

Classified Judgment on Same-different Based on Bayesian Inference

Zhi-Cheng Chen

State Key Laboratory of Mechanics and Control of Mechanical, Institute of Nano Science and Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, China
Department of Mathematics, Henan Institute of Science and Technology, Xinxiang, Henan, China

Xin-Sheng Liu

State Key Laboratory of Mechanics and Control of Mechanical, Institute of Nano Science and Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, China
E-mail: xsliu@nuaa.edu.cn

Abstract-The ability to detect objects precisely with the same orientations from random orientations is one of the most remarkable functions of perception and cognition. Many psychophysical evidences imply human's behavior follows near-optimal Bayesian inference in tasks of decision making. We explore the model of bars with the same orientation from the random orientations by the means of theories and experiments. In experiments, the numbers and locations of bars with same orientation are randomly assigned. The performances of subjects depend on the number of bars (set size), the number of bars with same orientations and reliabilities of measurements. We put forward a model based on Bayesian theory in visual search to judge how many bars with same orientation among a mass of orientations. To compare the effect of the model, we design two alternative models. We find our model on this kind of classified judgment matches well the experimental data, is superior to alternative models, and it provides a normative and mathematically quantitative description.

Keywords-variability discrimination; sameness-difference; Bayesian theory

I. INTRODUCTION

Classified judgment is important to human behavior and no arcane or trivial cognitive feat, and the one that is central to psychological science [1]. To discriminate the variability exists in a wide range of routine tasks, such as searching for mates, sorting letters in post office, putting goods on store shelves and so on. The prevalent method is to segment a visual scene, and they aggregate localized differences among objects, and comparing these finding differences of the objects [1-15]. As a special case, the same-different judgment attracts many researchers' interests [1-11]. Discriminating the same against the different is a simple case of the object categorization, which is an important basic concept in psychological perception and cognition and plays an effective role in the development of the abstract notion of equivalence [14]. Just as William James said, "sense of sameness is the very keel and backbone of our thinking" as well as "the most important of all the features of our mental structure" [15]. By far, there are plenty of studies on classified judgment the sameness from the differentness involving with humans [9-11], honey bees [2], pigeons [3], parrots [4], apes [16] and so on.

The computations based on the abundance of experimental data to the sameness judgment are not well

understood. On this account, researchers put forward many approaches to implement the proposition of the sameness [1-12]. Young and Wasserman found the information theoretic measure of categorical variability "entropy" systematically described the functional relationship between stimulus variability and discriminative responding [5,8]. Young, Wasserman and Ellefson considered a reasonable model for "finding differences", which judged the sameness by local differences between items. The finding differences model was effective in visual search tasks, and was applicable to investigations involving multidimensional variability [6] and quantitative item differences [7]. Wasserman et al. further illustrated discrimination of stimulus variability affected both human and animal action [12]. Ma et al. found that performances of human were near-optimal in a target detection task among a cluttered environment under the condition of the target hinted in advance [17,20]. They presented a neural implementation of near optimal visual search based on probabilistic population coding. Ronald et al. proposed that the optimal observer model accurately described human behavior, outperforms plausible alternatives in a rigorous model comparison. Their models provided a normative footing with the perfect form and rich connotation to study the sameness judgment based on Bayesian theory [11]. For all, previous studies have focused on the sameness and the differentness in visual search. To test whether human behavior in visual search best matches the performance predicted by the model based on Bayesian inference, we explore the problem of classified judgment on existing sameness against difference by the means of theoretical models and experiments. They concern with the sameness judgment, which has only two equal results: the sameness and the difference. In our classified judgment experiments, the possible results of experiments can be divided into three classes: sameness, difference and not all sameness. Our classified judgment is more difficult to judge than the sameness judgment, we need more time to accurately discriminate how many bars with the same orientation occur. To judge whether there exist bars with the same orientation among a mass of bars with random orientations and report the number of bars with the same orientation, we put forward a model of Bayesian optimal in visual search to implement the conjecture. Our models do not only extend the possible results in form, but also settle more complicated questions, including the rate of correct judgment

and false judgment mainly caused by random errors and perceived errors.

