USING WELL-BEHAVED THURSTONIAN-TYPE MODELS TO EMULATE REGULAR MINIMALITY

Janne V. Kujala[1] and Ehtibar N. Dzhafarov[2]
[1]*University of Jyväskylä, Agora Center, P.O.Box 35, FI-40014, Jyväskylä, Finland*
[2]*Purdue University, 703 Third Street, West Lafayette, IN 47907-2081, USA*
jvk@iki.fi; ehtibar@purdue.edu

Abstract

Well-behaved Thurstonian-type models are incompatible with psychometric functions that satisfy Regular Minimality and have nonconstant minima. We show, however, that even very simple Thurstonian-type models can provide tolerable approximation to empirical estimates of psychometric functions in which Regular Minimality holds within experimental error and the minima are nonconstant.

Theoretical Preliminaries

Let $\psi : X \times Y \to \; 0,1$ be a *psychometric function* for "same-different" discriminations,

$$\psi(x,y) = \Pr \; x \text{ and } y \text{ are judged to be different} \; ,$$

where the sets X and Y represent stimuli in two distinct observation areas (e.g., in every trial an X-stimulus is presented first, and a Y-stimulus is presented second). We will assume that X and Y are subsets of $\mathbb{R}^n$. We also assume that the psychometric function has unique *points of subjective equality* (PSE) for all $x \in X$ and $y \in Y$, defined by

$$h(x) := \operatorname*{argmin}_{y \in Y} \psi(x,y) \in Y, \qquad g(y) := \operatorname*{argmin}_{x \in X} \psi(x,y) \in X.$$

The psychometric function ψ is said to satisfy *Regular Minimality* (RegMin) if the PSE functions h and g are well-defined and continuous, and if $y = h(x) \Leftrightarrow x = g(y)$ for all $x \in X$ and $y \in Y$. In other words, RegMin means that $h = g^{-1}$ is a homeomorphism. The psychometric function ψ is said to satisfy *Constant Minimality* (ConstMin) if the PSE functions h and g are well-defined and continuous, and if at least one of the minimum level functions is constant (i.e., $\psi(x,h(x)) = c$ for all x or $\psi(g(y),y) = c$ for all y).

In Dzhafarov (2003a, b) and Kujala and Dzhafarov (2009) it was shown that an intuitively plausible class of models called Thurstonian-type models (see Fig. 1) — if they satisfy a very mild smoothness condition called "well-behavedness" — can satisfy RegMin only if ConstMin is also satisfied. However, while RegMin appears to be satisfied in known empirical data, ConstMin is, as a rule, prominently violated.

Thurstone's (1927) original models were formulated for "greater-less" discriminations only. The first Thurstonian model of "same-different" judgements was formulated by Luce and Galanter (1963). The model consists of a bivariate normal distribution $(P(x),Q(y)) \sim N \; (\mu(x),\mu(y)),\Sigma_{x,y}$ of perceptual images, with a covariance matrix

$$\Sigma_{x,y} = \begin{bmatrix} \sigma^2(x) & \rho(x,y)\sigma(x)\sigma(y) \\ \rho(x,y)\sigma(x)\sigma(y) & \sigma^2(y) \end{bmatrix}$$

and the decision rule $P(x) - Q(y) \in \; -c,d \; \Rightarrow$ "same", where μ and σ denote the dependence of the mean and standard deviation on the stimulus value (the same functions for both observation

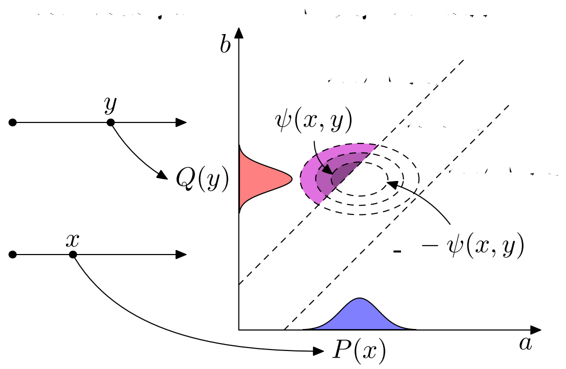

Figure 1: An illustration of a Thurstonian-type model. The physical stimuli x and y map into perceptual random images P and Q, which the "same-different" decision is assumed to be based on.

