Individual differences in great ape cognition across time and domains: stability, structure, and predictability

Manuel Bohn[1,2], Christoph Völter[2], Johanna Eckert[2], Daniel Hanus[2], Nico Eisbrenner[2], Jana Holtmann[3], & Daniel Haun[2]

[1] Institute of Psychology in Education, Leuphana University Lüneburg
[2] Department of Comparative Cultural Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
[3] Wilhelm Wundt Institute of Psychology, Leipzig University, Leipzig, Germany

Author Note

Manuel Bohn was supported by a Jacobs Foundation Research Fellowship (2022-1484-00). We are grateful to thank all children and caregivers for participating in the study. We thank the Max Planck Society for the Advancement of Science.

The authors made the following contributions. Manuel Bohn: Conceptualization, Formal Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Christoph Völter: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Johanna Eckert: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Daniel Hanus: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Nico Eisbrenner: Formal Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Jana Holtmann: Formal Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Daniel Haun: Conceptualization, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Manuel Bohn, Universitätsallee 1, 21335 Lüneburg, Germany. E-mail: manuel.bohn@leuphana.de

Abstract

200 words

Keywords: keywords

Individual differences in great ape cognition across time and domains: stability, structure, and predictability

Introduction

Bohn et al. (2023) studied

The current study extended previous work in two important aspects. First, we study a broader range of cognitive domains including social cognition, reasoning about quantities, executive functions and inferential reasoning. This allows us to assess whether the results obtained by Bohn et al. (2023) replicate within domains and generalize to others. Second, we explored the structure of great ape cognition in more depth: we pooled the data collected here with the data from Bohn et al. (2023) to study the correlations between cognitive traits within and across domains.

Methods

Participants

A total of 48 great apes participated at least in one tasks at one time point. This included 12 Bonobos (*pan paniscus*, 4 females, age 3.60 to 40.70), 24 Chimpanzees (*pan troglodytes*, 17 females, age 3.80 to 57.80), 6 Gorillas (*gorilla gorilla*, 4 females, age 4.40 to 24.40), and 6 Orangutans (*pongo abelii*, 5 females, age 4.70 to 43.10). The sample size at the different time points ranged from 34 to 45 for the different species (see supplementary material for details). All apes participated in cognitive research on a regular basis. Apes were housed at the Wolfgang Köhler Primate Research Center located in Zoo Leipzig, Germany. They lived in groups, with one group per species and two chimpanzee groups (group A and B). Research was noninvasive and strictly adhered to the legal requirements in Germany. Animal husbandry and research complied with the European Association of Zoos and Aquaria Minimum Standards for the Accommodation and Care of Animals in Zoos and Aquaria as well as the World Association of Zoos and Aquariums Ethical Guidelines for the

Conduct of Research on Animals by Zoos and Aquariums. Participation was voluntary, all food was given in addition to the daily diet, and water was available ad libitum throughout the study. The study was approved by an internal ethics committee at the Max Planck Institute for Evolutionary Anthropology.

Procedure

Apes were tested in familiar sleeping or observation rooms by a single experimenter. The basic setup comprised a sliding table positioned in front of a mesh or a clear plexiglas panel. The experimenter sat on a small stool and used an occluder to cover the table (see Figure 1).