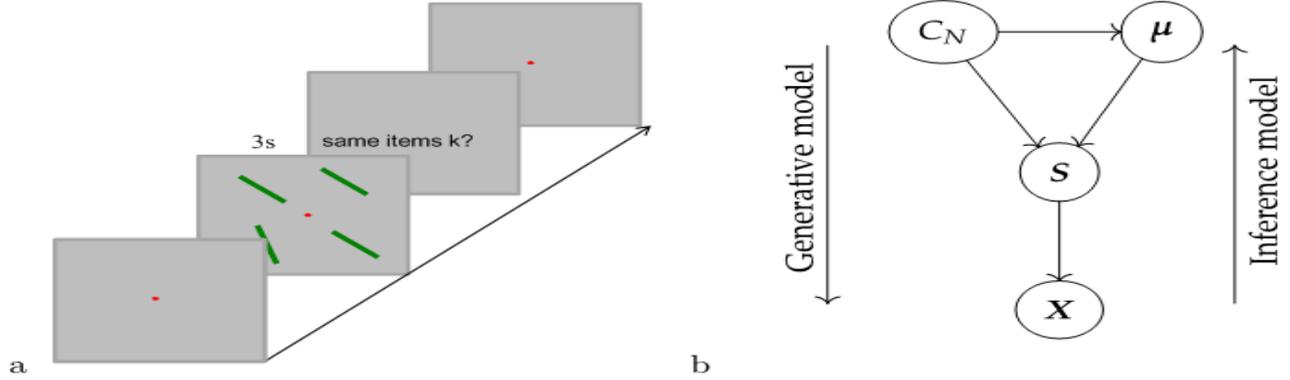


Figure 1. (a) Schematic of the trial procedure. Subjects reports the same terms k by numeric keyboard. Set sizes N are 3, 4, 5, 6. (b) Statistical structure of the task. Arrows indicates conditional dependencies and nodes have randomness. C_N indicates the number of the same orientation; μ , random direction vector; S , the vector of stimulus orientations and X the vector of sensory measurements.

II. EXPERIMENT

We conduct behavioral experiments (Fig.1a) to test whether human performance in visual search best matches the performance predicted by the our model. Subjects are shown briefly with an array of N random orientation bars and stimuli are characterized by orientations of bars, denoted by S . On each trial, N bars ($N=3, 4, 5, 6$) are presented on an imaginary circle. There are k bars with the same orientation u_0 and the others $N-k$ bars with orientations u_{k+1}, \dots, u_N respectively. The orientation variables u_0, u_{k+1}, \dots, u_N are randomly drawn respectively from a uniform distribution over orientation space $(0, \pi)$ and they are components of orientation vector $\mu = (u_0, u_{k+1}, \dots, u_N)$. The number of bars with the same orientation k takes the value randomly from the set $\{0, 2 \dots N\}$ with equal probability. The experiment consists of four parts for every subject, and the set size (3, 4, 5, 6) of each part is different input by keyboard in according to subject's intentions. Each part consists of six blocks of 150 trials.

III. MODELING

The statistical structure of the task is illustrated (Fig.1b). The observers have access to noisy measurements $X = (X_1, \dots, X_N)$ of the stimulus vector $S = (S_1, \dots, S_N)$ to judge the number of the bars with the same orientation. The sensory measurement X_i of the i th stimulus S_i , and $X_i \sim N(s_i, \sigma_i^2)$. We denote the number of bars with the same orientation among N bars by the variable C_N , which takes the value k . If k bars with the same orientation occurs, we compute the likelihood function by Bayesian theory over C_N , $p(x|C_N)$ the probability of the noisy observation X under the condition C_N . We obtain the likelihood function by marginalization in virtue of the true stimulus vector S and the orientation vector.

$$p(x | C_N) = \iint p(\mu | C_N) p(s | C_N, \mu) p(x | s) ds d\mu \quad (1)$$

When $C_N = k$, it means there are k stimuli $(S_{i_1}, \dots, S_{i_k}) \triangleq S_s^k$ with the same orientation μ_0 which is drawn from a uniform distribution on $(0, \pi)$, the other $N-k$ stimuli $(S_{i_{k+1}}, \dots, S_{i_N}) \triangleq S_d^{N-k}$, each of what is drawn from a uniform distribution $U(0, \pi)$ with mean μ_i and orientation vector $\mu = (u_0, u_{k+1}, \dots, u_N)$,

$$p(\mu | C_N) = \begin{cases} \pi^{-N}, & k=0 \\ \pi^{k-N-1}, & k=2, \dots, N. \end{cases} \quad (2)$$

