國立臺灣師範大學管理學院管理研究所

碩士論文

Graduate Institute of Management
College of Management
National Taiwan Normal University

Master's Thesis

以量化建模詮釋維度範圍重疊模型

A Tentative Mathematical Interpretation of the Dimensional Range Overlap Model

蔡昀泓

Tsài, Yún-Húng

指導教授：蕭中強　博士

Advisor: Hsiāo, Chūng-Chiáng, Ph.D

中華民國 111 年 9 月

September 2022

Acknowledgement

The author thanks Dr Hsiāo, the advisor, for his patience in guiding the author through the composition of this article, introducing the author to the academic literature on context effects.

This article was originally titled *A Probabilistic Representation of the Model of Dimensional Range Overlap: Perceptual Assimilation and Contrast Induced by Contexts*. The author thanks Dr Hsiāo for abridging it and proposing *Mathematical Modelling in the Interpretation of the Dimensional Range Overlap Model*, which the author then changed to the current title.

Abstract

The author proposes a set of mathematical equations attempting to simulate perceptual assimilation and contrast, which has rarely been quantitatively studied. The author bases the equations on the model of dimensional range overlap, which is the most integrative model available now. Underlying the model is the assumption of various concepts possibly describing a given object which is being evaluated, with differing salience on a dimension, which may be compared to the possible outcomes of a random variable. The author investigates the roles of the location (extremity), scale (ambiguity), skewness, peak and tail behaviour of the distribution of concepts in shifting the best representation of the given object. In addition to the distribution characteristics, the author introduces evaluative volatility, cognitive consumption, and attention to dissimilarities to the discussion.

By means of the equations, the author is hoping to allow for clearer academic communication, and more accurate predictions of people's perception after their exposure to contexts. The author urges future scholars to devise more sophisticated psychological measurement methods and tools, so as to empirically test the equations.

Keywords: assimilation, contrast, context effects, the dimensional range overlap model.

Table of Contents

List of Figures

Introduction

Studies concerning the influence of contextual stimuli on people's perception abound. Examples of such studies are given in the next chapter. The influence, as reported in these studies, is either assimilation or contrast. In either event, the influence is almost always explained in qualitative fashions in the existing literature.

The only exception, to the author's knowledge, is the model of dimensional range overlap, by Chiĕn et al. (2010), which will later be reviewed. However, even that exception did limited quantification. Quantification, or specifically, mathematical formulation, can contribute in that, first, it allows for clearer and consequently more efficient communication of scholarly concepts, which are often, especially in the domain of psychology, rather abstract; and that, second, it has potential to lend insights that can hardly be drawn otherwise, e.g., in this article, effects of skewness of the distribution of concepts on the probability and magnitude of perceptual assimilation and contrast, which have not ever been addressed, and effects of cognitive consumption, and so forth. These findings will mostly be presented on pages 36–66.

Mathematics has been introduced to psychology for decades and has helped produce such impactful work as prospect theory and the modification thereof (Tversky & Kahneman, 1992). The author is attempting to follow this mathematical approach in pursuing a possibly more accurate, integrative model accounting for perceptual assimilation and contrast.

The rest of this article is structured as follows: a compendious review of the accumulated literature, an overview of the model that the author proposes, extensive demonstrations of the model and then discussion.

Literature Review

Biased Evaluation

The very first postulate that Petty and Wegener (1999) made in proposing the elaboration likelihood model is that people are generally disposed to give evaluations that they think or feel that are accurate enough. Yet this postulate, as Petty and Wegener subsequently clarified, does not imply that people's evaluations are generally proofed against contextual stimuli; that is, it is still possible that contextual stimuli bias people's evaluations, leading them to give evaluations that they would find, on becoming conscious of the influence of the stimuli which they then regard as contextual and irrelevant to the focal object, to be inaccurate. This is context effects.

Set/Reset

Numerous studies have been conducted on the effects, which were summarised as assimilation and contrast. Different theories and models have been proposed, trying to explain why it is assimilation rather than contrast that takes place in certain cases and why it is the other that is observed in other cases. Martin (1986) argued that assimilation resulted from use of concepts relating to contextual stimuli in evaluating the focal object, and that contrast resulted from suppression of the use. This is the set/reset theory: setting for assimilation and resetting for contrast. As fair as it may appear, nevertheless, in reviewing the theory, a question crossed the author's mind: Why does suppression of the use necessarily result in contrast? The underlying assumption seems to be existence of an overlap between concepts relating to the focal object and those relating to contextual stimuli; the suppression then means not using or trying not to use concepts falling in the overlap. If the overlap were to be non-existent, then it would be implied, according to the theory, that no contrast could occur. Another question, by nearly the same token, follows: What if concepts relating to the focal object all were to coincide with concepts relating to contextual stimuli? That would imply no concepts left in evaluating the focal object after complete disuse of concepts relating to contextual stimuli unless new concepts were to be generated.

Diligent readers may go and find the reference paper, and come to a third question: Why would people suppress the use of concepts relating to contextual stimuli, if they were unconscious of the stimuli's influence? This is more fundamental than the previous two questions. In fact, in explaining contrast, Martin was knowingly or unknowingly theorising correction of bias, which requires people to be conscious of the bias, and for that reason, is beyond the scope of this article.

Ambiguity, Extremity, and Knowledge

Herein, context effects are the sole interest. It seems that Martin implicitly assumed assimilation to be the default: People set and then reset. If so, then his view appears to differ from that of Herr (1989). Herr hypothesised that assimilation took place if and only if, when people were unconscious of contextual stimuli in evaluating the focal object, all of the following conditions were met: First, the focal object was ambiguous when evaluated alone before any contextual stimuli took effect; second, when provided with focus and evaluated alone, the contextual stimuli aforementioned were at a position not far away from the focal object on the focal dimension; third and last, people had sufficient knowledge to give evaluations, rather than wild guesses. And the third condition was necessary for context effects to be observed, per Herr's hypotheses.

The foregoing hypotheses were based on the postulates that assimilation was observed because the focal object was unconsciously regarded as an existence or occurrence of the sort where contextual stimuli belonged, and that contrast was observed because the focal object was unconsciously compared with contextual stimuli for differences. Actually, Herr was not the first researcher holding this view; see the reference paper for a relatively comprehensive review. Such a view still holds and the author agrees. Herr's unique point was the third condition: the role of knowledge in contextual stimuli's effecting bias. In order to empirically substantiate this point, Herr conducted an experiment where participants were tested on knowledge about a particular product category pertinent to their subsequent evaluations. The participants were then randomly grouped, being primed with different contextual stimuli before evaluating two focal products, one existing and the other fictious. Multiple Student's t tests followed, and suggested no significant differences between the existing and/or fictitious product evaluations given by participants differently primed as long as the participants were relatively unknowledgeable. Herr thereby drew a conclusion.

