How to Model the Neurocognitive Dynamics of

Decision Making: A Methodological Primer With ACT-R

Cvetomir Dimov[1,3]

Patrick H. Khader*[2]

Julian N. Marewski*[3]

Thorsten Pachur*[4]

[1] Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue, Pittsburgh,

PA 15213, USA

[2] Brandenburg Medical School, Department of Psychology, Am Alten Gymnasium 1-3, D-

16816, Neuruppin, Germany

[3] University of Lausanne, Department of Organizational Behavior, Internef, CH-1015,

Lausanne, Switzerland

[4] Max Planck Institute for Human Development, Lentzeallee 94, D-14195, Berlin, Germany

*Alphabetical author order. The authors contributed equally to this work

Address correspondence to:

Cvetomir Dimov

345C Baker Hall

Carnegie Mellon University

5000 Forbes Avenue

Pittsburgh, PA 15213

Email: cdimov@andrew.cmu.edu

Abstract

Higher cognitive functions are the product of a dynamic interplay of perceptual, mnemonic, and other cognitive processes. Modeling such interplaying processes and generating predictions about both behavioral and neural data can be achieved with cognitive architectures. However, such architectures are still relatively rarely used, likely because working with them comes with high entry-level barriers. To lower these barriers, we provide a methodological primer for modeling higher cognitive functions and their constituent cognitive subprocesses with arguably the most developed cognitive architecture today – ACT-R. We showcase a principled method of generating individual response-time predictions and demonstrate how neural data can be used to refine ACT-R models. To illustrate our approach, we develop a fully specified neurocognitive model of a prominent strategy for memory-based decisions – the take-the-best heuristic – modeling decision making as a dynamic interplay of perceptual, motor and memory processes. This implementation allows us to predict the dynamics of behavior and temporal and spatial patterns of brain activity. Moreover, we show that comparing the predictions for brain activity to empirical BOLD data allows to differentiate competing ACT-R implementations of TTB.

Keywords: ACT-R, fMRI, response times, model comparison.

If a theory covers only one part or component, it flirts with trouble from the start.

A. Newell's (1990, p. 17)

1. Introduction

Research in cognitive psychology aims at identifying the mental processes that produce observable behavior – in Anderson's (1990) terms, to find "the function that maps input to output" (p. 24). Yet, perceptual, mnemonic, and other cognitive processes typically generate behavior in interplay with higher cognitive functions. A prime example of such higher cognitive functions are processes of decision making. When trying to understand the processes underlying decision making, cognitive scientists have been relying not only on data about the decision outcome, but also on various types of data potentially indicative of the decision process itself. Such process data include response times (RTs), measures of predecisional information search (e.g., eye tracking, Mouselab; Duchowski, 2002; Willemsen & Johnson, 2011), or verbal protocols (Ericsson & Simon, 1980). From those data, researchers try to infer what kind of decision mechanisms drove a persons' choices. Yet, in doing so, they face a conundrum: Due to the abovementioned interplay of cognitive subprocesses, the observed RTs and eye movements are not only a function of decisional, but also of perceptual, attentional, and mnemonic processes.

In addressing this conundrum many researchers try to keep all non-decisional processes constant in an experiment and rely on supplementary assumptions about these processes to evaluate the extent to which observations are consistent with a given decision mechanism. Those assumptions come in different forms. For example, some are explicit assumptions about memory, such as that the time to judge an object as recognized will vary as a function of frequency of occurrence in the media (e.g., Hertwig, Hertzog, Schooler, & Reimer, 2008). Others are assumptions about reading phases and key-stroke times, which are often assumed to be constant and then subtracted from observed RTs (e.g., Johnson, Schulte-

3

Mecklenbeck, & Willemsen, 2008; Pachur, Hertwig, Gigerenzer, & Brandstätter, 2013).

Moreover, there are assumptions within data analysis procedures, such as that the probability

of committing an error is equal for all decision trials (trembling-hand error, e.g., Bröder,

2003) or that RTs are log-normally distributed (e.g., Glöckner, 2009). Finally, approaches

differ in terms of their specificity and complexity, ranging from informal (i.e., verbal)

assumptions about the average duration of potential arithmetic and reading operations (Payne,

Bettman, & Johnson, 1993) to formal (i.e., mathematical or computational) theories of, for

example, memory (e.g., Dougherty, Gettys, & Ogden, 1999).

       While such approaches have been followed to identify specific regularities in the

observed data speaking to different cognitive processes of interest, very few researchers

actually model how those cognitive processes dynamically interplay and, consequently,

produce observable RTs, eye movements, and other process data. Yet, this can be

problematic when attempting to identify specific cognitive mechanisms. Is a model's

description of the data (in)adequate because of the main or because of the supplementary

assumptions (see epigram above)? What has been called the *irrelevant specification problem*

(Lewandowsky, 1993; A. Newell, 1990),[1] revolves around the dilemma of whether

assumptions make psychological claims or whether they merely serve to enable the

generation of predictions. In the worst case, incorrect assumptions will lead to patterns in

observed data being wrongly attributed to the cognitive process of interest (e.g., decisional

processes).

       Moreover, while behavioral data—including outcome data (i.e., overt decisions) and

non-physiological process data (i.e., RTs, eye movements, verbal protocols)—inform us

about at least some aspects of a cognitive task, they do not provide sufficient constraints to

---

[1] When translating an underspecified theory into a precise formal model, a lot of details have to be specified in the code, which are not part of the original theory. The irrelevant specification problem poses the question of which of those details can be considered part of a psychological theory and which not.

unequivocally identify the underlying cognitive subprocesses, because "there is an infinite number of mechanisms that compute the same input-output function[]" (Anderson, 1990, p. 24). In fact, Anderson's critique, known as the *identifiability problem*, puts into question any process model that is developed and tested only through behavioral data. Instead, he contends, we need the type of neural data that "traces out the states of computation in the brain" (Anderson, 1990, p. 25) to pin down the mental steps a participant goes through. Ideally, taking into account neural data would allow researchers to deduce the dynamics of these cognitive processes from physiological data on the temporal and spatial patterns of brain activity. However, how can task-related activity in specific brain regions be related to dynamically interplaying cognitive mechanisms?

Cognitive Architectures

A formal description of temporal and spatial patterns of brain activity in different tasks can be provided by *cognitive architectures*. A cognitive architecture is a quantitative model that applies to many different behaviors and that casts theories of memory, perception, action selection, and other components of cognition into a single mathematical or computational system (for an introduction to cognitive architectures, see, e.g., Gluck, 2010). At present, the most detailed cognitive architecture is *ACT-R* (Anderson, 2007; for other architectures, see e.g., Eliasmith, 2013; Meyer & Kieras, 1997; A. Newell, 1990). ACT-R has been continuously developed and updated over the last decades to incorporate current findings and theoretical ideas in a principled manner. ACT-R simultaneously generates multiple types of quantitative behavioral and neural data, ranging from RTs and eye movements to functional magnetic resonance imaging (fMRI) or electroencephalography (EEG) data. Indeed, ACT-R is being increasingly used in combination with neural data to link neural activation patterns with specific cognitive processes (see Anderson, Fincham, Qin, & Stocco, 2008). Moreover, neural data have been relied on to more rigorously test cognitive

models composed of those cognitive processes (see, Borst & Anderson, 2015, for an overview of these approaches). All this makes ACT-R an excellent tool to address the irrelevant specification and identifiability problems.

ACT-R is widely used by a large worldwide community and applied to areas as diverse as airplane flying (Byrne & Kirlik, 2005), intelligent tutoring (Ritter, Anderson, Koedinger, & Corbett, 2007), skill acquisition (Taatgen, Huss, Dickison, & Anderson, 2008), or list memory (Anderson, Bothell, Lebiere, & Matessa, 1998; see http://act-r.psy.cmu.edu/publication/ for a complete list of publications). Yet, building a model in ACT-R comes with at least three important entry-level barriers. First, users need to fully understand the theory. However, this is not trivial, because ACT-R consists of models of various aspects of cognition (e.g., memory, perception, procedural knowledge) and their interaction. Second, users need to know how to implement their hypotheses in this computational modeling framework, which is instantiated as a programming language with built-in human constraints, written in Common Lisp. Third, users need to be aware of methods for developing, calibrating, and testing complex architectural models.

Our goal in this article is to guide scientists modeling process data with ACT-R. In offering this methodological primer, we complement the extensive tutorial that comes with the ACT-R software (available at http://act-r.psy.cmu.edu/software/) and the excellent step-by-step tutorial of how to use the architecture with fMRI data (Borst & Anderson, 2017) in three ways. First, we illustrate how ACT-R model parameters can be empirically constrained on separate experimental tasks in a principled manner. Second, we showcase the capability of these constrained models to generate participant-specific behavioral predictions (e.g., RT distributions). Third, we illustrate how BOLD data can help to further refine a model beyond what can be achieved from behavioral data alone.

We will first introduce the formal underpinnings of ACT-R and demonstrate how to develop a neurocognitive model for a prominent strategy in memory-based decision making, the *take-the-best* (TTB) heuristic (Gigerenzer & Goldstein, 1996). To develop our model, we will rely on an fMRI experiment, in which participants were instructed to follow that decision strategy (Khader et al., 2011, Exp. 1). We will demonstrate how the parameters of our model can be empirically estimated in a behavioral task in that study that precedes the actual decision task. We will then generate individual predictions about RTs in the decision task and compare those to the empirical data. Finally, we will showcase how to use the BOLD predictions of our model to further refine it.

2. Overview of ACT-R

In order to model and predict behavior and brain activation with ACT-R, different cognitive processes are modeled by separate *modules*, which have been mapped onto different brain areas (see Figure 1). These modules include perceptual ones, namely a *visual* and an *aural* module, which model focused attention to perceptual input and are mapped to regions reflecting advanced perceptual processing: the secondary auditory cortex and fusiform gyrus, respectively. There are also *vocal* and *manual* modules, which model speech and typing on a keyboard, respectively. These output modules are mapped onto two regions in the central sulcus where the hand, and face and tongue are represented. Furthermore, there are three central cognitive modules: The *goal module* tracks an agent's goals; this module maps onto the anterior cingulate cortex. The *imaginary module* holds information relevant to the task and problem state at hand. This module corresponds to posterior parietal cortex—a region hypothesized to be involved in the transformation of mental representations. How information is stored in and retrieved from declarative memory is modeled by the *declarative module*. This module is associated with the lateral inferior prefrontal cortex. All modules can

operate in parallel, but within each module, information is processed serially (Byrne & Anderson, 2001).

Perceptual and central cognitive modules operate on declarative knowledge (i.e., explicit memory). Such knowledge is modeled by *chunks*. Chunks represent information input from the visual and aural modules, the current goal of the cognitive system, information relevant to the problem state as well as knowledge in long-term memory. Chunks are collections of attributes, called *slots*, and their corresponding values:

(**chunk-name**[2] *slot1* slot1-value *slot2* slot2-value *slot3* slot3-value …).

For example, we can represent factual knowledge, such as "Paris is the capital of France", or current states of the world, such as "the bird sings loudly", with the following two chunks:

(**capital-France** *name* Paris *role* capital *country* France)

(**loud-bird-by-my-office** *object* bird *action* sings *adverb* loudly).

A *procedural module* orchestrates the other modules and functions as the central bottleneck in information processing (Figure 1). The procedural module is associated with the basal ganglia—a system hypothesized to implement conditional information routing to the cortex (Stocco, Lebiere, & Anderson, 2010). This module is instantiated as a *production system* (i.e., A. Newell, 1973), that is, it consists of a collection of *production rules* (if-then rules). The productions' conditions (the "if"-parts of the rules) are matched against the current state of the other modules (e.g., whether something is retrieved from memory or whether an object is visually attended to). Production rules, whose conditions are met, can *fire*, that is, they can direct other modules to change their current state. Examples of production rules in natural language are:

1. **IF** an object is visually attended to and the goal is to look at it,

   THEN visually encode the object of attention.

---

[2] For convenience, chunk names (e.g., capital-France) are used in ACT-R to refer to chunks. Those names are not considered to form part of the chunk itself.

2. **IF** the goal is to guess the name of a country's capital and France is currently stored into the problem state,

THEN attempt to recall the name of France's capital.

Production rules do not access modules' contents directly, but via *buffers*. Buffers serve as communication channels between modules and productions, and as such can create a bottleneck for information transfer (Salvucci & Taatgen, 2008). For example, if the procedural module needs to access information in the visual field, the visual module has to first place that information (in the form of a chunk) into the *visual buffer*. Likewise, when a production rule sends a retrieval request to the declarative module, the retrieved information (i.e., a chunk) must first be placed in the *retrieval buffer* before another production rule can utilize it. In essence, when an ACT-R model is run, requests are sent to modules, which leads the modules to execute operations. After completing those operations, modules can place chunks into their respective buffers. Reversely, production rules await those chunks to be placed into specific buffers to match the conditions specified in their "if-part". Once those conditions are met, the rules can fire (i.e., execute their "then-part") and send further processing requests to modules. The serial operation of modules and the time that it takes them to complete their operations creates delays and bottlenecks in the system. Altogether, this complex interaction between the procedural and other modules (Figure 1) produces behavior.
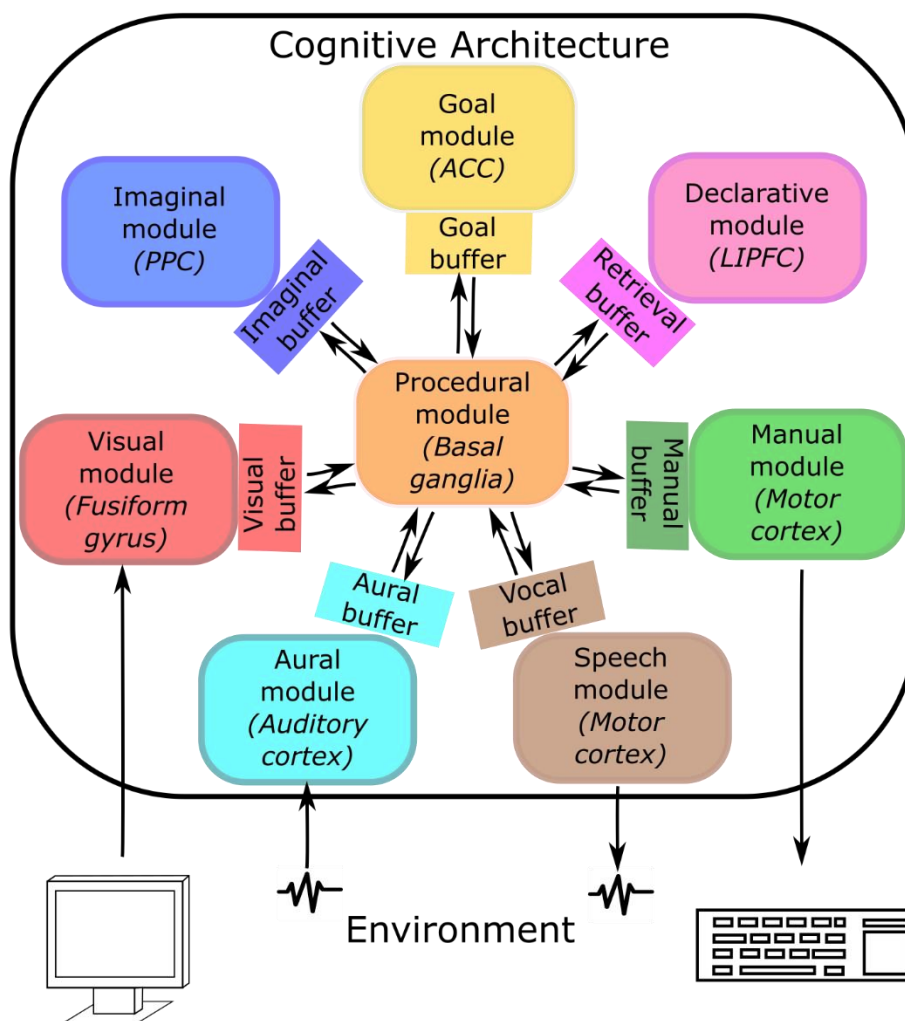
Figure 1. The modular organization of the cognitive system and module-to-brain mappings according to the ACT-R cognitive architecture. Seven modules interact with each other through an eighth, procedural, module. The procedural module communicates with the other modules through buffers, represented as small rectangles. The cognitive architecture interacts with the environment (e.g., in an experiment: with a computer screen and/or a keyboard) through its perceptual and motor modules. ACC = anterior cingulate cortex; LIPFC = lateral inferior prefrontal cortex; PPC = posterior parietal cortex. See the online article for the color version of this figure.

