Towards a Method for Evaluating Convergence Across Modeling Frameworks

Lorraine Borghetti (lorraine.borghetti.1.ctr@us.af.mil)
Air Force Research Laboratory
Wright Patterson AFB, OH USA

Christopher R. Fisher (christopher.fisher.27.ctr@us.af.mil)
Cubic Defense
Beavercreek, OH 45324 USA

Joseph W. Houpt (joseph.houpt@utsa.edu)
University of Texas at San Antonio
San Antonio, TX 78249

Leslie M. Blaha (leslie.blaha@us.af.mil)
Air Force Research Laboratory
Wright Patterson AFB, OH USA

Glenn Gunzelmann (glenn.gunzelmann@us.af.mil)
Air Force Research Laboratory
Wright Patterson AFB, OH USA

Christopher Stevens (christopher.stevens.28@us.af.mil)
Air Force Research Laboratory
Wright Patterson AFB, OH USA

Abstract

Model convergence is an alternative approach for evaluating computational models of cognition. Convergence occurs when multiple models provide similar explanations for a phenomenon. In contrast to competitive comparisons which focus on model differences, identifying areas of convergence can provide evidence for overarching theoretical ideas. We proposed criteria for convergence which require models to be high in predictive and cognitive similarity. We then used a cross fitting method to explore the extent to which models from distinct computational frameworks—quantum cognition and the cognitive architecture ACT-R—converge on explanations of the interference effect. Our analysis revealed the models to be moderately high in predictive similarity but mixed for cognitive similarity. Though convergence was limited, the analysis suggests that interference effects emerge from interactions between uncertainty and the degree to which an individual relies on typical cases to make decisions. This result demonstrates the utility of convergence analysis as a method for integrating insights from multiple models.

Keywords: ACT-R; Quantum cognition; Interference effects; Model convergence

Introduction

Model comparison often proceeds as a zero-sum game in which two or more models offering different explanations make opposing predictions. The winner of such competitions is assumed to offer a more convincing representation of the underlying cognitive processes. Although competitive comparisons can be useful to varying degrees, one potential limitation is that one may overlook areas of convergence by focusing exclusively on differences between models. Two models may point to similar conclusions for a particular empirical phenomenon even though they may differ in other regards. One important benefit of convergence is that confidence in an explanation will increase when two models are in agreement. As an example of convergence, two distinct computational frameworks, one based on the Adaptive Control of Thought-Rational (ACT-R) and the other based on the drift diffusion model—provided similar explanations for the deleterious effect of sleep loss on performance. Namely, they both explain a reduction in the signal-to-noise ratio and a reduction in response inhibition (Walsh et al., 2017).

Convergence offers an alternative approach for evaluating what models reveal about human cognition (Gunzelmann, 2019). The present study extends the existing work by elaborating upon the definition of convergence and its implications for theoretical correspondence. We then conduct an exploratory evaluation of the extent to which two distinct models of interference effects—an existing quantum cognition model and a model developed in ACT-R—converge on conclusions consistent with a single theoretical perspective.

Model Convergence

As shown in Figure 1, models can be compared along two orthogonal dimensions: predictive similarity and cognitive similarity. Predictive similarity is the degree to which the predictions of two models follow the same pattern. At minimum, we require the predictions to follow the same qualitative pattern, i.e., both models predict an effect in the same direction. Cognitive similarity is defined as the degree to which two models posit similar mental representations (i.e., the content and organization of information about the external environment) and/or cognitive processes (i.e., how information is transformed, manipulated, and combined) that are relevant for a particular empirical phenomenon. Although this space is continuous, it can be helpful to refer to prototypical examples or describe the space more coarsely as quadrants. Convergence occurs when two or more models are highly similar along both dimensions.

