Multinomial Processing Models for Syllogistic Reasoning: A Comparison

Hannah Dames (damesh@cs.uni-freiburg.de), Jan Ole von Hartz (hartzj@cs.uni-freiburg.de), Mario Kantz (mario.kantz@mars.uni-freiburg.de), Nicolas Riesterer (riestern@cs.uni-freiburg.de), Marco Ragni (ragni@cs.uni-freiburg.de)
Cognitive Conputation Lab, Department of Computer Science, Georges-Köhler-Allee 52, 79110 Freiburg, Germany

Abstract

To this day, a great variety of psychological theories of reasoning exist aimed at explaining the underlying cognitive mechanisms. The high number of different theories makes a rigorous comparison of cognitive theories necessary. The present article proposes to use Multinomial Processing Trees to compare two of the most prominent theories of syllogistic reasoning: the Mental Models Theory and the Probability Heuristics Model. For this, we reanalyzed data from a meta-analysis on six studies about syllogistic reasoning. We evaluate both models with respect to their overall fit to the data by means of G^2, AIC, BIC, and FIA, and on a parametric level. Our comparison indicates that a MMT-variant, though having more parameters, is slightly better on all criteria except of the BIC. Yet, none of the two models, realized as MPTs, is clearly superior. We outline the impact of the different theoretical principles and discuss implications for modeling syllogistic reasoning.

Keywords: Syllogistic Reasoning; Mental Models Theory; Probability Heuristics Model; Multinomial Processing Trees

Introduction

Consider the following two abstract statements:

(1) No researchers are gods.
Some gods are great reasoners.
What, if anything, follows?

When people are asked to draw a conclusion about *researchers* and *great reasoners* based on these two premises, some conclude that "some researchers are not great reasoners" or that "nothing follows". Yet, the only logically valid conclusion would be "Some great reasoners are not researchers". The two premises each have one of four possible quantifiers, called moods in their syllogistic combination: All (abbreviated by A), Some (I), Some...not (O), and None (E). Four different arrangements of the terms in the premises are possible. These are called figures and we use the numbering of the figures as in Bucciarelli and Johnson-Laird (1999):

Figure 1	Figure 2	Figure 3	Figure 4
A-B	B-A	A-B	B-A
B-C	C-B	C-B	B-C

Any syllogism can be described by the respective quantifiers of the first and second premise and its figure. We can write EI1 to uniquely characterize the syllogism given in (1). Hence, there are 64 distinct syllogisms.

Since 1908, more than twelve cognitive theories of human reasoning about syllogisms have been proposed. A meta-analysis conducted by Khemlani and Johnson-Laird (2012) compared the predictions of these theories based on three point estimates for each theory. First, the authors calculated the proportion of hits, measured using the sum of the proportion of the responses that were predicted by the theory and were also given within the experiments. Second, they used the proportion of correct rejections measured using the proportion of non-predicted responses that did not occur empirically. Last, they analyzed the overall proportion of correct predictions, which combines the hits and correct rejections. In general, these estimates allow for a comparison of different theories focusing on the responses given in an experiment. However, this approach is lacking rigorous quantitative criteria to compare the different models beyond participants' predicted conclusions. The predictive power of the proposed theories remains unclear. We intend to fill this gap by transforming two of the most prominent theories, the Probability Heuristics Model (PHM, Chater and Oaksford (1999)) and the Mental Models Theory (MMT, Johnson-Laird and Steedman (1978)) into probabilistic models. The PHM proposes a set of simple heuristics individuals use to draw a conclusion, whereas the MMT asserts the construction, usage, and modification of mental models.