ACT-R's Subsymbolic System

The modules and buffers can be best thought of as an "upper" symbolic layer of the architecture. ACT-R distinguishes that *symbolic system* from a "lower" layer, called the *subsymbolic system*. The subsymbolic system shapes the outcome of each module's and each buffer's operations. Specifically, the subsymbolic system describes memory retrieval, the

selection of different production rules, visual and other processes in terms of a series of mathematical equations. Those equations determine, for example, how likely and how quickly memories can be retrieved, which of several alternative courses of action will be executed, or how long a key press will take.

An important component of ACT-R's subsymbolic system are the equations governing the retrieval of memory traces (i.e., of chunks). Altogether, these equations cast memory as an information-processing device that systematically exploits the statistical patterns of occurrence of stimuli in the world. Specifically, memory can be thought of inferring the probability that a memory trace of stimuli will be needed based on the history of past encounters with those stimuli in order to achieve future processing goals. The history of past encounters, in turn, probabilistically hinges on patterns of occurrence of those stimuli in the world. For instance, we are more likely to learn about car brands and cities that occur more often in the media. In real world environments, patterns of past occurrence of stimuli are predictive of future ones (e.g., Anderson & Schooler, 1991; Schooler & Anderson, 1997). To illustrate this, the more often an object (e.g., a name) has been mentioned in the news in the past, the more likely it is that this object will be mentioned again in the future. Similarly, the longer it has been since an object has last been mentioned, the less likely it is that the object will be mentioned again in the future. These lawful relations in information occurrence in the environment allow memory to guide present information-processing demands, for example, by retrieving memory traces of recently encountered stimuli more quickly or by setting aside (i.e., forgetting) information that has been encountered infrequently or a long time ago.

Specifically, in ACT-R each chunk i has an *activation*, A_i, associated with it that quantifies the strength of that memory trace. Activation models the likelihood that a chunk is needed to achieve a given processing goal at the current moment. Activation itself is fed by

three subcomponents—the chunk's *base-level activation*, B_i, the *spreading activation*, SA_i, and *noise, ε*:

$$A_i = B_i + SA_i + \varepsilon. \tag{1}$$

The base-level activation is a function of the chunk's history:

$$B_i = \ln \sum_{k=1}^{n} t_k^{-d}, \tag{2}$$

where the *decay parameter*, d, specifies the rate forgetting over time, which is modeled in terms of a power function. The parameter n represents the number of encounters with the information that chunk i represents, and t_k is the time since the kth encounter. Those latter two parameters capture the history of encountering stimuli in the world.

SA_i, quantifies a chunk's relevance in the current context by assuming that chunks related to what is currently the focus of attention are more likely to be needed than those that are not. In ACT-R, context is modeled as all chunks currently stored in the buffers. Thus, spreading activation to chunk i in declarative memory is a function of the associations between that chunk and the chunks j currently in the buffers:

$$SA_i = \sum_j W_j S_{ji}. \tag{3}$$

The amount of spreading activation SA_i is determined by the associative strength, S_{ji}, between chunks i and j, which is weighted by the source activation, W_j, of chunk j in a buffer. The associative strengths, S_{ji}, between chunks is approximated by

$$S_{ji} = S - \ln(fan_j), \tag{4}$$

where S denotes the maximum associative strength and fan_j is the number of chunks associated with a chunk j. The more chunks are associated with a chunk in memory, the lower the associative strength between it and each of its associates becomes.

Finally, retrieval noise, ε, is added to the base-level and spreading activation components when a retrieval request is made. ε is generated from a logistic distribution with a mean of zero and a variance of

$$\sigma_d^2 = \frac{\pi^2}{3} s_d^2, \tag{5}$$

where s_d is a free parameter. The random nature of activation means that at each specific time a chunk can have an activation higher or lower than its mean.

A chunk can be retrieved only when its current activation value is above the *retrieval threshold, τ*. The retrieval probability, p_i, of a chunk is the probability that its activation exceeds this threshold:

$$p_i = \frac{1}{1+e^{-\frac{\mu_{A_i}-\tau}{s}}}, \tag{6}$$

where $\mu_{A_i} = B_i + SA_i$ is the mean of the activation distribution. The time required for retrieval is scaled by a *latency factor F*:

$$t_{retrieval} = Fe^{-A_i}. \tag{7}$$

Thus, more active chunks are more quickly retrieved. If no chunk matches a retrieval request or if the matching chunk with the highest activation is below the retrieval threshold, a retrieval failure will occur. The retrieval failure time is:

$$t_{retrieval\ failure} = Fe^{-\tau}. \tag{8}$$

Another important component of the subsymbolic system are the equations governing ACT-R's production rules. Various higher-order cognitive functions are implemented in ACT-R as sets of productions that specify, for instance, what the goals of the decision maker are (e.g., making inferences about car brands as accurately as possible), when and how that person will encode information in his/her environment (e.g., reading the names of different car brands in a catalogue or on a computer screen) and when and how the participant will search for information in memory (e.g., what facts about different cars will be recalled). These equations determine which production rules will be executed in case the conditions (i.e., the if-parts) of several of those if-then rules are met. According to these equations, the productions that have been most successful in the past are the ones that are most likely to be

chosen, with production success being quantified by its *utility*. A production's utility is

learned according to the Rescorla-Wagner learning rule (Wagner & Rescorla, 1972):

$$U_i(n) = U_i(n-1) + \alpha[R_i(n) - U_i(n-1)], \tag{9}$$

where $U_i(n)$ is the utility of the production after its nth application, $U_i(n-1)$ is its utility after

the $n-1^{th}$ application and $R_i(n)$ is the reward that it receives on the nth application. Basically,

upon each application of a production, its utility is updated in the direction of the reward that

it receives: If the reward is lower than its current utility, the utility will decrease, whereas if it

is higher, it will increase.

ACT-R and Brain Activation

To derive predictions regarding brain activation, ACT-R draws on the relationship

between brain activity and blood supply (e.g., Boynton, Engel, Glover, & Heeger, 1996):

metabolic demand in an active brain region leads to an increased blood supply to that region,

measured as a *hemodynamic response (HR)*. The HR is not immediate, but peaks around 6

seconds after the metabolic demand. Its temporal profile, labeled the *hemodynamic response*

function (*HRF*), is described by a gamma distribution or a mix of two gamma distributions.

Here, we will use the canonical *HRF* as implemented in the SPM fMRI analysis software

(*statistical parametric mapping*; Friston et al., 1998; see Figure 2 for a visualization of this

HRF):

$$HRF_{SPM}(t) = \frac{6t^5 e^{-t}}{\Gamma(6)} - \frac{1}{6}\frac{16t^{15}e^{-t}}{\Gamma(16)}, \tag{10}$$

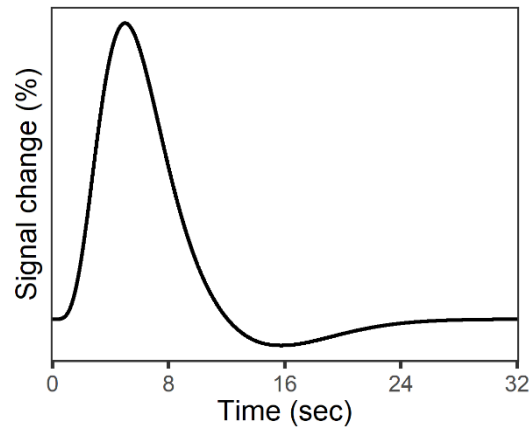where Γ is the gamma function.

Figure 2. SPM's canonical HRF.

Once an ACT-R model has been developed, generating BOLD response predictions involves the following steps: First, the activity of each module is described with a demand function $D(t)$. Here, we follow the standard assumption that whenever a module is active [$D(t) = 1$], the brain region associated with this module is active [else, $D(t) = 0$ and the brain region is inactive]. We then assume that at each moment when that brain region is active, it responds according to the *HRF* in Equation 10 (see Borst & Anderson, 2017). The resulting HR prediction is a convolution of the demand function over the entire experiment, and that *HRF*:

$$HR(t) = (D * HRF)(t). \qquad (11)$$

The HR predictions can be related to brain activation data in two ways: *model-based fMRI* analysis and *region-of-interest* (*ROI*) analysis (e.g., Borst & Anderson, 2015). Recall that ACT-R's modules are assumed to mediate different cognitive functions (e.g., declarative memory, vision). In model-based fMRI analysis, the *blood-oxygenation-level-dependent* (BOLD) response prediction of each ACT-R module is regressed against all voxels in the experimental data, which, in turn, allows identifying the brain correlates of the modules. This approach has been used to identify regions in the brain that strongly correlate with module activity.

15

While model-based fMRI analysis serves to establish module-to-brain mappings, ROI analysis uses already established mappings to evaluate cognitive models. Specifically, ROI analysis compares BOLD predictions associated with different cognitive processes to brain activation in predetermined regions.[3] In contrast to analyses involving behavioral data, an ROI analysis of neural data is especially useful for testing cognitive models (e.g., Anderson, Carter et al., 2008; Borst, Taatgen, Stocco, & van Rijn, 2010), because neural data relate directly to each individual module's activity.

Next, after introducing the decision strategy that will serve as a case-in-point – TTB – we will demonstrate how ACT-R models work and how predictions for that strategy can be derived and tested. We will use both behavioral data and neural data in a ROI analysis to test the detailed cognitive model of TTB.

3.   An Example of Decision Processes: The TTB Heuristic

TTB is a representative of an important class of decision strategies that implement sequential information search. TTB and similar lexicographic models, such as elimination-by-aspects (Tversky, 1972), stand in contrast to the classic assumption that people integrate and weight available evidence to make decisions (Chase, Hertwig, & Gigerenzer, 1998), such as in subjective expected utility theory (Edwards, 1954), or other compensatory weighted-additive strategies (e.g., Payne et al., 1993). TTB is a model of inference: It uses objects' (e.g., cell phones, car brands, cities) attributes (e.g., whether a cell phone is recommended by others) to infer which of two objects has the larger value on an unknown *criterion* (e.g., the phone's quality). To this end, TTB operates on attributes with binary attribute values that are coded as 1 if positive (e.g., phone is recommended) or 0 if unknown or negative (not

---

[3] For Talairach coordinates of those brain regions, see Anderson, Fincham et al. (2008). For MNI coordinates, see Borst, Taatgen, Stocco, and van Rijn (2010). Note that Borst, Nijboer, Taatgen, van Rijn, and Anderson (2015) have further refined those brain mappings through a data-driven model-based approach.

recommended). When making inferences, TTB inspects attributes in order of their importance. Once two objects have different values on an attribute *i* (i.e., one has a value of 1, the other of 0), that is, once a *discriminating* attribute is found, TTB makes a decision without considering further information (i.e., other attributes). In the literature, this decision process has been described in terms of three building blocks (e.g., Gigerenzer & Gaissmaier, 2010):

> *Search rule*: Search through attributes in the order of their validity.
>
> *Stopping rule*: Stop search as soon as an attribute is found that discriminates between the objects.
>
> *Decision rule*: Infer that the object with the positive attribute value has the higher value on the criterion of interest.

TTB has been shown to be spontaneously used in particular under conditions with high information cost, including memory-based (rather than screen-based) decisions (e.g., Bröder & Gaissmaier, 2007; Bröder & Schiffer, 2003). The notion underlying TTB that people sometimes ignore information has triggered a large number of empirical studies (Bergert & Nosofsky, 2007; Bobadilla-Suarez & Love, 2017; Bröder & Schiffer, 2003; Juslin, Jones, Olsson, & Winman, 2003; Pachur & Aebi-Forrer, 2013; Khader, Pachur, & Jost, 2013; Pachur & Marinello, 2013; Rieskamp & Otto, 2006; for an overview, see Pachur & Bröder, 2013). To test TTB against information integration models, such as compensatory weighted-additive strategies, those studies have made use of RTs (Bröder & Gaissmaier, 2007) and patterns of information search (B. R. Newell, Weston, & Shanks, 2003), often making the kind of supplementary assumptions about mnemonic, perceptual, and motor processes that we mentioned in the Introduction (Section 1). Indeed, as can be seen, the search, stopping, and decision rules used in the literature to describe TTB remain fully silent about how the decisional processes assumed by TTB nestle into the rest of the cognitive

architecture. Thus, developing an ACT-R implementation of TTB illustrates how to integrate TTB's key theoretical assumptions about decision making (such as sequential and limited search) with mnemonic, visual, and other information-processing activities.

When developing an ACT-R model of TTB—or any other decision mechanism, for that matter—to unveil the cognitive processes behind a people's decisions, it is important to ensure that the observed output on every instance (e.g., a trial in an experiment) is produced by that mechanism and not by another one. In the cognitive and decision sciences, many experiments present participants with a decision task and then, based on their responses, individual participants are classified, for instance, as "users" of TTB or alternative decision mechanisms (Bröder, 2000; B. R. Newell & Lee, 2011; Nosofsky & Bergert, 2007). While suited for testing competing models of decision making against each other, such data make it difficult to develop architectural models of TTB, because the observed data can, but need not be produced by TTB. A data set where participants' reliance on TTB is ensured has been provided by Khader et al. (2011). In their experiments, participants were instructed to rely on TTB for their decisions while their brain activity was recorded, which renders this data set an excellent basis for illustrating how neurocognitive ACT-R models can be developed.

4. Developing an ACT-R Implementation of TTB

Khader et al. (2011) employed a memory-based paradigm, where two objects (here: a pair of fictitious companies) are presented on a computer screen and participants have to rely on their memory to recall previously learned attribute values and make a decision according to TTB. An ACT-R implementation of TTB should perform the same operations that a participant would: For example, it needs to read the object names and then recall objects' attribute values in order to make a decision.

The attribute values of objects are stored as chunks in the model's declarative memory. What slots would these chunks consist of? In Khader et al.'s experiment, prior to

the decision task participants learned to associate objects (i.e., companies) with attributes

(e.g., where the company is located and which product it produces) and their values (i.e.,

whether the attributes are positively or negatively related to the decision criterion; see Khader

et al., 2011, for details). In our model, we rely on the following chunk structure to describe

object attributes stored in memory:

(**objectN-attributeM** *object-name* objectN *attribute-name* attributeM *attribute-value*

0/1).

How are such chunks used to develop a model of TTB? In modeling participants' declarative

memory, in a first step, the model's declarative memory is populated with all attribute values

that participants learned prior to working on the decision task.

Once we have defined the declarative chunk structure, we continue with outlining the

sequence of steps that a model needs to go through. In addition to TTB's search, stopping,

and decision rules (see Section 3), we need to include all steps that a participant in an

experiment would go through, such as visual and motor steps. For our task, these steps are:

(1) look at the company names, (2) retrieve the attribute, (3) retrieve the corresponding

attribute values, (4a) press the key on the keyboard that corresponds to the company with a

positive attribute value if the attribute values differ or (4b) retrieve the next attribute in the

hierarchy if the attribute values on the current attribute are the same.