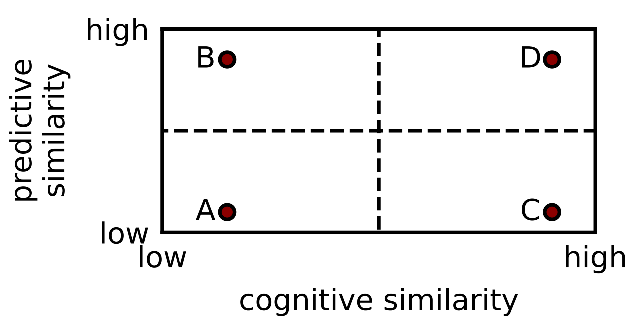


Figure 1: Four points in the space spanned by predictive similarity and cognitive similarity. Point A represents competitive comparisons and Point D represents convergence.

The left half of Figure 1 reflects so-called "zones of contention" where competing models propose different mental representations and/or cognitive processes to explain an empirical phenomenon (McClelland, 2009). At Point A in the bottom-left quadrant, critical tests can distinguish between competing models on the basis of their opposing predictions. By contrast, for Point B in which predictions are similar, different mental representations and cognitive processes cannot be distinguished on the basis of their predictions. Indeed, such ambiguity often leads to the development of the critical tests conducted in bottom-left quadrant, a cycle that can repeat itself many times (Gunzelmann, 2019). The right half represents cases in which models are cognitively similar and thus propose similar mental representations and/or cognitive processes. Point C in the bottom-right quadrant represents an unusual situation in which two models high in cognitive similarity yield differing predictions. In this case, the models provide contradictory evidence for a common explanation. Point D in the top-right quadrant represents the case where models converge on a common explanation: both models rely on similar mental representations and/or cognitive processes and make similar predictions. When convergence occurs, we find more evidence for an explanation than we would otherwise. We believe viewing model comparisons through the lens of convergence adds clarity to theoretical implications and may provide additional evidence for an overarching theory. By contrast, the competitive approach seeks to refute one of the models. Although both approaches have different goals, taken together, they offer complementary methods for evaluating theoretical support (Gunzelmann, 2019).

Current Application

We explore whether quantum cognition and ACT-R provide converging explanations of the interference effect. Interference effects emerge when uncertainty about an event changes the marginal probability of a subsequent decision, resulting in a violation of the law of total probability (Wang & Busemeyer, 2016). The models we investigated derive from highly disparate computational frameworks with strong empirical support. The belief-action entanglement model is based on the mathematical formalism of quantum probability which has been used to explain several empirical phenomena where models based on classical probability generally fail (Wang & Busemeyer, 2016). By contrast, ACT-R is a cognitive architecture in which cognition emerges from interactions between specialized information processing modules for declarative and procedural memory, perception, and action among others (Anderson et al., 2004). Given that both frameworks have withstood many rounds of empirical testing, one might expect points of convergence to emerge.

Categorization-Decision Paradigm

One popular paradigm for studying interference effects emerging from the interactions of categorization and decision making is the categorization-decision paradigm (Wang & Busemeyer, 2016). On each trial, a face from a "good" category or a "bad" category is presented, and participants must decide whether to attack or withdraw. Each face consisted of either a g-type or b-type feature, which were typically associated with the good category or bad category, respectively. Further, participants were typically rewarded for attacking a bad category and withdrawing from a good category. However, these associations were probabilistic, and atypical associations occurred in some trials.

Uncertainty about the category was manipulated across three conditions to elicit an interference effect. In the decision-only (d) condition, no category information was provided prior to the decision to act, and categorization was presumed to occur implicitly (Wang & Busemeyer, 2016). In the categorize-then-decide (cd) condition, participants were asked to self-categorize the feature then decide upon an action. In the third explicit-categorization (xd) condition, the true category was provided prior to the action decision.

Belief-Action Entanglement Model

The belief-action entanglement (BAE) model is a quantum cognition model of interference effects in categorization and decision making. A full mathematical description of the model can be found in Wang & Busemeyer (2016). In the BAE model, states evolve within a finite Hilbert space H (N-dimensional universal vector space) across a field of complex numbers. The potential of a state is given by the unit-length vector ψ. A defining feature of quantum systems, to include cognitive systems, is that a measurement changes the state. Consequently, transitions occur when ψ is measured, e.g., a feature is categorized or an action is selected.