areas) and ρ is the correlation coefficient depending on x and y in an arbitrary way. This yields the psychometric function

$$\psi(x,y) = 1 - \Phi\left(\frac{d - (\mu(x) - \mu(y))}{\sigma(x,y)}\right) + \Phi\left(\frac{-c - (\mu(x) - \mu_2(y))}{\sigma(x,y)}\right),$$

where $\sigma(x,y) = \sqrt{\sigma^2(x) + \sigma^2(y) - 2\rho(x,y)\sigma(x)\sigma(y)}$, and Φ is the standard normal cumulative distribution function.

If the correlation $\rho(x,y)$ between the images $P(x)$ and $Q(y)$ is allowed to be an arbitrary function, this model does not directly fit within the framework of Thurstonian-type models as defined in Dzhafarov (2003a, b): it bypasses the selective influence requirement (that P is selectively influenced by x and Q by y; see Dzhafarov, 2003c, and Kujala & Dzhafarov, 2008) and specifies directly the distribution of the composite image (P,Q) as depending on the composite stimulus (x,y). Moreover, with arbitrary $\rho(x,y)$ the model becomes universally fittable to data: it is demonstrable that it can be precisely fitted to any finite set $\{\psi(x_i,y_i)\}_{i\in\{1,\ldots,n\}}$. In Kujala and Dzhafarov (2008) it is shown that in a Thurstonian-type model satisfying selective influence the correlation function can be representable as $\rho(x,y) = \sum_{i=1}^m a_i(x)b_i(y)$ for some functions $a_i, b_i : \mathbb{R} \to [-1,1]$ satisfying $\sum_{i=1}^m a_i^2(x) \le 1$ and $\sum_{i=1}^m b_i^2(y) \le 1$.

In the following, we consider a very simple version of a modified Luce-Galanter model. The simplification is in assuming that the normally distributed random images $P(x)$ and $Q(y)$ are uncorrelated, $\rho(x,y) = 0$; the modification is in allowing the means and variances of these random variables to depend on observation area:

$$P(x) \sim N(\mu_1(x), \sigma_1(x)), \qquad Q(y) \sim N(\mu_2(y), \sigma_2(y))$$

for some functions μ_1, μ_2, $\sigma_1 > 0$, and $\sigma_2 > 0$. It is assumed that the observer answers "different" if $|P(x) - Q(y)| > \varepsilon$ (as we allow different mean and variance functions for different observation areas, no generality is lost in assuming the decision rule to be symmetric). This yields the psychometric function

$$\psi(x,y) = 1 - \Phi\left(\frac{\varepsilon - (\mu_1(x) - \mu_2(y))}{\sqrt{\sigma_1^2(x) + \sigma_2^2(y)}}\right) + \Phi\left(\frac{-\varepsilon - (\mu_1(x) - \mu_2(y))}{\sqrt{\sigma_1^2(x) + \sigma_2^2(y)}}\right). \tag{1}$$

To increase the predictive power of this model we impose a further restriction, that

$$\sigma_i^2 = a\mu_i + b, \quad i = 1,2, \tag{2}$$

i.e., we require that the means and variances of each random image be affine functions of each other.

Matrix 1: RegMin holds in a canonical form

	7	9	11	13	15	17	19	21	23
7			.940						
9			.655	.845					
11	.905	.650	.485	.520	.655				
13		.905	.570	.425	.445	.610			
15			.830	.535	.405	.450	.520		
17				.765	.500	.400	.405	.505	
19					.720	.480	.375	.385	.430
21						.680	.485		
23							.660		

Matrix 3: RegMin holds in a noncanonical form

	7	9	11	13	15	17	19	21	23
7			.740						
9			.505	.770					
11	.985	.895	.570	.460	.750				
13		.995	.870	.595	.425	.675			
15			.980	.870	.570	.365	.670		
17				.985	.860	.525	.365	.635	
19					.950	.820	.500	.315	.580
21						.950	.805		
23							.920		