Both MMT and PHM assume the existence of latent cognitive processes that occur while solving reasoning problems. For a fair evaluation and to explicitly encode the assumptions in the model, we used *Multinomial Processing Trees* (MPTs). MPTs are a family of cognitive models for the analysis of categorical data (Riefer & Batchelder, 1988; Moshagen, 2010). The usage of MPTs in cognitive science and psychology has grown in the past two decades (Erdfelder et al., 2009). Most published models refer to various memory paradigms, such as recognition, source monitoring, and process dissociation, but also other fields, for instance reasoning and, recently, implicit attitude measurement, just to name a few (for an overview see Erdfelder et al., 2009). In this work, we make use of the `MPTinR` package (Singmann & Kellen, 2013) in R. Within this framework, models are fitted as binary trees, where each node in the tree represents a cognitive state or process. Therefore, each branch represents a theoretically motivated, assumed sequence of cognitive processes that take place between an input (presented premises) and a response, here possible conclusions (Erdfelder et al., 2009). The probability that a latent stage is reached thus depends on the successful occurrence of other, associated processes. The rigorous comparison of the MMT and PHM requires a weighting between the ability of each model to account for the observed data and to be generalized to other datasets (its flexibility). In order to find the model with the best trade-off between goodness of fit and flexibility, we calculated four measures. The smaller each of the following measures' value the better it is. First, the goodness of fit was measured using the G^2 statistic, which maximizes the likelihood of the frequencies of observations given the parameter values. Second, the *Akaike Information Criterion* (AIC, Akaike, 1974) and the *Bayesian Information Criterion* (BIC, Schwarz et al., 1978) were calculated. Both information criteria penalize models according to their number of free parameters and indicate how much information was lost when a model represented the process that generates the data. Last, the *Fisher Information Approximation* (FIA, Wu, Myung, & Batchelder, 2010) was used to measure the flexibility of the models. The FIA estimates the amount of information that an observed frequency carries about a parameter which models the observation. This allows us to compare models of different sizes. In the following, we briefly outline the core principles of both theories and their implementation as MPTs.

The Theories

Mental Models Theory According to the MMT, syllogistic reasoning involves the construction of iconical representations of situations, for instance, sets of people that belong to different subgroups (e.g., Bucciarelli & Johnson-Laird, 1999; Johnson-Laird, 1975; Johnson-Laird & Steedman, 1978). These representations embody mental models of the world, from which conclusions can be inferred. The MMT postulates that first, humans build an initial model of the two premises. Then, a preliminary conclusion is drawn from this initial model. Next, using specific operations (see Figure 1), individuals are assumed to seek for counterexamples to this conclusion analogous to an attempt to falsify it (Johnson-Laird & Steedman, 1978). If no conflicting models are constructed, the conclusion based on the initial model is maintained, otherwise it is refuted and another conclusion is sought (see Johnson-Laird & Steedman, 1978). The more models that are necessary to draw a conclusion, the harder the inference becomes (Johnson-Laird & Khemlani, 2013): Syllogisms with just one valid model are called *One Model Problems* (OMP) and are considered to be easier than *Multi Model Problems* (MMP) and *No Valid Conclusion Problems* (NVCP) (Johnson-Laird & Steedman, 1978). While the earlier stages involve heuristics, the testing step makes the whole process logically valid, if executed correctly. Thus, the MMT can account for differences between logically trained and untrained individuals.

Probability Heuristics Model The PHM as proposed by Chater and Oaksford (1999) is inspired by a Bayesian approach and builds on the idea that naive logical reasoners employ heuristics that often yield probabilistically valid ("p-valid") rather than logically rigorous conclusions. The application of these heuristics has two prerequisites: First, a total order of informativeness of quantifiers, which is obtained by complementing the *existential presupposition* (assuming that "All A are (not) B" implies "Some A are (not) B", i.e. that universal claims are not made on empty sets) with a rarity assumption (attributes in descriptions using natural language rarely overlap, so E-statements are usually true and therefore less informative than I-statements), yielding: A > I > E > O.

[A]			[A]			[A]		C	[A]		C
[A]			[A]			[A]			[A]		C
	[B]	C		[B]	C		[B]	C		[B]	C
	[B]			[B]			[B]			[B]	
					C						

Figure 1: Four illustrative mental models for "No A are B. Some B are C." Each row represents the properties of an individual. The square brackets signify that the set of As and Bs are represented exhaustively. When the left model is the initial model, the second model can be created by adding an individual entailing only property C. The third and fourth model can be built by merging the first and the fifth entity, respectively. The four models enable the read-off of different conclusion (e.g., for the third model "Some A are C", which is refuted by model one and two). Only "Some C are not A" is a valid conclusion. Note that it is also assumed that reasoners are able to construct fully explicit mental models (Khemlani & Johnson-Laird, 2012) representing what is false in addition to what is true. This is done by using mental footnotes, often symbolized with the token "¬B" for negation, that prohibit the existence of co-occurrences with other entities.

Second, quantified assertions can entail others (so called p-entailments), either due to the already mentioned existential presupposition, or due to Gricean Implications (Grice, 1975): the usage of a particular (I, O) instead of universal (A, E) statement is taken to imply that the universal statement is wrong. Hence, I and O p-entail each other.