The density function of stimulus S is a the Dirac Delta distribution denoted by $\delta(S)$. When $C_N=0$, the N stimuli are different, each $S_i \sim N(\mu_i, \sigma_i^2)$ and the corresponding noisy measurement $X_i \sim N(s_i, \sigma_i^2)$. When $k=2 \dots N$,

$$p(x | C_N) = \frac{1}{\pi^{N-\frac{k-1}{2}} \frac{k-1}{2} \sqrt{A_k}} \left(\prod_{j=1}^k \frac{1}{\sigma_{i_j}} \right) \exp \left(-\frac{1}{2} \left(D_k - \frac{B_k^2}{A_k} \right) \right) \quad (3)$$

$$\text{where } A_k = \sum_{j=1}^k \frac{1}{\sigma_{i_j}^2}, \quad B_k = \sum_{j=1}^k \frac{x_{i_j}}{\sigma_{i_j}^2}, \quad D_k = \sum_{j=1}^k \frac{x_{i_j}^2}{\sigma_{i_j}^2}.$$

For Eq.(1), repeating the calculation as $C_N=0$, it yields the following, $p(x | C_N=0) = \frac{1}{\pi^N}$.

The log likelihood ratio, denoted by d_k , could be evaluated using Bayesian rule as below,

$$\begin{aligned} d_k &= \log \frac{P(C_N = k | X)}{P(C_N = 0 | X)} = \log \frac{P(X | C_N = k)}{P(X | C_N = 0)} + \log \frac{P(C_N = k)}{P(C_N = 0)} \\ &= -\frac{1}{2} ((X^{(k)})' Q_s X^{(k)} - (k-1) \log \frac{\pi}{2} + \log \sum_{j=1}^k w_j - \sum_{j=1}^k \log w_j) + \\ &\quad \dots + \log \frac{P(C_N = k)}{P(C_N = 0)}, \end{aligned}$$

where Q_s is a k -order symmetric and non-negative definite matrix, the components of the matrix Q_s are $Q_{ij} = \delta_{ij} - \frac{1}{k}$. $X^{(k)} = (X_{i_1}, \dots, X_{i_k})$ follows an k -dimensional multivariate normal distribution with mean vector $s_s^{(k)}$ and covariance matrix $\Sigma = \text{diag}(\sigma^2, \dots, \sigma^2)$, that is $X^{(k)} \sim N(s_s, \sigma^2 I_k)$. The matrix Q_s has $k-1$ eigenvalues equal to 1, and an eigenvalue equals to 0. There exists an orthogonal matrix T satisfying $T'Q_sT = \Lambda = \text{diag}(1, \dots, 1, 0)$ for non-negative definite Q_s . Let Y be $Y^{(k)} = TX^{(k)}$. As T is an orthogonal matrix, $Y^{(k)} \sim N(u^{(k)}, \sigma^2 I_k)$, with $u^{(k)} = T's_s^{(k)}$. Since

$$\sum_{i=1}^k \frac{y_i^2}{\sigma^2} = \frac{1}{\sigma^2} Y^{(k)'} \Lambda Y^{(k)} = \frac{1}{\sigma^2} x^{(k)'} Q_s x^{(k)} = w x^{(k)'} Q_s x^{(k)}.$$

It follows that $w x^{(k)'} Q_s x^{(k)}$ complies with a non-central chi-squared distribution with $k-1$ degrees of freedom. The variable is equal to $wk \text{Var}(X^{(k)})$. The noncentrality parameter of the sum of above the formula is $wk \text{Var}(s_s^{(k)})$. Then, we obtain

$$wk \text{Var}(X^{(k)}) \sim \chi_{k-1}^2(wk \text{Var}(s_s^{(k)})) \quad (4)$$

We denote the subjects' estimation of the variable C_N by \hat{C}_N , and the probability of the subjects responding \hat{C}_N when the subjects are presented with a specific stimuli $s_s^{(k)}$, denoted by $P(\hat{C}_N | s_s^{(k)})$, where $s_s^{(k)} = (s_s^{(k)}, s_d^{(N-k)})$. All reliabilities are equal $\sigma_i = \sigma$, then $w_i = \sigma_i^{-2} = w$, we have

$$P(\hat{C}_N = k | s_s^{(k)}) = P\left(\chi_{k-1}^2(wk \text{Var}(s_s^{(k)})) < (k-1) \log \frac{\pi}{2} - \log kw + k \log w + 2 \log \frac{P_k}{P_0}\right) \quad (5)$$

where $P_k = P(C_N = k)$, $k = 0, 2, \dots, N$.