At this point, the author would like to invite readers to ponder over a question: Did the insignificance as reported necessarily nullify context effects? In reviewing Herr's work, it was found that no experimental manipulation checks were reported; the checks would be obligatory by present academic publishing standards. The lack thereof allowed for the following alternative explanation of the results: It was the diverging within-group product evaluations given by those relatively unknowledgeable that statistically trivialised the between-group significance of the tested effects.

People's knowledge does really matter, but what it does is potentially allowing for more cognitive investment which entails evaluations that are more temporally persistent, as postulated by Petty and Wegener (1999); and potentially reducing the ambiguity of the focal object and contextual stimuli, or equivalently, narrowing their dimensional ranges as defined by Chiĕn et al. (2010), which will be detailed immediately. People are generally supposed to be more knowledgeable about existing than fictitious products.

Dimensional Range Overlap

Thus far two factors deciding whether it is assimilation or contrast that takes place when people are exposed to contextual stimuli have been revealed: the ambiguity of the focal object and the extremity of contextual stimuli. Note that ambiguity and extremity are in the mind of the evaluator. Intuitively, a question arises: Is it possible that the ambiguity of contextual stimuli, in a way much the same as that of the focal object, also have an effect on people's evaluations? According to the model of dimensional range overlap proposed by Chiĕn et al. (2010), the answer is affirmatory. The term *dimensional range* refers to the range of accessible concepts relating to an object on a certain dimension. The model predicts that substantial overlap between the focal and contextual dimensional ranges will lead to assimilation, and that when the overlap is unsubstantial or non-existent, it will be contrast observed instead. Comparing the model with Herr's hypotheses, one should find overall consistence: Ambiguity allows for substantial overlap, and extremity pulls the ranges apart. The main contribution of Chiĕn et al., by developing the model and subsequently conducting a sequence of four manipulation-checked experiments substantiating it, followed by analyses, to the stream of literature on context effects, was perhaps providing an integrated view; meanwhile, for the first time, measuring the magnitude of the effects and tentatively describing by linear regression the quantitative relationship, which was found to be positive, between the signed magnitude of the effects (with positive for assimilation and negative for contrast) and the signed magnitude of the overlap (with positive for existence and negative for non-existence) between the focal and contextual dimensional ranges, both implicitly assumed to be intervals.

Salience as Probability Density

Explicit mentions of the model of dimensional range overlap can actually be traced to Chiĕn and Hsiāo's de facto preliminary work (2001), which may interest studious readers. In the currently reviewed reference paper, Chiĕn et al. noted the possibly unequal nature of the salience of concepts relating to an object but did not proceed to investigate the role of the differing salience in assimilation and contrast. Such an investigation, to the author's knowledge, has not ever been academically conducted. The gap in research notwithstanding, mathematical formulation was already attempted: Hsiāo (2002) compared the salience to probability density and introduced a threshold of salience for concepts to be included in dimensional ranges.

Reciprocity Hypothesis

Before concluding the literature review, it may be worth taking note of the reciprocity hypothesis, by Hsiāo (2002), which was unprecedented. Hsiāo hypothesised and empirically showed via experiments that, on people's evaluations of the focal object being assimilated to or contrasted from their evaluations of contextual stimuli, people's evaluations of the contextual stimuli, on the other hand, nett of their interaction effects, would be assimilated to or contrasted from their evaluations of the focal object, respectively. This result will be taken into consideration in the author's work proposed and presented below.

Research Framework

The author will be resuming part of Hsiāo's work on the mathematical formulation aforementioned. Hsiāo's work, admittedly, was skeletal, and needs development, which the author attempts.

To begin with, the author is making what might have been deemed trivial and implicitly assumed in Hsiāo's work explicitly documented here. First of all, for any object, on any dimension, concepts, if exists, are distributed exactly on a single interval. Second, there exist uncountably many concepts everywhere on the interval. Third, the threshold of salience is equal everywhere on the interval. Fourth and last, Kolmogorov's axioms of probability, which have three, are all satisfied.

Hereafter the author will be working on these assumptions. Of them, it is acknowledged that the second assumption is questionable, but it can be justified, most of the time; see *General Discussion*.

In the present article, the author additionally imposes uniform continuity everywhere and infinite differentiability almost everywhere on the salience of concepts.

This is the end of the list of assumptions on which the following modelling will be based. The developed mathematical model shall be capable of reproducing the experimental observations reported in the past studies and conceptually consistent with the model of dimensional range overlap.

Methods

Distribution Distortion

Let X and Y denote two random variables representing the context-free evaluations of two different objects; X' and Y' their respective, context-biased versions; $\mathbf{D}(\cdot)$ probability density, $\mathbf{E}(\cdot)$ expected value, $\mathbf{P}[\cdot]$ the probability of the specified event, $\mathbb{V}$ the interval supporting both X and Y. The author then defines

$$\mathbf{D}_{X'}\left(x \middle| I_{Y|X}\right) = \frac{\mathbf{D}_X(x)\left(1 + \delta\, I_{Y|X}\, \mathbf{M}_Y(x)\right)}{1 + \delta\, I_{Y|X} \int_{\mathbb{V}} \mathbf{D}_X(v)\, \mathbf{M}_Y(v) \partial v} \quad (1)$$

X describes the focal object, and Y describes some contextual stimulus. $0 < \delta < 1$; the parametre δ may be interpreted as evaluative volatility.

$I_{Y|X}$ is a binary, random indicator:

$$\ln\left(\mathbf{P}\left[I_{Y|X} = 1\right]\right) = \lambda \int_{\mathbb{V}} \mathbf{D}_X(v) \ln\left(\frac{\mathbf{D}_Y(v)}{\mathbf{D}_X(v)}\right) \partial v \quad (2)$$

$I_{Y|X} = 1$ indicates assimilation, and $I_{Y|X} = -1$ indicates contrast. $\lambda > 0$; the parametre λ may be interpreted as attention to dissimilarities, in relation to attention to similarities. Dissimilarity is measured by, within the scope of this article, relative information entropy as introduced by Kullback and Leibler (1951), which is the integral on the right-hand side. Note its asymmetry. Taking the equation presented right above, one may interpret the entropy as the expected extent of surprise resulting from testing the hypothesis that the focal object were an instance or occurrence of the sort where the contextual stimulus belongs. (The hypothesis testing is done in unconsciousness.) The more surprise the hypothesis is expected to bring out, the less likely it is that assimilation will be observed.

$\mathbf{M}(\cdot)$ is defined by the integral below:

$$\mathbf{M}_Y(x) = \int_{\mathbb{V}} \mathbf{D}_Y(v)\, e^{-|v-x|/\alpha} \partial v \tag{3}$$

This integral is indicative of the relative magnitude of the influence of the contextual stimulus on $\mathbf{D}_X(x)$. $\alpha > 0$; the parametre α may be interpreted as the cognitive resources consumed in evaluating the focal object. The cognitive resources available for the evaluation is naturally constrained by knowledge, as aforesaid.