The proposed heuristics comprise three for generating conclusions and two for testing them:

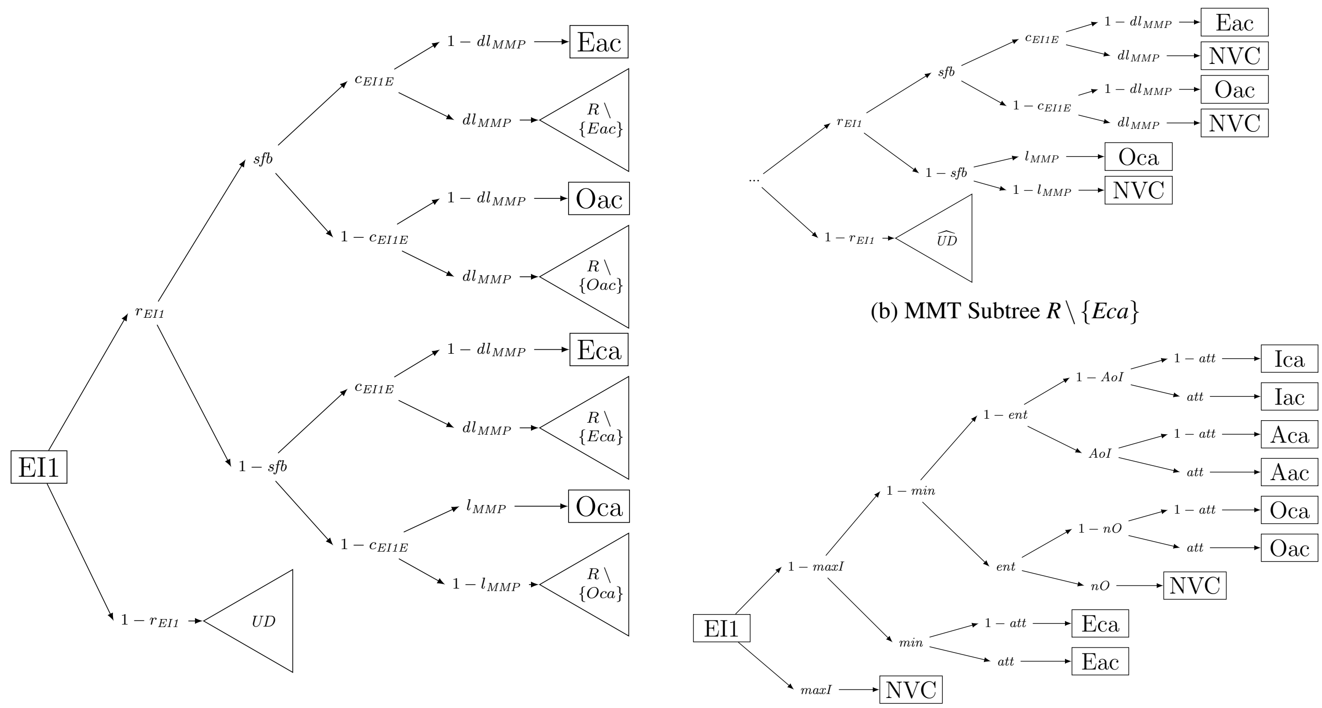
- *Min*: The preferred conclusion quantifier is that of the less informative premise (min-premise).
- *Entailment*: The second preferred conclusion quantifier is that of the min-premise's p-entailment.
- *Attachment*: The end term in the min-premise retains its position as either subject or predicate in the conclusion.
- *Max*: Confidence in the generated conclusion is proportional to the informativeness of the more informative premise (max-premise).
- O: Avoid drawing O-conclusions.

The PHM assumes these heuristics to be a rather complete description of the underlying processes in syllogistic reasoning. Additionally, it does not make a statement on preference for the remaining two quantifiers once neither the min- nor the entailment-heuristic were accepted. Note that in the classical version of the PHM, the answer NVC is not predicted. We return to this later as it needs to be included in the MPT.

