An Online Database of ACT-R Parameters: Towards a Transparent Community-based Approach to Model Development

Tsunhin John Wong[1] (jwong@mpib-berlin.mpg.de)

Edward T. Cokely[1,2] (cokely@mpib-berlin.mpg.de)

Lael J. Schooler[1] (schooler@mpib-berlin.mpg.de)

Max Planck Institute for Human Development, Lentzeallee 94 14195 Berlin, Germany[1]
Michigan Technological University, Cognitive & Learning Sciences, Townsend Drive, Houghton, MI 49931[2]

Abstract

We present a database that provides an interface for the ACT-R modeling community to interact with each other (http://www-abc.mpib-berlin.mpg.de/actrdb/). The database includes estimated values of ACT-R parameters from a wide range of ACT-R modeling studies, selected from the studies available on the ACT-R website. It serves as a tool to query studies and estimated values for ACT-R parameters, providing the exact range of values for each of the available free numerical parameters. In short, the database supports an alternative community-based approach to manage the challenges associated with parameter estimation for complex cognitive architectures like ACT-R.

Keywords: ACT-R; modeling; parameter.

Managing Parameters for ACT-R Models

Unified theories of cognition allow us to approach mechanisms of human cognition in a holistic, cumulative manner (Simon & Newell, 1973). Among the existing unified theories of cognition, ACT-R is one of the most widely used architectures, producing the largest body of sustained research and application. In order to study a wide range of cognitive mechanisms, ACT-R includes a variety of modifiable parameters. While these parameters enable flexibility they also result in fundamental challenges.

Wexler (1978) criticized the early framework of the ACT research program (Anderson, 1976), stating that "There is no explanatory power in ACT because there are no restrictions on human abilities". He also posited that "the general problem with ACT is (its flexibility), it is simply so weak that there is no way to find evidence for or against it". About twenty years later, Pashler and Roberts (2000, 2002) again brought these concerns to the fore, arguing that the practice of using good fits as major evidence for complex theories is "rotten to core". Indeed, goodness-of-fit metrics remain a very common means of model validation. These concerns not only hold when criticizing ACT-R and some other unified models, but also address a wide-spread misuse of goodness-of-fits as key evidence in psychology. Sound scientific theory requires that models not only fit but also predict data (Gigerenzer, 1998; Gigerenzer & Brighton, 2009). How can modelers of the ACT-R architecture deal with these concerns about parameter estimation and model fitting?

There have been some attempts to understand the relation among ACT-R parameters that result from parameter fitting. For example, Anderson, Bothell, Lebiere, and Matessa (1998) suggested that there is a systematic linear relationship between the estimated values of activation thresholds and the logarithm of estimated latency factors. Their data also implied that estimated values of these parameters are exceedingly regular. To date, however, there has been no meta-analytic assessment to evaluate whether there is any sustained regularity of these estimated parameters for ACT-R models across other published studies.

Computational cognitive models are often evaluated by their fit and generalizability. These properties of a model are related to two aspects of model complexity: (1) number of parameters and (2) the functional forms of computation. In part, such evaluations seek to evaluate the extent to which noise is unnecessarily captured (Pitt, Myung, & Zhang, 2002; Oaksford, 2002). Using cross-validation, Taatgen, van Rijn, and Anderson (2007) estimated parameters of a base-model once and then made use of these estimated values throughout subsequent models. This study exemplifies a strict practice that allows minimal parameter estimation; however, like many ACT-R studies, the work of Taatgen et al. still relied on superior goodness-of-fits as the major support for their proposed models.

The latest ACT-R architecture version 6.0 has 62 free parameters with numerical values, together with the flexibility of mapping these parameters to tailor-made handlers and tens of other non-numerical parameters. Different instantiations of specific ACT-R models do not typically require setting and optimizing all these numerical parameters, as default values are provided. However, our analyses of a large and representative sample of ACT-R studies indicates that on average each ACT-R model modifies nearly six free numerical parameters for better model fitting. Moreover, many of these studies added task-specific parameters.