Matrix 11: Apparent RegMin violations

	7	9	11	13	15	17	19	21	23
7			.240						
9			.120	.120					
11	.740	.400	.160	.065	.085				
13		.500	.205	.075	.060	.085			
15			.300	.105	.055	.050	.105		
17				.180	.055	.020	.070	.130	
19					.105	.040	.035	.060	.170
21						.050	.040		
23							.030		

Experiment design

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
y_1			n						
y_2			n	n					
y_3	n	n	n'	$2n$	$2n$				
y_4		n	$2n$	n'	$2n$	$2n$			
y_5			$2n$	$2n$	n'	$2n$	$2n$		
y_6				$2n$	$2n$	n'	$2n$	n	
y_7					$2n$	$2n$	n'	n	n
y_8						n	n		
y_9							n		

Table 1: Top left and right and bottom left: three empirical matrices with the estimated PSE relations $y = h(x)$ and $x = g(y)$ indicated. Bottom right: the experimental design matrix which indicates the number of replications for each stimulus pair (x, y). In ExpA, ExpB, and ExpC, $n' = 2n$ and $n \in \{100\} \cup 250, 300$. In ExpFB, $n = 150$, and balancing of the "same" and "different" trials yields $n' = 8n = 1200$. In ExpA, ExpB, and ExpC, $x_k = y_k = 5.\ \ 2k$ pixels; in ExpFB, $x_k = y_k = 10.\ \ k$ pixels.

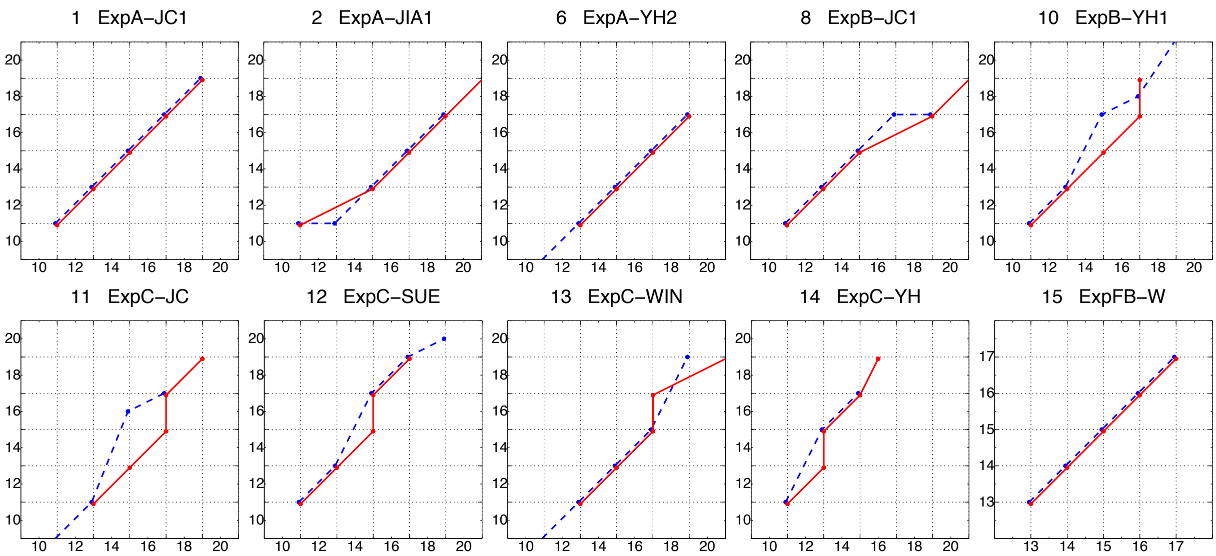

Figure 2: PSE curves $y = h(x)$ (dashed lines) and $x = g(y)$ (solid lines) estimated as the locations of column and row minima (averaged when adjacent cells have the same minimum value) of a representative sample of 10 of the 16 empirical discrimination probability matrices. Simulations indicate that the deviations from coincidence of the two curves are no more frequent than what is expected to happen due to sampling error.