Construction of the MPTs

Mental Models Theory

We implemented the MMT as an MPT based on the following stages proposed by the MMT: First, we modeled individual reasoning parameters estimating the probability with which a



(a) MMT Tree

(b) MMT Subtree $R \setminus \{Eca\}$

(c) PHM Tree

Figure 2: The MMT and PHM trees for "No A are B. Some B are C."(EI1). Capital letters refer to the quantifiers ($A = All$, $I = Some$, $O = Some...not$, $E = None$). Manifest, observable variables are drawn as squares (on the left: the premises, the number indicates the figure; on the right: responses as conclusions, *ac* indicates A-C and *ca* C-A conclusions e.g., Eac = "None of the A is C"; NVC="no valid conclusion"). In between, the assumed latent states (not observable) are printed borderless. The MMT tree begins with the reasoning parameter r_i ($i =$ Syllogism), followed by the Figural Bias parameter for applicable syllogisms (e.g., for IE1 the Strong Figural Bias *sfb*), and (if necessary) the decision between the predicted quantifier (e.g., C_{EI1E}). Finally, the preliminary conclusion is confirmed with l_j if correct, with $1 - dl_j$ otherwise, or rejected with the respective converse probability (here: $j = MMP$). After rejecting, a new conclusion is sought (see Figure 2b, subtree: $R \setminus \{c\}$, with c being the discarded conclusion). *UD* indicates the uniform distribution over all nine conclusions, $\widehat{UD}$ over the remaining eight. The PHM tree begins with the max-heuristic in the case of max-premise being I (maxI), selects quantifier E from min-heuristic (min), or quantifier O from entailment-heuristic (ent), or either of the remaing two (AoI). In the case of O, the O-heuristic is applied (nO). Finally, the order of end terms is set using or refusing the attachment-heuristic (att).

response would be produced by a general reasoning or guessing process (Ragni, Singmann, & Steinlein, 2014). The probabilities of the reasoning branch were unrestricted, since we assumed differences in reasoning for each syllogism based on difficulty and content effects. The first stage can be referred to as the "reasoning" stage. For the guessing parameters we assumed a uniform distribution over all possible conclusions. In sum, for the implementation of the MPT we had to make a number of assumptions: A subject will (i) not build more than one revised model (to limit the depth of the MPT model), (ii) will neither conclude nor guess an already refuted conclusion, (iii) only answer NVC through refuting two distinct conclusions or guessing, and (iv) possibly return to guessing after discarding one conclusion. Normally, the "reasoning"-subtree should begin with representing the construction of an initial model on the basis of heuristics (Johnson-Laird & Steedman, 1978). However, this process is thought to take place rapidly and intuitively (e.g., Johnson-Laird & Steedman, 1978; Khemlani & Johnson-Laird, 2016) and failing to create a model forces a subject to guess. We therefore subsume this stage in the "reasoning"-parameter, which is already dependent on the type of syllogism. It is assumed that building one initial model from both premises is subject to a heuristic bias towards linking up end items by way of middle items (Johnson-Laird & Bara, 1984), causing an effect known as the *figural bias*. The inspection of this initial model, and thereby the drawing of a preliminary conclusion, is influenced by this bias. We include this effect by letting the figural bias model the direction of reading off. Depending on their figure, syllogisms are divided into three types based on their affinity towards figural bias: strong, weak and no figural bias. If the conclusion matches this bias, the respective parameter is appended to the path, otherwise its converse probability. Sometimes, multiple conclusions with the same figure are predicted. Subsequently, another stage with free variables specific to each syllogism determines which quantifier is concluded.

Finally, a falsification attempt is made. Because this stage is considered to be easier for OMPs than for MMPs or NVCPs, this parameter is dependent on this classification of the syllogisms. A valid conclusion could be given either by successful testing or not testing at all. In contrast, the refutation of an invalid conclusion definitely requires some logical assessment of the validity of putative conclusions, i.e., testing. Therefore, we use distinct variables here. Valid conclusions are refuted and invalid ones confirmed with the respective converse probability. Confirming a conclusion will make a subject answer respectively, while refuting one is likely to initiate the search for an alternative conclusion. If the participant again fails to find a valid conclusion, the response is assumed to be that there is no valid conclusion (NVC). Note that, although it is reasonable to assume that the participant may give up after two unsuccessful attempts, concluding that there is NVC, there is no empirical evidence for this assumption as of yet. Moreover, this constrain was chosen for modeling reasons and simplicity. See Figure 2a and 2b for an example of a tree constructed as detailed above.