Parameter Estimations by "Wisdom of the crowd"

To provide another path to parameter estimation-free modeling, we developed a database to collect estimated and modified ACT-R parameters from the ACT-R modeling community. With this database, we hope to facilitate comparisons of ACT-R parameters by drawing on the wisdom of the crowd. Accordingly, we catalog previous studies that have provided estimates together with corresponding parameters. The database makes it relatively easy to determine whether a particular newly estimated value falls within a reasonable range according to previous related studies. Moreover, this database also serves to provide some meta-analytical data on the variety and ranges of selected parameters across a large representative set of studies.

Taatgen, et al. (2007) have argued that the ideal goal for an ACT-R modeler is to fix all parameters: A modeler should not estimate any parameter during modeling. One key goal of this current project is to collect and compile data from a representative range of published ACT-R models (with exact values for estimated parameters) in a sustainable database to assist ongoing modeling projects. The online database provides several potentially useful functions including updated information about the means and the medians of the existing free numerical parameters. Moreover, the database has been designed to be scalable so as to be readily extended to other models and tasks. In what follows, we describe the database and briefly review some functions and findings. We close with a discussion of potential applications and implications.

Method

We started with the studies and models that made use of the ACT-R architecture listed on the ACT-R website (http://act-r.psy.cmu.edu/). From this online repository of ACT-R studies, we selected all studies that have made both their ACT-R models and manuscripts available; a total of 44 studies were included at the time of data collection. From these models, we collected the information about the version of ACT-R architecture that was used as well as the particular ACT-R parameters that were modified. We also collected information about the deprecated parameters from previous versions of ACT-R and other task-specific parameters that these models made use of.

Overview of functions of the database

The database can be accessed through an Internet-interface at the URL:

http://www-abc.mpib-berlin.mpg.de/actrdb/

The Internet-interface was tested and works with most of the popular website browsers, such as Firefox 3+, Safari 4, Internet Explorer 8+, and Opera. Along with the information about ACT-R parameters, our database serves at least four main functions:

1. Monitoring how frequently parameters are modified.
2. Obtaining parameter means, medians and distributions.
3. Searching the keyword descriptions of ACT-R studies in the database.
4. Collecting fields of study and other information related to ACT-R parameter estimations.

Below we describe the basic functions of the database.

Frequency graph In the middle of the frontpage, there is a frequency graph listing all the numerical parameters that have been modified by at least one study in the database. Layout of the frequency graph is arranged so that the modified parameters are listed in descending order of frequency from the bottom to the top (Figure 2).

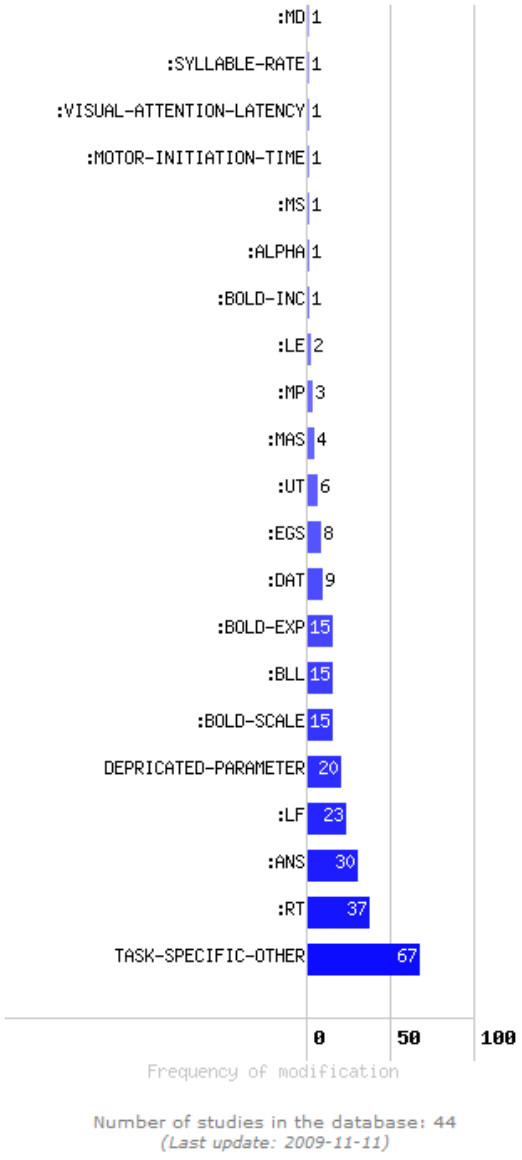

Figure 1: Frequency of modifications of ACT-R parameters.