To translate these steps into a sequence of productions, it is necessary to consider

ACT-R's architectural constraints. For example, the visual module can only process objects

serially. This means that the visual system needs to attend the first company name on the

screen—an action guided by a production—and then encode that company name guided by a

second production. Only then can it attend and encode the second company name. Similar

buffer capacity and temporal constraints exist for other modules. For instance, the imaginal

module can also only perform one operation at a time (e.g., store a chunk or modify the

chunk it currently holds) and that operation also incurs a time cost (i.e., 200 ms). Similarly, the retrieval module can only be attempting to retrieve one chunk from long-term memory and the time it takes to perform the retrieval is determined by Equation 7.

Figure 3 shows a process trace of a run of the entire ACT-R implementation of TTB for a decision on which the most valid attribute in the hierarchy discriminates between the two companies.[4] The components of the cognitive architecture are active at different points in time, with eight production rules coordinating the modules' actions. In this example, the model first starts by comparing the two objects on the screen: It first looks on the left part of the screen (guided by production 1), reads the name of the company present on that part of the screen (production 2) and stores it in the imaginal buffer, while shifting its gaze to the right part of the screen (production 3). After reading (production 4) and storing the name of the company present there (production 5), the model checks if the companies are different and, if so, it executes TTB, starting with the most valid attribute (production 6). For the most important attribute, the model recalls the attribute value of the left company (also production 6), then it recalls the attribute value of the right company (production 7) and finally, it compares them (production 8). Then, the model chooses the company to which that attribute is pointing and finishes with an overall RT of 2 s. Note that on other trials retrieval might be faster or slower, or more attributes might need to be examined prior to making a decision, which will lead to a different RT and a different relative activity of each module.

---

[4] See model "TTB_v1.lisp" in online materials at osf.io/25pt8 for the model code.
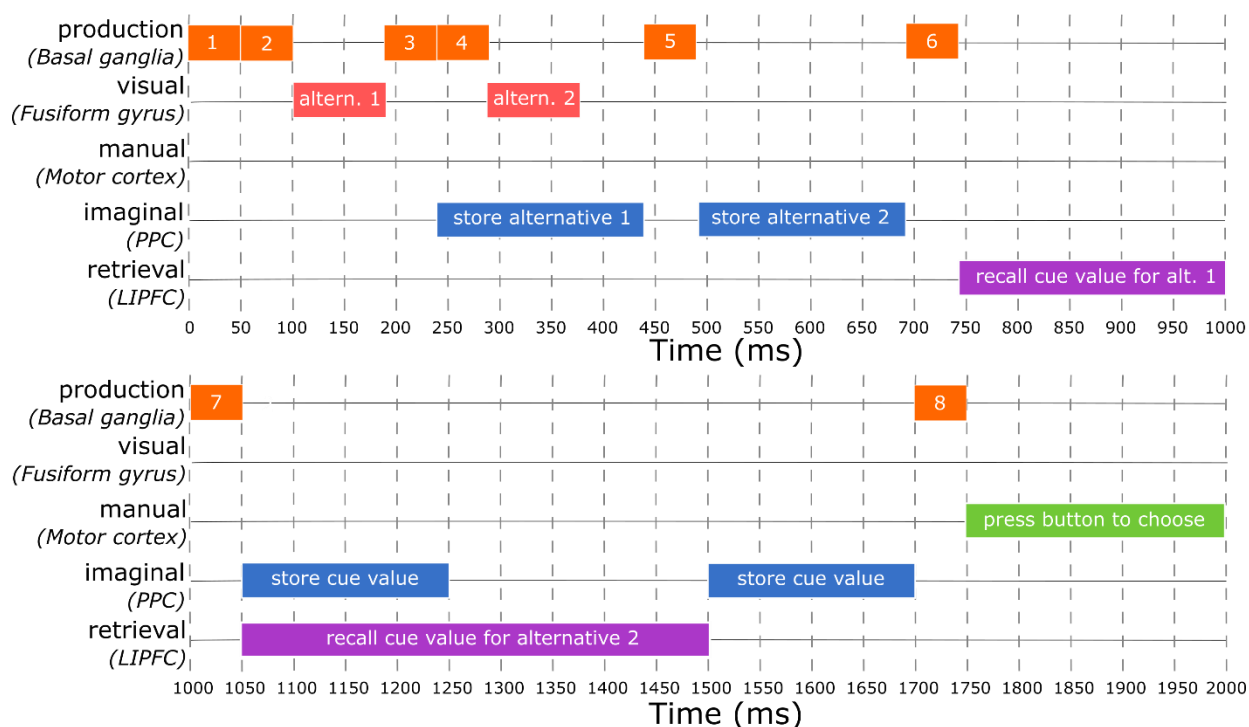
Figure 3. Schematic process trace of an ACT-R implementation of TTB. For the sake of illustration, the model trace schematically depicts arbitrary recall times (for actual times run the model in online materials at osf.io/25pt8). The y-axis denotes various ACT-R modules and associated brain regions. *LIPFC* = Lateral Inferior Prefrontal Cortex. *PPC* = Posterior Parietal Cortex. Eight production rules control the behavior of this model. Production 1 directs visual attention to the location of the first object (company 1) and production 2 requests that the visual module encodes it. Productions 3 and 4 repeat the same steps for the second object (company 2) and also request storing the first object in the imaginal buffer. Production 5 starts storing the name of the second object in the imaginal buffer. Productions 6 and 7 request that the declarative module retrieves the attribute values of the first attribute for each object. Finally, production 8 selects the object with a positive attribute value my requesting a key press. See the online article for the color version of this figure.

5. Testing the ACT-R Model of TTB

In an initial learning task, the 17 participants memorized the values on 4 attributes about 16 fictitious companies; a total of 64 attribute values. In a subsequent strategy-training task, participants were instructed how use TTB. They also practiced to apply this heuristic, using a decision task different from that in the main decision task (a fictitious job-scenario). Finally, participants learned the hierarchy of the four attributes by indicating their importance repeatedly. In the decision task, participants' responses and the associated RTs and BOLD

signals were recorded. One hundred thirty-two company pairs were presented (on the left and right side of the screen, respectively) in three blocks of 44 trials each. Participants were instructed to use TTB and the acquired attribute-knowledge to decide which of the two companies is more likely to be successful in the future. The inter-trial interval (ITI) was 2, 4 or 6 s (varied randomly), and each trial started with a fixation cross presented for 2 s.

There were five types of decision trials, which differed in the number of attributes that TTB would need to consider prior to making a decision (i.e., none, one, two, three, or all four attributes). This is relevant because it is assumed that, due to TTB's stopping rule, the time it takes to make decisions with this strategy depends on how many attributes have been considered before a discriminating attribute is retrieved from memory (e.g., Bröder & Gaissmaier, 2007). In control trials (i.e., where no attributes need to be considered) the same company name was presented on both sides of the screen and participants were instructed to respond directly, without retrieving any attributes. For further details on the experimental methodology, see Khader et al. (2011).

A Roadmap for Model Testing

We used data from Khader et al.'s (2011; Experiment 1) to develop and test an ACT-R implementation of TTB. Figure 4 provides an overview of the various steps and stages of the procedure.[5] First, we calibrated the model by fitting its free parameters in the learning task. To this end, we developed a recall model in ACT-R. We then used the estimated parameters from the learning task to generate distributional RT predictions for the decision task with the TTB implementation. Note that the parameters were not fitted to the data of the decision task, so these are genuine predictions. We also performed a ROI analysis to compare brain activation predictions to BOLD data. Finally, in an iterative process, we used the ROI analysis to further refine our model. Specifically, we constructed seven alternative

---

[5] Data, analysis scripts, and model files are available under osf.io/25pt8.

implementations of TTB and selected the implementation whose processing steps generated

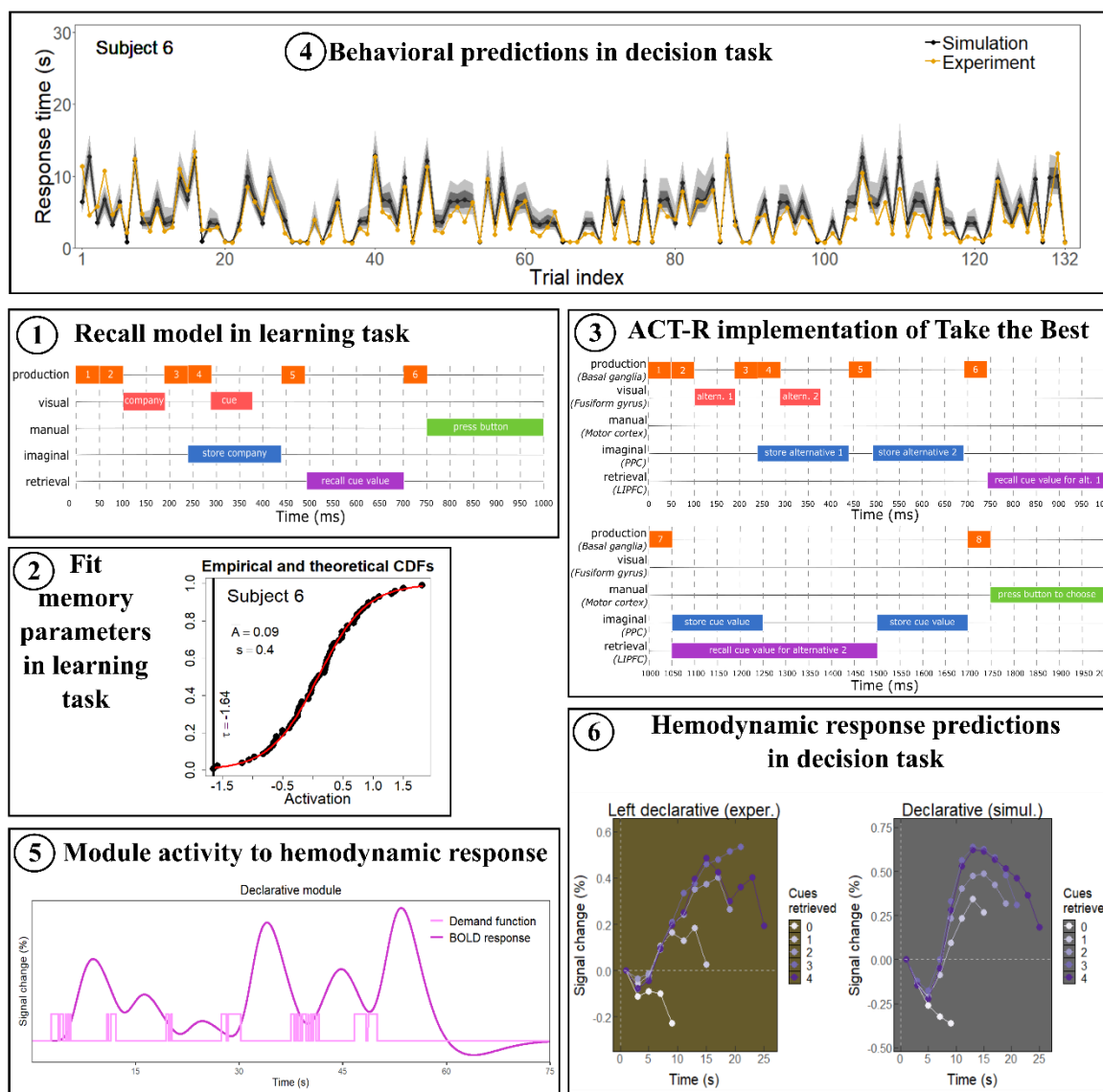BOLD predictions best corresponding to observed data.

Figure 4. Overview of the different steps of model development and testing. (1) An ACT-R model of recall is developed to estimate perceptual-motor times in the learning task. (2) Memory parameters are estimated from the learning task. (3) An ACT-R model of TTB is developed, which uses as input the memory parameters estimated in the learning task. (4) Distributional RT predictions are generated using this model and compared with experimental RTs. (5) Module activity is mapped onto HR. (6) The predicted HR is compared to the experimental fMRI data. See the online article for the color version of this figure.

Model Calibration in the Learning Task

In order to rigorously test a model's descriptive power, it is important to ensure that it performs well in predicting data out-of-sample rather than being able to fit data (e.g., Pitt, Myung, & Zhang, 2002). In order to derive predictions about RTs and BOLD responses in the *decision* task, we estimated ACT-R's memory parameters for each participant from his/her RTs in the *learning* task (Figure 4, steps 1 & 2). During the learning task, different attribute values are remembered increasingly better across various rounds of learning (see the description of the learning task above), until the corresponding chunks are activated strongly enough to be retrieved with a probability of almost 1 (see Equation 6) in the last round of learning. Thus, the last learning round defines the peak activation of each attribute-value chunk. Moreover, it is temporally closest to the decision task. This renders a chunk's activation at this point in time a reasonable approximation to that chunk's activation at the beginning of the decision task.

ACT-R relates activation to RT as per Equation 7 (Section 2). We transformed this equation to estimate the activation of chunks representing attribute values from the retrieval time of each attribute-value chunk in the last round of learning:

$$A = -\log \frac{t_{retrieval}}{F} \ , \tag{12}$$

where the latency factor, F, is left fixed at its *default value* ($F = 1$). This implies that estimating activation as per Equation 12 requires no parameter fitting.[6] To assess retrieval time, we then assumed that

$$t_{\text{retrieval}} = \text{RT} - t_{\text{non-retrieval}}, \tag{13}$$

that is, that the total RT consists of separable retrieval and non-retrieval components. We then estimated $t_{\text{non-retreival}}$ by constructing an ACT-R model of recall in a learning trial, whereby we

---

[6] ACT-R includes parameters which shape the workings of its subsymbolic system. For example, these parameters determine how long retrieval takes, or low quickly one would attend an object in the visual field. These parameters have default values, which ensure consistency between models constructed within this architecture.

relied entirely on ACT-R's default parameters (Figure 4, step 1). This model, shown in

Figure 5, starts by looking at the company name. It then stores that name in the imaginal

buffer and looks at the attribute. Once both attribute and company are available in the

model's buffers, the model attempts recalling the attribute value and responds by pressing a

key on the keyboard. We computed the median duration of non-retrieval processes over 100

runs[7] of that model to estimate non-retrieval time. This was necessary because ACT-R

assumes a certain variability in the operation time of its various cognitive components, such

as visual attention and motor action. As a result, ACT-R makes predictions about the

distributional characteristics of the RTs. Our model estimated a mean perceptual-motor time

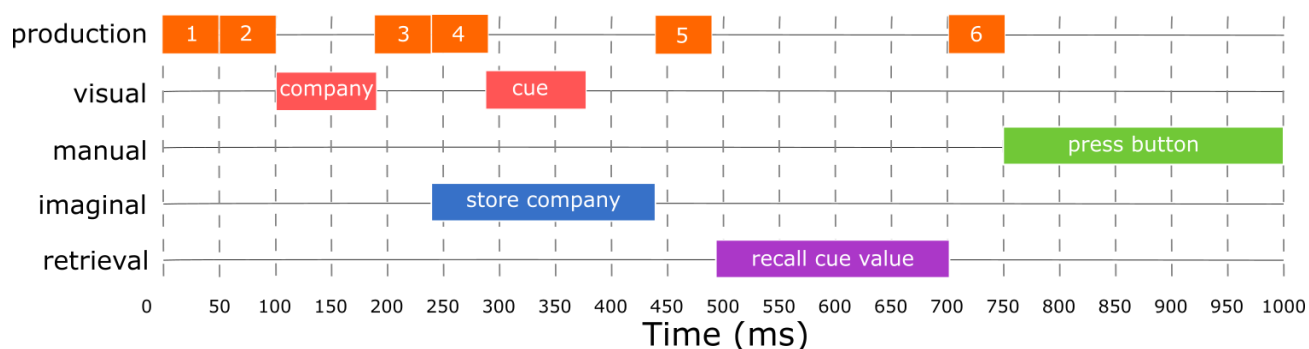of 780 ms for the first attribute of a company and 495 ms for the remaining three attributes.[8]

Figure 5. Schematic process trace of ACT-R model of recall. The y-axis denotes various
ACT-R modules. This model operates through six productions. Productions 1 and 2 request
that the visual module attends to and encodes the company name, while productions 3 and 4
direct the visual module to attend and encode the attribute. Production 5 tell the declarative
module to recall the value on that attribute and, finally, production 6 makes the appropriate
response by requesting a key press from the manual module. For the sake of illustration, the
model trace schematically depicts somewhat arbitrary recall times, and not those that a
participant would likely need. See the online article for the color version of this figure.