Probability Heuristics Model

To construct an MPT representing the PHM, we first need to look at the dependencies of the heuristics. For a given syllogism, the PHM identifies the quantifier of the *max-premise*, the quantifier of the *min-premise*, the *p-entailment* of the latter, and the end-term ordering suggested by the *attachment-heuristic* following the principles outlined above. In the following, parameter names are given in parentheses: Since the confidence in the preliminary conclusion according to the *max-heuristic* depends only on the premises and not on the actual choice of quantifier or ordering, we decided to prepend the decision whether or not to discard the preliminary conclusion and answer with NVC instead. Hence, this uses one of four parameters corresponding to each *max*-quantifier (maxA, maxI, maxE, maxO) that is applicable to the given syllogism. If the choice to give a supposedly valid conclusion is made, quantifier selection for the conclusion is initialized by first trying the *min-heuristic* (min). If it is refused, its respective *p-entailment* is considered (ent). If this is refused as well, a binary decision between the remaining two quantifiers yields a final quantifier choice (AoI, EoO, AoE). If the quantifier selected for the conclusion is O, the O-heuristic is applied through another trial whose failure leads again to NVC (nO). If it passed or the quantifier is different, the tree ends with a binary choice for the order of the end terms in the conclusion either in accordance or contradictory to the attachment heuristic (att). Figure 2c shows the thereby constructed tree for the same example syllogism as before, EE1.

Method

We used the data from the meta-analysis on syllogistic reasoning provided by Khemlani and Johnson-Laird (2012) in order to fit the models created for both the MMT and PHM. The data set consists of six empirical studies (Johnson-Laird & Steedman, 1978; Johnson-Laird & Bara, 1984; Bara, Bucciarelli, & Johnson-Laird, 1995; Roberts, Newstead, & Griggs, 2001) with a total sample size of $n = 156$ and we used the aggregated results (see Table 6 in Khemlani & Johnson-Laird, 2012). In all experiments participants were presented with two premises and instructed to draw their own conclusions to all 64 syllogisms (i. e., participants were, for instance, asked what followed necessarily from the premises).

MPT Analysis for Model Comparison

A model selection analysis was used to evaluate the two discussed cognitive theories. Each of the proposed models was fitted to the aggregated data via Maximum Likelihood Estimation using `MPTinR` (Singmann & Kellen, 2013). The package also makes use of the four introduced measures (G^2, AIC, BIC, and FIA). The following approach was taken to systematically evaluate the theoretical assumptions of each theory.

First, we fitted a model that only consists of the guessing subtree, modeling a uniform distribution over all possible conclusions, as a standalone model. This *Guessing-Model*, having no reasoning path for any conclusion, served as a baseline to evaluate the MPT implementations of the MMT and PHM. If the reasoning subtrees of the theories contribute in explaining the data considerably, the information criteria should be better (lower) than those for the Guessing-Model. Second, we fitted an unrestricted model for each theory, that does not include any restrictions with the exception of the guessing parameters. This model served as a reference model for the following models, that included restrictions proposed by the theoretical framework. If the assumptions raised in the theory hold true, the fit of the restricted models should not be considerably worse than for this unrestricted model. Third, for the MMT, restrictions were added using a hierarchical, stepwise approach. Last, the full models were fitted using all restrictions. In summary, the discussed theoretical assumptions of the MMT can be represented by the following parameter restrictions in the MPT: One Model Problems should be easier to solve correctly than both Multiple Model and NVC Problems, since no alternative models are needed to verify a conclusion. Thus: $l_{nvcp} < l_{omp}, l_{mmp} < l_{omp}$. These constraints are included in the full models and the “Number MM” model. For syllogisms that are subject to the strong figural bias (*sfb*) or the weak figural bias (*wfb*) the corresponding branching should be taken with a probability higher than 0.5. Also we should observe that $wfb < sfb$. For all other syllogisms no figural bias (*nfb*) is expected, thus $nfb = 0.5$. These constraints are included in the full models and the “Figural Bias” model. The suitability of the different parameter restrictions can be compared by evaluating the relative performance of the models instantiating them.

For the PHM, we implemented three different sets of possible restrictions: first, no restrictions, second, restricted order of the four max-parameters, and, third, restricted choice between quantifiers, after min and entailment have failed, to

Table 1: Results of MPT fits to the aggregated data set of Khemlani and Johnson-Laird (2012).

