10-2012

Refuting data aggregation arguments and how the IBL model stands criticism: A reply to Hills and Hertwig (2012)

Cleotilde Gonzalez
Carnegie Mellon University, conzalez@andrew.cmu.edu

Varun Dutt
Carnegie Mellon University

Published In
Psychological Review, 119, 4, 573-579.

In press: Psychological Review

Refuting data aggregation arguments and how the IBL model stands criticism: A reply to Hills and Hertwig (2012)

Cleotilde Gonzalez and Varun Dutt

Dynamic Decision Making Laboratory

Department of Social and Decision Sciences

Carnegie Mellon University

Word count (main text): 3,312

Address correspondence to:

Professor Cleotilde Gonzalez
Dynamic Decision Making Laboratory
Department of Social and Decision Sciences
Carnegie Mellon University. Porter Hall 208.
Pittsburgh, PA 15213
Phone: +1-412-268-6242
E-mail: coty@cmu.edu

Authors' Note

Varun Dutt is now at Schools of Computing and Electrical Engineering and Humanities and Social Sciences, Indian Institute of Technology, Mandi. This research is supported by the Defense Threat Reduction Agency (DTRA) grant number: HDTRA1-09-1-0053 to Dr. Cleotilde Gonzalez. The second author The authors would like to thank Hau-yu Wong for her editorial help in proofreading this manuscript. The authors would like to thank Ralph Hertwig and Ido Erev for providing the human data used in our analyses, and to Noam Ben-Asher, Michael Yu, and Katja Mehlhorn, Dynamic Decision Making Laboratory, for providing valuable comments on a draft of this paper. We would also like to thank the John Wixted, Ralph Hertwig, Mike Oaksford, and one anonymous reviewer for their helpful comments.

Abstract

Hills and Hertwig (2012) challenge the proposed similarity of the exploration-exploitation transitions found in Gonzalez and Dutt (2011) between the two experimental paradigms of decisions from experience (sampling and repeated-choice), which was predicted by an Instance-Based Learning (IBL) model. The heart of their argument is that in the sampling paradigm, an impression of reduced exploration over time (alternation rate, A-rate) is produced by an inverse relationship between the sample size and the A-rate, and the aggregation of participants with different sample sizes. They suggest a normalization of the A-rate, which produces constant A-rate curves during sampling, and conclude with certain "ensuing problems for the IBL model." We show that: the reduction of A-rate during sampling occurs even when sample length is controlled for; that regardless of the sampling length, the maximization behavior during sampling predicts the final choice; and that the IBL model accounts for a negative correlation between sample size and the predicted A-rate. Furthermore, when the IBL model's data is normalized following the procedure specified by Hills and Hertwig (2012), it results in similar flattened exploration curves as those found in human data. These results indicate that the transition from exploration to exploitation in the sampling paradigm (which has also been found in the repeated-choice paradigm) is not an illusion resulting from data aggregation: The same data with or without normalization may be interpreted differently, but such interpretations do not invalidate the mechanisms of the IBL model.

Keywords: IBLT, decisions from experience, risky choice, exploration, exploitation, sampling.

Gonzalez and Dutt (2011) demonstrated how a cognitive model based on the Instance-Based Learning Theory (IBLT; Gonzalez, Lerch, & Lebiere, 2003) accounts for human behavior in two experimental paradigms of decisions from experience, sampling and repeated-choice, through a comprehensive model comparison and generalization process. The sampling paradigm involves trying alternatives before making a consequential decision, while the repeated-choice paradigm involves making repeated consequential decisions. Regardless of the similar behavior observed in both paradigms, research often sets them apart. Gonzalez and Dutt (2011) demonstrated that decision making in these two paradigms can be explained with the same instance-based learning (IBL) model: through the selection of the option with the highest utility (blended value) or from inertia (i.e., repeating the last choice). Blended value is a weighted average of experienced outcomes where the weights are a function of the recency and frequency of observed outcomes. Furthermore, the IBL model also predicts the learning process that is characterized by a gradual transition from exploration to exploitation of alternatives: There is a gradual decrease in alternation rate between options (A-rate), which is shown to be remarkably similar in the sampling and repeated-choice paradigms.