---

[7] We ran each model 100 times, which strikes a good balance between number of data points generated and
required simulation time.
[8] In the learning task of the experiment, attributes were grouped per company. Thus, a participant only needs to
encode the company name for the first attribute.

After subtracting the average $t_{\text{non-retrieval}}$ from the overall RT, we removed outliers from the resulting $t_{\text{retrieval}}$: Specifically, we first, we removed outliers on the left side of the distribution by eliminating all negative values; this amounted to 4.3% of all observations. Second, we removed outliers on the right side of the distribution by removing the 2.5% most extreme values (see Ratcliff, 1993, for general recommendations of outlier removal). The 97.5$^{\text{th}}$ percentile of the distribution of all participants' RTs was at 6.2 s.

To account for memory retrieval being inherently noisy, ACT-R models the noise on activation with a logistic distribution (see Equations 1 and 5). We assume that all 64 attribute-value chunks stored in a participant's memory are characterized by the same parameter values (logistic distributions with equal means and scales; Figure 4, step 2). This assumption is plausible for three reasons. First, all attributes have approximately the same learning history (i.e., they were presented similarly often and similarly long ago) and hence the same base-level activation. Second, all attribute-values receive the same amount of spreading activation, because every attribute is related to 16 attribute values (an associative fan of 16) and each company to 4 attribute values (a fan of 4). And third, by definition, there is a single activation noise parameter per participant. Starting from Equation 1, this assumption means that all chunks i for a participant have an activation:

$$A_i \sim Logistic(\mu_A, s). \qquad (14)$$

In essence, three parameters have to be estimated per participant: the two parameters of the logistic distribution (mean activation μ_A, and activation noise s) and, also, a retrieval threshold τ (see Equations 6 and 8).

Figure 6 shows, for three representative participants, the resulting fit of a logistic distribution to the 64 samples from the activation distribution from the last round of learning (corresponding to the 64 attribute values). These 64 data points are RT distributions transformed with Equations 12 and 13. The parameters of such a theoretical cumulative

distribution function provide an estimate of the mean activation and activation noise for all 64 attribute-value chunks of a participant. To estimate each participant's retrieval threshold, we set it equal to the activation of the least active attribute value (also, see Figure 6).
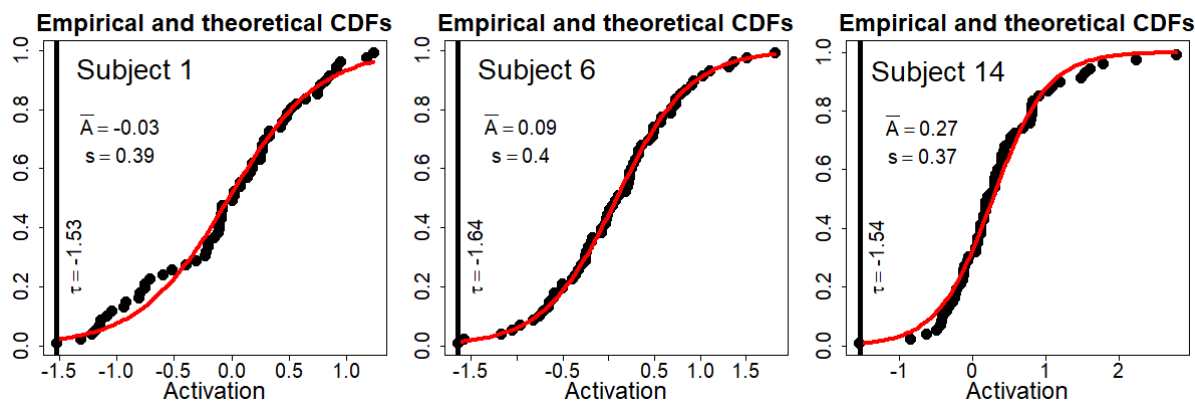

Figure 6. Examples of fitting a logistic distribution to learning data of three participants. Points show 64 empirically estimated activations from the last round of the learning task for each of the 64 attributes. The red line shows the fitted logistic curve, while the black line designates the value of the retrieval threshold. See the online article for the color version of this figure.

RT Predictions

The memory parameters, estimated individually for each participant, are used to derive predictions for the decision task: activation and other components feed into the TTB model described above (Figure 4, step 3).[9] When generating predictions, all other parameters were set to their ACT-R default values. To mimic the exact experimental conditions, we added the timing details of an ITI of 2, 4, or 6 s, a fixation cross presented for 2 s at the beginning of each trial, and a presentation of a company pair until a response is made (see the description of the experimental procedure above) to the model presented in Figure 3. When the fixation cross is drawn, the model looks at it. When an ITI is presented, the model does nothing. Finally, when the two companies are presented, the model executes TTB: it reads the

---

[9] Given that the time elapsed between last round of the learning task and the beginning of the decision task is minimal, we assume negligible memory decay between those tasks.

company names sequentially, recalls attributes and attribute values and responds with a keypress (see Figure A1 in the Appendix for a trace of the complete final model, which also includes screen events and mental events for clarity).

As mentioned in Section 2, ACT-R assumes that cognitive processes are inherently noisy, implying that the same process can produce different patterns of data. To model this variation, ACT-R models are typically run multiple times in a computer simulation. In our case, the simulation of the complete decision task, consisting of 132 trials, was repeated 100 times for each participant. In so doing, in each simulation run and for each participant, stimuli (i.e., pairs of company names) were presented in the same order as the participants saw them. The resulting distributional RT predictions are compared against the empirical RTs (Figure 4, step 4). Figure 8 offers a snapshot of such a comparison by plotting median RTs and RT percentiles, together with empirical RTs for three participants (see the Supplementary Online Materials for plots for all participants).
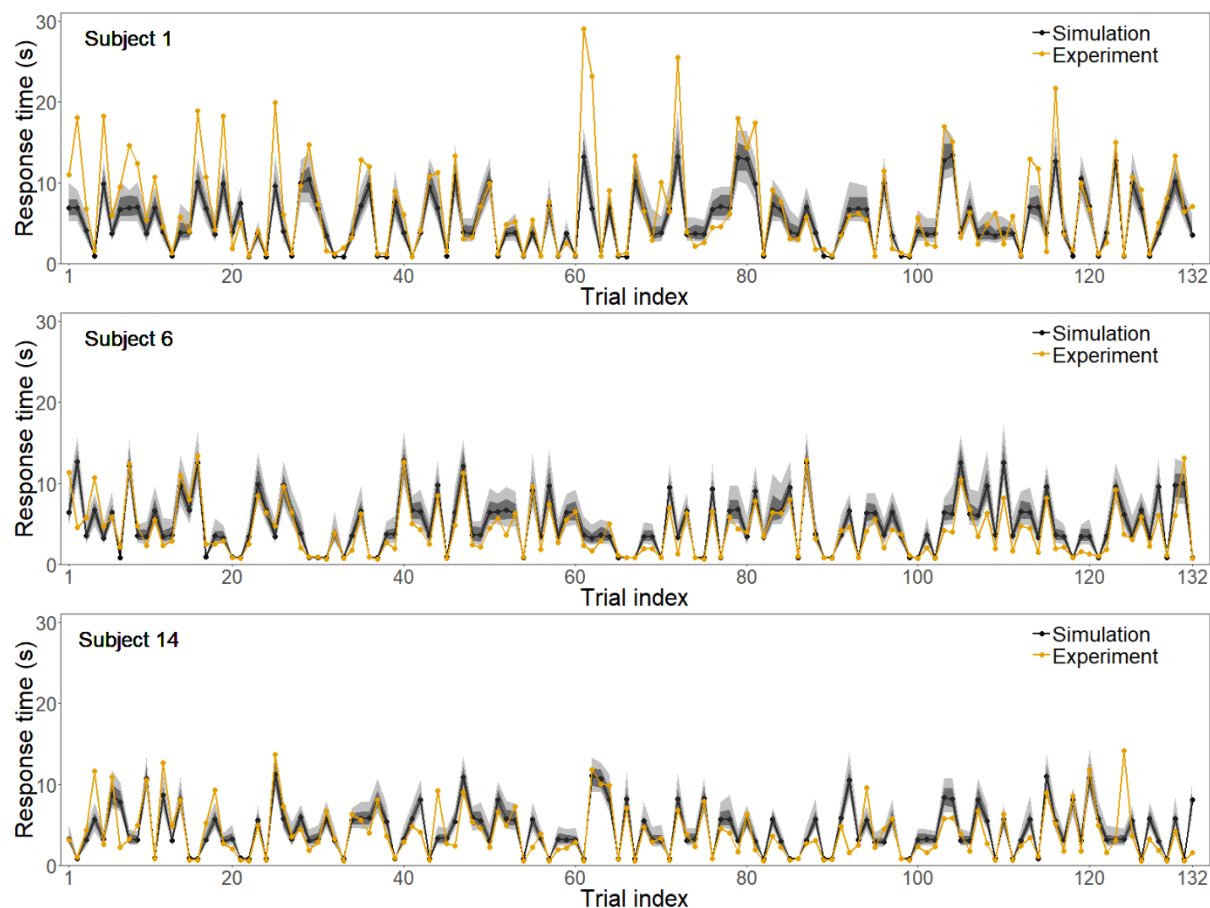
Figure 8. Distributional RT predictions for three participants (over 100 runs) of the ACT-R model of TTB and observed RTs of the corresponding participants in Khader et al. (2011, Exp. 1). The trial index is shown on the x-axis. That index can be thought of as timeline, with the first trial corresponding to the first and the last trial to the last comparison of two companies in the decision task. The black line represents the median predicted RTs across trials; the dark grey strip is the region between the 25th and 75th RT percentile; the light grey strip is the region between the 10th and 90th RT percentile. The observed RTs are presented in yellow. There were 132 paired comparisons of companies in total. See the online article for the color version of this figure.

The trial-by-trial RT predictions of participant 1 are summarized as a function of the number of attributes that need to be retrieved to make a decision and compared to experimental data in Figure 9[10] (see Supplementary Online Materials for such plots for all participants). When the same company name is presented on both sides of the screen and,

---

[10] There are few items, on which 3 attributes (16 items) or 4 attributes (8 items) need to be retrieved. As a consequence, we have grouped those items together to have a better representation of their distribution. Note that there are also only 8 items, on which no attributes need to be retrieved, but these exhibit a very small variability in response time and are thus not grouped with other items.

consequently, no attributes need to be retrieved, the model almost always responds within 1 s (Figure 9a; $median_{sim}$: 918 ms), while this subject more frequently needs between 1 and 2 s ($median_{exp}$: 1237 ms). Most participants (11 out of 17) also more frequently respond within 1 s on such trials. Moreover, as can be seen, for this participant the more attributes need to be retrieved (1 attribute: Figure 9b, $median_{sim}$: 3788 ms, $median_{exp}$: 3816 ms; 2 attributes: Figure 9c, $median_{sim}$: 6923 ms, $median_{exp}$: 7886 ms; 3 or 4 attributes: Figure 9d, $median_{sim}$: 11026 ms, $median_{exp}$: 14854 ms), the more likely it is that his/her RTs deviate from the model predictions. This trend can be seen among most subjects with some subjects' data aligning better with our predictions and some worse. We suspect that three factors contribute to this. First, sample sizes are smaller the more attributes need to be retrieved and, thus, each sample is more variable. Second, RT variability also increases with increasing RTs. Finally, the probability of not precisely following the prescribed strategy (e.g., by getting distracted or wrongly remembering an attribute value) increases the longer the execution of that strategy. Interestingly, some subjects exhibit very fast RTs (in some cases within 2 s) even on trials that require 3 or 4 attributes to be considered, which supports our last hypothesis. When developing ACT-R models, detecting such deviations by plotting predictions and data in different ways is important. This aids to gauge the overall performance of a model, and can uncover where further model refinements are warranted.
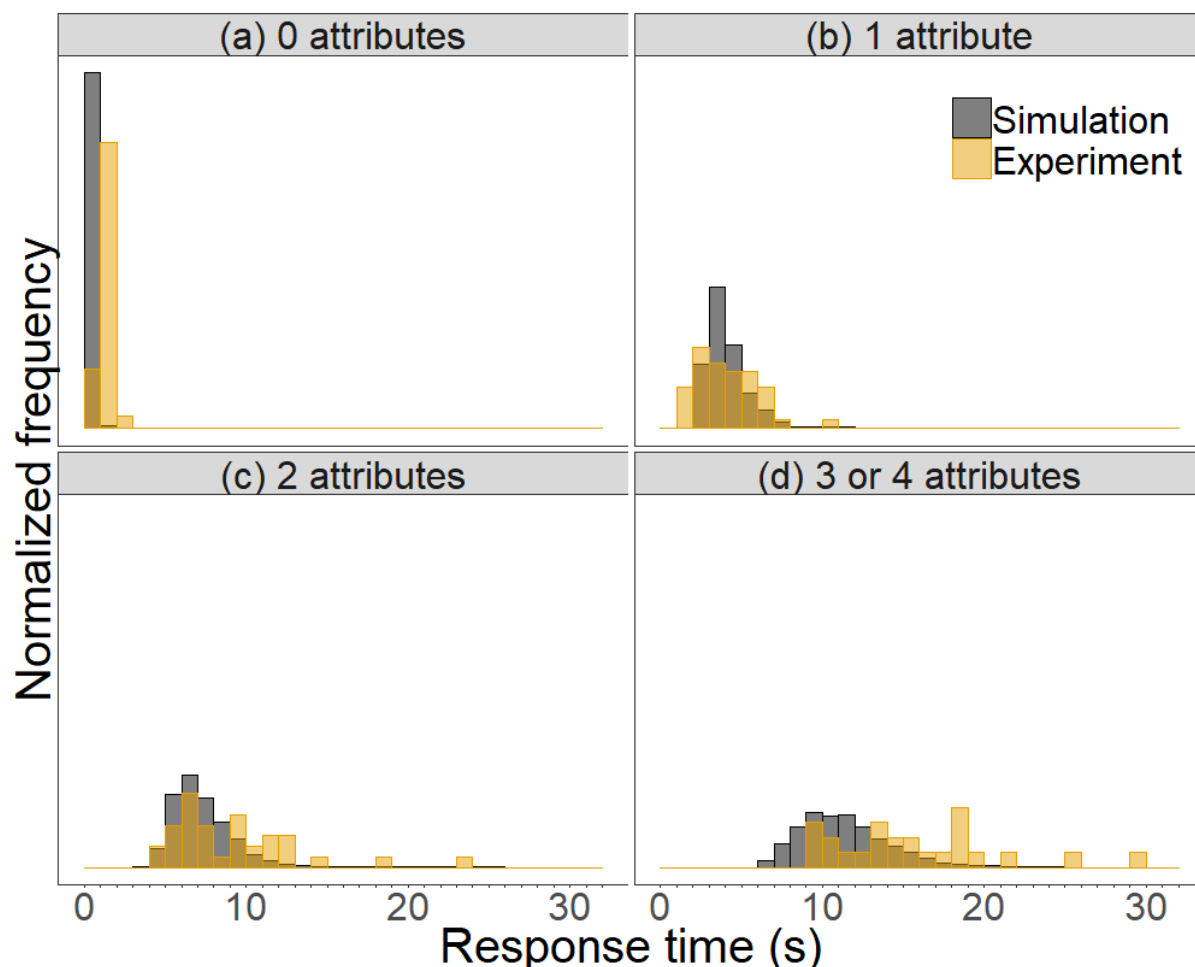
Figure 9. Comparison of predicted and observed RT distributions in the decision task as a function of the number of attributes considered for participant 1 from Khader et al. (2011, Exp. 1). Each count corresponds to a trial of that participant. See the online article for the color version of this figure.