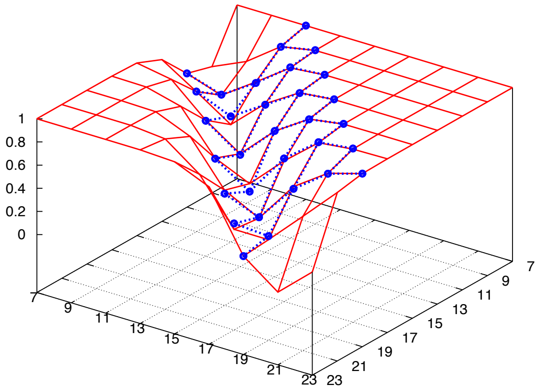

Figure 3: Empirical discrimination probabilities, shown by circles and dotted lines, and the probabilities predicted by the best fit of Model (1)-(2), solid lines. Shown are the matrices that had the largest error of fit (left, $\|\Psi_{\mathrm{pred}} - \Psi_{\mathrm{obs}}\|_{\mathrm{sup}} = 0.070$) and the smallest error of fit (right, $\|\Psi_{\mathrm{pred}} - \Psi_{\mathrm{obs}}\|_{\mathrm{sup}} = 0.015$).

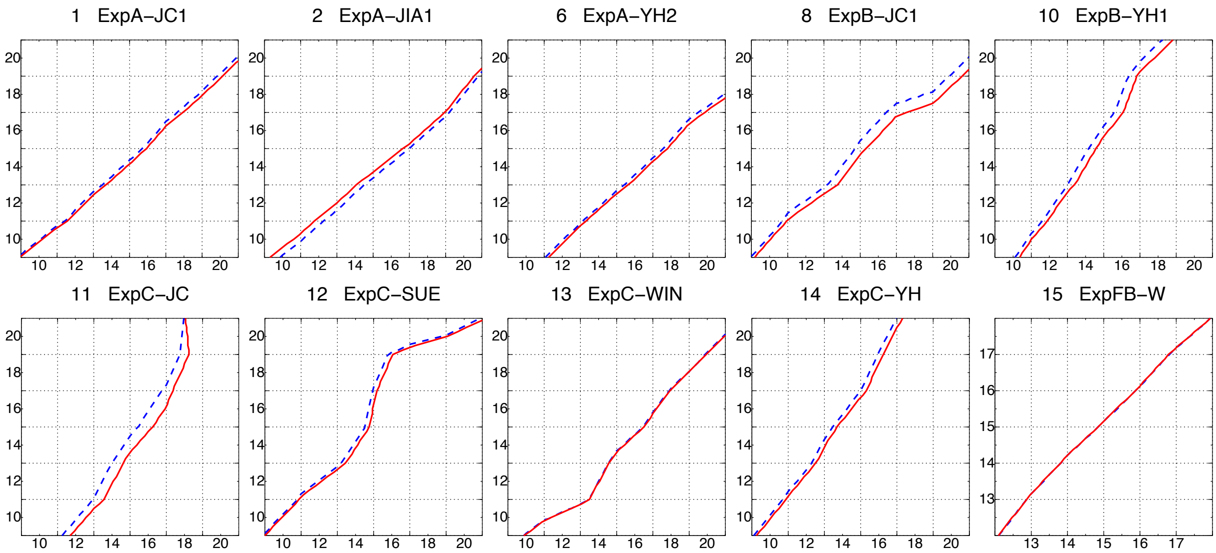

Figure 4: PSE curves $y = h(x)$ (dashed lines) and $x = g(y)$ (solid lines) predicted from the best fits of Model (1)-(2). The model parameters are linearly interpolated between the fitted values at the sample points.

X^2	df	p-value	experiment	subject	#
24.55	12	.017*	ExpA	JC1	1
30.25	12	.003**	ExpA	JIA1	2
22.50	12	.032*	ExpA	YH1	3
18.77	12	.094	ExpA	JC2	4
12.16	12	.433	ExpA	JIA2	5
37.46	12	.000**	ExpA	YH2	6
145.68	72	.000**			

X^2	df	p-value	experiment	subject	#
10.82	12	.544	ExpB	JC2	7
6.98	12	.859	ExpB	JC1	8
14.68	12	.259	ExpB	YH2	9
15.97	12	.193	ExpB	YH1	10
48.46	48	.454			

X^2	df	p-value	experiment	subject	#
11.68	12	.472	ExpC	JC	11
20.35	12	.061	ExpC	SUE	12
16.10	12	.187	ExpC	WIN	13
12.97	12	.371	ExpC	YH	14
61.10	48	.097			

X^2	df	p-value	experiment	subject	#
42.82	12	.000**	ExpFB	W	15
10.77	12	.549	ExpFB	P	16
53.58	24	.000**			

Table 2: Statistical testing of the fits of Model (1)-(2): * denotes $p < .05$ and ** denotes $p < .01$.