Hills and Hertwig (2012) challenge the similarity of the learning process across both experimental paradigms. They suggest that an illusion of a reduced A-rate across sample sizes in the sampling paradigm is produced by aggregating participants with different sample sizes. They normalized the A-rates, which resulted in flattened A-rate curves in the sampling paradigm in contrast to the reduction of A-rates over time shown in both, sampling and repeated-choice paradigms in Gonzalez and Dutt (2011). From this demonstration they discuss a set of problems for the IBL model, assuming that the same IBL model would not be able to account for the different patterns of A-rate curves across the two paradigms. Hills and Hertwig's critique has resulted in a positive and productive exchange of ideas that challenge the IBL model with compelling arguments. However, we show next that the reduced A-rate during sampling is neither an illusion resulting from data aggregation, nor does it pose a problem for the IBL model.

In retrospect, the gradual transition from exploration to exploitation that is common across both paradigms was not given enough attention in Gonzalez and Dutt (2011). Part of this lack of attention is due to the fact that the learning process during sampling has not received much attention in the literature, and to the fact that the exploration phase in the sampling paradigm is often considered to be a random process that occurs during sampling while exploitation occurs only during the final choice after sampling (Erev, Ert, Roth et al., 2010; Hertwig & Erev, 2009; Hills & Hertwig, 2012; Rakow & Newell, 2010). However, as discussed in Gonzalez and Dutt (2011), this random pattern has been challenged by several findings like the existence of two types of individual search strategies: high alternators and low alternators (Hills & Hertwig, 2010). Similarly, Gonzalez and Dutt (2011) demonstrated that the search process during sampling is not random. They showed that even when samples are small, people can learn in the same alternation pattern that manifests in a costly search: a decrease in the A-rate across time. The similar learning pattern found in Gonzalez and Dutt (2011) across sampling and repeated-choice is extremely important, because it is this same learning process that has been used as a "theoretical divide" between the two paradigms (Hills & Hertwig, 2012).

In this reply, we will concentrate on addressing Hills and Hertwig's (2012) two major criticisms toward Gonzalez and Dutt (2011) that focus on the sampling paradigm:

1. That a gradual reduction of exploration (A-rate) during sampling mischaracterizes what many individuals actually do as it is a result of aggregating A-rate for participants with different sample lengths.

2. That normalization of the A-rate across the proportion of samples is necessary and it results in *constant* exploration across the samples, not in a reduction of exploration over trials. This finding divides the two paradigms again and it creates problems for the IBL

model.

The reduction of exploration during sampling is NOT a result of aggregating A-rates for participants with different sample sizes

The key issue in Hills and Hertwig's (2012) commentary is that an inverse relationship between the sample size and exploration (A-rate) and the aggregation of participants with different search lengths produces an erroneous impression of a gradual reduction of exploration. Here, we first analyze the A-rate during sampling in an experimental dataset where all participants were required to sample the two alternatives for a fixed number of times before making a final choice (Hau, Pleskac, Kiefer, & Hertwig, 2008). Second, we analyze two large human datasets in the sampling paradigm from the Technion Prediction Tournament (TPT), where participants were free to sample in any order with different sample sizes (Erev et al., 2010). To analyze participants with the same sample size in the TPT datasets, we create three different groups with the same number of participants that sampled a minimum of 6, 10, and 18 times, and then study the alternations of all participants in each group. Analyzing these datasets demonstrates a gradual reduction of A-rate during sampling when the sample size is controlled for and when aggregation is done across the same number of participants.

Hau et al.'s (2008) Experiment 3

In Hau et al. (2008) (Experiment 3), participants were asked to sample for a fixed number of times (=100 samples) before making a consequential choice. We first labeled a participant’s sampled choices in two successive trials as an alternation (=1) or as a no-alternation (=0). Then, we calculated the A-rate across samples as in Gonzalez and Dutt (2011), by averaging the alternations over all participants and problems for a sample (=sum of alternations across participants and problems/(number of participants * number of problems), for each sample from 2 to 100). Figure 1A plots the resulting A-rate over samples.