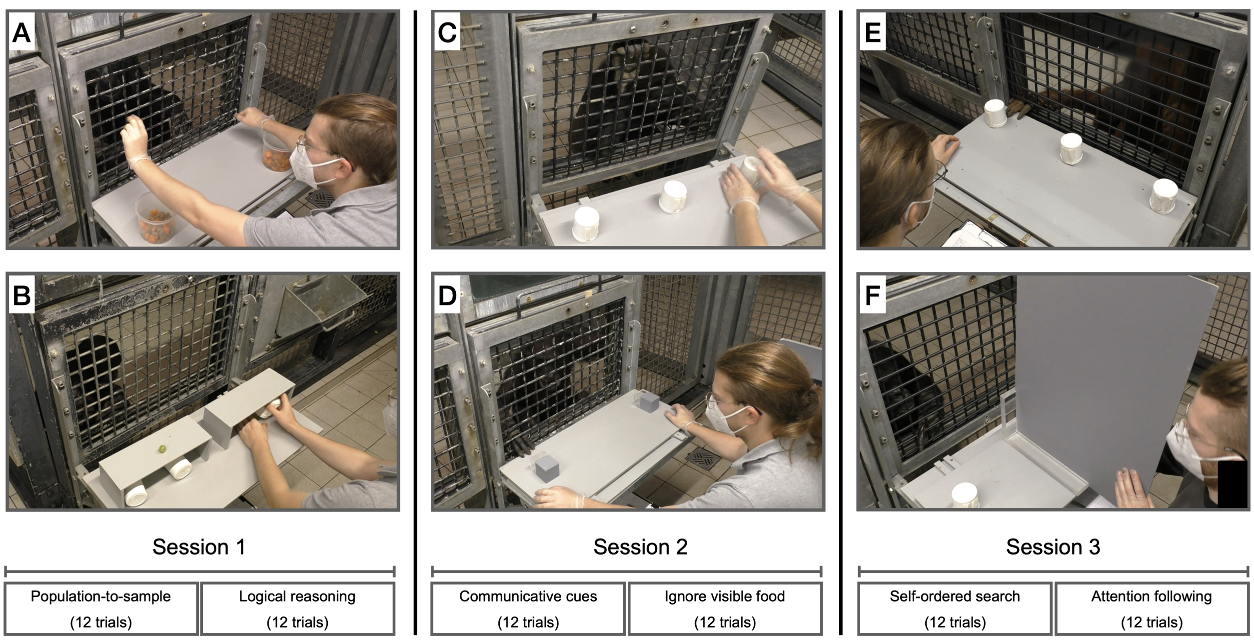

Figure 1. Setup used for the six tasks. A) population-to-sample, B) logical-reasoning, C) communicative-cues, D) ignore-visible-food, E) self-ordered-search and F) attention-following. Text at the bottom shows order of task presentation and trial numbers

The study involved a total of six cognitive tasks. These were based on published procedures in the field of comparative psychology. The original publications often include control conditions to rule out alternative, cognitively less demanding ways to solve the tasks.

We did not include such controls here and only ran the experimental conditions. For each task, we refer to these papers if they want to know more about control conditions and/or a detailed discussion of the nature of the underlying cognitive mechanisms. Example videos for each task can be found in the associated online repository. In the following, we give a brief description of each task. Additional details can be found in the supplementary material.

Attention-following. The Attention-following task was loosely modeled after Kaminski, Call, and Tomasello (2004). The setup consisted of two identical cups placed on the sliding table and a large opaque screen that was longer than the width of the sliding table (Supplementary Figure 1F). The experimenter placed both cups on the table and showed the ape that they were empty. Then, the experimenter baited both cups in view of the ape and placed the opaque screen in the center between the two cups, perpendicular to the mesh. Next, the experimenter moved to one side and looked at the cup in front of them. Then, the experimenter pushed the sliding table forward and the ape was allowed to choose one of the cups by pointing at it. If the ape chose the cup that the experimenter was looking at, they received the food item. If they choose the other cup, they did not. We coded whether the ape chose the side the experimenter was looking at (correct choice) or not. Apes received twelve trials. The side at which the experimenter looked was counterbalanced with same number of looks to each side and looks to the same side not more than two times in a row. We assumed that apes follow the experimenters focus of attention to determine whether or not their request could be seen and thus be successful.

Communicative-cues. This task was modeled after Schmid, Karg, Perner, and Tomasello (2017). Three identical cups were placed equidistantly on a sliding table directly in front of the ape (Figure 1C). In the beginning of a trial, the experimenter showed the ape that all cups are empty. After placing an occluder between the subject and the cups, the experimenter held up one food item and moved it behind the occluder, visiting all three cups but baiting only one. Next, the occluder was lifted and E looked at the ape (ostensive cue), called the name, and looked at one of the cups, while holding on to it with one hand and tapping it with the other (continuous looking, 3 times tapping). Finally, the experimenter pushed the sliding table forward for the ape to make a choice. If the ape chose the baited cup, they received the reward – if not, not. We coded as correct choice if the ape chose the indicated cup. Apes received twelve trials. The location of the indicated cup was counterbalanced, with each cup being the target equally often and the same target not more than two times in a row. We assumed that apes use the experimenter's communicative cues to determine where the food is hidden.