Thus far, Equation (1) has been fully explained. Its psychological implications should be self-explanatory, but still, will be detailed in *General Discussion*. When giving evaluations, a dimensional range will be formed spontaneously. Let $W \in \{X, Y, X', Y'\}$. The dimensional range, as previously defined, is $\mathbb{R}_W = \{w \mid \mathbf{D}_W(w) > \varepsilon\}$. The threshold of salience $\varepsilon > 0$. The parametre ε is individual- and situation-specific.

Best Representation on a Measurement Scale

As demonstrated later, $\mathbb{V}$ does not necessarily have an upper or a lower bound, but in practice, people, when asked to indicate their evaluations on a measurement scale (e.g., a semantic differential scale), denoted by $\widehat{\mathbb{V}}$, are rarely provided with an unbounded interval to consider and select from. Which point, then, will people consider to be the best representation of their evaluation? As far as the author is concerned, there are two different possible mechanisms: The first one is that people will select the most representative concept on the dimension, denoted by w_*, then mapping it to the scale, indicating $\widehat{w_*}$; the second one is that people will map all concepts on the dimension to the scale and decide the representativeness of each point, selecting the most representative on the scale, denoted by $\widehat{w}_*$. The mental map

$W \mapsto \widehat{W}$ under either mechanism preserves the order of concepts on the dimension, but is not necessarily linear; in fact, it cannot be linear if $\widehat{\mathbb{V}}$ is bounded but $\mathbb{V}$ is not, or the other way around. It is the non-linearity that makes $\widehat{w_*} \neq \widehat{w}_*$ in general even if the scale is continuous.

Hereafter the author will be assuming the first. The reason for the choice is that the first is substantially less mentally demanding than is the second, and that people, according to the elaboration likelihood model (Petty & Wegener, 1999), are generally unwilling to ponder over decisions that involve minimal risk. Indicating one's evaluation on a measurement scale is presumably not risky enough for people to be motivated to try investing the cognitive resources needed for the second; even if it is, people may not have all those resources at their disposal because of the time limit or because of distracting contextual stimuli.

It naturally follows, when asked to indicate a range estimate of their evaluations, people, in general, will map the dimensional range $\mathbb{R}_W$ to the scale, indicating $\widehat{\mathbb{R}_W}$.

Best Representation on the Latent Dimension

Assuming the first mental mechanism as just described, the author is now tentatively defining the representativeness of concepts. Let $\mathbb{T}_w$ denote a testing, inspection interval such that $w \in \mathbb{T}_w$, that $\sup(\mathbb{T}_w) - \inf(\mathbb{T}_w) = \beta$ where $\beta > 0$, and that $\mathbf{P}[W \in \mathbb{T}_w \mid W \in \mathbb{R}_W]$ is maximum for the concept w. Then $\mathbf{P}\left[W \in \mathbb{T}_{w_*} \mid W \in \mathbb{R}_W\right] \geq \mathbf{P}[W \in \mathbb{T}_w \mid W \in \mathbb{R}_W]$ for any $w \neq w_*$. The parametre β may be interpreted as the cognitive resources consumed in selecting the most representative concept on a dimension, which shall be distinguished from the parametre α. While people may be motivated to find such a β that w_* is unique, by trial and error (e.g., when they are explicitly asked to mark a single discrete point on a measurement scale), they may not always manage to do so. In the case of multiple w_*, each candidate w_* is equally likely to be selected as the best representation of one's evaluation.

In the upcoming chapter, the author will be considering a number of different probability distributions. In the first section *Analytical Approach*, the author will be exploring some general properties shared by unimodal, symmetric distributions supported on the whole real number set. As will be shown, unimodality may not be preserved after the distortion. The parametre β will be temporarily assumed to be minimal, so the global mode will be taken.

Ideally, the author should find that concept and track it to see what and how model parametres influence the expected sign and magnitude of context effects:

$$d_{Y|X} = \text{sgn}(y_* - x_*) \cdot \mathbf{E}(x'_* - x_*) \tag{4}$$

$d_{Y|X} > 0$ indicates assimilation, and $d_{Y|X} < 0$ indicates contrast; if $y_* = x_*$, then $d_{Y|X}$ is inconclusive. However, the author simply cannot, in general. This is because of the transcendental equations and special functions involved (e.g., the error function, which naturally arises from the cumulative Gaussian probability).

In the second section *Numerical Simulations*, the author, with the aid of computer programmes, will numerically solve for the best representations of the focal object and check the results previously obtained, meanwhile experimenting with various distributions (mostly accommodated by the skewed generalised Student's t distribution). The parametre β will be touched in studying bimodal distributions. The reciprocity hypothesis (see *Literature Review*) can and will be addressed by interchanging symbols.

Results

Analytical Approach

The purpose here is to analytically demonstrate the model proposed and presented in this article. To ensure mathematical tractability, much generality will be lost, inevitably.

As having been announced, unimodal, symmetric distributions supported on the whole real number set will be the interest of the following investigation. The author expects readers to have learnt elementary differential and integral calculus already.

General Properties

Let $\mathbf{D}_X(x) = \mathbf{D}_X(-x)$ and $\mathbf{D}_Y(\xi + y) = \mathbf{D}_Y(\xi - y)$. For $v > 0$, $\partial \mathbf{D}_X(x)/\partial x < 0$, and $\partial \mathbf{D}_Y(\xi + y)/\partial y < 0$. Also, $\mathbb{V} = (-\infty, \infty)$. Without loss of generality, let $\alpha = 1$. Then with little effort, one can derive

$$\begin{aligned}
\mathbf{M}_Y(\xi + x) &= \int_{-\infty}^{\infty} \mathbf{D}_Y(v)\, e^{-|v-\xi-x|} \partial v \\
&= \int_{-\infty}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{-|v-x|} \partial v \\
&= \int_{x}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{x-v} \partial v + \int_{-\infty}^{x} \mathbf{D}_Y(\xi + v)\, e^{v-x} \partial v \\
&= \int_{x}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{x-v} \partial v + \int_{-x}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{-v-x} \partial v \qquad (5) \\
&= (e^{x} + e^{-x}) \int_{0}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{-v} \partial v \\
&\qquad - \int_{0}^{x} \mathbf{D}_Y(\xi + v)\, e^{x-v} \partial v + \int_{-x}^{0} \mathbf{D}_Y(\xi + v)\, e^{-v-x} \partial v \\
&= (e^{x} + e^{-x}) \int_{0}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{-v} \partial v - \int_{0}^{x} \mathbf{D}_Y(\xi + v)\, (e^{x-v} - e^{v-x}) \partial v
\end{aligned}$$

Equation (5) implies that $\mathbf{M}_Y(\xi + x) = \mathbf{M}_Y(\xi - x)$. For $x \geq 0$,

$$\begin{aligned}
\frac{\partial}{\partial x} \mathbf{M}_Y(\xi + x) &= (e^{x} - e^{-x}) \int_{0}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{-v} \partial v - \int_{0}^{x} \mathbf{D}_Y(\xi + v)\, (e^{x-v} + e^{v-x}) \partial v \\
&= (e^{x} - e^{-x}) \int_{x}^{\infty} \mathbf{D}_Y(\xi + v)\, e^{-v} \partial v - e^{-x} \int_{0}^{x} \mathbf{D}_Y(\xi + v)\, (e^{v} + e^{-v}) \partial v \\
&\leq \mathbf{D}_Y(\xi + x) \left((e^{x} - e^{-x}) \int_{x}^{\infty} e^{-v} \partial v - e^{-x} \int_{0}^{x} (e^{v} + e^{-v}) \partial v \right) = 0 \qquad (6)
\end{aligned}$$

This suggests that $\mathbf{M}_Y(x) < \mathbf{M}_Y(\xi)$ for all $x \neq \xi$; that is, the salience of the focal concept ξ (which is the most salient contextual concept) is more seriously influenced than is the salience of any other focal concepts.