To further illustrate this point and examine inter-participant variability, one can also ask how well our model performs across all participants. In Figure 10 we compare the extent to which the empirically observed RTs deviate from the median predicted RTs (in terms of mean absolute deviation, MAD) relative to how much the model deviates, on average, from median predicted RTs on individual runs. How much the model deviates on each individual run from its median RT predictions provides us with an estimate of model variability. If a participant shows a smaller deviation from median model predictions than individual model runs, our model overestimates the variability in RT. On the other hand, if a participant shows

a larger MAD, our model either underestimates a participant's RT variability or it systematically deviates from the participant's RT. Figure 10 demonstrates that participants and model exhibit similar variability, although the empirical RTs typically depart more from the median predictions than individual runs of the simulation. Moreover, as noted above, as a participant needs to retrieve more attributes (0 attribute: Figure 10a; 1 attribute: Figure 10b; 2 attributes: Figure 10c; 3 or 4 attributes: Figure 10d), his/her RTs depart more strongly from the predictions.
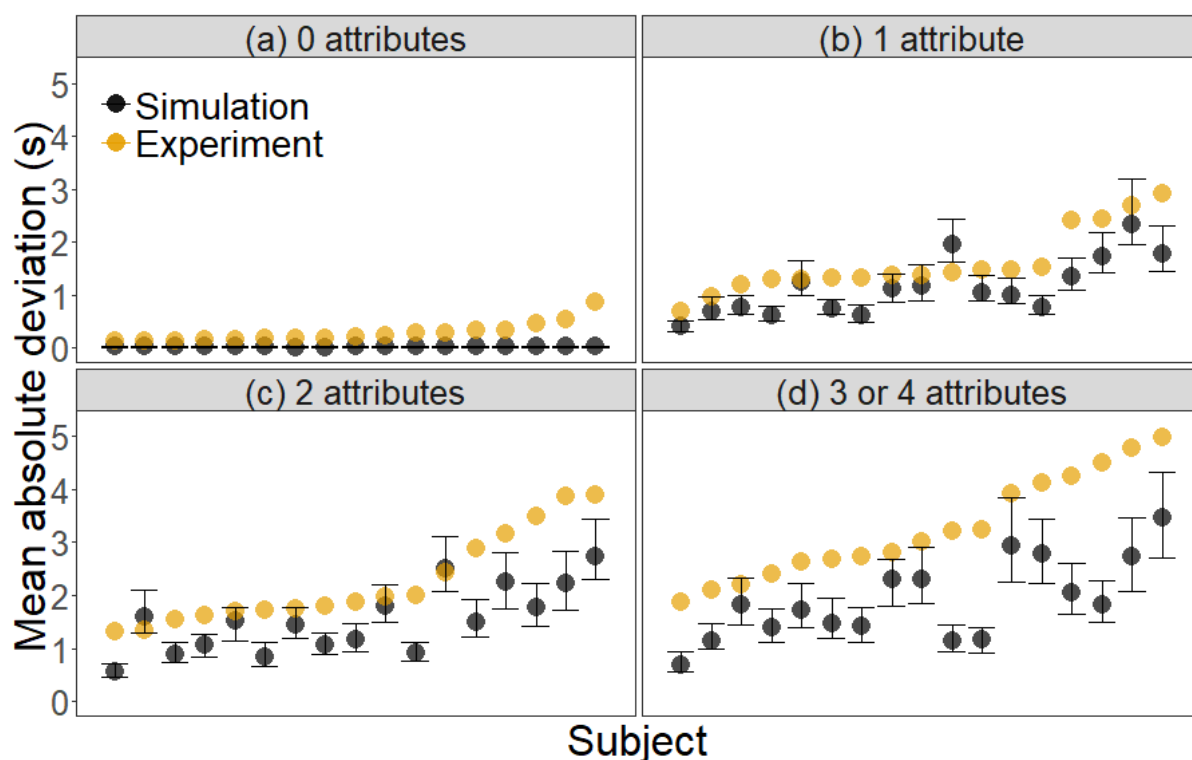

Figure 10. Comparison of MAD of individual model runs and of experimental data from median predicted RTs of model as a function of number of attributes retrieved before deciding. RT absolute deviations were averaged over participant and trial type (i.e., number of attributes that need to be recalled). Error bars represented minimum and maximum MAD from the 100 runs of the model.

Neural Activation Predictions

In Section 2, we explained how HR predictions are derived in ACT-R. Figure 11 illustrates, in practice, how the HR resulting from a module's activity emerges (Equation 11). In this figure, the activity of two modules (declarative and visual modules), as described by their demand functions (light color), is transformed into predicted BOLD responses (dark color) for the first 50 s of a model run for a participant. Following the same procedure as for these 50 s, we derive BOLD-response predictions for each of the 100 model runs of each participant for the entire decision task (Figure 4, step 5). Thus, we are effectively able to specify the expected pattern of brain activation related to each of ACT-R's modules, given the sequence of cognitive steps assumed by our model.

When engaging in such ACT-R modeling, the details matter: For instance, we model the repeated presentation of fixation crosses and ITIs in the experimental procedure (see Figure A1 in the Appendix, for a process trace of the complete final model). This is important for accurately generating BOLD predictions, because the repeated presence of fixation crosses and ITIs shape the time course of the HR. As can be seen in Figure 2, the HRF needs more than 20 s to settle back to its baseline level. Given that the fixation cross duration is 2 s and that the ITI is at most 6 s, there will always be some residual HR from the previous trial in the current trial (see the dashed lines in Figure 11, which represent the time point at which a new trial begins).
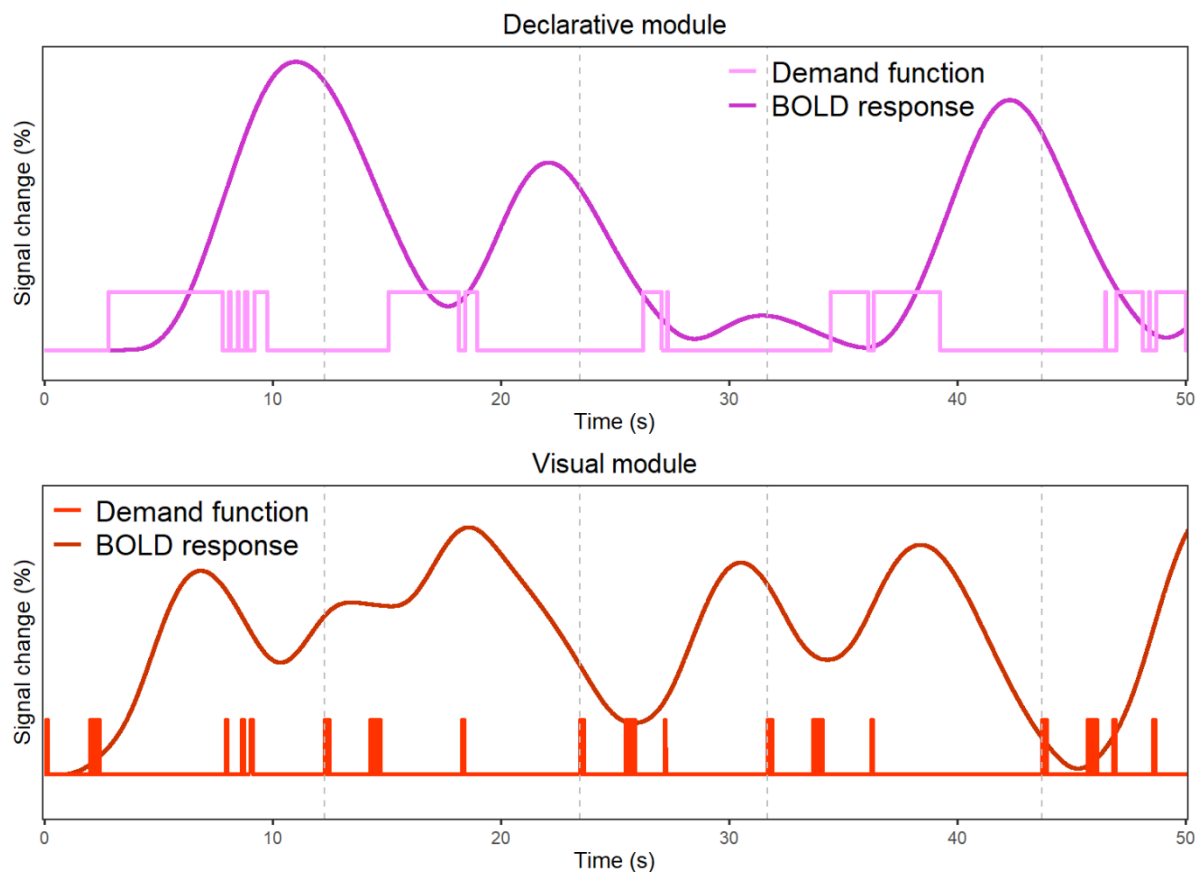
Figure 11. Transformation of visual and retrieval modules activity, as describe by the demand function, to HR for first 50 s of a run of an ACT-R implementation of TTB. The beginning of each trial is denoted with a dashed line. The HR often has not decayed to baseline before the beginning of the subsequent trial. See the online article for the color version of this figure.

Following the ROI procedure, these HR predictions, associated with different cognitive processes specified in the model were compared to the observed brain activity. BOLD signals were extracted from module-specific areas based on the center coordinates and ROI sizes provided in Anderson (2007).[11] To mimic an fMRI scan, model predictions were averaged every 2 s. Then, for both predictions and observations, the first scan of each trial for that participant (and that model run in the case of the predictions) served as baseline – BOLD response was estimated relative to its magnitude. Finally, both predicted and observed BOLD

---

[11] All coordinates are in Talairach space: declarative module: x = +/- 43, y = 23, z = 24; imaginal module: x = +/-23, y = -63, z = 40 ; procedural module: x = +/-14, y = 10, z = 7; manual module: x = +/-41, y = -20, z = 50; visual module: x = +/-42, y = -61, z = -9 (from Anderson, 2007, p. 189).

responses for all participants were grouped according to the number of attributes that have to be considered before a decision can be made and averaged over bins of 2 s. Figure 12 compares predictions for the five modules of interest[12] and observations from the corresponding regions in both the left and right hemisphere. The first column of Figure 12 plots recordings from the left hemisphere, the second from the right hemisphere, and the third column plots our model predictions.

---
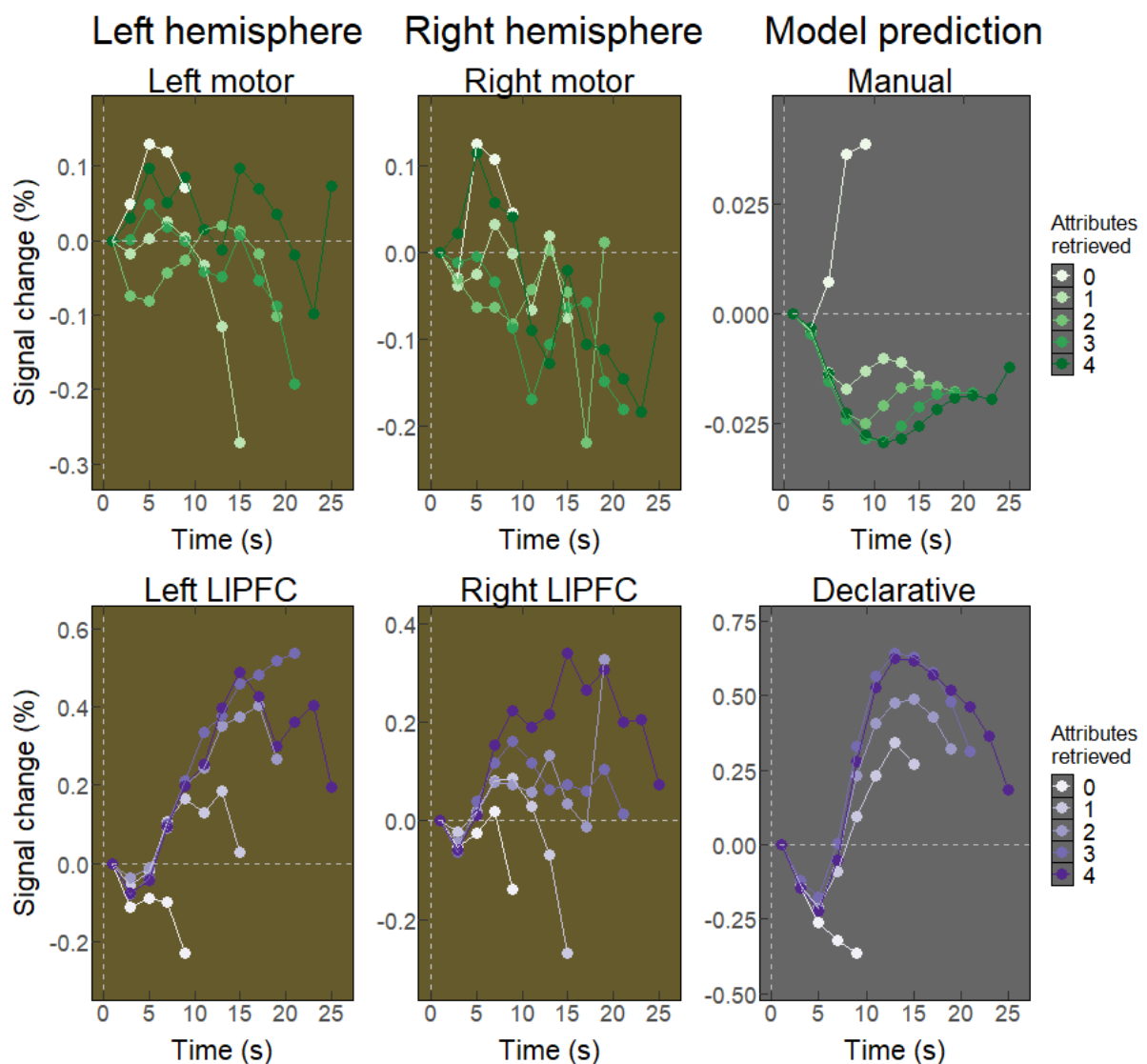
[12] ACT-R currently does not generate separate predictions for each hemisphere.
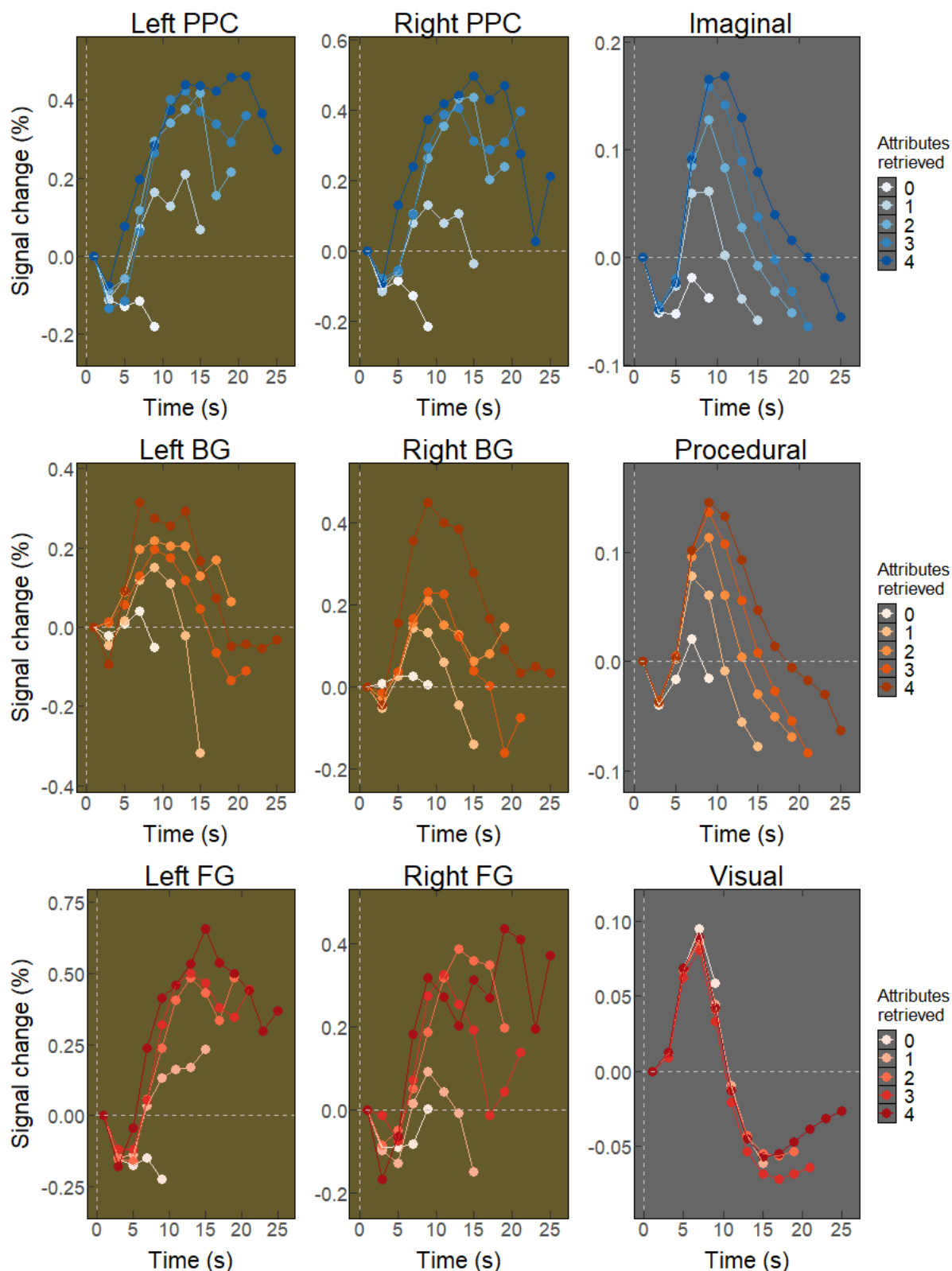
Figure 12. Observed and predicted BOLD responses for ten brain regions, associated with the manual, procedural, declarative, visual, and imaginal modules. *LIPFC* = Lateral Inferior Prefrontal Cortex. *PPC* = Posterior Parietal Cortex. *BG* = Basal Ganglia. *FG* = Fusiform Gyrus. The x-axis represents the time point in the trial at which the BOLD signal is measured