Ignore-visible-food. The task was modeled after Völter, Tinklenberg, Call, and Seed (2022). The task involved two opaque cups with an additional, sealed but transparent, compartment attached to the front of each cup (facing the ape). For one cup, the compartment contained a preferred food item that was clearly visible, for the other cup, the compartment was empty (Figure 1D). In the beginning of the trial, the two cups were placed upside down on the sliding table so that the ape could see that the opaque compartments of both cups were empty. Next, the experimenter baited one of the cups in full view of the subject. In non-conflict trials, the baited cup was the cup with the food item in the transparent compartment. In conflict trials, the baited cup was the cup with the empty compartment. After baiting the experimenter pushed the sliding table forwards and the ape could chose by pointing. If the baited cup was chosen, the ape received the food. Apes received 14 trials, twelve conflict trials and two non-conflict trials (1st and 8th trial). Only conflict trials were analyzed. The location of the cup with the baited compartment was counterbalanced, with the cup not being in the same location more than twice in a row. We assumed that apes need to inhibit selecting the visible food item and instead use their short-term memory to remember where the food was hidden.

Logical-reasoning. The task was modeled after Hanus and Call (2014). Three identical cups were presented side-by-side on a sliding table, with the cup in the middle sometimes positioned closer to the left cup and sometimes closer to the right. (Supplementary Figure 1B). Two half-open boxes served as occluders to block the ape's view when shuffling the cups. Each trial started by showing the ape that all three cups (one on one side of the table, two on the other) were empty. After placing the occluders over both sides of the table, the experimenter put one piece of food on top of each occluder. Next, the experimenter hid each piece of food under the cup(s) behind the occluders. In case of the occluder with the two cups, the food was randomly placed under one of the two cups while both cups were visited and even shuffled. Finally, both occluders were lifted and the table pushed forwards, allowing the ape to choose one of the three cups, from which they then received the content. We coded whether the ape chose the certain cup (i.e. the cup from the side of the table with only one cup). Apes received 12 trials. The side with one cup was counterbalanced, with the same constellation appearing not more than two times in a row on the same side. We assumed that apes would infer that the cup from the tray with only one cup certainly contains food while the other cups contain food only in 50% of cases.

Population-to-sample. The task was modeled after Eckert, Call, Hermes, Herrmann, and Rakoczy (2018). During the test, apes saw two transparent buckets filled with pellets and carrot pieces (the carrot pieces had roughly the same size and shape as the pellets). Each bucket contained 80 food items. The distribution of pellets to carrot pieces was 4:1 in bucket A, and 1:4 in bucket B. Pellets are preferred food items in comparison to carrots. The experimenter placed both buckets on a table, one left, one right (Figure 1A). In the beginning of a trial, the experimenter picked up the bucket on the right side, tilted it forward so the ape could see inside, placed it back on the table and turned it around 360°. The same procedure was repeated with the other bucket. Next, the experimenter looked at the ceiling, inserted each hand in the bucket in front of it and drew one item from the bucket without the ape seeing which type (E picked always of the majority type). The food items remained hidden in the experimenter's fists. Next, the experimenter extended the arms (in parallel) towards the ape who was then allowed to make a choice by pointing to one of the fists. The ape received the chosen sample. In half of the trials, the experimenter crossed arms when moving the fists towards the ape to ensure that the apes made a choice between samples and not just chose the side where the favorable population was still visible. In between trials, the buckets were refilled to restore the original distributions. Apes received twelve trials. We coded whether the ape chose the sample from the population with the higher number of high quality food items. The location of the buckets (left and right) was counterbalanced, with the buckets in the same location no more than two times in a row. The crossing of the hands was also counterbalanced with no more than two crossings in a row. We assumed that apes reasoned about the probability of the sample being a high quality item based on observing the ratio in the population.

Self-ordered-search. The task was modeled after Völter, Mundry, Call, and Seed (2019; Diamond, Prevor, Callender, and Druin, 1997; see also Petrides, 1995). Three identical cups were placed equidistantly on a sliding table directly in front of the ape (Supplementary Figure 1E). The experimenter baited all three cups in full view of the ape. Next, the experimenter pushed the sliding table forwards for the ape to choose one of the cups by pointing. After the choice, the table was pulled back and the ape received the food. After a 3s pause, the table was pushed forward again for a second choice. This procedure was repeated for a third choice. If the ape chose a baited cup, they received the food, if not, not. We coded the number of times the ape chose and empty cup (i.e. chose a cup they already chose before). Please note that this outcome variable differed from the other tasks in two ways: first, possible values were 0, 1, and 2 (instead of just 0 and 1) and second, a lower score indicated better performance. Apes received twelve trials. No counterbalancing was needed. We assumed that apes use their working memory abilities to remember where they had already searched and which cups still contained food.