For the record, the author presents

$$\frac{\mathbf{M}_Y^2(x)}{\alpha} \leq \int_{\mathbb{V}} \mathbf{D}_Y^2(v)\partial v \tag{7}$$

This holds for all possible distributions, not necessarily unimodal or symmetric. It immediately follows the Cauchy–Schwarz inequality.

$$\frac{\partial}{\partial x}\mathbf{D}_{X'}\left(x \middle| I_{Y|X}\right) \equiv \left(1 + \delta\, I_{Y|X}\, \mathbf{M}_Y(x)\right) \cdot \frac{\partial}{\partial x}\mathbf{D}_X(x) + \delta\, I_{Y|X}\, \mathbf{D}_X(x) \cdot \frac{\partial}{\partial x}\mathbf{M}_Y(x) \tag{8}$$

Without loss of generality, let $\xi > 0$. Then Equivalence (8) can be derived. The equivalence herein is in the sense of both sides being positive, negative, or zero. In the event $I_{Y|X} = 1$, if $x < 0$, then $\partial\mathbf{D}_{X'}(x)/\partial x > 0$; if $x > \xi$, then $\partial\mathbf{D}_{X'}(x)/\partial x < 0$. It is therefore decided that if $\mathbf{D}_X(x)$ is continuously differentiable everywhere, and if $\partial^2\mathbf{D}_{X'}(x)/(\partial x)^2 < 0$ for $x \in (0,\ \xi)$, then $\partial\mathbf{D}_{X'}(x)/\partial x = 0$ at $x = x'_*$ for some unique $x'_* \in (0,\ \xi)$ where the subscripted asterisk is as defined in *Methods*.

In the event $I_{Y|X} = -1$, the author, however, cannot be certain of the behaviour of the context-distorted salience-concept curve if $x < 0$ or $x > \xi$. If $0 \leq x \leq \xi$, then $\partial\mathbf{D}_{X'}(x)/\partial x < 0$; if $x \to -\infty$, then $\mathbf{D}_{X'}(x) \to 0$. Taking these together, the author deduces that for some $x < 0$, $\partial\mathbf{D}_{X'}(x)/\partial x = 0$. Nevertheless, uniqueness of modes is unguaranteed; the distorted distribution can be multimodal.

The possibility of multiple modes of the distorted distribution adds complicacy to the analysis of perceptual contrast. The author does not have the confidence to claim that it is always that $x'_* < 0$. The author is expecting that but that is a mere educated guess, which needs a proof or a disproof. The author attempted both only to realise that neither is within the author's mathematical capability. Thus, hereafter, the author will be silent about contrast and focus the discussion on assimilation till *Illustrative Examples*.

Let $X_\theta = X/\theta; \theta > 0$. The parametre θ may be interpreted as latent ambiguity, which is not to be confused with perceived ambiguity, their difference explained later.

If $\mathbf{D}_{X'}(x)$ is continuously differentiable everywhere and $\partial^2\mathbf{D}_{X'}(x)/(\partial x)^2 < 0$ for $x \in (0,\ \xi)$, then in the event $I_{Y|X} = 1$,

$$\left(\frac{\partial^2}{\partial\theta\partial x}\ln\left(\mathbf{D}_{X_\theta}(x)\right) \lesseqgtr 0, \quad \forall x \in (0, \xi)\right) \quad \Rightarrow \quad \frac{\partial x'_{\theta_*}}{\partial\theta} \lesseqgtr 0 \tag{9}$$

The converse argument is not implied. In the premise of the argument presented right above, the left-hand side of the inequality may be expanded:

$$\begin{aligned}
\frac{\partial^2}{\partial\theta\partial x}\ln\left(\mathbf{D}_{X_\theta}(x)\right) &= \frac{\partial^2}{\partial\theta\partial x}\ln\left(\frac{1}{\theta}\ \mathbf{D}_X(x/\theta)\right) \\
&= \frac{\partial^2}{\partial\theta\partial x}\ln\left(\mathbf{D}_X(x/\theta)\right) \\
&= \frac{\partial}{\partial\theta}\left(\frac{1}{\theta}\cdot\frac{\partial}{\partial u}\ln\left(\mathbf{D}_X(u)\right)\right)\bigg|_{u=x/\theta} \\
&= -\frac{1}{\theta^2}\cdot\frac{\partial}{\partial u}\left(u\cdot\frac{\partial}{\partial u}\ln\left(\mathbf{D}_X(u)\right)\right)\bigg|_{u=x/\theta} \\
&= -\frac{1}{\theta^2}\cdot\left(\frac{\partial}{\partial u}\ln\left(\mathbf{D}_X(u)\right) + u\cdot\frac{\partial^2}{(\partial u)^2}\ln\left(\mathbf{D}_X(u)\right)\right)\bigg|_{u=x/\theta} \\
&= -\frac{u}{\theta^2}\cdot\frac{\partial}{\partial u}\ln\left(\mathbf{D}_X(u)\right)\cdot\left(\frac{1}{u} - \frac{\partial\mathbf{D}_X(u)/\partial u}{\mathbf{D}_X(u)} + \frac{\partial^2\mathbf{D}_X(u)/(\partial u)^2}{\partial\mathbf{D}_X(u)/\partial u}\right)\bigg|_{u=x/\theta}
\end{aligned} \tag{10}$$

Accordingly, if $\partial^2\mathbf{D}_X(x)/(\partial x)^2 < 0$ for all $x \in (0,\ \xi/\theta)$, then $\partial x'_{\theta_*}/\partial\theta > 0$; otherwise, additional calculations are needed. Similarly,

$$\left(\frac{\partial^2}{\partial\theta\partial x} \ln\left(1 + \delta\, \mathbf{M}_{Y_\theta}(x)\right) \lesseqgtr 0, \quad \forall x \in (0,\xi) \right) \quad \Longrightarrow \quad \frac{\partial x'_{\theta_*}}{\partial\theta} \lesseqgtr 0 \tag{11}$$

$$\left(\frac{\partial^2}{\partial\xi\partial x} \ln\left(1 + \delta\, \mathbf{M}_{Y_\theta}(x)\right) \lesseqgtr 0, \quad \forall x \in (0,\xi) \right) \quad \Longrightarrow \quad \frac{\partial x'_{\theta_*}}{\partial\theta} \lesseqgtr 0 \tag{12}$$