Frontpage function buttons At the top of the frontpage of the database there are functional buttons labeled "Query parameters", "list studies", and "Enter your ACT-R study". These buttons provide access to the major ways to interact with the database (See Figure 2). A 'Home' button returns the user to the portal frontpage.

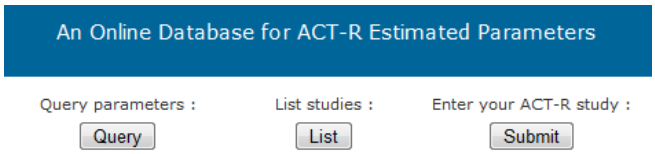

Figure 2: Major functional buttons at the top of frontpage.

Performing keyword search in the database At the bottom of the frontpage there is a search box where users can perform keyword searches or exact title searches (See Figure 3).

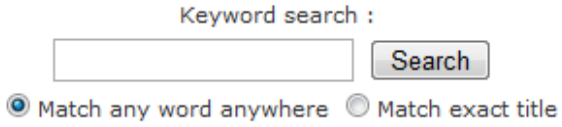

Figure 3: Keyword search or title search in the database.

Query parameters By pressing the query button a user can query any ACT-R parameter for any particular version of ACT-R in the database, using a drop-down menu (See Figure 4).

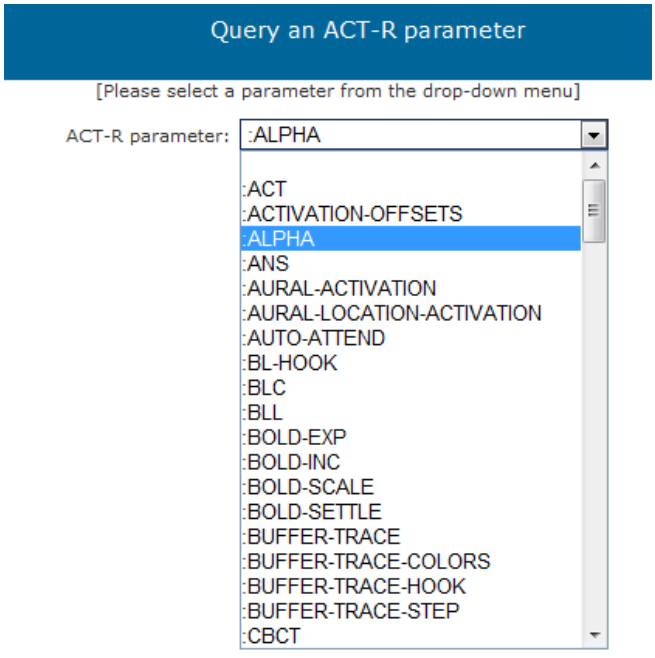

Figure 4: A drop-down menu to query ACT-R parameters.

After a parameter is chosen, the studies in the database that modified the particular parameter are displayed together with the mean and the median. The database also provides a graph describing the distribution of its modification among studies listed as well as the default values and the equation(s), if any (See Figure 5).

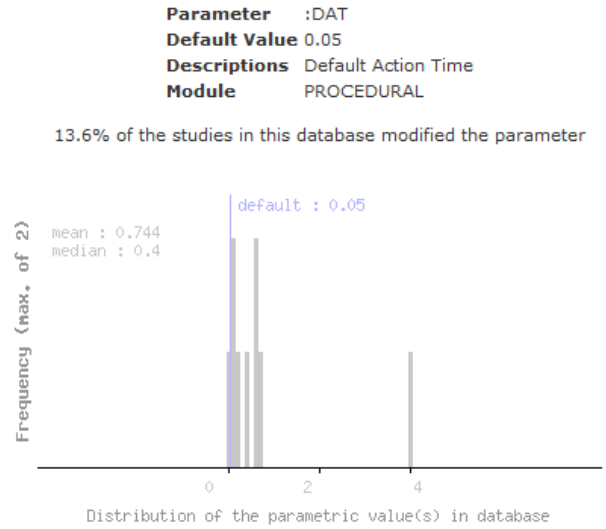

Figure 5: Information about the ACT-R parameter :DAT.

Results of listing or searching the database By pressing the 'List' button, or by performing a search on the frontpage, a user will reach a list of studies (See Figure 6).

