (grid, standard FL, summary FL) for Bin. The figure displays tests for statistical significance of the difference as well as effect sizes.

linear extrapolations (Busemeyer et al., 1997) more strongly reflects a difficulty to extrapolate non-linearly than a more basic difficulty to recognize non-linear processes as non-linear. However, these results need to be interpreted with caution as functions were displayed in a fixed order. Any performance differences between the functions might hence potentially be due to learning effects, rather than differences in function slope. While we cannot rule out this option based on the present results, learning effects do appear less likely, because of the observed *decline* in extrapolation accuracy for the last (and steepest) function, Cin.

The generally high proportions of participants who could identify the correct shape and slope of exponentially decreasing functions is therefore particularly telling in the present experiment using exponentially declining functions since these are among the function types with the strongest deviation from participants' expectation of positive linearity.

For the first process, the relative majorities of participants displaying exemplar-based or simple exemplar-based extrapolations (67% and 52%, respectively) could identify the correct function shape in the rule-selection task, while considerably smaller proportions of participants (19% and 27%, respectively) believed the trained functions to be actually linear. Subsequent processes showed a similar pattern of results, with one exception: for the process Cin the majority of simple exemplar-based learners believed the process to be linear (47%). For the exemplar-based learners results were the same as for the previous functions, that is the majority identified the correct function shape (53%) even though extrapolating linearly. That is, approximately half of participants extrapolating linearly were well-aware that extrapolations should not in fact be linear. For participants displaying rule-based extrapolations, the pattern was reversed in that the majority believed their extrapolation style to be accurate. These results suggest that participants, up to a certain extent, are aware of what accurate extrapolations should look like, and that around half of exemplar-based and simple exemplar-based extrapolators employed linear extrapolations *despite* their understanding of non-linearity.

Such a disconnect between rule knowledge and rule application has also been found in the category-learning literature, where participants tended to use exemplar-based classifications even when explicitly given the correct rules before training, and instructed to apply those rules when categorizing (Allen & Brooks, 1991; Regehr & Brooks, 1993). One possible explanation for a disconnect in this direction (participants being aware of a rule, but applying exemplar-based strategies) is provided by the connectionist model of Erickson and Kruschke (1998) stating that even though knowledge about the correct rule exists, attention is shifted to the learned associations between exemplars.

Future research could investigate whether a disconnect between rule knowledge and rule application can be experimentally enhanced, for example by explicitly providing the last two training points. If salience of the last training points is high, even participants who proved to know the correct rule in a rule-selection, or grid paradigm might be more likely to use simple exemplar-based strategies by extrapolating through these points. It might also be worthwhile to investigate whether such an experimental enhancement of exemplar-based extrapolation depends on individual difference parameters such as working memory capacity, which has been shown to be relevant for both rule-abstraction, and application (Fischer & Holt, 2016; McDaniel et al., 2014). Similarly to the present results demonstrating that rule knowledge is not necessarily applied when predicting non-linear processes (and how this has previously been shown in the category-learning literature, Allen and Brooks (1991); Regehr and Brooks (1993)), moreover, future research could explore the conditions under which rule-application given rule knowledge is enhanced in more applied settings. When making predictions about the development of climate change or greenhouse gas emissions, for instance, application of the correct function rule might depend on predictors' prior beliefs about the issue, or political attitude. This is because strong motivational factors might exist to downplay the severity of future developments in politicized domains. Hence, when framing an extrapolation task as a "climate change prediction task" (as opposed to more neutral bacteria growth frames, for instance), participants might be differentially willing to apply a learned rule.

It is a common assumption of many function-learning studies that given that participants acquired an understanding of the function rule, they also apply that rule during extrapolation. The present results suggest, however, that a considerable proportion of participants who had acquired an understanding of the accurate function displayed exemplar-based or even simple exemplar-based extrapolation in the classical function-learning paradigm. We conclude that rule-learning is not tantamount to rule application and that the proportion of rule-based learners in the current function-learning literature likely represents an underestimation.

Declaration of competing interest

(1) Hereby we confirm that, this manuscript is original, not previously published, and not under concurrent consideration elsewhere.
(2) The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge support by the Excellence Initiative, Institutional Strategy ZUK 5.4 (Scientific Computing in the Social and Behavioral Sciences), Heidelberg University.

https://figshare.com/s/6be582d99deee687c126

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120*(1), 3–19.

Bott, L., & Heit, E. (2004). Nonmonotonic extrapolation in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 38–50.

Busemeyer, J. R., Byun, E., Delosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In *Studies in cognition knowledge, concepts and categories* (pp. 408–437).

Carroll, J. D. (1963). Functional learning: The learning of continuous functional mappings relating stimulus and response continua. *ETS Research Bulletin Series, 1963* (2), i–144.

Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition, 35*(6), 1445–1455.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 968.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*(2), 107–140.

Fischer, H., & Holt, D. V. (2016). When high working memory capacity is and is not beneficial for predicting nonlinear processes. *Memory & Cognition, 45*(3), 404–412.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

Kalish, M. L. (2013). Learning and extrapolating a periodic function. *Memory & Cognition, 41*(6), 886–896.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review, 111*(4), 1072–1099.

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Kwantes, P. J., & Neal, A. (2006). Why people underestimate y when extrapolating in linear functions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(5), 1019–1030.

Lewandowsky, S., Kalish, M., & Ngang, S. K. (2002). Simplified learning in complex situations: Knowledge partitioning in function learning. *Journal of Experimental Psychology: General, 131*(2), 163–193.

MacKinnon, A. J., & Wearing, A. J. (1991). Feedback and the forecasting of exponential change. *Acta Psychologica, 76*(2), 177–191.

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestr: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software, 4*(40), 1541.

McDaniel, M. A., & Busemeyer, J. R. (2005). The conceptual basis of function learning and extrapolation: Comparison of rule-based and associative-based models. *Psychonomic Bulletin & Review, 12*(1), 24–42.

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*(2), 668–693.

Regehr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General, 122*(1), 92–114.

Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, inhibition, and analogical reasoning in older adults. *Psychology & Aging, 19*(4), 581–591.