$Y_\theta - \xi = (Y - \xi)/\theta$. Again, the author is expanding the derivatives; see the following for the results. For the sake of compactness, the author introduces

$$\mathbf{I}_{Y_\theta}(x) = (e^x - e^{-x}) \int_0^\infty v\, \mathbf{D}_Y(\xi + v)\, e^{-\theta v} \partial v - \int_0^{x/\theta} v\, \mathbf{D}_Y(\xi + v)\, \left(e^{x-\theta v} - e^{\theta v - x}\right) \partial v$$

$$\mathbf{J}_{Y_\theta}(x) = (e^x - e^{-x}) \int_0^\infty e^{-\theta v} \cdot \frac{\partial}{\partial\xi} \mathbf{D}_Y(\xi + v)\, \partial v - \int_0^{x/\theta} \left(e^{x-\theta v} + e^{\theta v - x}\right) \cdot \frac{\partial}{\partial\xi} \mathbf{D}_Y(\xi + v)\, \partial v$$

These functions are auxiliary. Find them in the lines below:

$$\begin{aligned} \frac{\partial^2}{\partial\theta\partial x} \ln\left(1 + \delta\, \mathbf{M}_{Y_\theta}(x)\right) \equiv &-\Bigg(\frac{1}{\delta} \cdot \mathbf{I}_{Y_\theta}(u) - \frac{2u}{\theta^2}\, \mathbf{D}_Y(\xi + u/\theta) \left(\frac{1}{\delta} + \mathbf{M}_{Y_\theta}(u) \right) \\ &+ \left(\int_0^{u/\theta} \mathbf{D}_Y(\xi + v)\, \left(2e^{\theta v} + e^{2u - \theta v} + e^{-\theta v}\right) \partial v \right) \left(\int_{u/\theta}^\infty v\, \mathbf{D}_Y(\xi + v)\, e^{-\theta v} \partial v \right) \\ &+ \left(\int_{u/\theta}^\infty \mathbf{D}_Y(\xi + v)\, e^{-\theta v} \partial v \right) \left(\int_0^{u/\theta} v\, \mathbf{D}_Y(\xi + v)\, \left(2e^{\theta v} - e^{2u - \theta v} - e^{-\theta v}\right) \partial v \right) \Bigg) \Bigg|_{u = x - \xi} \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{\partial^2}{\partial\xi\partial x} \ln\left(1 + \delta\, \mathbf{M}_{Y_\theta}(x)\right) \equiv &\Bigg(\frac{1}{\delta} \cdot \mathbf{J}_{Y_\theta}(u) \\ &+ 2\left(\int_0^{u/\theta} \mathbf{D}_Y(\xi + v)\, \left(e^{\theta v} + e^{-\theta v}\right) \partial v \right) \left(\int_{u/\theta}^\infty e^{-\theta v} \cdot \frac{\partial}{\partial\xi} \mathbf{D}_Y(\xi + v)\, \partial v \right) \\ &- 2\left(\int_{u/\theta}^\infty \mathbf{D}_Y(\xi + v)\, e^{-\theta v} \partial v \right) \left(\int_0^{u/\theta} \left(e^{\theta v} + e^{-\theta v}\right) \cdot \frac{\partial}{\partial\xi} \mathbf{D}_Y(\xi + v)\, \partial v \right) \Bigg) \Bigg|_{u = x - \xi} \end{aligned} \tag{14}$$

Deriving these is conceptually easy but computationally tedious. For the sake of brevity, the author has factored out and omitted several non-negative terms that are irrelevant to the sign of the second-order partial derivatives, and that, when included, would make the equations a mass of symbols that is effectively unreadable. The derivations are not provided since they are potentially distracting and are not mathematically interesting, but they will be given on readers' request. The point of deriving (9)–(14) is to see the fundamentals and help future scholars, if this article is of their interest, advance the author's work on unimodal, symmetric distributions; see *Illustrative Examples*.

The following fact is simple but noteworthy:

$$\mathbf{M}_Y(x) \lesseqgtr \mathbf{E}\big(\mathbf{M}_Y(X)\big) \quad \Leftrightarrow \quad I_{Y|X} \cdot \frac{\partial}{\partial \delta}\mathbf{D}_{X'}\big(x\big|I_{Y|X}\big) \lesseqgtr 0 \tag{15}$$

The proof is straightforward and is left to readers. The argument is valid not only for unimodal, symmetric distributions, but for multimodal, asymmetric distributions as well. The mathematical argument psychologically states that greater evaluative volatility brings about stronger enhancement or weaker suppression of the salience of a focal concept in the event of assimilation, and stronger suppression or weaker enhancement in the event of contrast, if and only if the contextual influence on the salience of the focal concept being considered is greater than the expected contextual influence on the salience of a focal concept.

The author is now stopping the discussion on the locus shift of the best representation, since it is exceedingly difficult, if possible at all, for the author to draw more insights from the general formulae having been derived thus far in this subsection. Discussion on the indicator $I_{Y|X}$ is left to *Illustrative Examples*, too, since no general properties were found in the author's analytical search, and no computational short cuts were discovered. The end of this subsection is a light study of latent and perceived ambiguity, as follows:

Without loss of generality, the author is taking the focal dimensional range: $\mathbb{R}_X \neq \emptyset$.

To begin with, the author bisects the distribution:

$$\mathbf{D}_{X_\theta}(x) = \begin{cases} \mathbf{D}_{X_\theta,+}(x), & \forall x \geq 0 \\ \mathbf{D}_{X_\theta,-}(x), & \forall x \leq 0 \end{cases}$$

$\sup(\mathbb{R}_{X_\theta}) = \mathbf{D}_{X_\theta,+}^{[-1]}(\varepsilon)$ and $\inf(\mathbb{R}_{X_\theta}) = \mathbf{D}_{X_\theta,-}^{[-1]}(\varepsilon)$. The bracketed superscript $[-1]$ indicates inversion. Perceived ambiguity is given by

$$\varphi_X = \mathbf{E}(|X - x_*||X \in \mathbb{R}_X) \cdot \mathbf{P}[X \in \mathbb{R}_X] \tag{16}$$