In the first 11 trials, the A-rate falls by 44.11% (from 34% in trial 2 to 19% in trial 12) and for the last 89 trials, the curve flattens to about 19%. Thus, participants do a little exploration and then a lengthy exploitation favoring one of the alternatives during sampling, a pattern similar to that known in the repeated-choice paradigm and under free sampling, as predicted by the IBL model (see Figure 5B in Gonzalez and Dutt (2011) where a similar flattening at about 20% occurs).

Next, we calculated the initial A-rate (the average A-rate in the first 9 samples from 2 to 10) and the final A-rate (the average A-rate in the last 9 samples from 92 to 100) (similar to Hills and Hertwig's (2012) Figure 3B). Figure 1B shows the initial and final A-rates at the individual level. This figure shows that out of a total of 40 participants, only 4 (10%) kept their initial and final A-rates constant; 12 participants (30%) showed an increase in the A-rate; and a majority of 24 participants (60%) fell below the diagonal, showing a decrease in the A-rate from initial to final samples. Thus, consistent with Figure 1A, these findings demonstrate the reduction of exploration as predicted by Gonzalez and Dutt (2011) over increased sample size at the aggregated (Figure 1A) and individual (Figure 1B) levels.

Furthermore in agreement with the IBL model's predictions, we also find that the final choice in Hau et al.'s (2008) data set is explained by the learning during sampling: The proportion of participants selecting the maximizing alternative during sampling is positively correlated to their final choices for the same maximizing alternative ($r(38) = 0.36$, $p < .05$). The magnitude of this correlation is large given that it measures the agreement between choice rates during sampling (that are close to 0.5) and a single binary final choice made after the sampling ends. In addition, we estimated the final choices according to the experienced expected value in the two options (an experienced expected value was estimated in each option for each participant in a problem based upon the participant’s sampling of the option). The final choice can be predicted by the higher of the two experienced expected values, one for each option (just like the IBL model would do with blended values except that the model more accurately represents the frequency and recency of outcomes): 60% of the predicted final choices are consistent with the actual choices.

The TPT sampling dataset

Experiment 3 in Hau et al. (2008) used a fixed sample size (= 100 samples) for all participants. This manipulation of fixing the sample size is different from the majority of sampling paradigm studies (Gonzalez & Dutt, 2011; Hertwig & Erev, 2009), where the sample size is free and determined by the participant rather than fixed by the experimenter. Thus, as an additional demonstration, we performed similar analyses with the TPT dataset that involved free sampling, where the end of sampling was decided by participants themselves (Erev et al., 2010). The TPT dataset consists of two sets, each with 60 problems and 40 participants. In each set, 40 participants were randomly assigned to two different groups, where each group contained 20 participants who were presented with a representative sample of 30 problems (out of a total of 60 problems). This setup resulted in a total of 2,400 observations (a combination of a set, problem, and participant) across the two sets. These observations were used for the purposes of analyses, as it is expected that the sample size depends on particular characteristics of the problem (Lejarraga, Hertwig, & Gonzalez, 2012) and because the problems were assigned in random order to each participant. Three groups were formed from the TPT dataset by selecting a subset of the 2,400 observations such that all observations in a group had a minimum sample size of 6-, 10-, or 18-samples (based upon the observations selected in different groups, these minimum sample sizes corresponded to 75, 72, and 62 participants in the 6-, 10-, and 18-samples groups, respectively). For the observations in these three groups, we analyzed their first 6, 10 and 18 samples, respectively, which ensured that all observations within a group had the same sample size.

Figure 2A shows the A-rate curves for observations in each of the three groups. For the 6-samples group, the A-rate decreased by 34.61% (from 52% in sample 2 to 34% in sample 6); for the 10-samples group, the A-rate decreased by 46.94% (from 49% in sample 2 to 26%