or predicted, while the y-axis represents the signal change relative the first scan of that trial. The different degrees of brightness are associated with a different number of attributes that need to be retrieved before TTB makes a decision. The BOLD response is averaged over all participants and trials that require that number of attributes to be retrieved. Only points averaged from least 20 observations are included in the plot, because the empirical data become very noisy with just a few observations. See the online article for the color version of this figure.

To quantify the degree of correspondence between predicted and observed fMRI patterns, we used the Tucker Congruence Coefficient (TCC)[13] and the coefficient of determination R^2 (similar to Borst et al., 2015). In addition, we used a weighted coefficient of determination R_w^2 (computed as the square of the weighted correlation between predictions and observations) which weights each point by the number of observations that were averaged to produce that point, whereby the averaging took place over participants and trial time. Table 1 compares predictions and observations on these three measures for the 10 brain regions of interest. The measures of correspondence between model and data are comparably good to work of others (e.g., the fit Borst et al., 2015, which established new mappings of modules to brain regions, shows TCCs in the range between .86 and .96 and R^2s in the range of .67 and .93). There were also two important deviations: First, the visual regions failed to match the observed increase in BOLD response amplitude with number of attributes considered. Second, the two motor regions correlated more weakly with the predicted BOLD response than the remaining regions.

When developing ACT-R models, such discrepancies need to be explained and resolved so that they can, potentially, inform future model refinements. For instance, the weaker match between predictions and observations for the manual module was probably due to low motor activity: Motor activity (i.e., pressing a key on the response device) was

---

[13] TCC is used to assess the similarity of two quantities. Just like the correlation coefficient, it ranges between -1 and 1. Unlike the correlation, TCC is based on the degree to which the quantities deviate from 0 as opposed to from their means.

necessary only once, at the end of each trial. This, in turn, probably resulted in a low signal

relative to the BOLD noise. We can probably do nothing to further improve the motor

predictions of our model. However, predictions based on the visual module and BOLD data

recorded in the corresponding brain region exhibited low correlation and negative TCC. The

negative TCC in particular means that the observed signal tends to follow the opposite sign

from our predictions: Whenever the observed signal is positive, our predictions tend to be

negative and vice versa. A visual inspection of Figure 12 further corroborates that our model

fails to predict activity in the visual region. This might be due to a possible mismatch

between our model and the sequence of cognitive processes that participants executed, calling

for a further refinement our model to better match participants' visual activity.

Table 1. *Comparison of empirical and model brain activity for the brain regions*

corresponding to the 5 modules active in our TTB implementation.

Region	TCC	R^2	R_w^2
Left manual	.26	.08	.14
Right manual	.62	.21	.33
Left visual	-.51	.48	.29
Right visual	-.33	.20	.18
Left procedural	.80	.61	.73
Right procedural	.84	.73	.80
Left imaginal	.72	.52	.75
Right imaginal	.75	.58	.76
Left declarative	.94	.86	.87
Right declarative	.66	.30	.40

Note. TCC = Tucker's Congruence Coefficient, R_w^2 = Weighted Coefficient of

Determination. R_w^2 is computed by squaring the weighted correlation between predictions

and observations.

Model Refinement

There are often multiple ways of translating a model into a detailed architectural implementation. This holds also true when translating TTB's search, stopping, and decision rules into ACT-R. Our first ACT-R implementation of TTB relied on a set of assumptions about how people would execute this strategy. Yet, the ROI analysis outlined above demonstrated that some of those assumptions may not hold. This provides us with an opportunity to further refine our model. What other ways are there to translate TTB into ACT-R?

Currently, our model maintains the company names in a short-term store (i.e., the imaginal buffer). Yet, a participant might avoid the burden of storing this information that is readily available on the computer screen by reading one or both company names off the screen upon the inspection of each attribute. Additionally, our model stores both attribute values in the imaginal buffer before comparing them. Yet, the second attribute value is immediately accessible after being recalled (i.e., in the retrieval buffer) and it might not be necessary to first move it to the imaginal buffer before comparing the attribute values. Finally, when we attempt to retrieve knowledge from long-term memory, we often shift our visual attention (i.e., we look away) and so then need to reallocate it back to its initial position (i.e., towards the screen), while for our model visual attention is always on the last object that it looked at.

We have created 7 additional models (Table 2) that vary according to these dimensions. TTB_1 is the model that we have examined up to here. TTB_2 reflects the intuition that people may access an attribute's value directly after they recalled it instead of first moving it to a short-term store. Further reducing the amount of information maintained in short-term memory, TTB_3 and TTB_4 read company names off the computer screen instead of storing them in the imaginal buffer. TTB_5 lies at the intersection between TTB_1/TTB_2 and

TTB_3/TTB_4, because it only maintains one company name in the imaginal buffer. Similarly, TTB_6 and TTB_7 are a combination of TTB_1/TTB_2 and TTB_3/TTB_4 as they not only look at both company names upon the inspection of each attribute, but also store those company names in the imaginal buffer. Finally, TTB_8 is the model that loses the attentional focus once a retrieval starts and so attention needs to be reallocated after retrieval is completed.

Table 2. *Summary descriptions of the 8 implementations of TTB.*

Assumed processing steps	TTB_1	TTB_2	TTB_3	TTB_4	TTB_5	TTB_6	TTB_7	TTB_8
Store company names in imaginal buffer	X	X				X	X	
Look back and forth upon inspecting each attribute			X	X		X	X	X
Store second attribute value in imaginal buffer	X		X			X		X
Store one company name in imaginal buffer					X			
Look at company name again after attribute is retrieved								X

We submitted these eight implementations of TTB to a competitive test against each other. While all models generated comparable behavioral predictions (see sections S2 & S3 in Supplementary Online Materials), their neural predictions (see section S4 in Supplementary Online Materials for figures) pointed towards the most plausible among the models. Table 3 reports goodness-of-fit measures for the neural predictions of our initial TTB implementation (TTB_1), the best-fitting implementation (TTB_8), and one with an intermediate fit (TTB_4).

These results illustrate the merits of an ROI analysis: Even though the behavioral measures produced by the models were very similar, the unique module-to-brain mappings afforded neural data to measure the contribution of each cognitive process separately. Put differently, variations in cognitive activity that lead to a noticeable change in a brain region's activity can only be identified with the help of neural data. Identifying the degree of activity

of each cognitive function then allows us to point at the cognitive processes that most likely generated observed BOLD signals. In our case, that is TTB$_8$, which assumes object names are not stored into the imaginal buffer. Instead, they are read off the computer screen every time they are needed. This model also assumes that visual attention is lost upon each retrieval attempt and then recovered once retrieval is completed.

Table 3. *Goodness-of-fit measures of 3 of the 8 TTB ACT-R implementations.*

Region	TTB$_1$			TTB$_4$			TTB$_8$		
	TCC	R^2	R_w^2	TCC	R^2	R_w^2	TCC	R^2	R_w^2
Left manual	.26	.08	.14	.26	.08	.15	.24	.07	.13
Right manual	.62	.21	.33	.63	.22	.34	.62	.22	.34
Left visual	**-.51**	**.48**	**.29**	**.08**	**.02**	**.01**	**.75**	**.45**	**.70**
Right visual	**-.33**	**.20**	**.18**	**.20**	**.00**	**.03**	**.73**	**.42**	**.66**
Left procedural	.80	.61	.73	.77	.56	.68	.68	.38	.55
Right procedural	.84	.73	.80	.84	.72	.80	.78	.54	.65
Left imaginal	.72	.52	.75	.55	.25	.54	.77	.56	.77
Right imaginal	.75	.58	.76	.60	.32	.56	.80	.61	.78
Left declarative	.94	.86	.87	.94	.85	.88	.95	.87	.89
Right declarative	.66	.30	.40	.66	.30	.41	.66	.29	.41

Note. Models differ primarily in how well they predict the signal in the fusiform gyrus. The winner of the model comparison is TTB$_8$ which predicts well brain activity related to the visual module. Neural predictions for the other 5 models can be found in section S4 in the Supplementary Online Materials.

At the same time, these results illustrate another important aspect of ACT-R modeling: Cumulative theory building (see Marewski & Olsson, 2008). Specifically, the ACT-R architecture has been cumulatively refined, updated, and extended over the past decades, based on thousands of data points from experimental research from all over the world (current version: ACT-R 7). If one develops isolated models (e.g., of decision making or other cognitive processes) in ACT-R, a similar cumulative processes of theory building can take place. Typically, the model code is publicly shared, allowing for models developed on one data set (or a series of data sets) to be re-used and tested by researchers from other

labs on other data sets. The next step of what can be thought of an iterative research process of model development and continuous testing would be to submit our TTB implementations to tests on new data sets. Conducting such tests would be particularly important for TTB_{2-8}, because those seven implementations emerged after the fact, as a result of a refinement of TTB_1. Hence, in a strict sense, Table 3 reports actual model predictions from foresight only for TTB_1, but not for the other implementations for which we adapted the model structure after the fact (although we did not re-estimate model parameters). The next competitive test of those models should therefore be conducted on new data.

Summary

We outlined how to develop an ACT-R model based on behavioral and neural data, using the TTB heuristic in memory-based decision making as a case in point. Importantly, the parameters of our model were constrained in a separate (i.e., learning) task, which as independent from and preceded the actual decision task (cf. Khader et al., 2011). The model was used to generate predictions about RTs and brain activation for a decision task, in which participants are instructed to use TTB to decide between two options, for which they had to retrieve decision-relevant information from memory. Overall, the RT predictions matched well both central tendencies and variability of individual participants' data. The brain activation predictions of the first model corresponded well to observed fMRI signals in regions associated with the manual, retrieval, imaginal, and procedural buffers, but failed to predict activity in the visual region. Pushed by this failure, we further refined our model by generating alternative hypotheses about the sequence of processing steps, implemented them as ACT-R models and selected the model whose predictions best corresponded to observed fMRI data.

6. General Discussion

Cognitive activities, such as decision making, are the result of various interplaying cognitive resources. When trying to understand such activities, researchers are faced with the challenge of separating the contribution of each cognitive capacity. In this article, we provide a step-by-step methodological primer of how to separate the contribution of each cognitive capacity by implemented models in ACT-R, and how such models can be tested with both behavioral and neural data, and further refined. As an illustrative example, we focused on a commonly studied model of decision making, the TTB heuristic (Gigerenzer & Goldstein, 1996). After estimating the free parameters of an ACT-R implementation of TTB using an independent learning task, we generated predictions about RTs and brain activation for a decision task, and tested those predictions on an fMRI dataset. Overall, both predicted RT distributions and temporal and spatial patterns of brain activity of our final model corresponded well to the observed data. Our results demonstrate that a properly specified and constrained model can predict both of these types of data, without any further adjustment of its parameters. As expected, if the components of that model are refined in an iterative process, the resulting models' fit can further increase. More broadly, our results illustrate how decision making research can be grounded in more general cognitive theories (Dimov, 2018; Dimov & Link, 2017; Dimov, Marewski, & Schooler, 2013, 2017; Dougherty et al., 1999; Fechner et al., 2016; Fechner, Schooler, & Pachur, 2018; Fechner, Pachur, & Schooler, in press; Gonzalez, Lerch, & Lebiere, 2003; Link, Marewski, & Schooler, 2016; Marewski & Mehlhorn, 2011; Marewski & Schooler, 2011; Schooler & Hertwig, 2005; Thomas, Dougherty, Sprenger, & Harbison, 2008). In fact, this approach might become a trend that, so we and others think, has the potential to ultimately revolutionize the field – once entry-level barriers to complex architectural modeling tools such as ACT-R break away.

In what follows, we (a) compare how our models would do in predicting participants'
RT data if we used a common set of parameters for all participants, (b) explicitly outline the
four modeling principles that we followed in our model developing efforts, (c) refer readers
to material relevant to ACT-R.