By the chain rule of differentiation, the author writes

$$\begin{aligned} \frac{\partial}{\partial\theta}\mathbf{D}_{X_\theta,\pm}^{[-1]}(\varepsilon) &= \left(\frac{\partial}{\partial\theta}\mathbf{D}_{X_\theta,\pm}(x)\right)^{-1}\Bigg|_{x=\mathbf{D}_{X_\theta,\pm}^{[-1]}(\varepsilon)} \\ &= \left(\frac{\partial}{\partial\theta}\frac{1}{\theta}\,\mathbf{D}_{X,\pm}(x/\theta)\right)^{-1}\Bigg|_{x=\mathbf{D}_{X_\theta,\pm}^{[-1]}(\varepsilon)} \\ &= -\theta^2\left(\frac{\partial}{\partial x}x\,\mathbf{D}_{X,\pm}(x)\right)^{-1}\Bigg|_{x=\mathbf{D}_{X_\theta,\pm}^{[-1]}(\varepsilon)/\theta} \\ &= -\theta^2\left(\mathbf{D}_{X,\pm}(x) + x\cdot\frac{\partial}{\partial x}\mathbf{D}_{X,\pm}(x)\right)^{-1}\Bigg|_{x=\mathbf{D}_{X_\theta,\pm}^{[-1]}(\varepsilon)/\theta} \end{aligned} \tag{17}$$

Equation (17) suggests that $\partial\mathbf{D}_{X_\theta,+}^{[-1]}(\varepsilon)/\partial\theta \lesseqgtr 0$ if and only if at $x = \mathbf{D}_{X_\theta,+}^{[-1]}(\varepsilon)/\theta$, $\partial\mathbf{D}_{X,+}(x)/\partial x \gtreqless -\mathbf{D}_{X,+}(x)/x$, and that $\partial\mathbf{D}_{X_\theta,-}^{[-1]}(\varepsilon)/\partial\theta \lesseqgtr 0$ if and only if at $x = \mathbf{D}_{X_\theta,-}^{[-1]}(\varepsilon)/\theta$, $\partial\mathbf{D}_{X,-}(x)/\partial x \lesseqgtr -\mathbf{D}_{X,-}(x)/x$. From these it can be deduced, for sufficiently large θ, it is possible that $\partial\mathbf{P}[X_\theta \in \mathbb{R}_{X_\theta}]\,/\,\partial\theta < 0$ and that $\partial\varphi_{X_\theta}/\partial\theta < 0$; this should be intuitive.

Illustrative Examples

Considering the mathematical difficulty in abstracting additional psychologically interesting properties shared by all possible unimodal, symmetric distributions, the author is taking three different special cases: the Laplace, logistic, and Gaussian distributions, which are sorted in the order of decreasing kurtosis, i.e., decreasing allowance for extreme concepts on the dimension.

Hereafter the assumption that $\alpha = 1$, which was made in the previous subsection, is removed. Recall that $\xi > 0$. For the sake of brevity, and also in fear of potential distraction, certain mathematical details will be given only on reader's request.

Laplace Distribution. Consider the following probability distributions:

$$\mathbf{D}_X(x) = \frac{1}{2\theta_X} e^{-|x|/\theta_X} \quad \wedge \quad \mathbf{D}_Y(y) = \frac{1}{2\theta_Y} e^{-|y-\xi|/\theta_Y} \tag{18}$$

Apply Equation (5). For $x \geq 0$,

$$\begin{aligned}
\mathbf{M}_Y(\xi + x) &= \frac{1}{2\theta_Y}\left(\int_x^\infty e^{-(v-x)/\alpha - |v|/\theta_Y} \partial v + \int_{-x}^\infty e^{-(v+x)/\alpha - |v|/\theta_Y} \partial v\right) \\
&= \frac{1}{2\theta_Y}\left(\int_x^\infty e^{-(v-x)/\alpha - v/\theta_Y} \partial v + \int_0^\infty e^{-(v+x)/\alpha - v/\theta_Y} \partial v + \int_{-x}^0 e^{-(v+x)/\alpha + v/\theta_Y} \partial v\right) \\
&= \frac{1}{2\theta_Y}\left(\left(e^{-x/\alpha} + e^{-x/\theta_Y}\right)\int_0^\infty e^{-v/\alpha - v/\theta_Y} \partial v + \int_0^x e^{-v/\alpha + (v-x)/\theta_Y} \partial v\right) \\
&= \frac{1}{2\theta_Y}\left(\frac{e^{-x/\alpha} + e^{-x/\theta_Y}}{1/\theta_Y + 1/\alpha} + \frac{e^{-x/\alpha} - e^{-x/\theta_Y}}{1/\theta_Y - 1/\alpha}\right) \\
&= \alpha \cdot \frac{\alpha e^{-x/\alpha} - \theta_Y e^{-x/\theta_Y}}{\alpha^2 - \theta_Y^2}
\end{aligned} \tag{19}$$

By symmetry, the author immediately sees

$$\mathbf{M}_Y(x) = \alpha \cdot \frac{\alpha e^{-|x-\xi|/\alpha} - \theta_Y e^{-|x-\xi|/\theta_Y}}{\alpha^2 - \theta_Y^2} \tag{20}$$

The magnitude reaches its maximum at $x = \xi$:

$$\mathbf{M}_Y(x) \leq \frac{\alpha}{\alpha + \theta_Y} \tag{21}$$

Below is the regularising integral in the denominator of the right-hand-side fraction in

Equation (1), which has an elementary mathematical closed form:

$$\int_{-\infty}^{\infty} \mathbf{D}_X(v)\,\mathbf{M}_Y(v)\partial v$$

$$= \frac{\alpha}{2\theta_X(\alpha^2-\theta_Y^2)} \int_{-\infty}^{\infty} e^{-|v|/\theta_X}\left(\alpha e^{-|v-\xi|/\alpha} - \theta_Y e^{-|v-\xi|/\theta_Y}\right)\partial v$$

$$= \frac{\alpha}{2\theta_X(\alpha^2-\theta_Y^2)}\left(\int_{2\xi}^{\infty} e^{-v/\theta_X}\left(\alpha e^{-(v-\xi)/\alpha} - \theta_Y e^{-(v-\xi)/\theta_Y}\right)\partial v\right.$$

$$+\int_{\xi}^{2\xi} e^{-v/\theta_X}\left(\alpha e^{-(v-\xi)/\alpha} - \theta_Y e^{-(v-\xi)/\theta_Y}\right)\partial v$$

$$+\int_{0}^{\xi} e^{-v/\theta_X}\left(\alpha e^{(v-\xi)/\alpha} - \theta_Y e^{(v-\xi)/\theta_Y}\right)\partial v$$

$$\left.+\int_{-\infty}^{0} e^{v/\theta_X}\left(\alpha e^{(v-\xi)/\alpha} - \theta_Y e^{(v-\xi)/\theta_Y}\right)\partial v\right)$$

$$= \frac{\alpha}{2\theta_X(\alpha^2-\theta_Y^2)}\left(\left(e^{\xi/\theta_X}+e^{-\xi/\theta_X}\right)\int_{\xi}^{\infty} e^{-v/\theta_X}\left(\alpha e^{-v/\alpha} - \theta_Y e^{-v/\theta_Y}\right)\partial v\right.$$

$$\left.+e^{-\xi/\theta_X}\int_{0}^{\xi}\left(e^{v/\theta_X}+e^{-v/\theta_X}\right)\left(\alpha e^{-v/\alpha} - \theta_Y e^{-v/\theta_Y}\right)\partial v\right)$$