in sample 10); and for the 18-samples group, the A-rate decreased by 52.17% (from 46% in sample 2 to 22% in sample 18). Thus, the A-rate gradually decreases in groups containing the same sample size for each observation. The greater the sample size, the larger the decrease in the A-rate. Figure 2B shows the initial A-rate (calculated for the first 2, 3, and 6 samples) and the final A-rate (calculated for the last 2, 3, and 6 samples) at the individual (participant) level for the 6-, 10-, and 18-samples groups, respectively. Out of 75 participants among observations in the 6-samples group, only 6 (9.7%) kept their initial and final A-rates constant; 30 participants (48.4%) showed an increase in their A-rates; and a majority of 39 participants (62.9%) fell below the diagonal (showing a decrease in their A-rates). A similar pattern was observed in the 10-samples group: out of 72 participants among the observations in this group, only 9 (14.5%) kept their initial and final A-rates constant; 23 participants (37.1%) showed an increase in their A-rates; and a majority of 40 participants (64.5%) showed a decrease in their A-rates. The same pattern also holds for the 18-samples group: out of 62 participants, only 11 (17.7%) kept their initial and final A-rates constant; 15 participants (24.2%) showed an increase in their A-rates; and a majority of 36 participants (58.1%) showed a decrease in their A-rates. The similarity of the A-rate patterns in these three groups demonstrates a reduction of exploration similar to that predicted by Gonzalez and Dutt (2011), over increased sample size at the aggregated (Figure 2A) and the individual level (Figure 2B).

Finally, in agreement with the IBL model's predictions, the final choice is predicted by the maximization choice during sampling: The proportion of selecting the maximizing alternative during sampling is positively correlated to the final choice for the same maximizing alternative in the three groups, regardless of their sample sizes: $r(73) = .26$, $p < .05$ for the 6-samples group, $r(70) = .34$, $p < .01$ for the 10-samples group, and $r(60) = .40$, $p < .01$ for the 18-samples group. Also, 84% of the choices predicted by the experienced expected value during sampling are consistent with the observed choices.

The resulting "ensuing problems for the IBL model" are invalid

Given that we refuted the main point of Hills and Hertwig's (2012) commentary in the previous section, the problems they describe for the IBL model are only speculative: the gradual transition from exploration to exploitation that is common in the sampling and repeated choice paradigms is not an illusion from aggregating A-rates across participants with different sample sizes during sampling. Furthermore, we believe that Hills and Hertwig's proposed A-rate normalization is an incorrect representation of the learning process that occurs during sampling.

Hills and Hertwig (2012) proposed normalizing the A-rate over the proportion of samples for each participant in order to address the inverse relationship between the sample size and the A-rate when aggregated across participants with different sample sizes. They found a constant exploration pattern in the A-rate in the sampling paradigm after normalization, but a reduction of exploration in the repeated-choice paradigm. This disparity led them to describe a set of problems for the IBL model, as the same model could not predict the different trends across paradigms. However, the normalization equalizes all participants in a way so that the earlier and later samples mean “the same” for participants with different sample sizes. For example, consider participant A, who sampled three times and participant B who sampled 50 times. Normalization of the A-rate for participant A contributes to increase the count in two bins: the 60-70% bin and the 90-100% bin, while normalization of the A-rate for participant B contributes to increase the count in all of the 10% bins from 0% to 100%. The problem is that the last alternation after 49 trials for participant B is considered to carry the same weight as the alternation after 2 trials for participant A.

In learning research, earlier and later samples cannot have the same weight. The well-known practice effect (e.g., Newell & Rosenbloom, 1981) occurs as people sample more. Furthermore, in the sampling paradigm, small samples amplify the difference between the expected earnings, making the choice easier; while larger samples reduce this difference (the "amplification effect;" Hertwig & Pleskac, 2008). Thus, the later samples do not mean the same for Participant A and B in the example above and the flattened curves generated by normalizing the A-rate are an incorrect representation of the learning process during sampling. Furthermore, these curves should not be used to extrapolate conclusions on the IBL model's mechanisms.

Hills and Hertwig (2012) conclude that the IBL model would fail to predict the inverse relationship between the A-rate and the sample size because to determine each simulated participant's sample size, the model draws from a distribution function derived empirically from a human dataset. While we agree that the IBL model in Gonzalez and Dutt (2011) does not "predict" the sample size, the model indeed shows a negative correlation between the sample size and the predicted A-rate (r = -0.18, p < .001), similar to what is found in human data (Hills & Hertwig, 2012). This negative correlation shows that as sample size increases, the A-rate decreases, representing a decrease in the exploration across samples.