The BAE represents category-action events as basis states where GW symbolizes the combined event of categorizing a feature as good then deciding to withdraw. The initial state ψ_f is uncertain and superposed over the four possible basis states, $\psi_f = [\mathrm{GW}, \mathrm{GA}, \mathrm{BW}, \mathrm{BA}]^\top$. Basis states are assigned amplitudes such that the square magnitude gives its probability: $|\psi_{\mathrm{GW}}|^2 = \Pr(\mathrm{GW})$. The parameter j governs the probability a b-type or g-type feature will be judged as belonging to either category, e.g., for a b-type feature, $\psi_f = \psi_b = \frac{1}{2}\left[\sqrt{1-j}, \sqrt{1-j}, \sqrt{j}, \sqrt{j}\right]^\top$.

Prior to action evaluation, the state remains in the superposed ψ_f in the d condition. In cd and xd, the state transitions to either being in the good or bad category. After transitioning to the bad category, as an example, the state is updated to $\psi_f \rightarrow \psi_b = \frac{1}{2}[0,0,1,1]^\top$, where the latter values represent BW and BA and the state is only superposed over the actions.

During action evaluation, the state transitions according to the reward rate and utility parameters which influence the probability of a action given a feature and category. For example, $\mu_{b,b}$ is the utility for attacking a b-type feature categorized as bad. The transition to the final action state is computed using a separate unitary matrix for each feature type. When the category is ambiguous as in cd and d, the transition includes the entanglement parameter γ which amplifies amplitudes for typical category-action events, e.g., GW and BA, and attenuates amplitudes for atypical events, e.g. GA and BW. Alternately, γ has no effect in xd because the true category is known. Consequently, the BAE model predicts that interference effects emerge from differences in the utilities for each feature type and the influence of γ on uncertain states.

ACT-R Model

We developed a memory-based ACT-R model of the interference effect. and focus our description on the declarative memory system.

Declarative Memory

In ACT-R, the basic unit of declarative knowledge is a set of slot-value pairs called a chunk: $\mathbf{c}_m = \{(s_i, v_i)\}_{i \in I_m}$, where s_i and v_i are the slot and value of pair i, and I_m is the index set for slot-value pairs of chunk m. We will use the set $Q_m = \{s_i\}_{i \in I_m}$ to denote a set of slots (e.g., domain) in $\mathbf{c}_m$. The mapping from slots to values is defined as $c_m(s) = v$, where v is null if the chunk does not include s.

The set of slots for each chunk is defined as $Q = \{\text{feature}, \text{category}, \text{action}\}$, where the feature can be b-type or g-type, the category can be good or bad and the action can be attack or withdraw. Declarative memory M consists of $2^3 = 8$ chunks formed by permuting the possible values for feature, category and action. For example, $\mathbf{c}_{\text{gba}} = \{(\text{feature}, \text{g-type}), (\text{category}, \text{bad}), (\text{action}, \text{attack})\}$ is a chunk for attacking a g-type face in the bad category. We will use a three letter abbreviation, such as gba, to denote the feature, category, and action values of a chunk.

Memory Activation

Each chunk is associated with an activation value representing its ability to be retrieved. As activation increases, the probability of retrieval increases. We omit the base-level learning mechanism because learning was not observed in Wang & Busemeyer (2016). Activation is defined as $a_m = \beta_m + \rho_m + \varepsilon_m$ where β is the base-level constant, ρ is the partial matching term, and $\varepsilon \sim \text{logistic}(0, s)$ is logistically distributed noise with scalar parameter s. The partial matching mechanism allows chunks that do not match the retrieval request $\mathbf{r}$ to be retrieved as a decreasing function of mismatch. The retrieval request is treated as a chunk with slot-value pairs. We use a binary mismatch penalty function: $\rho_m = -\delta \sum_{q \in Q_r} I(c_m(q), r(q))$, where δ is the mismatch penalty parameter, Q_r is the set of slots in the request, and I is an indicator function which returns 1 when both inputs are not equal and returns 0 otherwise.