$$= \frac{\alpha}{2\theta_X(\alpha^2-\theta_Y^2)}\left(\left(e^{\xi/\theta_X}+e^{-\xi/\theta_X}\right)\left(\alpha\cdot\frac{e^{-\xi/\alpha-\xi/\theta_X}}{1/\theta_X+1/\alpha} - \theta_Y\cdot\frac{e^{-\xi/\theta_Y-\xi/\theta_X}}{1/\theta_X+1/\theta_Y}\right)\right.$$

$$+e^{-\xi/\theta_X}\left(\alpha e^{-\xi/\alpha}\left(\frac{e^{\xi/\theta_X}}{1/\theta_X-1/\alpha} - \frac{e^{-\xi/\theta_X}}{1/\theta_X+1/\alpha}\right) - \frac{2}{1/\theta_X^2-1/\alpha^2}\right.$$

$$\left.\left.-\theta_Y e^{-\xi/\theta_Y}\left(\frac{e^{\xi/\theta_X}}{1/\theta_X-1/\theta_Y} - \frac{e^{-\xi/\theta_X}}{1/\theta_X+1/\theta_Y}\right) + \frac{2}{1/\theta_X^2-1/\theta_Y^2}\right)\right)$$

$$= \frac{\alpha}{\alpha^2-\theta_Y^2}\left(\frac{\alpha^2}{\alpha-\theta_X}e^{-\xi/\alpha} - \frac{\theta_Y^2}{\theta_Y-\theta_X}e^{-\xi/\theta_Y}\right) + \frac{\alpha\theta_X^3}{(\alpha^2-\theta_X^2)(\theta_Y^2-\theta_X^2)}e^{-\xi/\theta_X} \qquad (22)$$

Plugging this into Equation (1) yields

$\mathbf{D}_{X'}\left(x \middle| I_{Y|X}\right)/\mathbf{D}_X(x) =$

$$\frac{\dfrac{\alpha^2-\theta_Y^2}{\alpha}+\delta\, I_{Y|X}\left(\alpha e^{-|x-\xi|/\alpha}-\theta_Y e^{-|x-\xi|/\theta_Y}\right)}{\dfrac{\alpha^2-\theta_Y^2}{\alpha}+\dfrac{\alpha^2}{\alpha-\theta_X}e^{-\xi/\alpha}-\dfrac{\theta_Y^2}{\theta_Y-\theta_X}e^{-\xi/\theta_Y}+\dfrac{\theta_X^3(\alpha^2-\theta_Y^2)}{(\alpha^2-\theta_X^2)(\theta_Y^2-\theta_X^2)}e^{-\xi/\theta_X}} \tag{23}$$

By Equation (23), if $x < \xi$,

$$x < 0 \quad \Longrightarrow \quad \frac{\partial}{\partial x}\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)$$

$$\equiv \quad \frac{\alpha+\theta_Y}{\alpha\delta}+I_{Y|X}\cdot\frac{(\alpha+\theta_X)e^{(x-\xi)/\alpha}-(\theta_Y+\theta_X)e^{(x-\xi)/\theta_Y}}{\alpha-\theta_Y} \tag{24}$$

$$x > 0 \quad \Longrightarrow \quad \frac{\partial}{\partial x}\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)$$

$$\equiv -\frac{\alpha+\theta_Y}{\alpha\delta}-I_{Y|X}\cdot\frac{(\alpha-\theta_X)e^{(x-\xi)/\alpha}-(\theta_Y-\theta_X)e^{(x-\xi)/\theta_Y}}{\alpha-\theta_Y} \tag{25}$$

From (24) and (25), it can be deduced

$$\lim_{x\to 0_+}\frac{\partial}{\partial x}\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)=\lim_{x\to 0_-}\frac{\partial}{\partial x}\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)$$

$$\Leftrightarrow \quad \frac{\alpha+\theta_Y}{\alpha\delta}+I_{Y|X}\cdot\frac{\alpha e^{-\xi/\alpha}-\theta_Y e^{-\xi/\theta_Y}}{\alpha-\theta_Y}=0 \tag{26}$$

The argument right above implies existence of a jump discontinuity of $\partial\mathbf{D}_{X'}(x)/\partial x$ at $x = 0$, in the event $I_{Y|X} = 1$, unless $\theta_Y = \alpha$, since $\partial\theta e^{-\xi/\theta}/\partial\theta > 0$ for $\xi > 0$; the equation on the right-hand side of the argument does not hold in general.

In the event $I_{Y|X} = 1$, from Equivalence (25), the author deduces that for sufficiently ambiguous focal object, i.e., sufficiently large θ_X, and for sufficient volatility, i.e., sufficiently large δ, $\lim_{x\to 0_+}\partial\mathbf{D}_{X'}(x)/\partial x > 0$. Since $\partial\mathbf{D}_{X'}(x)/\partial x > 0$ for $x < 0$, and $\partial\mathbf{D}_{X'}(x)/\partial x < 0$ for $x > \xi$, if $\lim_{x\to 0_+}\partial\mathbf{D}_{X'}(x)/\partial x > 0$, then it is decided that $x'_* > 0$; this means that the best representation of the focal object does shift as predicted by the past research on context effects. If $x'_* > 0$, then x'_* satisfies the following equation: $\partial\mathbf{D}_{X'}(x)/\partial x = 0$.

The equation $\partial \mathbf{D}_{X'}(x)/\partial x = 0$ does not have a general solution; in fact, in certain cases, it may not even be solvable at all. For example, if $\theta_Y = \theta_X = \theta$, then in the event $I_{Y|X} = 1$, for $x \in (0,\ \xi)$, it is decided that $\partial \mathbf{D}_{X'}(x)/\partial x > 0$, so $x'_* = 0$:

$$\frac{\partial}{\partial x}\mathbf{D}_{X'}(x) \equiv -\frac{\alpha+\theta}{\alpha\delta} - e^{(x-\xi)/\alpha} \tag{27}$$