Also, the IBL model is able to predict the gradual transition from exploration to exploitation found in the analyses above. First, we did a generalization of the model that was calibrated to the six problems in Hertwig et al. (2004) (with d = 0.29; σ = 0.27; and *pInertia* =0.22; see Table 2 in Gonzalez and Dutt, 2011), but fixed the sample size to 100 samples as was done in Hau et al.'s (2008) Experiment 3 to predict the A-rate in this experiment. The IBL model makes fairly accurate predictions about the A-rate in this dataset (MSD between model and human data = 0.0101 and r = 0.72). Second, we reanalyzed the data produced by the model in the TPT datasets (Gonzalez & Dutt, 2011). These analyses involved the same grouping as in the previous section involving 6-, 10-, and 18-samples groups. The model again provided accurate predictions of the A-rate observed in human data. The MSD and r values between the model and human data were: MSD= 0.0115 and r = 0.90 (6-samples group); MSD= 0.0058 and 0.92 (10-samples group); and MSD= 0.0041 and r= 0.93 (18-samples group).

Finally, we normalized the data produced by the IBL model (Gonzalez & Dutt, 2011) in the manner proposed in Hills and Hertwig (2012). Figure 3 shows a flattened A-rate curve produced by the IBL model, similar to the one from human data reported in Hills and Hertwig (2012) (MSD between the model and human data = .0084 and r = 0.65). Thus, the same data set may be presented in two ways (with or without normalization), but these different representations do not invalidate the mechanisms and predictions from the IBL model.

Conclusion

We have shown that the gradual transition from exploration to exploitation during sampling is not a result of aggregating A-rates for participants with different sample sizes. This transition occurs in cases where sample length is experimentally controlled for the same number of participants (Hau et al., 2008) and also when analyses of the TPT datasets are limited to groups with fixed sample size. Furthermore, this transition occurs at the average as well as at the individual level. We also demonstrate that the final maximizing choice in human data is predicted by the learning process during sampling as expected from the IBL model (Gonzalez & Dutt, 2011); that the model explains the inverse relationship between sample size and the A-rate; and that it predicts the transition from exploration to exploitation in the new datasets.

Hills and Hertwig's (2012) critique relies on an assumption that the individual learning curves should be normalized so that the initial and final samples "mean the same" for people with different sample sizes. We believe, however, that this normalization procedure is an incorrect representation of learning during sampling: The final samples are different for someone that samples extensively compared to someone that samples very little. Our evidence together refute Hills and Hertwig's (2012) main arguments and suggest that the same data may be represented and analyzed in different ways (with or without normalization), but that a normalization of the data does not invalidate the mechanisms and predictions from the IBL model. In fact, the model's data produces similar flattened curves after normalizing the A-rate as proposed in Hills and Hertwig (2012).

Although Hills and Hertwig (2012) have challenged the IBL model in important and interesting ways, our new analyses provide evidence against their critique. We suggest that a more precise way to challenge models is through comparison with alternative models, by using data generated from alternative models and analyzing their performance against human data. Gonzalez and Dutt (2011) demonstrated the robustness of the IBL model by following such a process across two simple experimental paradigms of decisions from experience. We are aware that the proposed IBL model might still be incomplete and that it might not be able to account for behavior in more complex and dynamic situations (see discussion in Gonzalez & Dutt, 2011), but the way Hills and Hertwig (2012) challenged this model led to additional important theoretical developments which we hope will continue to be questioned in the future.

References

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., et al. (2010). A choice prediction competition: Choices from experience and description. *Journal of Behavioral Decision Making*, *23*(1), 15-47. doi: 10.1002/bdm.683

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118*(4), 523-551. doi: 10.1037/a0024558.

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

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description–experience gap in risky choice: the role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, *21*(5), 493-518. doi: 10.1002/bdm.598

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. doi:10.1016/j.tics.2009.09.004

Hertwig, R. & Pleskac, T. J. (2008). The game of life: How small samples render choice simpler. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 209-235). Oxford: Oxford University Press.

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science*, *21*(12), 1787–1792. doi:10.1177/0956797610387443

Hills, T., & Hertwig, R. (2012). Two distinct exploratory behaviors in decisions from experience: Comment on Gonzalez & Dutt, 2011. *Psychological Review*.

Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). Decisions from experience: Do losses and risk affect information search? Manuscript in press. *Cognition*.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Earlbaum.

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*(1), 1–14. doi:10.1002/bdm.681

Figure Captions

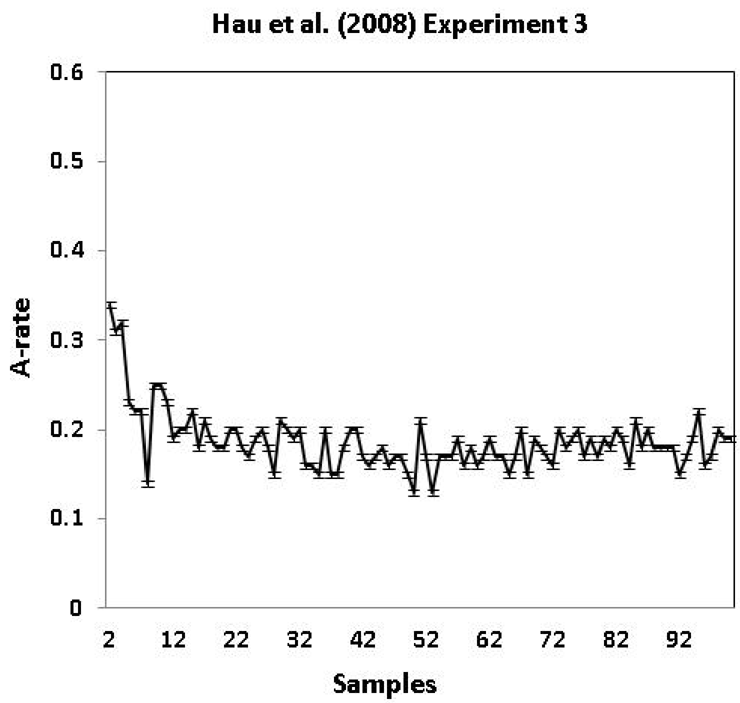
Figure 1. (A) The A-rate learning curves for the human data in Hau et al. (2008), Experiment 3. The error bars represent 95% confidence interval. (B) The A-rate for the first 9 samples and the final 9 samples for human data in Hau et al. (2008)'s Experiment 3. Circles on the diagonal represent people whose initial and final A-rates are identical.

Figure 2. (A) The A-rate for three groups with the same number of participants, showing their first "n" samples: n=6-samples, n=10-samples, n=18-samples from the TPT's sampling paradigm datasets. (B) The A-rate for the first 2, 3, and 6 samples and the final 2, 3, and 6 samples for human data in the 6-, 10-, and 18-samples groups. Circles on the diagonal represent people whose initial and final A-rates are identical.

Figure 3. The normalized A-rate in the sampling paradigm for the TPT's human dataset and that for the IBL model reported in Gonzalez and Dutt (2011). The error bars show 95% confidence interval around the point estimate.

Figure 1.

(A)


Hau et al. (2008) Experiment 3
0.6
0.5
0.4
0.3
0.2
0.1
0
A-rate
2
12
22
32
42
52
62
72
82
92
Samples

(B)