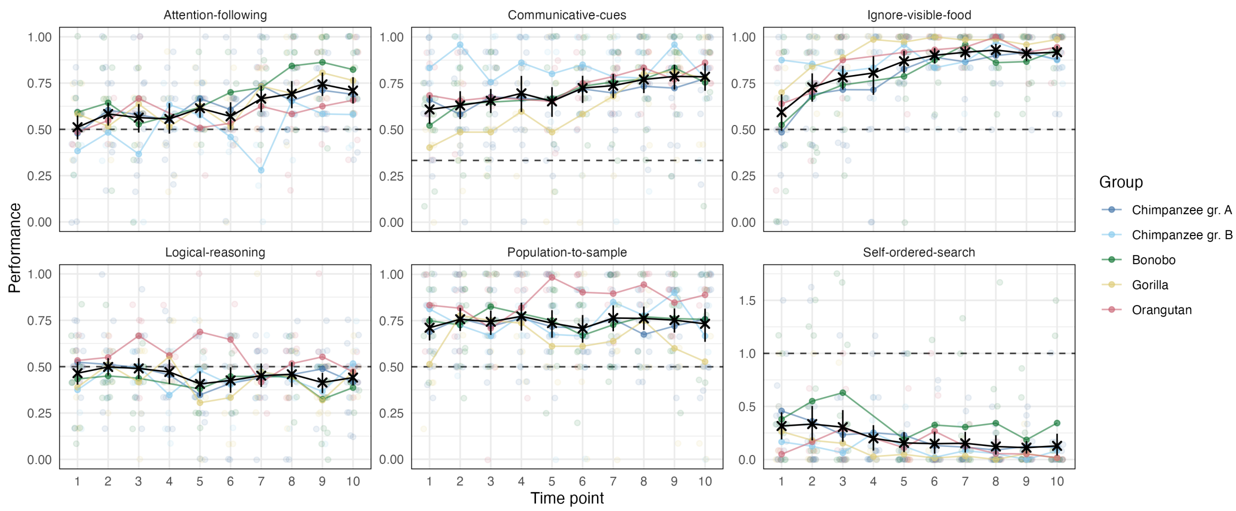

Figure 2. Results from the six cognitive tasks across time points. Black crosses show mean performance at each time point across species (with 95% CI). The sample size varied between time points and can be found in Supplementary Figure 1. Colored dots show mean performance by species. Dashed line shows chance level performance.

Data collection

Analysis, results and discussion

Stability and reliability

Structure

Figure4 shows the correlations between trait estimates for the different tasks reported in the present study as well as those reported in Bohn et al. (2023). For the tasks reported in Bohn et al. (2023) we used the data from phase 2 because it was closer in time. Overall, most correlations were not significantly different from zero (i.e. the 95% CI did include zero). Because of this low average level of correlations, we decided not to explore models with higher-order factors and will only interpret the qualitative patterns.

Conceptually, the tasks can be clustered in the following broader domains: *social cognition* (attention-following, gaze-following, communicative-cues), *reasoning about*