IV. ALTERNATIVE MODEL

To verify our model fits observer's data best, we give the following models to compare. In the Minimum Arbitrarily Variance(MAV) model, The decision rule is to respond k items with the same orientation if the minimum of variance of arbitrary $k+1$ stimuli of all orientations about stimulus S is greater than η_n , but the minimum of variance of arbitrary k stimuli is smaller than η_n , a decision criterion for every block. By means of the intuition, the sensory measurements of the same orientation should be closer than the different. To compare the fitted optimal model, we ponder over maximum-minimum range (MMR) model in the text. The observer uses the decision rule, obtaining the maximum

value in experiments about the minimum of all ranges on k sensory measurements of stimuli S as the range of k sensory measurements. The observer responds k same items occurrence if the range of being regarded k same orientation is less than a decision criterion for every block.

V. RESULTS

The proportions of "Correct" and "False" responses are presented as functions of k , N and σ in Fig. 2. As a matter of fact, the set size N and the same term k have an effect on the proportion. As expected, performances of observers become better at detecting a change as the same term k increases when set size N fixes. Meanwhile, we find a significant effect of the number of bars with the same orientation k on the "Correct" rate (one-way ANOVA, $F(2,8)=271.84>3.5546$; $F(3,24)=667.2>3.0088$; $F(4,30)=1758.72>2.6896$; $F(5,36)=2034.7>2.4772$; $p < 0.001$). This indicates it is easier to detect for a larger k as set size N is fixed and k has positive effect on the "Correct" rate. On the contrary, performances of subjects become worse as the set size N increases when k is fixed. The effect of set size N is highly significant as k is fixed (one-way ANOVA, $F(3,24)=11.95$; $F(3,24)=181.66>3.0088$; $F(3,24)=571.44$; $F(3,24)=1766.77$; $p < 0.001$). This shows the hit rates decrease as the number of noise $N-k (> 0)$ increases when k is fixed at $k = 0, 2, 3$. The "Correct" rate is a function of N , which has the negative effect on it. The perceived error is one of the key factors influencing the judgment rate (The error judgment is named "false" judgment), which we can't neglect in classified judgment (one-way ANOVA, $F(1,12) = 300.11>4.7472$; $F(2,18)=2160.69>3.5546$; $F(3,24) = 1428.94>3.0088$; $F(4, 30)=1653.7>2.6896$; $p < 0.001$). The factors about "false" judgment rate also are affected by the set size N and the number of same terms k . It reveals the "false" rates decrease with the same item k increasing as the set size N keeps constant. By simulation parameters using maximization the parameter likelihood, we count the probability of all subjects' based on a given model. For every subject and every model, we make use of the models fitting subjects data are shown in Fig.2. According to the fitting effect using the RMS errors, our model (optimal model) (0.1326) fits approximately better than that of MAV model (0.8659). Surprisingly, we notice the proportions of "Correct" that MAV model and MMR model are higher than the optimal model, which does not mean that optimal model worse than MAV model and MMR model. MAV model and MMR model are ideal model, which does not think over the perceived errors. According to the reality and the previous analysis, we have realized perceive error is one of the important factors affecting judgment result. While MAV model and MMR model do not get rid of the effect of the perceived errors. With a view to effect of perceived errors, the optimal modes fit subjects' data best among all the models.

VI. CONCLUSION

To test variability discrimination under sensory noise still satisfy near-optimal Bayes rule, we put forward a model,

which depends on the parameters set size N , the item of the same orientation k and reliabilities of the measurements σ based on Bayesian theory in visual search to judge how many bars with same orientation among a mass of orientations. We find our model of variability discrimination on existing the sameness against the difference matches well the experimental data, is superior to alternative models, and it provides a normative and mathematically quantitative description.