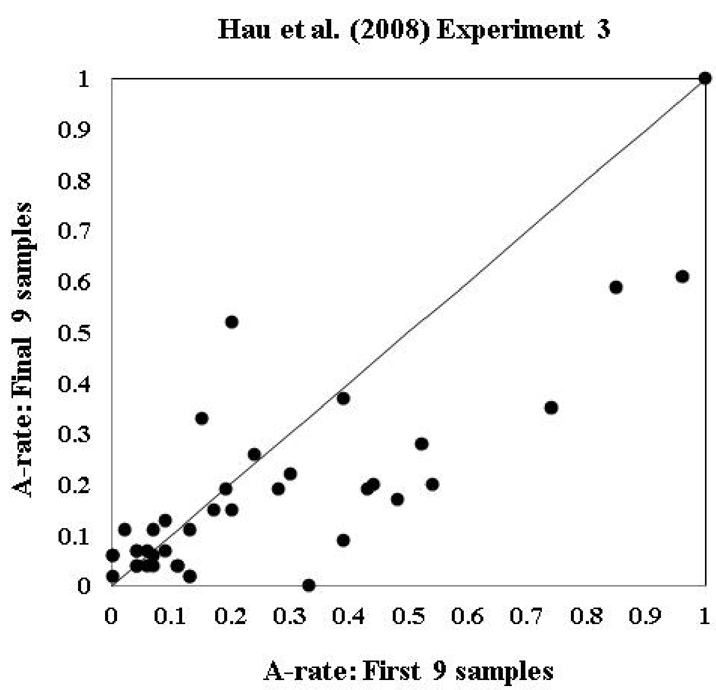
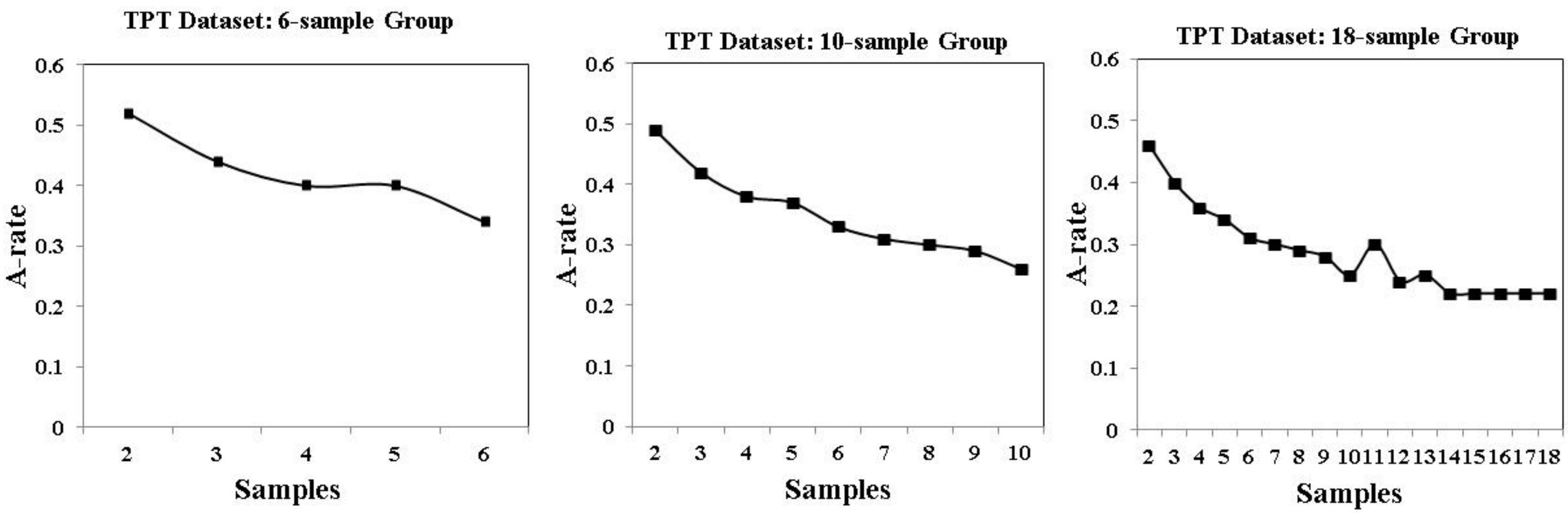

Hau et al. (2008) Experiment 3
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
A-rate: Final 9 samples
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
A-rate: First 9 samples

Figure 2.