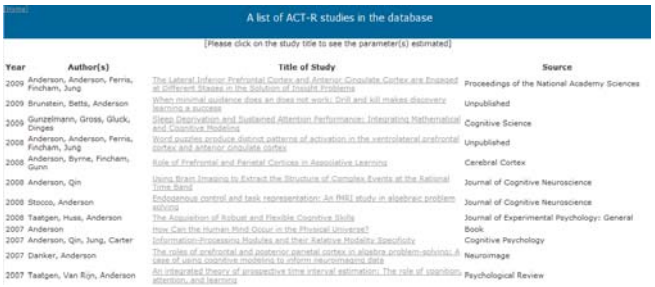

Figure 6: A list of ACT-R studies displayed after pressing the 'List' button or performing a search.

By further clicking on the title of a study specific information about parameter modifications of that study will be displayed (See Figure 7).

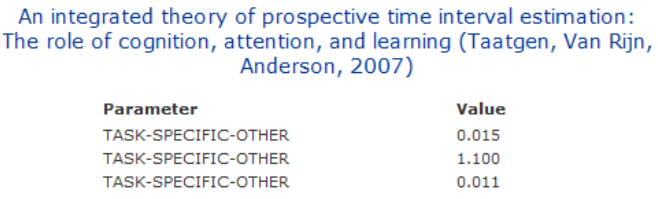

Figure 7: Information about parameter modifications.

Submit your model By pressing the 'Submit' button a user will reach the interface for entering information about parameter modifications of an ACT-R study (See Figure 8).

The interface is designed so as to guide the user through reporting their study in a step-by-step manner. By allowing ACT-R modelers to interact through our online database (providing their own estimated parametric values, modified values, and comments on their entries and models) we provide a more sustainable 'living' archive that benefits from the 'wisdom of crowds'. Readers are welcome to try out the database and provide feedback.

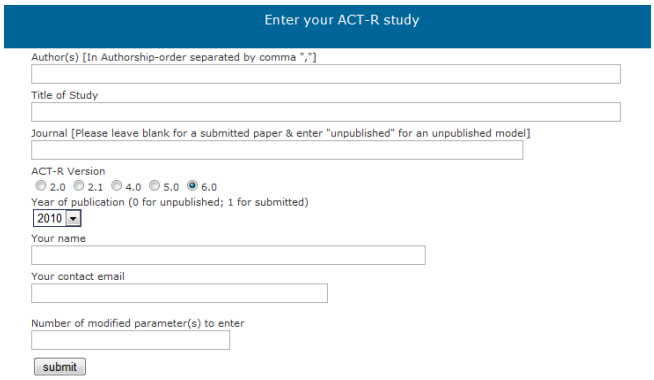

Figure 8: An interface to enter information about an ACT-R modeling study.

Results and Discussion

A brief report of some notable ACT-R parameters

At the time of publication, 44 studies were included in the database Together with a total of 261 instances of parameter modifications or estimations. On average each study modified 5.93 parameters. Among the ACT-R parameters that were modified in these studies, the three most frequently modified were :RT, :ANS, and :LF. The two ACT-R parameters :BOLD-EXP and :RT have the widest ranges of modified values among all parameters (See Table 1).

Table 1: The Most Frequently Modified ACT-R Parameters and Parameters with The Widest Range of Values

ACT-R Parameter	Default value	Description	Frequency modified
:ANS	Nil	Activation noise of chunks	30
:LF	1	Latency factor of chunks retrieval	23
:RT	0	Retrieval threshold of chunks	37
:BOLD-EXP	6	Exponential parameter for computing the BOLD response.	15

Note: For detailed descriptions of all the ACT-R parameters, we refer the interested reader to the ACT-R website (http://act-r.psy.cmu.edu/), Anderson (2007), and Anderson & Lebiere (1998).

Applications

Anderson et al. (1998) demonstrated that there are systematic variations between τ (:RT) and F (:LF) across studies. Unfortunately, not all the parameters in ACT-R have received this level of attention. As ACT-R continues to develop, it will acquire even more parameters. To help manage obstacles and challenges associated with such growth, our online database may provide a useful and convenient way for the ACT-R community to interact with each other and monitor these parameters. In the long run, by flagging frequently monitored parameters the database may point to weaknesses in the theory. In the short run, the database provides an overview of the parameter space.