ACKNOWLEDGMENT

This research was supported by the National Natural Science Foundation of China (No.:61374183), the Humanistic and Social Science Foundation of Ministry of Education of China (No.: 12YJA630122), and the Higher Education Institution Key Research Project Plan of Henan Province, China (No.:15A110021).

REFERENCES

- [1] E. A. Wasserman, M. E. Young. "Same-different discrimination: the keel and backbone of thought and reasoning," *Journal of Experimental Psychology Animal Behavior Processes*, vol. 36(1), pp.3-22, 2010.
- [2] M. Giurfa, S. Zhang, A. Jenett, R. Menzel, and M. V. Srinivasan. "The concepts of 'sameness' and 'difference' in an insect," *Nature*, vol. 410, pp. 930-933, 2001.
- [3] J. S. Katz, and A. A. Wright. "Same/different abstract-concept learning by pigeons," *Journal of Experimental Psychology Animal Behavior Processes*, vol.32(1), pp. 80-6, 2006.
- [4] I. M. Pepperberg. "Acquisition of the same/different concept by an african grey parrot: Learning with respect to categories of color, shape, and material," *Learning Behavior*, vol.15(4), pp. 423-432, 1987.
- [5] M. E. Young, E. A. Wasserman. "Entropy detection by pigeons: response to mixed visual displays after same-different discrimination training," *Journal of Experimental Psychology Animal Behavior Processes*, vol. 23(2), pp. 157-70, 1997.
- [6] M. E. Young, M. R. Ellefson, E. A. Wasserman. "Toward a theory of variability discrimination: finding differences," *Behavioural Processes*, vol. 62, pp. 145-155, 2003.
- [7] M. E. Young, E. A. Wasserman, and M. R. Ellefson. "A theory of variability discrimination: Finding differences," *Psychonomic Bulletin Review*, vol. 14(5), pp.805-22, 2007.
- [8] M. E. Young, E. A. Wasserman, Garner KL. "Effects of number of items on the pigeon's discrimination of same from different visual displays," *Journal of Experimental Psychology Animal Behavior Processes*, vol. 23(4), pp. 491-501, 1997.
- [9] M. E. Young, M. R. Ellefson. "The joint contributions of shape and color to variability discrimination," *Learning and Motivation*, vol. 34(1), pp. 52-67, 2003.
- [10] E. A. Wasserman, M. E. Young, R. G. Cook. "Variability discrimination in humans and animals: implications for adaptive action," *American Psychologist*, vol. 59(9), pp. 879-90, 2004.
- [11] V. D. B. Ronald, V. Michael, Kresimir J, and Ma Wj. "Optimal inference of sameness," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109(8), pp. 3178-83, 2012.
- [12] B. J. Gibson. "The Perception of the Visual World," Greenwood Publishing Group, 1953.
- [13] M. W. Daehler and D. Bukatko. "Cognitive Development," Ablex Pub. Corp., 1985.
- [14] W. James. "Principles of Psychology," Wiley, 1984.
- [15] R. A. Rensink. "Change detection," *Annual Review of Psychology*, vol. 53(1), pp. 245-277, 2002.
- [16] W. J. Ma, V. Navalpakkam, J. M. Beck, V. D. B. Ronald, and A. Pouget. "Behavior and neural basis of near-optimal visual search," *Journal of African History*, vol.14(6), pp. 783-790, 2011.
- [17] J. Hawkes. "On the asymptotic behaviour of sample spacings," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 90(2), pp. 293-303, 1981.
- [18] H. Liu, Y. Tang, H. H. Zhang. "A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables," *Computational Statistics and Data Analysis*, vol. 53(4), pp.853-856, 2009.
- [19] W. J. Ma, J. M. Beck, P. E. Latham, A. Pouget. "Bayesian inference with probabilistic population codes," *Nature Neuroscience*, vol. 9(11), pp. 1432-1438, 2006.
- [20] W. J. Ma, M. Rahmati. "Towards a neural implementation of causal inference in cue combination," *Multisensory Research*, vol. 26(1-2), pp. 159-76, 2013.