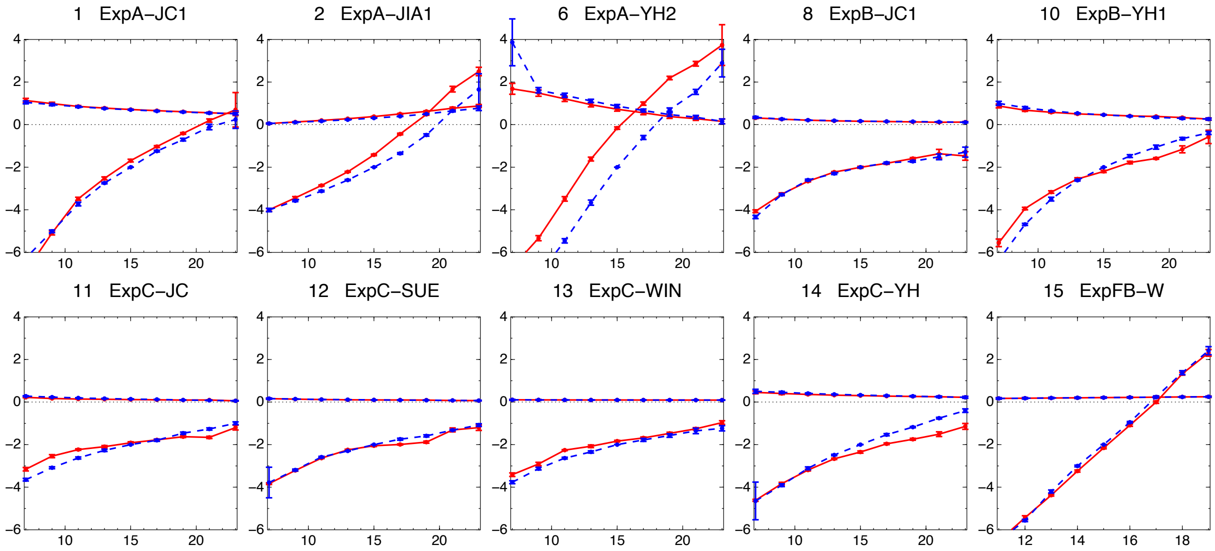

Figure 5: Fitted parameter values of Model 2. Dashed lines indicate the mean $\mu_1(x_i)$ and variance $\sigma_1^2(x_i)$ of $P(x_i)$ and solid lines indicate the mean $\mu_2(y_i)$ and variance $\sigma_2^2(y_i)$ of $Q(y_i)$. For clarity, the parameters are normalized so as to yield $\mu_1(x_5) = -2$. The (very small) error bars indicate the estimated standard deviation of the parameter values over 10 parametric bootstrap replications, providing an estimate of the accuracy of the fitted parameter values (see text).

Fitting a Special Form of Luce-Galanter Model to Empirical Data

We use the experimental data that were presented in Dzhafarov and Colonius (2005) as empirical evidence for RegMin (ExpA, ExpB, ExpC), to which we add the results of an unpublished experiment, with feedback (ExpFB). The stimuli were either line segments (ExpA, ExpFB) or apparent motions (ExpB, ExpC). A subset of a 9×9 matrix of different stimulus pairs (x, y) was used in each experiment. The number of trials per stimulus pair varied between 100 and 1200, depending on the experiment and the position within the matrix (see the bottom right matrix in Table 1).

The data sets being analyzed consist of 16 matrices indicating the participants' discrimination frequencies for each stimulus pair in each experiment run. In every matrix, ConstMin is prominently violated, but RegMin appears to hold within experimental error. Table 1 shows three representative examples of these matrices. Fig. 2 depicts the estimated PSE relations for a representative set of 10 of the 16 matrices. The estimates are crude, so the apparent RegMin violations may seem more significant than they are.