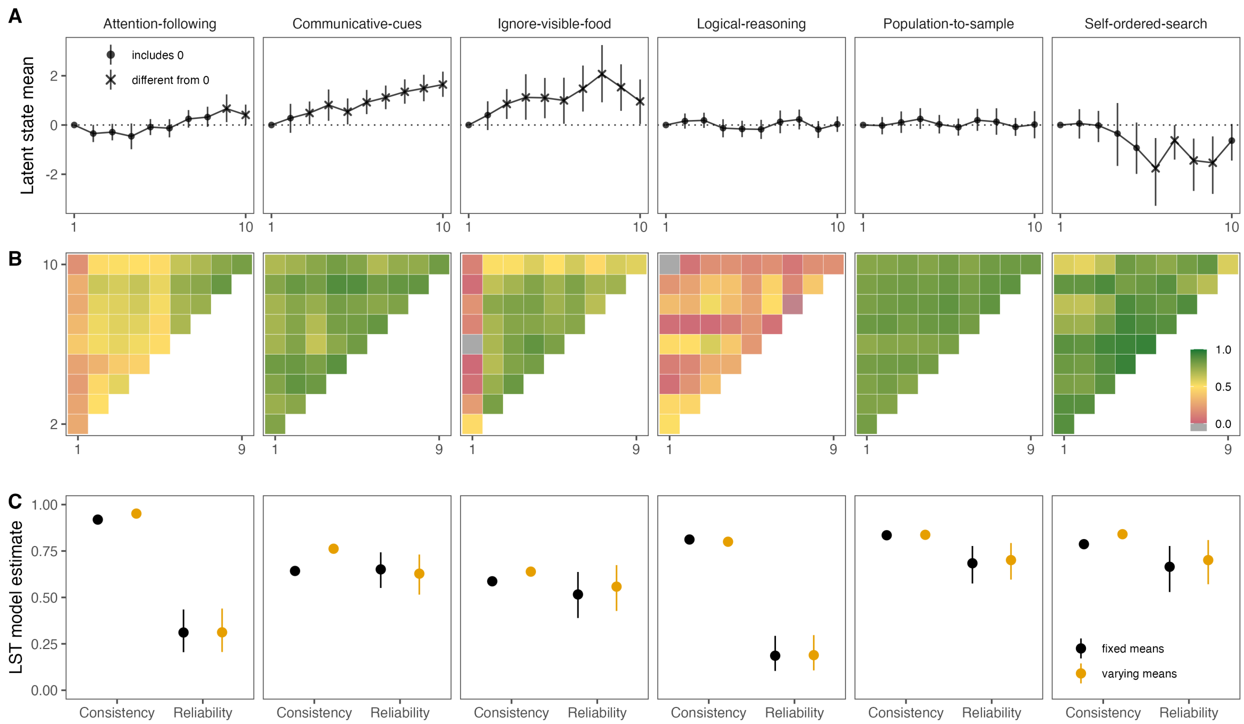

Figure 3. A) Latent mean estimates for each time point by task based on latent state model. Means at time point 1 are set to zero. Shape denotes whether the 95% CrI included zero (dashed line). The sample size varied between time points and can be found in Supplementary Fig. 1. B) Correlations between subject-level latent state estimats for the different time points by task. C) Mean estimates from latent state-trait models with fixed and varyin means (color codeed) with 95% CrI. Consistency refers to the proportion of (measurement-error-free) variance in performance explained by stable trait differences. Reliability refers to the proportion of true score variance to variance in raw scores.

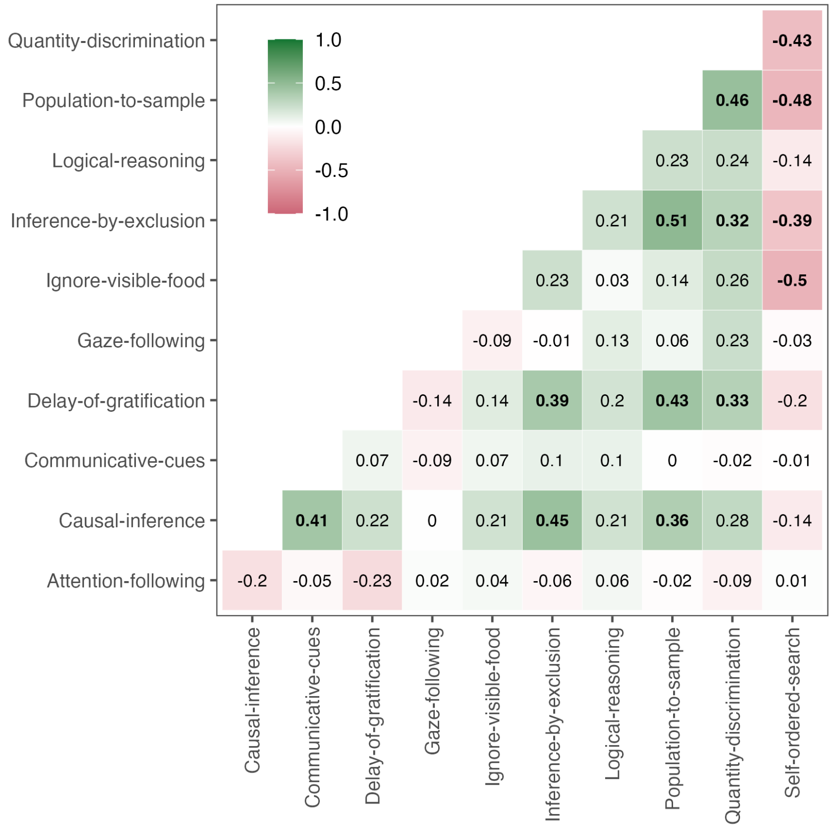

Figure 4. Correlations between ... trait estimates. Bold correlations have 95% CrI nont overlapping with zero.