Retrieval Process

Upon stimulus presentation, a retrieval request $\mathbf{r}$ based on all available information is submitted to declarative memory. For example, in the d condition, only the feature is available, but in the xd condition both the feature and the category are available. The chunk with the highest activation value above the retrieval threshold τ is retrieved and determines the eventual response. To simplify the model, we set the retrieval threshold to -10 under the assumption that chunks are sufficiently active to be retrieved.

Model Predictions

In the predictions for each condition below, we use A to denote a random variable for the action, F to denote a random variable to denote the feature, and C as a random variable to denote the category.

d condition Participants decided to attack or withdraw from a face with feature f. The retrieval request is $\mathbf{r} = \{(\text{feature}, f)\}$. We will define R_d as the set of chunks that map to a decision to attack $R_d = \{c_m \in M : c_m(\text{feature}) = r(\text{feature}), c_m(\text{action}) = \text{attack}\}$. In other words, R_d is the set of chunks that match feature f and have a value "attack" for the action slot. The approximate probability of attacking is computed using the soft max function (Weaver, 2008):

$$\Pr(A = a \mid F = f) = \frac{\sum_{k|\mathbf{c}_k \in R_d} e^{\mu_k/\sigma}}{\sum_{j|\mathbf{c}_j \in M} e^{\mu_j/\sigma}} \tag{1}$$

where $\sigma = s\sqrt{2}$ and the expected activation is $\mathbb{E}[a_m] = \mu_m$.

xd condition Participants were told the true category v for a face with feature f then decided to attack or withdraw, leading to the retrieval request $\mathbf{r} = \{(\text{feature}, f), (\text{category}, v)\}$. The set of chunks that map to the decision to attack is defined as: $R_{\text{xd}} = \{c_m \in M : c_m(\text{feature}) = r(\text{feature}), c_m(\text{category}) = r(\text{category}), c_m(\text{action}) = \text{attack}\}$. The probability of attacking a face with feature f in category v is given by:

$$\Pr(A = a \mid F = f, C = v) = \frac{\sum_{k|\mathbf{c}_k \in R_{\text{xd}}} e^{\mu_k/\sigma}}{\sum_{j|\mathbf{c}_j \in M} e^{\mu_j/\sigma}} \tag{2}$$

cd condition Participants categorized a face with feature f as good or bad followed by a separate response to attack or withdraw. The retrieval request for the categorization is $\mathbf{r}_c = \{(\text{feature}, f)\}$. The set of chunks that map to a category response v is defined as $R_{\text{cd,c}} = \{c_m \in M : c_m(\text{feature}) = r_c(\text{feature}), c_m(\text{category}) = v\}$. The probability of categorizing face with feature f as v is given by:

$$\Pr(C = v \mid F = f) = \frac{\sum_{k|\mathbf{c}_k \in R_{\text{cd,c}}} e^{\mu_k/\sigma}}{\sum_{j|\mathbf{c}_j \in M} e^{\mu_j/\sigma}} \tag{3}$$

The judged category v is incorporated into the retrieval request for the subsequent decision: $\mathbf{r}_d = \{(\text{feature}, f), (\text{category}, v)\}$. The set of chunks that map to the decision to attack is the same as in the cd condition: $R_{\text{xd}} = R_{\text{cd,d}}$, which implies that the probability of attacking a face with feature f categorized as v is equal to equation 2 from the xd condition.

Cross Fitting

To measure predictive and cognitive similarity, we used a cross fitting method inspired by Donkin et al. (2011). In our study, predictive similarity is measured by comparing the qualitative predictions of the two models, whereas cognitive similarity is measured by assessing the mapping of parameters from one model to another. Our cross fitting method entails two steps. First, we generated predictions from the BAE model by varying one parameter at a time while holding the others constant at their best fitting values reported in Wang & Busemeyer (2016). Second, we fit the ACT-R model to the predictions of the BAE by minimizing Kullback-Leibler divergence (KLD; Kullback & Leibler, 1951). KLD is the amount of information lost by using one distribution in place of another, i.e., how much information is lost when using the best fit ACT-R model to represent the BAE mode. One advantage of comparing two probability distributions using KLD instead of fitting a model to a finite sample of simulated data is that it eliminates the role of noise in the mapping.