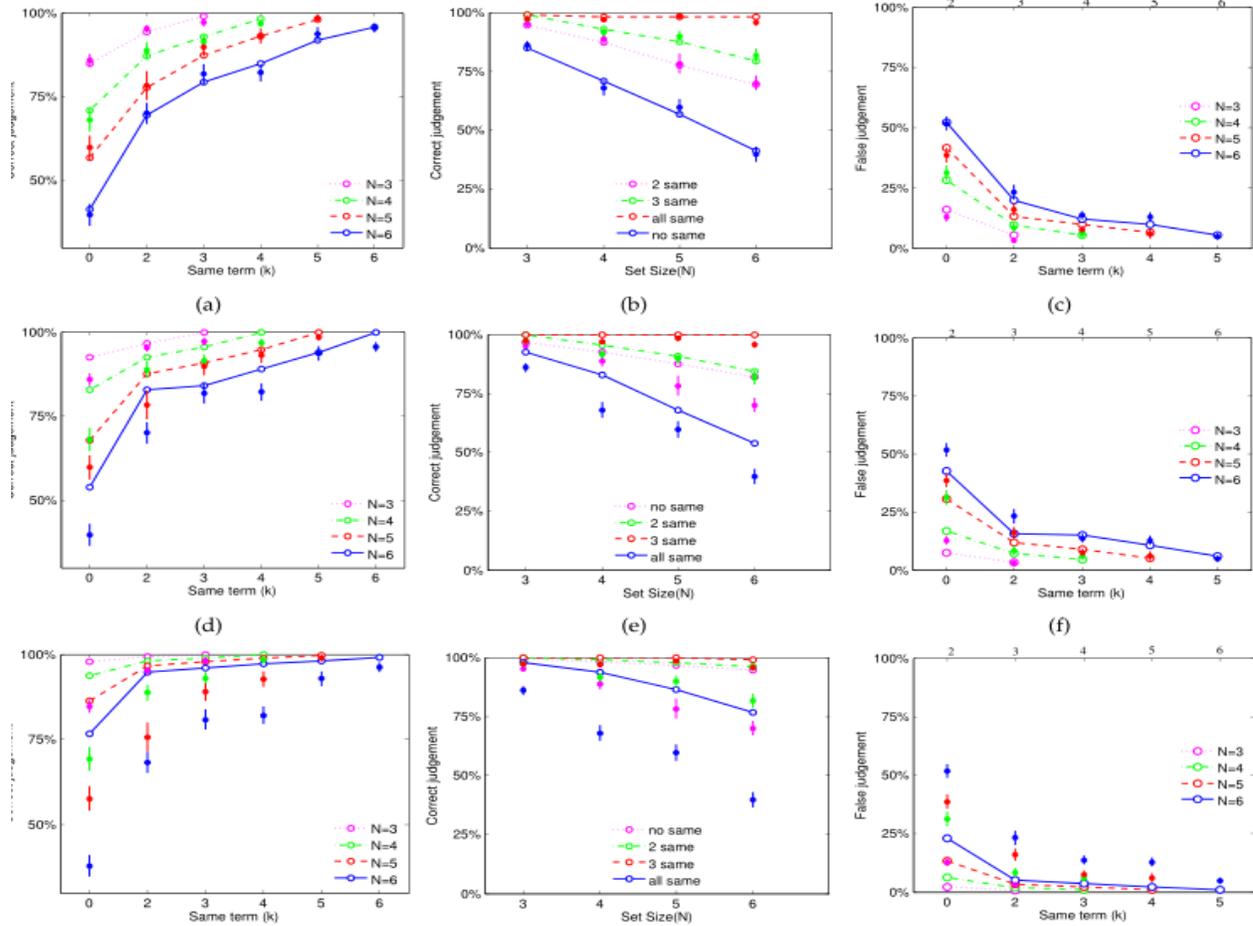


Figure 2. Comparison of models. Circles and error bars denote mean and SEM of subject data. Lines represent model fits. 1. The first line, second line and third line are the fit cases about our model, MAV and MMR models; 2. The first row is the proportion "correct judgment" responses as a function of same items k when set sizes $N=3, 4, 5, 6$; The second row is proportion "correct judgment" responses as a function of set size N when same items $k=0, 2, 3, N$; The third row proportions "False" response is a function of the same terms k when set sizes $N=3, 4, 5, 6$.