Model		***k***	G^2	**AIC**	**BIC**	**FIA**	**CFIA**
Guessing	Baseline	1	17581.69	17583.69	17590.84	8795.645	4.80
MMT	Unrestricted	96	**4224.94**	**4416.94**	5103.32	2323.32	210.85
	Figural Bias	95	4323.87	4513.87	5193.11	2367.02	205.09
	Number MM	96	**4224.94**	**4416.94**	5103.32	**2319.20**	206.73
	Full Model	95	4323.87	4513.87	5193.11	2364.56	202.62
	Unrestricted, Global r	33	4820.98	4886.98	5122.92	2494.74	84.26
	Full Model, Global r	32	4934.86	4998.86	5227.66	2546.24	78.80
PHM	O, none	11	4844.29	4866.29	**4944.94**	2463.13	40.99
	O, max	11	4844.75	4866.75	4945.40	2461.08	38.70
	O, uniguess Q	8	5183.02	5199.02	5256.22	2619.88	28.38
	not O, none	10	5052.8	5072.80	5144.29	2565.48	39.08
	not O, max	10	5044.17	5064.17	5135.67	2558.68	36.59
	not O, uniguess Q	7	5393.04	5407.04	5457.09	2722.67	**26.15**

Note. The resulting model parameters. k indicates the number of parameters. The total number of degrees of freedom is $df = 512 - k$, all $p < .001$. The smallest value per column is printed in bold. CFIA: penalty term for FIA.

be uniform. Additionally, we implemented two variants of the model: the one described above, and - after initial results hinted at some conflict between the max- and O-heuristic - one where the O-heuristic was omitted.

All models were fitted using 10 optimization runs. FIA was estimated using 200.000 Markov Chain Monte Carlo samples. The full dataset had $2 \times 4 \times 64 = 512$ available degrees of freedom.

Results

First, we evaluated the MMT and PHM respective to their predictive power. Table 1 shows the results of the hierarchical, stepwise fitting approach for both the MMT and PHM as well as the pure Guessing-Model. As expected, the Guessing-Model has the worst fit. Therefore, the reasoning subtrees for both models add a substantial amount of predictive power. Considering all information criteria, the MMT fits the data best. This is not too surprising given that it predicts a larger set of conclusions for any syllogism than the PHM and also has the largest number of parameters.

In the next step of the analysis we looked at the reasoning-parameters of the MMT models. Assuming that most of the participants reason instead of randomly guessing when giving a syllogistic task, the reasoning parameters should be larger than 0.5 if the MPT makes reasonable predictions. For all models of the MMT, the mean probability for a reasoning-based response was 90% ($M = 0.90$ over all r_i, $SD = 0.08$). The high probability of reasoning processes postulated by the theory indicates an overall satisfying model fit. Furthermore, the low standard deviation indicates only minor differences in r_i. Hence, participants' probability to reason does not differ greatly for individual syllogisms. Based on this finding, we constructed another set of MPTs identical to the presented MPT with the exception that this time, we assumed a global reasoning parameter r, equal for all of the 64 syllogisms, resulting in a more parsimonious tree (number of parameters $k = 32$ for the global r-model compared to $k = 95$ for MPTs that assume individual r_i). Although this approach reduced the amount of parameters dramatically resulting in a lower FIA penalty estimate, the FIA estimate as well as the other information criteria increased (see Table 1).

Likewise, inspecting the parameter fits for the PHM trees without restrictions, we find that the predicted probabilities to choose in accordance with the min- (76%), entailment- (70%) and attachment- (68%) heuristics all constitute significant preferences in line with the theory. The order predicted by the max-heuristic is generally matched well by the fit results, as evident from both the resulting values and the improved FIA of the restricted over the unrestricted model. The only exception to this is the order between max-premise being E and it being O, where in the unrestricted case with O-heuristic, we find the order barely swapped (46% for E and 44% for O), though this corresponds to only a few syllogisms. In the case where the O-heuristic is omitted from the model, the expected order is restored (55% for O and 53% for E). All in all, this suggests that the theory is well reflected in the given MPT implementation on the given data set.