We selected three parameters on the basis of their qualitatively distinct roles in the model: the entanglement parameter, γ, the category judgement parameter, j, and a utility parameter, $\mu_{b,b}$. Each parameter was varied across 20 equally spaced values: $\gamma \in [0,2]$, $j \in [.01,.99]$, and $\mu_{b,b} \in [-1,1]$. We set $s = .2$ and base level constants $\beta_{\text{bbw}} = 0.0$ and $\beta_{\text{ggw}} = 0.2$ to ensure identifiability of the model parameters. We used differential evolution to minimize KLD.

Convergence Predictions

Psychologically, interference effects can result from increased on reliance typical associations in the absence of certain information (Fiske & Taylor, 1991). For example, β_{bba} and $\mu_{b,b}$ represent the influence of typical associations between a face with a b-type feature in the bad category and the decision to attack. The strength of influence varies with certainty about the category. If convergence is present, the BAE and ACT-R accounts of these processes should be relatable.

First, we expect typical associations to strengthen the probability to attack, $\Pr(A)$, for b-type features in both models, irrespective of category certainty. In the BAE, this should be most evident as $\mu_{b,b}$ increases. In ACT-R, we expect to observe a comparable increase β_{bba} with a commensurate decrease in βs for atypical associations according to equations 1 and 2. Second, we expected category uncertainty in the d condition to moderate the $\Pr(A)$. In the BAE, changes in j should vary the influence of typical associations. Because the retrieval request only contains the feature, a comparable process in ACT-R should systematically influence βs for typical categories and actions according to equation 1.

Third, we expect γ and δ parameters to modulate the influence of utility and β parameters, respectively. In particular, we expect γ to amplify the effect of typical associations for the $\Pr(A)$, but only with category uncertainty in cd and d. In ACT-R, the analogous effect should occur at higher values of δ which increase the probability of selecting an exact match.Consequently, we expect the influence of β_{bba} to be amplified while attenuating the influence of atypical βs.

Results

We assess predictive and cognitive similarity between the BAE and ACT-R for each of the three BAE parameters to determine

Figure 2: Best fitting values for the base level constant (β) parameters and the mismatch penalty parameter (δ) for ACT-R as a function of γ from the BAE model.

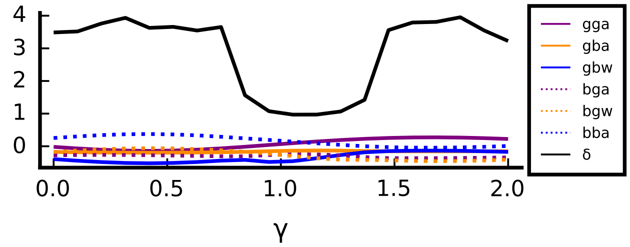
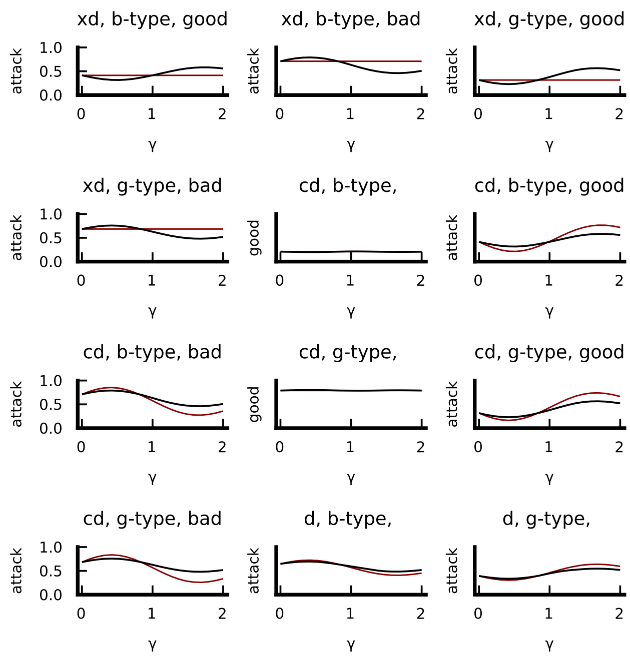


Figure 3: Response probabilities for the BAE (red) and ACT-R (black) models as a function of γ paneled by response category. Subplot titles give condition, face type, category.



whether the models converge on similar explanations of the interference effect.