To illustrate, when a modeler wants to study a phenomenon that requires estimation of an ACT-R parameter, this database serves as a portal to get an overview of the parameter in question with just a few mouse-clicks. With a keyword search about the phenomenon one can get a list of related modeling studies. When directly querying the parameter, the database provides studies that have modified the parameter from its default value, alongside with the means, medians, and default value (if any) on a distribution graph. This provides the modeler with a transparent window onto what was previously opaque information about what parameter values other ACT-R modelers were using.

Beyond fixing exact parameters, we also expect that the database can simplify much of the procedure used to estimate ranges of ACT-R parameters. The database can provide information about ways and approaches for capturing individual differences (e.g. age, abilities), environmental differences, and task differences (e.g. vigilance). For example, the default action time (:DAT), which is set at 0.05 second, dictates the basic firing speed of a procedure in an ACT-R model. While it is standard to use to default values when possible, there are indications that age (Mata, Schooler, & Rieskamp, 2007) and environmental factors (Gunzelmann, Gross, Gluck, & Dinges, 2009) may alter this basic firing speed. Another example is the retrieval threshold parameter (:RT), which is normally set to zero but can be expressed instead as a logistic function with a range of possible values, reflecting forgetting (Schooler & Hertwig, 2005). In these instances, using the interactive database to gather information provides a way to better monitor and estimate the most reasonable (or common) parameters of variation for human speed of processing.

Implications for parameter estimations in ACT-R

We setup the database in response to some important concerns stemming from the general problems of parameter estimation associated with a framework as complex as ACT-R. By setting up this database, we appeal to the 'wisdom of the crowd' among ACT-R modelers. In ongoing work we are testing the *median parameter hypothesis:* The parameterization of ACT-R based on the median estimated

ACT-R parameter values across all studies will fare better in predicting performance when compared to the parameterization that was used for each particular study. We could also imagine searching for a set of parameters that gives the best fit to all the studies in the database. It is our hope that these efforts may bring ACT-R modelers closer to true "zero-parameter fits".

Setting parameter values *a priori* to plausible values constrains overly flexible models by restricting the range of a model's predictions. This should lead to more accurate and perhaps more useful predictions of human performance patterns. A possible further development is to estimate a recommended range of values for every ACT-R parameter that correspond to human cognitive limitations. Finding such a correspondence would be in line with practices used in the human factor community, where limits of human performance are essential inputs for system design.

Conclusions

The major aim of this database is to provide a collaborative interface for ACT-R modelers to document and monitor values of ACT-R parameters in an efficient and sustainable way. By making use of the "wisdom of the crowd", ACT-R modelers can minimize model flexibility and increase the generalizability of their models. This can also be seen as a natural experiment concerning how best to estimate parameters in a social manner. By using a database of parameters that encourages generalizability and penalizes flexibility the ACT-R community might move closer to answering Newell's beautiful call for a truly unified theory of cognition.

References

Anderson, J. R. (1976). *Language, Memory, and thought*. New Jersey: Lawrence Erlbaum.

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

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Lawrence Erlbaum.Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341–380.

Gigerenzer, G. (1998). We need statistical thinking, not statistical rituals. *Behavorial and Brain Sciences, 21*, 199-200.

Gigerenzer, G., & Brighton, H. J. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*, 107-143.

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep Deprivation and Sustained Attention Performance: Integrating Mathematical and Cognitive Modeling. *Cognitive Science, 33(5)*, 880-910.

Mata, R., Schooler, L., & Rieskamp, J. (2007). The aging decision maker: cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging, 22(4)*, 796-810.

Oaksford, M. (2002). How does it fit? *Trends in Cognitive Sciences, 6*(10), 412-413.

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*(3), 472–491.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107(2)*, 358–367.

Roberts, S., & Pashler, H. (2002). Reply to Rodgers and Rowe (2002). *Psychological Review, 109(3)*, 605–607.

Schooler, L. J., & Hertwig, R. (2005). How Forgetting Aids Heuristic Inference. *Psychological Review*, 112(3), 610–628.

Taatgen, N. A., Van Rijn, H., & Anderson, J. (2007). An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning. *Psychological Review, 114(3)*, 577–598.

Wexler, K. (1978). A review of John R. Anderson's Language, Memory, and Thought. *Cognition, 6*, 327-351.