General discussion

Multinomial Processing Trees (MPTs) provide a powerful mathematical framework to model cognitive theories of reasoning and to quantify the impact of cognitive processes. We have developed MPTs for two prominent theories of syllogistic reasoning making implicit processes explicit: One that assumes the use of mental models (MMT) and one that uses a heuristics model inspired by a probabilistic, Bayesian approach (PHM). Our findings are in line with previous results (e.g., Khemlani & Johnson-Laird, 2012; Ragni et al., 2014) regarding the ability to explain a sufficient amount of empirical data. In addition to G^2, our implementation of the

MMT outperforms PHM in respect to AIC and FIA despite its higher number of parameters, whereas the PHM scores better in the BIC. An analysis of the contribution of different processes (figures or heuristics) demonstrates that some formalizations are better than others. This aspect of MPTs, allowing for a systematic evaluation of the contribution of processes is more important than an overall comparison, as it helps to identify the contribution of specific processes. In order to support the processes, the MPTs can inspire a sequence of experimental studies to systematically manipulate these cognitive states, so that the manipulation should then, in turn, also be reflected in the parameters of the MPTs. By doing so, the various model extensions or modifications proposed in this work could be further supported. In a next step, the MPT framework allows to analyze combinations of theories, e.g. joint MPTs replacing the guessing part of the MMT-MPT with the PHM-MPT. As a result, we may be able to infer what individual components may lead a person to either use an heuristic or a mental model approach when reasoning. Furthermore, future research could focus on existing computational frameworks for the modeling of reasoning processes. For instance, a well-developed framework is available in form of the computational implementation of the MMT, called `mReasoner` (Khemlani & Johnson-Laird, 2016). The implementation also depends on the construction and manipulation of mental models operating stochastically based on four separate parameters (Khemlani & Johnson-Laird, 2016): the size of a mental model, the model's contents, a counterexample search mechanism, and a nested parameter describing what happens when a counterexample is found. Similar to the proposed MPT, `mReasoner` has found to provide a close match to aggregated data from syllogistic reasoning studies (Khemlani & Johnson-Laird, 2016). Yet, the parameters of importance differ from the parameters of our MPT. A closer examination and comparison of these differences on a parameter-level is a next step. The great variety of cognitive theories that coexist bear the issue of lacking research on comparing and evaluating these theories in a unified framework. To solve this issue, we took a MPT modeling approach, which incorporates some promising features to compare these theories both in terms of their overall predictive power as well as their assumed cognitive states. As our work showed, this methodology is an excellent approach to disentangle latent processes.

Acknowledgements

This paper was supported by DFG grants RA 1934/3-1, RA 1934/2-1 and RA 1934/4-1 to MR.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716–723.

Bara, B. G., Bucciarelli, M., & Johnson-Laird, P. (1995). Development of syllogistic reasoning. *The American journal of psychology*, 157–193.

Bucciarelli, M., & Johnson-Laird, P. (1999). Strategies in Syllogistic Reasoning. *Cognitive Science*, *23*(3), 247–303.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive psychology*, *38*(2), 191–258.

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology*, *217*(3), 108–124.

Grice, H. P. (1975). Logic and conversation. *1975*, 41–58.

Johnson-Laird, P. (1975). Models of deduction. *Reasoning: Representation and process in children and adults*, 7–54.

Johnson-Laird, P., & Bara, B. G. (1984). Syllogistic inference. *Cognition*, *16*(1), 1–61.

Johnson-Laird, P., & Khemlani, S. S. (2013). Toward a unified theory of reasoning. In *Psychology of learning and motivation* (Vol. 59, pp. 1–42). Elsevier.

Johnson-Laird, P., & Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology*.

Khemlani, S., & Johnson-Laird, P. (2012). Theories of the Syllogism: A Meta-Analysis. *Psychological Bulletin*, *138*(3), 427.

Khemlani, S., & Johnson-Laird, P. (2016). How people differ in syllogistic reasoning. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.*

Moshagen, M. (2010). multitree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*(1), 42–54.

Ragni, M., Singmann, H., & Steinlein, E.-M. (2014). Theory comparison for generalized quantifiers. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36).

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, *95*(3), 318.

Roberts, M. J., Newstead, S. E., & Griggs, R. A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking & Reasoning*, *7*(2), 173–204.

Schwarz, G., et al. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461–464.

Singmann, H., & Kellen, D. (2013). MPTinR: Analysis of multinomial processing tree models in R. *Behavior Research Methods*, *45*(2), 560–575.

Wu, H., Myung, J. I., & Batchelder, W. H. (2010). Minimum description length model selection of multinomial processing tree models. *Psychonomic bulletin & review*, *17*(3), 275–286.