The entanglement parameter γ. Predictive similarity: for both models, the response probabilities follow qualitatively similar patterns in cd and d, a distinction more pronounced in cd than d (see Figure 3). However, $\Pr(A)$ patterns are qualitatively dissimilar in xd. Specifically, the BAE model is invariant to γ, as intended, whereas as ACT-R simply reproduces pattern of probabilities in cd. This is not surprising as equation 2 computes the $\Pr(A)$ in both xd and cd. The results indicate predicative similarity is moderately high for ambiguous category knowledge but low for unambiguous categorization.

Cognitive similarity: for simplicity, we focus on mappings where γ is less than 1 (see Figure 2). Though the pattern is not strictly linear, decreases in γ and increases in δ favor typical associations, as predicted. In ACT-R specifically, the

Figure 4: Best fitting values for the base level constant (β) parameters and the mismatch penalty parameter (δ) for ACT-R as a function of j from the BAE model.

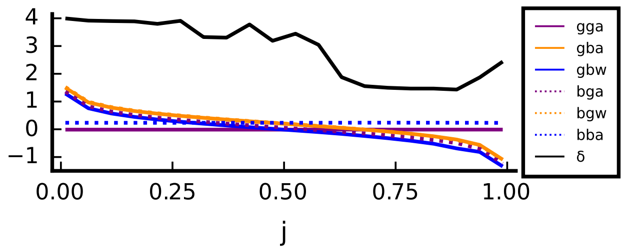
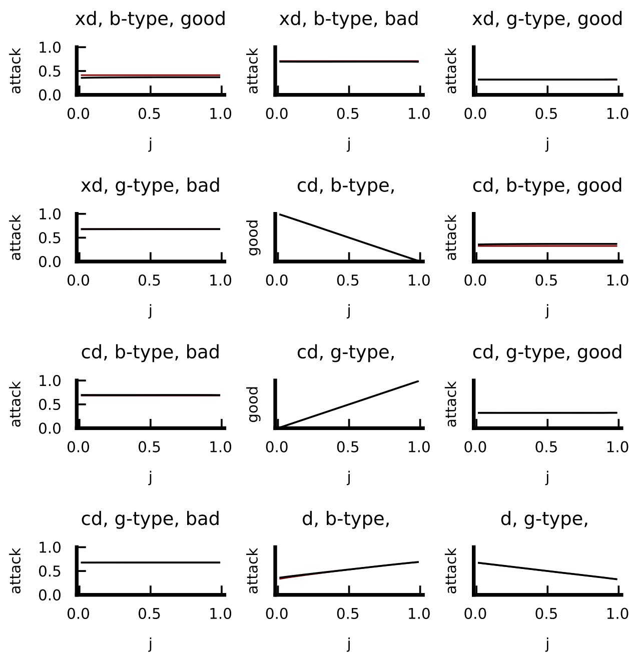


Figure 5: Response probabilities for the BAE (red) and ACT-R (black) models as a function of j paneled by response category. Subplot titles give condition, face type, category.



process entails expected interactions with β_{bba} as well as the atypical β_{gbw}. As a result, we conclude the models exhibit high cognitive similarity for modulating bias.