quantities (quantity-discrimination, population-to-sample), *executive functions* (delay-of-gratification, self-ordered-search, ignore-visible-food) and *inferential reasoning* (logical-reasoning, causal-inference, inference-by-exclusion). As a first step, we will evaluate whether we find evidence for such a clustering in the data.

There was no significant correlation between any of the social cognition tasks. Furthermore, attention-following and gaze-following did not correlate significantly with any of the other tasks and communicative-cues correlated only with causal-inference – a result we will discuss below. Thus, and in line with previous work (Herrmann, Hernández-Lloreda, Call, Hare, & Tomasello, 2010), we found no evidence for shared cognitive processes in tasks measuring different aspects of social cognition.

The two tasks measuring reasoning about quantities did correlate significantly. Both tasks require discriminating between different quantities, directly in the case of quantity-discrimination and as part of the decision making process in the case of population-to-sample. Deciding between the samples from the two populations requires discriminating between the relative quantities within each bucket from which the samples were drawn.

Within the executive functions measures, self-ordered-search and inhibit-visible-food were significantly correlated but none of the two correlated with delay-of-gratification. The significant correlation can be explained by the need to inhibit a premature response (selecting visible food or a cup that was previously rewarded) in both tasks. It has been argued that delay-of-gratification requires self control (tolerating a longer waiting time to gain a more valuable reward) over and above behavioral inhibition (Beran, 2015). From this point of view, individual differences in the delay-of-gratification task might be due to differences in self control and less due to differences in inhibition.

Finally, for the three inferential reasoning measures we found a correlation between inference-by-exclusion and causal-inference. Logical-reasoning did not correlate with either (neither did it with any other task). This is not surprising given the results reported above: the observed variation in the logical-reasoning task was largely noise and did not reflect systematic individual differences. The correlation between causal-inference and inference-by-exclusion is most likely due to the fact that both tasks involve making inferences about the location of food based on reasoning about its physical properties.

Next we turn to the correlations across domains. Perhaps the most surprising finding is the correlation between causal-inference and communicative-cues. On a closer look, the origin might be task impurity in that there are two ways to solve the causal-inference task: first, as hypothesized, by using the rattling sound to infer the location of the food. Second, by interpreting the experimenter's shaking of the cup as a communicative cue, which is very similar to the communicative-cues task. Thus, we suspect that at a least some individuals solved the task via the second route.

Finally, there was a cluster of significant correlations between delay-of-gratification, self-ordered-search, inference-by-exclusion, causal-inference, population-to-sample and quantity discrimination. Of the 15 possible correlations, only four were non-significant. One commonality between these tasks that might – in part – explain this pattern is that they all benefit from sustained attention to the task. Sustained attention facilitates the processing of the experimenter's demonstrations (population-to-sample, inference-by-exclusion, causal-inference, delay-of-gratification), ones one actions on the setup (self-ordered-search) or visually complex stimuli (quantity discrimination). Tentative support for this idea comes from the analysis of relevant predictors (see Bohn et al., 2023 and below) in which `time spent in research` was selected as a relevant predictor of performance for all of these tasks except causal-inference. This predictor reflects individual's experience with experimental studies, which often involve sustained attention to distributions of food items, actions of conspecifics and/or demonstrations by experimenters.