Once a certain parameterization ψ_θ, $\theta \in \mathbb{R}^m$, is chosen, the model parameters θ are fitted by minimizing the discrepancy

$$X^2 = \sum_{x,y} \frac{n(x,y)\ \psi_{\text{obs}}(x,y) - \psi_\theta(x,y)\ ^2}{\psi_\theta(x,y)\ 1 - \psi_\theta(x,y)},$$

where $n(x,y)$ is the number of Bernoulli trials for each data point (x,y). We have fitted Model (1)-(2) to each of the 16 matrices. Fig. 3 shows two examples of the fits, the worst and the best ones. Fig. 4 shows the predicted PSE curves. Table 2 presents the results of statistical testing of the fits. There are a few clear misfits, but considering the great simplicity of the model, its overall fitting power is surprisingly good. Fig. 5 depicts the fitted parameter values and a parametric bootstrap estimate of their accuracy. An unusual feature of these fits is that, except for one subject (data set 2 in Fig. 5), the variance of the perceptual image generally decreases as its mean increases, i.e., the parameter a has a negative value. However, this does not indicate a violation of Weber's law, as the dependence of the means on the stimulus values generally compresses the larger values, counteracting the effects of the decreasing variance.[1]

Looking at Table 2 and Fig. 2, it can be seen that the matrices that were fitted the worst were also those that had the fewest apparent RegMin violations: there are 5 matrices with p-value $< .05$ and 5 matrices with no apparent RegMin violations and these two sets of five matrices have 4 matrices in common. The incompatibility with exact RegMin is not, however, a sufficient explanation of this correlation, as Fig. 4 reveals that in several other cases the model predicts a very close agreement with RegMin.

Acknowledgements

This research has been supported by NSF grant SES 0620446 and AFOSR grant FA9550-09-1-0252 to Purdue University, as well as by the Academy of Finland grant 121855 to University of Jyväskylä.

References

Dzhafarov, E. N. (2003a). Thurstonian-type representations for "same-different" discriminations: Deterministic decisions and independent images. *Journal of Mathematical Psychology*, 47, 208–228.

Dzhafarov, E. N. (2003b). Thurstonian-type representations for "same-different" discriminations: Probabilistic decisions and interdependent images. *Journal of Mathematical Psychology*, 47, 229–243.

Dzhafarov, E. N. (2003c). Selective influence through conditional independence. *Psychometrika*, 68, 7–25.

Dzhafarov, E. N., & Colonius, H. (2005). Psychophysics without physics: A purely psychological theory of Fechnerian Scaling in continuous stimulus spaces. *Journal of Mathematical Psychology*, 49, 1–50.

Kujala, J. V., & Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, 52, 128–144.

Kujala, J. V., & Dzhafarov, E. N. (2009). Regular Minimality and Thurstonian-type modeling. *Journal of Mathematical Psychology*, 53, 486–501.

Luce, R. D., & Galanter, E. (1963). Discrimination. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology* : Vol. 1 (pp. 103–189). New York: Wiley.

Thurstone, L. L. (1927). A law of comparative judgments. *Psychological Review*, 34, 273–286.

[1] The fitting was done with a trust-region based modified Newton algorithm utilizing (automatically computed) first and second derivatives. The minimization was repeated from several different initial values in order to find the global minimum. The statistical testing was done assuming that X^2 is asymptotically distributed as χ^2 with $df =$ #data points − #effective parameters under the null hypothesis that ψ_θ is the true model. In our model we have 9 means in each of the two observations area, plus the parameters a and b in the relations $\sigma_i^2 = a\mu_{i\cdot}\ \ b$ for $i = 1,2$. This yields 20 parameters, of which one can be eliminated (one of the 18 means can be chosen arbitrarily). As the number of data points in each of the 16 matrices is 31, X^2 is distributed as χ^2 with $df = 12$. For additional verification, we have done maximum likelihood estimation, too, and in every case, the results were practically identical. The fits were also evaluated by parametric bootstrap. That is, replicated experiment results were simulated using the observed frequencies as the true probabilities. The replications were then fitted in the same way as the true observations. The resulting distribution of fitted parameter values then was used as an estimate of the expected random error in the fitted parameter values of the true data.