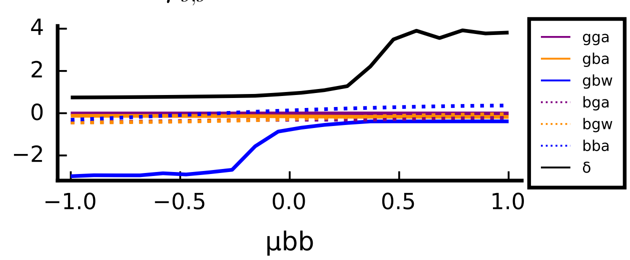
The category judgement parameter j. Predictive similarity: The BAE and ACT-R produced identical distributions for $\Pr(A)$ in d (see Figure 5). By contrast, $\Pr(A)$ in xd and cd remained invariant in both models, indicating they were constrained to the d condition, as expected. All told, these patterns indicate high predictive similarity.

Cognitive similarity: The relatively linear decreases in δ and atypical β values with increases in the j parameter reveal an unexpected mapping between the two models (see Figure 4). In the BAE model, the $\Pr(A)$ derives from an uncertain superposition state over possible outcomes and is systematically modulated by j. In contrast, the ACT-R model is less systematic and it is not clear the mental states represented by parameter interactions. Specifically, the Pr(A) increases at high values of δ, which approximates increasing bias in the decision, and also at low values of δ, ostensibly representing indecision between alternatives. Because ACT-R's varied account cannot easily be reconciled with the BAE account, we conclude the models are low in cognitive similarity when category is uncertain and not made explicit.

The utility parameter $\mu_{b,b}$. Predictive similarity: Visual inspection of Figure 7 indicates that predictive similarity is high when $\mu_{b,b}$ is varied. The predictions exhibit some discrepancy for b-type faces in the bad category in the xd and cd conditions. Nonetheless, the predictions are qualitatively similar throughout.

Cognitive similarity: The varied behavior of ACT-R parameters across the range of $\mu_{b,b}$ was surprising (see Figure 6). In the BAE, $\mu_{b,b}$ exerts a relatively linear effect on the $\Pr(A)$, as expected. In ACT-R, the $\Pr(A)$ varies with parameter interactions when $\mu_{b,b}$ is above versus below 0. Specifically, when $\mu_{b,b} > 0$, β_{bba} amplifies the $\Pr(A)$ when an exact match is more probable (e.g., at higher δ values), in line with our expectations. Alternatively when $\mu_{b,b} < 0$, the atypical β_{gbw} increasingly attenuates the $\Pr(A)$ when a mismatch becomes more likely (e.g., at lower δ values) which was neither expected nor a predictable function of δ. Because only a portion of ACT-R interactions are analogous $\mu_{b,b}$'s function, cognitive similarity between the models was deemed moderate, at best, for the influence of typical associations.

Figure 6: Best fitting values for the base level constant (β) parameters and the mismatch penalty parameter (δ) for ACT-R as a function of $\mu_{b,b}$ from the BAE model.

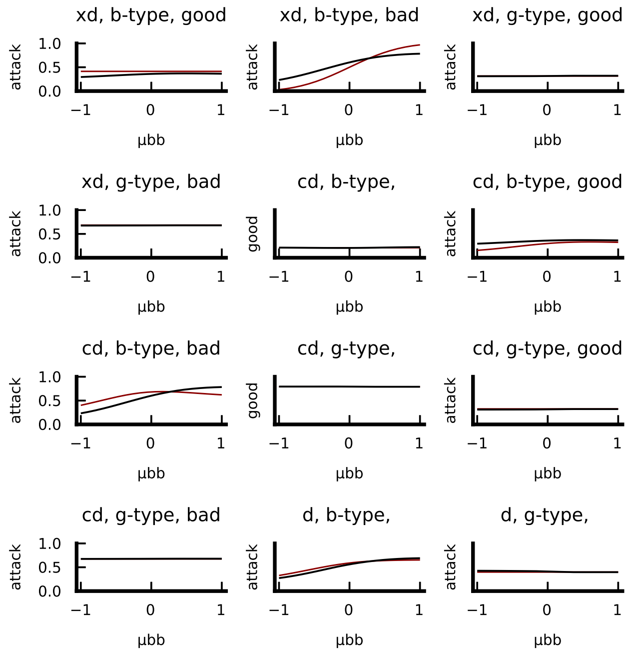


Discussion

The goal of the present study was two-fold. First, we elaborated upon the definition of model convergence. Second, we explored the extent the BAE, a model based in quantum cognition, and a model based in ACT-R provide converging explanations of the interference effect. Our criteria for convergence required models be both high in predictive and cognitive similarity. For interference effects, we expected similarities to emerge from interactions between category certainty and the influence of typical associations on decisions.