Predictability

General Discussion

Conclusion

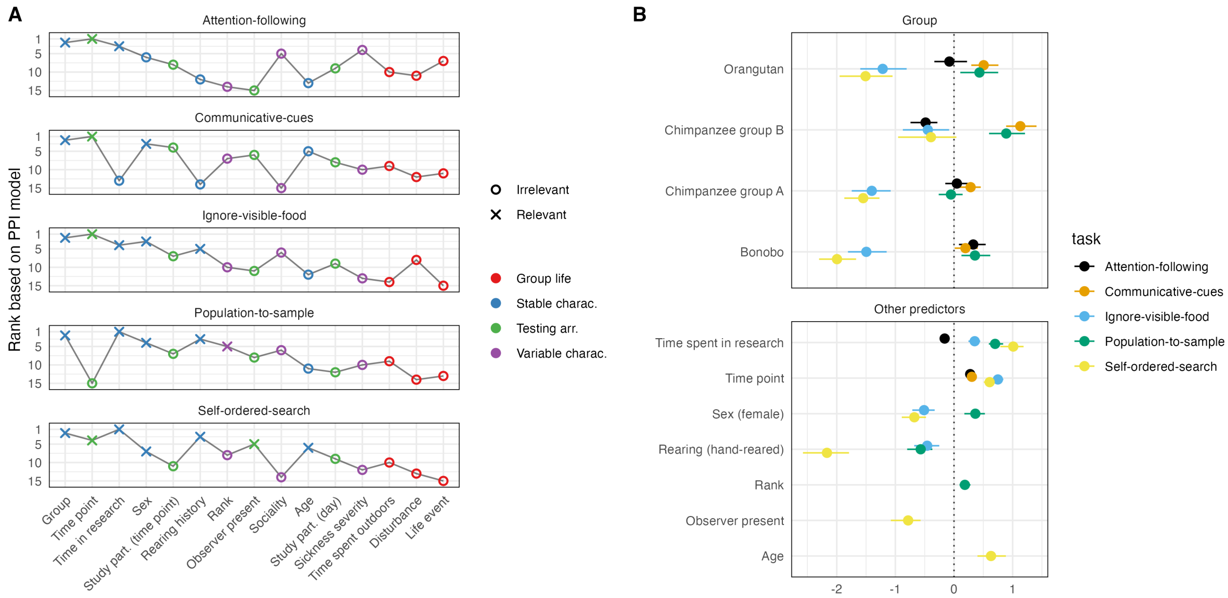

Figure 5. A) Predictor ranking and selection based on PPI models. Crosses mark predictors that were selected to be relevant based on the PPI models. Color shows the broader category each predictor belongs to. The x-axis is sorted by the average rank across tasks. B) Posterior model estimates for the selected predictors for each task based on data. Points show means with 95% Credible Interval. Color denotes task. For categorical predictors, the estimate gives the difference compared to the reference level (Gorilla for group).

References

Beran, M. J. (2015). The comparative science of "self-control": What are we talking about? *Frontiers in Psychology*, *6*, 51.

Bohn, M., Eckert, J., Hanus, D., Lugauer, B., Holtmann, J., & Haun, D. B. (2023). Great ape cognition is structured by stable cognitive abilities and predicted by developmental conditions. *Nature Ecology & Evolution*, *7*(6), 927–938.

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development*, i–206.

Eckert, J., Call, J., Hermes, J., Herrmann, E., & Rakoczy, H. (2018). Intuitive statistical inferences in chimpanzees and humans follow weber's law. *Cognition*, *180*, 99–107.

Hanus, D., & Call, J. (2014). When maths trumps logic: Probabilistic judgements in chimpanzees. *Biology Letters*, *10*(12), 20140892.

Herrmann, E., Hernández-Lloreda, M. V., Call, J., Hare, B., & Tomasello, M. (2010). The structure of individual differences in the cognitive abilities of children and chimpanzees. *Psychological Science*, *21*(1), 102–110.

Kaminski, J., Call, J., & Tomasello, M. (2004). Body orientation and face orientation: Two factors controlling apes' begging behavior from humans. *Animal Cognition*, *7*, 216–223.

Petrides, M. (1995). Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience*, *15*(1), 359–375.

Schmid, B., Karg, K., Perner, J., & Tomasello, M. (2017). Great apes are sensitive to prior reliability of an informant in a gaze following task. *PLoS One*, *12*(11), e0187451.

Völter, C. J., Mundry, R., Call, J., & Seed, A. M. (2019). Chimpanzees flexibly update working memory contents and show susceptibility to distraction in the self-ordered search task. *Proceedings of the Royal Society B*, *286*(1907), 20190715.

Völter, C. J., Tinklenberg, B., Call, J., & Seed, A. M. (2022). Inhibitory control and cue relevance modulate chimpanzees'(pan troglodytes) performance in a spatial foraging task. *Journal of Comparative Psychology*, *136*(2), 105.