Individual vs. Group Parameters

Two key features of our approach are, first, that we generate predictions (in our case,
in the decision task) by estimating model parameters (i.e., mean activation, activation noise
and retrieval threshold) on a different task from the task of interest (a learning task) and,
second, that we do so separately for each participant. Constraining parameters in different
tasks lends credence to architectural models (Newell, 1973) and, more generally, reducing the
number of free parameters is a widely accepted good practice in psychology. Thus, we call
for fitting free parameters on tasks separate from the main task whenever possible. Yet, how
important is the estimation of individual parameter values for participants?

At least since Estes (1956) it is known that artifacts in parameter estimation can
emerge if model parameters are fitted to group data. For example, only when fitting to
individual data Estes and Maddox (2005) could recover reasonable parameter values. Others
argue that averaged data can change the underlying functional form, for example, from
exponential at the individual level to a power function at the group level (Anderson &
Tweney, 1997; Heathcore, Brown, & Mewhort, 2000; Myung, Kim, & Pitt, 2000). In our
case, we will address this question by estimating common parameters from the learning data
of all participants and generate new predictions with our model. To this end, we pool the RTs
of all participants from the last round of the learning task, remove outliers, and fit memory
parameters to those.

Figure 13 demonstrates the group parameter fit and compares them to the individual
parameter estimates. Not surprisingly, the common parameters fall in-between the individual

parameters: While individual mean activations range between -0.55 and 1.02, the common

mean activation is 0.26; while estimated retrieval noise ranges between 0.3 and 0.88,
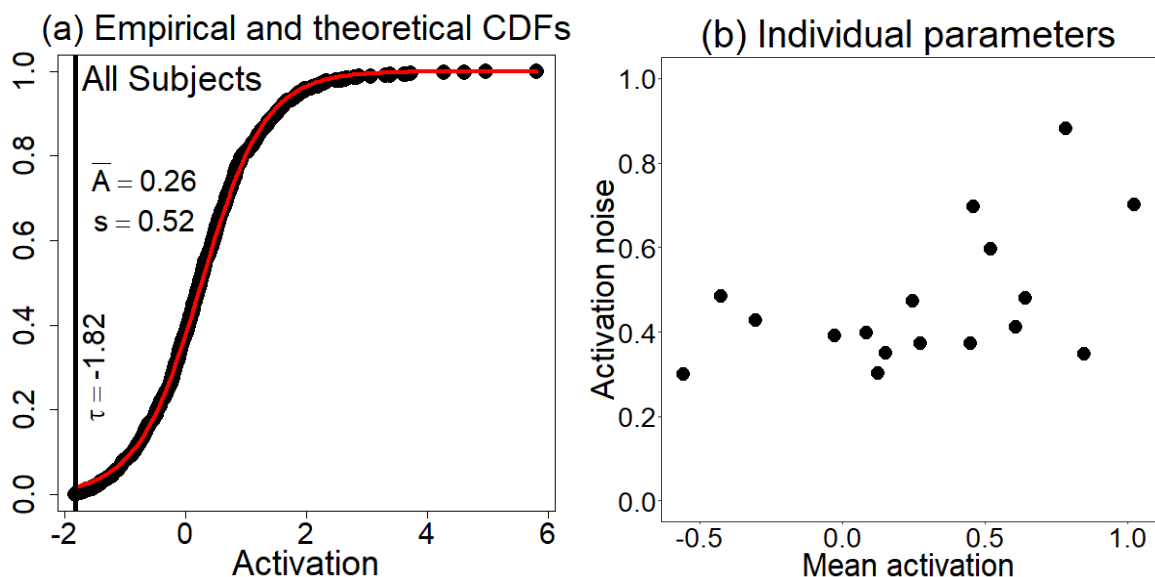
common retrieval noise is 0.52.

Figure 13. (a) Activation fit of all observed RTs in learning task to generate common memory parameters for all participants. $\bar{A}$ denotes estimated mean activation, $\underline{s}$ the retrieval noise and τ denotes the retrieval threshold. (b) Individual parameter estimates of mean activation and activation noise.

It can be expected that for participants with parameters close to the mean parameters

it would not matter much whether we use the individual or group parameters, while for

participants with parameters at the extreme it will matter more. Figure 14 shows how much

participants' RTs deviate from model predictions with common vs. individual parameters.

For more than half of the participants, it does not matter which parameters we use. For one

participant (the right-most on the graph), the common parameters do better than the

individual parameters that we estimated. Finally, for 5 of the 17 participants (5 of the 6

leftmost) using individual parameters improves the predictive power of our models. The

predictions for those participants are roughly between 0.5 s and 1.7 s more accurate with the individual than with the common parameters.
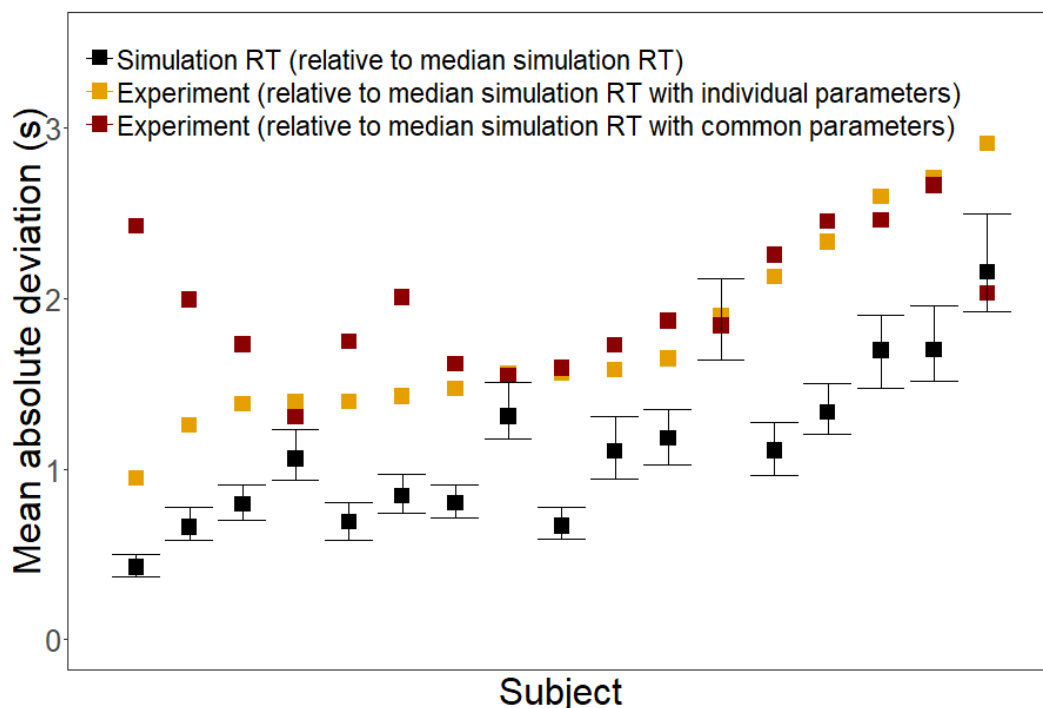

Figure 14. RT deviation from median predicted RTs for model TTB_1. With black is plotted MAD of individual runs of TTB_1 (with memory parameters fitted separately for each participant) from median RTs of that model. The MAD of each participant's RTs from median predicted RTs of TTB_1 (with memory parameters fitted separately for each participant) is plotted in yellow. Red plots the MAD of each participant's RTs from median predicted RTs of TTB_1 (with memory parameters fitted on all participant's data from the learning task). The black dots estimate how much the simulation varies from run to run. Participants' RTs (in yellow) typically deviate more from median model predictions than any individual model run (see black error bars, which delineate the region between 10th and 90th percentile of RTs on each model run). Moreover, participants' RTs deviate more for some participants from a model with a common set of memory parameters (in red) than from a model with individual memory parameters (in yellow).

Which procedure – individual or average parameters – should be adopted when developing ACT-R models? In general, we recommend working with individualized parameter values, because people differ in terms of their ability to, for instance, remember information. People also differ in terms of the statistical structure of real-world environments

they have encountered. For instance, if our experimental data set had not entailed fictitious, but realistic stimuli names, one might expect that people with different prior real-world exposure histories to company names would have different activation levels for the same company names (see section 2 on ACT-R's subsymbolic system). Note that computing individualized parameter values does not lead to the problem of overfitting if, as we have done, the resulting models are tested with fixed parameter values on a different task for each participant. Pooling data, in turn, can became an interesting option when the number of observations available for each individual participant is sparse and when one has reasons to assume that individual differences across participants are negligible. That said, when working with pooled data the modeler should be aware of the problem that the average parameter values used might not correspond to any individual and also of the artifacts that might arise from estimating parameters from group data.

Individual differences in decision strategy

While individual parameter values lead to predictions which are at least as good as a common set of parameters for all participants, our model still does not capture all the variance in participants' data. This might be partially due to the limited size of the learning data sample, which does not allow us to perfectly estimate the values of memory parameter. Other, human, factors such as accumulating fatigue or distraction might also play a role. Yet another possibility for the unaccounted variance might lie in inter-participant variability in strategy selection. In decision making, it is well known that participants vary widely in the strategies they adopt when facing a decision problem and results typically consist of tendencies of switching from one strategy to another as certain experimental factors are manipulated rather than a unanimous adoption of a single decision strategy (see, e.g., Bröder, 2012). Even in a task as the one that we are modeling, in which participants are instructed to follow a particular strategy, each participant might still be executing TTB in their specific

47

manner. In this case, different participants would be best described by different TTB implementations.

Another potential source of variability that we do not include in our analysis is that potential for strategy switching as the experiment progresses. Factors such as fatigue, exploration or reinforcement (Rieskamp & Otto, 2006) might slowly increase the likelihood that one implementation of TTB is executing over another. For example, a participant might start by storing all information in the short-term store (i.e., the imaginal module) similar to TTB_1 and then discover that it is more efficient to not store information readily available to visual attention, as in TTB_8. Exploring such individual differences is another potential pathway of model refining.

Principles for Model Development and Testing in Cognitive Architectures

Modeling frameworks as complex as ACT-R, which consists of multiple interplaying components and many adjustable parameters, are sometimes criticized for being capable of fitting everything (Pohl, 2011; Rieskamp & Otto, 2006). Yet, such critique is misleading. ACT-R models make extremely precise multi-dimensional predictions (i.e., about different variables, including RT, BOLD signals, eye-movements, etc.) that can be easily proved wrong in experiments. Our analysis of our ACT-R implementations of TTB serve to illustrate, by means of a practical example, methods for creating strong tests beds for architectural process models (see Figure 4 for a roadmap of the steps that we followed). Specifically, we followed four modeling principles (Dimov & Marewski, 2018).

First, while the ACT-R cognitive architecture has a certain number of free parameters, the calibration of those parameters is neither arbitrary nor is the careful modeler fully free to "choose" his/her parameter values. A common standard in the ACT-R community is to use, wherever possible, ACT-R's default parameter values. If parameters need to be estimated, those parameters ought to be constrained on different data sets and tasks (see A. Newell,

1990; see also Anderson, 2007). In developing our model, we used the specified default parameter values if feasible and calibrated all models' free declarative parameters on the learning task, and then carried over those parameter values unchanged to the decision task (Figure 4, steps 1 & 2).

Second, ACT-R models ought to be tested in *predicting* new, unseen data rather than in *fitting* existing data. Fitting refers to situations where a model's free parameters are estimated from the human data by minimizing the difference between the human and the model. In contrast, testing a model's predictions entails evaluating a model's ability to reproduce human data to which the model has not been *calibrated*, that is, out-of-sample with fixed parameters (Marewski & Olsson, 2008; Pitt, Myung, & Zhang, 2002; Roberts & Pashler, 2000). After having developed our ACT-R implementation of TTB and calibrated its memory parameters on an independent behavioral task (see Figure 4, steps 1 & 2), we predicted, out of sample, participants' decisions, RTs, and neural patterns in the decision task (Figure 4, steps 3, 4 & 6).

Third, ACT-R models allow making detailed quantitative predictions about distributions of human data, rather than merely predicting means, medians or other point estimates (for a related approach, see Smith & Ratcliff, 2004). Predicting the complexities of multiple distributions creates very strong tests for models, and often even merely trying to fit (rather than to predict) those distributions, poses a serious challenge. For example, Marewski and Mehlhorn (2011) specified 39 ACT-R models of decision making with none of them being able to fit and predict human RT distributions perfectly in two experiments. Yet, if only median RTs would have served as a criterion for model selection, several of those models would have been wrongly judged as being able to account equally well for human data. In our current analyses, we used our ACT-R implementation of TTB to predict distributional patterns (Figure 4, step 4).

Fourth, models are not evaluated in isolation, but in their ability to make predictions, out of sample, *relative to each other*. In such model comparisons, models' predictions about dynamic distributional data, and not just means, should be contrasted. This way, we might discover that no model accounts perfectly for both central tendencies and variabilities of various types of data, but, instead, we can establish the degree to which one model better predicts the data than another. The best model can then serve as a benchmark in future evaluations.

ACT-R Resources

In closing, we provide resources to get the reader started with ACT-R. First, the software package for the appropriate operating system can be downloaded from the ACT-R website (http://act-r.psy.cmu.edu/software/). In addition to ACT-R, that website contains tutorials, a reference guide and other documentation, which can aid the beginner modeler in immersing into the theory. Moreover, the website contains additional resources, such as a list of publications related to ACT-R and a list of researches working with this architecture. Publications cover a broad range of topics and usually come together with ACT-R models, which can be freely downloaded. For those willing to experiment with ACT-R without engaging with Common Lisp, versions in Java (jACT-R, http://jact-r.org/; Java ACT-R, http://cog.cs.drexel.edu/act-r/) and Python (pyact, https://github.com/jakdot/pyactr/; Python ACT-R, https://sites.google.com/site/pythonactr/) exist, while the related *ACTransfer* theory and software (Taatgen, 2013) is written in Swift (to be downloaded from https://github.com/ntaatgen/ACTransfer). However, the most complete and up-to-date version of ACT-R is the one developed in Lisp and, since version 7.5, ACT-R includes a new interface that allows interacting with the architecture through any programming langue. Additionally, we refer those interested in modeling neural data with ACT-R to the excellent

tutorial provided by Borst and Anderson (2017). Finally, we recommend future (ACT-R) modelers to seek out the literature on methods for model selection. A short overview (pertinent to ACT-R) is provided in Marewski and Olsson (2008).

Author note

References

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Psychology Press.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341-380. doi: 10.1006/jmla.1997.2553

Anderson, J. R., Carter, C. S., Fincham, J. M., Qin, Y., Ravizza, S. M., & Rosenberg-Lee, M. (2008). Using fMRI to test models of complex cognition. *Cognitive Science, 32*, 1323-1348. doi:10.1080/03640210802451588

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences, 12*, 136-143. doi:10.1016/j.tics.2008.01.006

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396–408. doi:10.1111/j.1467-9280.1991.tb00174.x

Anderson, R. B., & Tweney, R. D. (1997). Artifactual power curves in forgetting. *Memory & Cognition, 25*, 724-730. doi:10.3758/BF03211315

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 107-129. doi:10.1037/0278-7393.33.1.107

Bobadilla-Suarez, S., & Love, B. C. (2017). Fast or frugal, but not both: Decision heuristics under time pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi:10.1037/xlm0000419

Borst, J. P., & Anderson, J. R. (2015). Using the ACT-R Cognitive Architecture in combination with fMRI data. In B. Forstmann & E.-J. Wagenmakers (Eds.), *An*

Introduction to Model-based Cognitive Neuroscience (pp. 339-352). New York: Springer.