Both models exhibited moderately high predictive similarity. Predictions were more similar when the category was uncertain (cd and d) but diverged when the category was certain (xd). In ACT-R, the divergence can be attributed to the partial

Figure 7: Response probabilities for the BAE (red) and ACT-R (black) models as a function of $\mu_{b,b}$ paneled by response category. Subplot titles give condition, face type, category.



matching mechanism which is constrained to implementing penalties for mismatches in slot-value pairs. The architecture does not permit penalizing mismatches at the condition level. Hence, when the category was included in the retrieval request, the model was unable to differentiate between an uncertain category in cd and a true one in xd. The BAE can account for unambiguous category knowledge because the entanglement parameter γ was not applied to state transitions in xd (Wang & Busemeyer, 2016).

Cognitive similarity between the two models was mixed. The BAE's γ parameter and ACT-R's mismatch penalty δ modulated the influence of typical and atypical associations in comparable ways. Overall, we found the expected relationship between $\mu_{b,b}$ and β_{bba}. However, for $\mu_{b,b}$ and j, ACT-R parameter mappings were by determined by the ratio of βs (see equations 1, 2, 3) and varying values of δ which at times appeared unsystematic and difficult to predict. The variability is surprising given that both the BAE and ACT-R models produce interference effects and can account for violations of total probability. One explanation for the unexpected mappings may be due to the idiosyncrasy of a particular implementation rather than the function of partial matching. If so, then cognitive similarity may be higher than assessed.

Indeed, while useful, our cross fitting analysis may have obscured areas of cognitive similarity. First, our mappings were asymmetrical such that ACT-R parameters were cross fitted as a function of the BAE parameters but not the other way around. However, ACT-R's fluctuating parameter interactions pose a challenge for symmetrical mappings, and it is unclear whether mapping to a single parameter would be sufficient to evaluate convergence. Second, our mappings centered on best fitting values, ergo limiting the scope of our analysis; the full space of potential convergence was not explored. Evidence for similarity would be greater if the relationships hold across a larger sub-space of parameters. Even so, our approach of evaluating parameters near the best fitting is a reasonable starting point.

With respect to supporting a single theoretical perspective, our analysis was informative, even as convergence was limited. Had we conducted a competitive comparison, the theoretical contribution of the losing model might have been eclipsed. As it stands, not only have we accumulated evidence for the psychological processes underlying interference effects, but our analysis identified areas where future research can further elucidate how and when the human mind is influenced by the strength of beliefs in uncertain situations.

Acknowledgments

The opinions expressed herein are solely those of the authors and do not necessarily represent the opinions of the United States Government, the U.S. Department of Defense, the Department of the Air Force, or any of their subsidiaries or employees. This research was supported by Air Force Office of Scientific Research grant 21RHCOR080. Approved for public release. Case number: AFRL-2022-2018.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion vs linear ballistic accumulation: Different models but same conclusions about psychological processes? *Psychonomic Bulletin & Review*, *18*(1), 61–69.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. Mcgraw-Hill Book Company.

Gunzelmann, G. (2019). Promoting cumulation in models of the human mind. *Computational Brain & Behavior*, *2*(3), 157–159.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*(1), 11–38.

Walsh, M. M., Gunzelmann, G., & Van Dongen, H. (2017). Computational cognitive modeling of the temporal dynamics of fatigue from sleep loss. *Psychonomic Bulletin & Review*, *24*(6), 1785–1807.

Wang, Z., & Busemeyer, J. R. (2016). Interference effects of categorization on decision making. *Cognition*, *150*, 133–149.

Weaver, R. (2008). Parameters, predictions, and evidence in computational modeling: A statistical view informed by ACT–R. *Cognitive Science*, *32*(8), 1349–1375.