(A)


TPT Dataset: 6-sample Group
A-rate
Samples
TPT Dataset: 10-sample Group
A-rate
Samples
TPT Dataset: 18-sample Group
A-rate
Samples

(B)

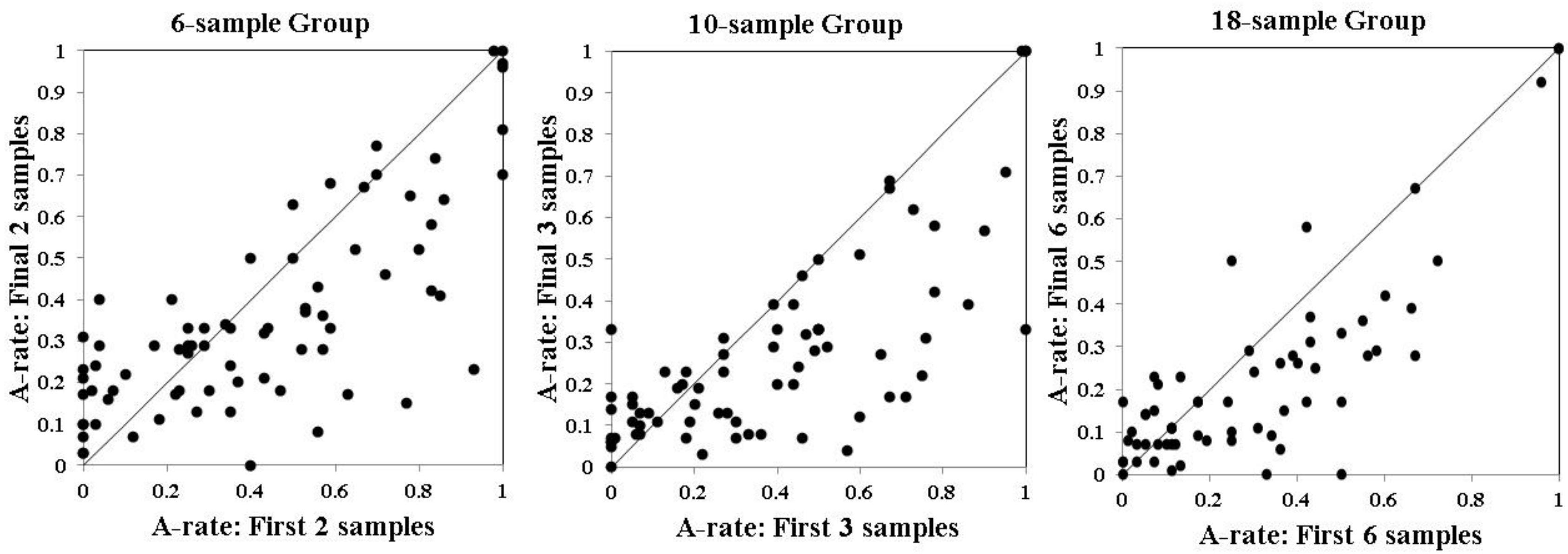
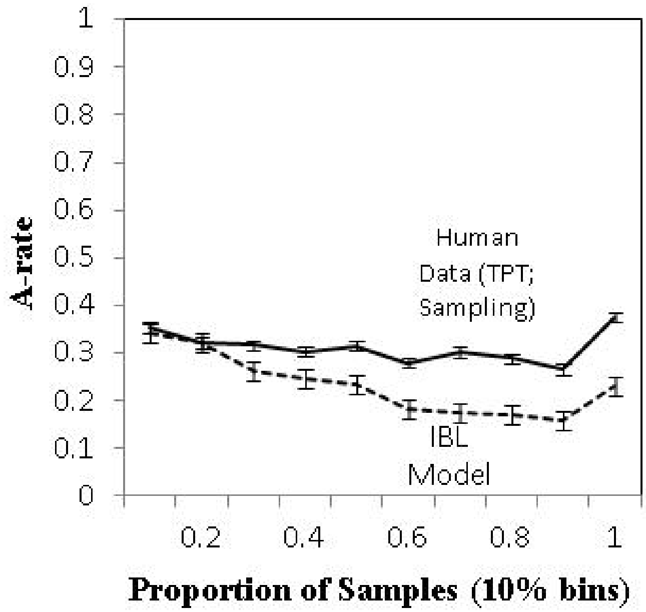

6-sample Group
A-rate: Final 2 samples
A-rate: First 2 samples
10-sample Group
A-rate: Final 3 samples
A-rate: First 3 samples
18-sample Group
A-rate: Final 6 samples
A-rate: First 6 samples

Figure 3.


A-rate
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Human Data (TPT; Sampling)
IBL Model
0.2
0.4
0.6
0.8
1
Proportion of Samples (10% bins)