Borst, J. P., & Anderson, J. R. (2017). A step-by-step tutorial on using the cognitive architecture ACT-R in combination with fMRI data. *Journal of Mathematical Psychology, 76*, 94-103. doi:10.1016/j.jmp.2016.05.005

Borst, J. P., Nijboer, M., Taatgen, N. A., van Rijn, H., & Anderson, J. R. (2015). Using data-driven model-brain mappings to constrain formal models of cognition. *PLoS One, 10*, e0119673. doi:10.1371/journal.pone.0119673

Borst, J. P., Taatgen, N. A., Stocco, A., & Van Rijn, H. (2010). The neural correlates of problem states: Testing fMRI predictions of a computational model of multitasking. *PLoS One, 5*, e12966. doi:10.1371/journal.pone.0012966

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience, 16*, 4207-4221. doi:10.1523/JNEUROSCI.16-13-04207.1996

Bröder, A. (2000). Assessing the empirical validity of the" Take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1332-1346. doi:10.1037/0278-7393.26.5.1332

Bröder, A. (2003). Decision making with the" adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 611-625. doi:10.1037/0278-7393.29.4.611

Bröder, A. (2012). The quest for take the best - Insights and outlooks from experimental research. In P. Todd, G. Gigerenzer, & the ABC Research Group, *Ecological rationality: Intelligence in the world* (pp. 216-240), New York: Oxford University Press.

Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review, 14*, 895-900. doi:10.3758/BF03194118

Bröder, A., & Schiffer, S. (2003). Take The Best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General, 132*, 277-293. doi:10.1037/0096-3445.132.2.277

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108*, 847-869. doi:10.1037/0033-295X.108.4.847

Byrne, M. D., & Kirlik, A. (2005). Using computational cognitive modeling to diagnose possible sources of aviation error. *The International Journal of Aviation Psychology, 15*, 135-155. doi:10.1207/s15327108ijap1502_2

Chase, V. M., Hertwig, R., & Gigerenzer, G. (1998). Visions of rationality. *Trends in Cognitive Sciences, 2*, 206-214. doi:10.1016/j.tics.2004.11.005

Dimov, C. M. (2018). How to implement HyGene into ACT-R. *Journal of Cognitive Psychology, 30*, 163-176. doi:10.1080/20445911.2017.1394863

Dimov, C. M., & Link, D. (2017). Do people order cues by retrieval fluency when making probabilistic inferences? *Journal of Behavioral Decision Making*, *4*, 843-854. doi:10.1002/bdm.

Dimov, C. M., & Marewski, J. N. (2018). Cognitive Architectures as Scaffolding for Risky Choice Models. In M. Raue, E. Lermer, B. Streicher (Eds.), *Psychological Perspectives on risk and risk analysis* (pp. 201-216). Springer, Cham.

Dimov, C. M., Marewski, J. N., & Schooler, L. J. (2013). Constraining ACT-R models of decision strategies: An experimental paradigm. In M. Knauff, M. Pauen, N. Sebanz,

& I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2201–2206). Austin, TX: Cognitive Science Society.

Dimov, C. M., Marewski, J. N., & Schooler, L. J. (2017). Architectural process models of decision making: Towards a model database. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 1931-1936). Austin, TX: Cognitive Science Society.

Dougherty, M. R. P., Franco-Watkins, A. N., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, *115*, 199-213. doi:10.1037/0033-295X.115.1.199

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). Minerva-DM: A memory processes model for judgments of likelihood. *Psychological Review, 106*, 180– 209. doi:10.1037/0033-295X.106.1.180

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers, 34*, 455-470. doi:10.3758/BF03195475

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*, 380-417. doi:10.1037/h0053870

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition.* Oxford University Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251. doi:10.1037/0033-295X.87.3.215

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin, 53*, 134-140. doi:10.1037/h0045156

Estes, W. K., & Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychonomic Bulletin & Review, 12*, 403-408. doi:10.3758/BF03193784

Fechner, H. B., Pachur, T., Schooler, L. J., Mehlhorn, K., Battal, C., Volz, K. G., & Borst, J. P. (2016). Strategies for memory-based decision making: Modeling behavioral and neural signatures within a cognitive architecture. *Cognition, 157*, 77-99. doi:10.1016/j.cognition.2016.08.011

Fechner, H. B., Pachur, T., & Schooler, L. J. (in press). How does aging impact decision making? The contribution of cognitive decline and strategic compensation revealed in a cognitive architecture. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Fechner, H. B., Schooler, L. J., & Pachur, T. (2018). Cognitive costs of decision-making strategies: A resource demand decomposition analysis with a cognitive architecture. *Cognition, 170*, 102-122. doi:10.1016/j.cognition.2017.09.003

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A.P., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterising differential responses. *NeuroImage, 7*, 30 – 40. doi:10.1006/nimg.1997.0306

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *104*, 650-669. doi:10.1037/0033-295X.103.4.650

Glöckner, A. (2009). Investigating intuitive and deliberate processes statistically: The multiple-measure maximum likelihood strategy classification method. *Judgment and Decision Making, 4*, 186-199.

Gluck, K. A. (2010). Cognitive architectures for human factors in aviation. In E. Salas & D. Maurino (Eds.) *Human Factors in Aviation, 2ⁿᵈ Edition* (pp. 375–400). New York, NY: Elsevier.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27*, 591-635. doi:10.1016/S0364-0213(03)00031-4

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*, 185-207. doi:10.3758/BF03212979

Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review, 115*, 263-272. doi:10.1037/0033-295X.115.1.263

Juslin, P., Jones, S., Olsson, H., & Winman, A. (2003). Cue abstraction and exemplar memory in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 924-941. doi:10.1037/0278-7393.29.5.924

Khader, P H., Pachur, T., Meier, S., Bien, S., Jost, K., & Rösler, F. (2011). Memory-based decision making with heuristics involves increased activation of decision-relevant memory representations. *Journal of Cognitive Neuroscience, 23*, 3540-3554. doi:10.1162/jocn_a_00059

Khader, P. H., Pachur, T., & Jost, K. (2013). Automatic activation of attribute knowledge in heuristic inference from memory. *Psychonomic Bulletin and Review, 20* (2), 372-377, doi:10.3758/s13423-012-0334-7

Lewandowsky, S. (1993). The rewards and hazards of computer simulations. *Psychological Science, 4*, 236–243. doi:10.1111/j.1467-9280.1993.tb00267.x

Link, D., Marewski, J. N, & Schooler, L. J. (2016). An ecological model of memory and inferences. In A. Papafragou, D. Grodner, D. Mirman, & J.C. Trueswell (Eds.),

Proceedings of the 38th Annual Conference of the Cognitive Science Society (pp. 1883-1888). Austin, TX: Cognitive Science Society.

Marewski, J. N. & Mehlhorn, K. (2011). Using the ACT-R architecture to specify 39 quantitative process models of decision making. *Judgment and Decision Making*, *6*, 439-519.

Marewski, J. N., & Olsson, H. (2009). Beyond the null ritual: Formal modeling of psychological processes. *Zeitschrift für Psychologie/Journal of Psychology, 217*, 49-60. doi:10.1027/0044-3409.217.1.49

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review*, *118*, 393-437. doi:10.1037/a0024143

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*, 749-791. doi:10.1037/0033-295X.104.4.749

Myung, I. J., Kim, C., & Pitt, M. A. (2000). Toward an explanation of the power law artifact: Insights from response surface analysis. *Memory & Cognition, 28*, 832-840. doi:10.3758/BF03198418

Newell, A. (1973). Production systems: Models of control structures. In Chase, W. G. (Ed.)*, Visual information processing* (pp. 463-526). New York: Academic Press.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–310). New York: Academic Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, B. R., & Lee, M. D. (2011). The right tool for the job? Comparing an evidence

accumulation and a naive strategy selection model of decision making. *Journal of*

Behavioral Decision Making, 24, 456-481. doi:10.1002/bdm.703

Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal

heuristic: Not everyone "takes-the-best". *Organizational Behavior and Human*

Decision Processes, 91, 82-96. doi:10.1016/S0749-5978(02)00525-3

Nosofsky, R. M., & Bergert, F. B. (2007). Limitations of exemplar models of multi-attribute

probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, 33, 999-1019. doi:10.1037/0278-7393.33.6.999

Pachur, T., & Aebi-Forrer, E. (2013). Selection of decision strategies after conscious and

unconscious thought. *Journal of Behavioral Decision Making, 26*, 477-488.

doi:10.1002/bdm.1780

Pachur, T., & Bröder, A. (2013). Judgment: A cognitive processing perspective. *Wiley*

Interdisciplinary Reviews: Cognitive Science, 4, 665-681. doi:10.1002/wcs.1259

Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process

predictions of models of risky choice: A quantitative model comparison approach.

Frontiers in Psychology, 4. doi:10.3389/fpsyg.2013.00646

Pachur, T., & Marinello, G. (2013). Expert intuitions: How to model the decision strategies of

airport customs officers? *Acta Psychologica, 144*, 97-103.

doi:10.1016/j.actpsy.2013.05.003

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New

York: Cambridge University Press.

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among

computational models of cognition. *Psychological Review, 109*, 472-491.

doi:10.1037/0033-295X.109.3.472

Pohl, R. F. (2011). On the use of recognition in inferential decision making: An overview of the debate. *Judgment and Decision Making, 6*, 423-438.

Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207-236. doi:10.1037/0096-3445.135.2.207

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review, 14*, 249-255. doi:10.3758/BF03194060

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107*, 358-367. doi:10.1037/0033-295X.107.2.358

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological Review, 115*, 101-130. doi:10.1037/0033-295X.115.1.101

Schooler, L. J. & Anderson, J. R. (1997). The role of processes in the rational analysis of memory. *Cognitive Psychology, 32*, 219 – 250. doi:10.1006/cogp.1997.0652

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112*, 610-628. doi:10.1037/0033-295X.112.3.610

Schulte-Mecklenbeck, M., Kühberger, A., Johnson, J. G., & Ranyard, R. (Eds.). (2011). *A handbook of process tracing methods for decision research: A critical review and user's guide*. Psychology Press.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*, 161-168. doi:10.1016/j.tins.2004.01.006

Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review, 117*, 541-574. doi:10.1037/a0019077

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review*, 120, 439-471. doi:10.1037/a0033138

Taatgen, N. A., Huss, D., Dickison, D., & Anderson, J. R. (2008). The acquisition of robust and flexible cognitive skills. *Journal of Experimental Psychology: General*, *137*, 548-565. doi:10.1037/0096-3445.137.3.548

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review, 115*, 155-185. doi:10.1037/0033-295X.115.1.155

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299. doi:10.1037/h0032955

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts

Willemsen, M. C., & Johnson, E. J. (2011). Visiting the decision factory: Observing cognition with MouselabWEB and other information acquisition methods. In M. Schulte-Mecklenbeck, A., Kühberger, A. & R. Ranyard (Eds.). *A handbook of process tracing methods for decision research: A critical review and user's guide* (pp. 21-42). New York: Taylor & Francis

Appendix

Figure A1 is a description of a run of the final model, which extended the model described on Figure 3 by adding an inter-trial interval and a presentation cross. The presentation cross leads to additional visual activity when the model attends to the presented cross, while the inter-trial interval is a time of inactivity, during which the BOLD signal can decay. Thus, adding these two components to our model modifies its BOLD predictions. To simplify the understanding of the extended model, we have added screen events, which depict what is happening on the screen at this point in time, and mental events, which are higher level descriptions of what the model does at this point in time. For example, 2000 ms after the trial starts, the two company names are presented on the screen (see screen event). At this point, the model starts encoding the two company names off the screen (mental events "read company 1" and "read company 2") and, after determining that the names are different, starts executing TTB (mental event "names different; start TTB").

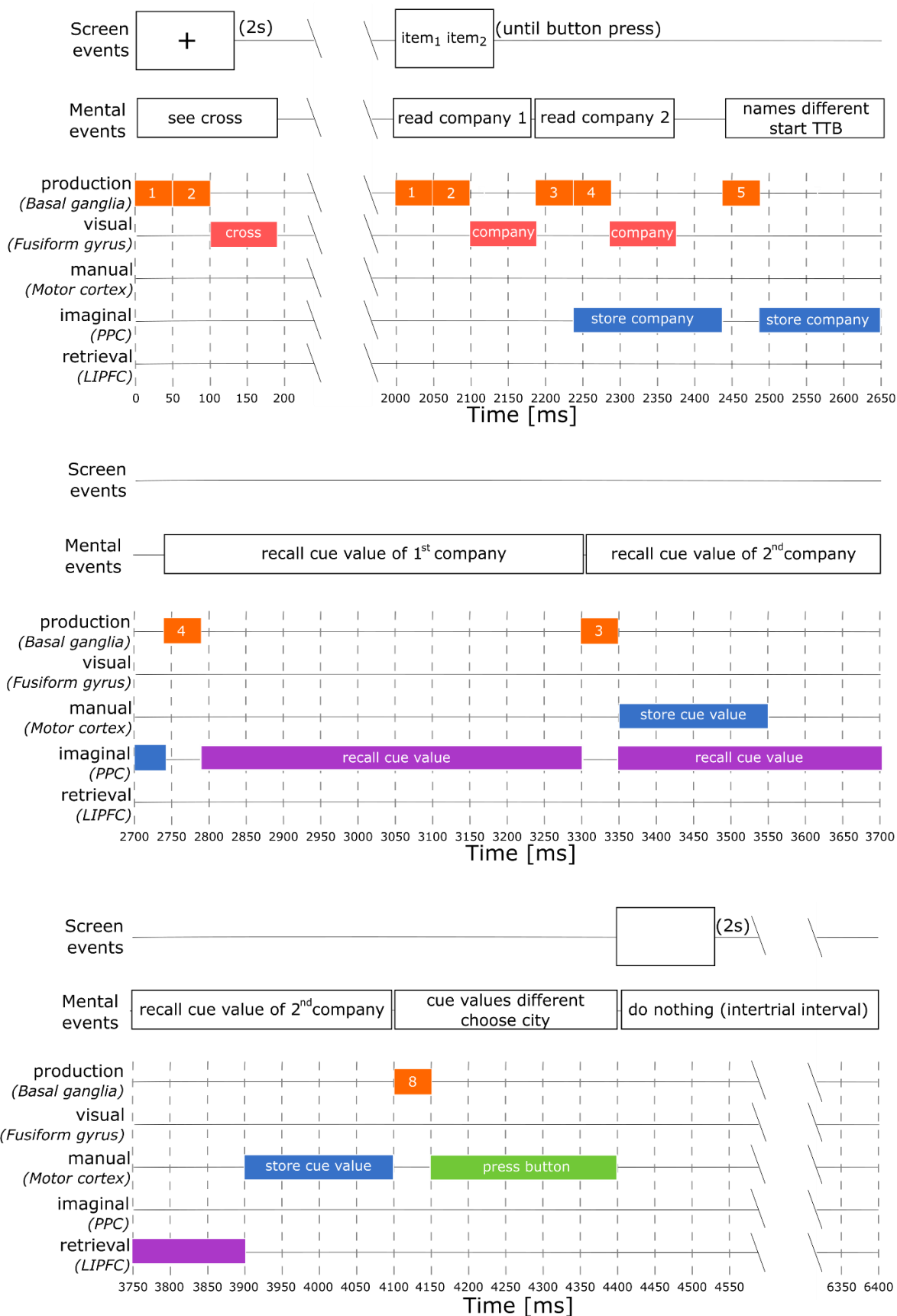*Figure A1*. Schematic process trace of an ACT-R implementation of TTB as applied to this

experiment. The y-axis denotes various ACT-R modules and associated brain regions. *LIPFC* = Lateral Inferior Prefrontal Cortex. *PPC* = Posterior Parietal Cortex. See the online article for the color version of this figure.