Presented below are four special cases where in the event $I_{Y|X} = 1$, for $x \in (0,\ \xi)$, the equation $\partial \mathbf{D}_{X'}(x)/\partial x = 0$ may have elementary solutions in radicals:

$$\frac{\theta_Y}{\alpha} = \frac{1}{2} \quad \Longrightarrow \quad e^{(x_{c,\pm}-\xi)/\alpha} = \frac{\alpha-\theta_X}{\alpha-2\theta_X} \pm \sqrt[2]{\left(\frac{\alpha-\theta_X}{\alpha-2\theta_X}\right)^2 + \frac{3\alpha/2\delta}{\alpha-2\theta_X}} \tag{28}$$

$$\frac{\theta_Y}{\alpha} = 2 \quad \Longrightarrow \quad e^{(x_{c,\pm}-\xi)/2\alpha} = \frac{\alpha-\theta_X/2}{\alpha-\theta_X} \pm \sqrt[2]{\left(\frac{\alpha-\theta_X/2}{\alpha-\theta_X}\right)^2 + \frac{3\alpha/\delta}{\alpha-\theta_X}} \tag{29}$$

$$\left(\frac{\theta_Y}{\alpha} = \frac{1}{3} \quad \wedge \quad \frac{9\delta^2}{16\alpha^2}\cdot\frac{(\alpha-\theta_X)^3}{\alpha-3\theta_X} < 1\right) \Longrightarrow$$

$$e^{(x_{c,\#}-\xi)/\alpha} = \sqrt[3]{\frac{4\alpha/3\delta}{\alpha-3\theta_X} + \sqrt[2]{\left(\frac{4\alpha/3\delta}{\alpha-3\theta_X}\right)^2 - \left(\frac{\alpha-\theta_X}{\alpha-3\theta_X}\right)^3}}$$

$$+\sqrt[3]{\frac{4\alpha/3\delta}{\alpha-3\theta_X} - \sqrt[2]{\left(\frac{4\alpha/3\delta}{\alpha-3\theta_X}\right)^2 - \left(\frac{\alpha-\theta_X}{\alpha-3\theta_X}\right)^3}} \tag{30}$$

$$\left(\frac{\theta_Y}{\alpha} = 3 \quad \wedge \quad \frac{\delta^2}{16\alpha^2}\cdot\frac{(\alpha-\theta_X/3)^3}{\alpha-\theta_X} < 1\right) \Longrightarrow$$

$$e^{(x_{c,\#}-\xi)/3\alpha} = \sqrt[3]{\frac{4\alpha/\delta}{\alpha-\theta_X} + \sqrt[2]{\left(\frac{4\alpha/\delta}{\alpha-\theta_X}\right)^2 - \left(\frac{\alpha-\theta_X/3}{\alpha-\theta_X}\right)^3}}$$

$$+\sqrt[3]{\frac{4\alpha/\delta}{\alpha-\theta_X} - \sqrt[2]{\left(\frac{4\alpha/\delta}{\alpha-\theta_X}\right)^2 - \left(\frac{\alpha-\theta_X/3}{\alpha-\theta_X}\right)^3}} \tag{31}$$

$$\left(\frac{\theta_Y}{\alpha}=\frac{1}{3} \quad \wedge \quad \frac{9\delta^2}{16\alpha^2}\cdot\frac{(\alpha-\theta_X)^3}{\alpha-3\theta_X}\geq 1\right) \quad \Rightarrow$$

$$e^{(x_{\mathrm{c},k}-\xi)/\alpha}=-2\cdot\sqrt[2]{\frac{\alpha-\theta_X}{\alpha-3\theta_X}}\cos\left(\frac{2k\pi}{3}+\frac{1}{3}\arccos\left(\frac{4\alpha}{3\delta}\cdot\sqrt[2]{\frac{\alpha-3\theta_X}{(\alpha-\theta_X)^3}}\right)\right) \tag{32}$$

$$\left(\frac{\theta_Y}{\alpha}=3 \quad \wedge \quad \frac{\delta^2}{16\alpha^2}\cdot\frac{(\alpha-\theta_X/3)^3}{\alpha-\theta_X}\geq 1\right) \quad \Rightarrow$$

$$e^{(x_{\mathrm{c},k}-\xi)/3\alpha}=-2\cdot\sqrt[2]{\frac{\alpha-\theta_X/3}{\alpha-\theta_X}}\cos\left(\frac{2k\pi}{3}+\frac{1}{3}\arccos\left(\frac{4\alpha}{\delta}\cdot\sqrt[2]{\frac{\alpha-\theta_X}{(\alpha-\theta_X/3)^3}}\right)\right) \tag{33}$$

Let x_{c} denote the plausible solutions; $k\in\{1,2,3\}$. Deriving formulae (28) and (29) should have been covered in secondary school. (30)–(33) are for the talented.

The author is now studying these cases in the hope of gaining more insights. The author begins with the cases where $\theta_Y=\alpha/2$ and $\theta_Y=2\alpha$. For analytical convenience, assume that $\theta_X\neq\alpha/2$ and that $\theta_X\neq\alpha$. In the event $I_{Y|X}=1$, (35)–(44) can be deduced. With hindsight, the author begins by deriving

$$\frac{\partial^2}{(\partial x)^2}\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)\equiv\frac{I_{Y|X}}{\alpha-\theta_Y}\left(\left(1-\frac{\theta_X}{\theta_Y}\right)e^{(x-\xi)/\theta_Y}-\left(1-\frac{\theta_X}{\alpha}\right)e^{(x-\xi)/\alpha}\right) \tag{34}$$

This directly follows Equivalence (25). In the case where $\theta_Y=\alpha/2$, if $\lim_{x\to 0_+}\partial\mathbf{D}_{X'}(x)/\partial x>0$, then Inequality (35) is implied. (35) mathematically implies (36); this can be proved by minimising the right-hand side of (35) while fixing the parametre δ. Inequality (36) implies (37), which guarantees that x_{c} is real. (38) directly follows (35). From (38), it can be deduced that $x_{\mathrm{c},-}<0$; hence, $x_{\mathrm{c},-}$ is invalid. $0<x_{\mathrm{c},+}<\xi$ by Inequalities (36) and (38), so $x_{\mathrm{c},+}$ is valid. By Equivalence (34), $\partial^2\mathbf{D}_{X'}(x)/(\partial x)^2\lesseqgtr 0$ at $x=x_{\mathrm{c},\pm}$; accordingly, $x'_*=x_{\mathrm{c},+}$. In summary, in the case where $\theta_Y=\alpha/2$, if $\lim_{x\to 0_+}\partial\mathbf{D}_{X'}(x)/\partial x>0$, then it is guaranteed that $x'_*=x_{\mathrm{c},+}$.

$$\left(\frac{\theta_Y}{\alpha}=\frac{1}{2} \qquad \wedge \qquad \lim_{x\to 0_+}\frac{\partial}{\partial x}\mathbf{D}_{X'}(x)>0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha}>\frac{(3/2\delta)/2+e^{-\xi/\alpha}-e^{-2\xi/\alpha}/2}{e^{-\xi/\alpha}-e^{-2\xi/\alpha}} \tag{35}$$

$$\geq 1+\frac{3}{2\delta}+\sqrt[2]{\frac{3}{2\delta}\left(1+\frac{3}{2\delta}\right)}=\frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{36}$$

$$\Longrightarrow \quad \delta \geq \frac{3}{2}\alpha\cdot\frac{2\theta_X-\alpha}{(\alpha-\theta_X)^2} \tag{37}$$

$$\wedge \quad e^{(x_{\mathrm{c},-}-\xi)/\alpha}<e^{-\xi/\alpha}<e^{(x_{\mathrm{c},+}-\xi)/\alpha} \tag{38}$$

$$\Longrightarrow \quad x'_*=x_{\mathrm{c},+} \tag{39}$$

In the case where $\theta_Y = 2\alpha$, the reasoning is analogous:

$$\left(\frac{\theta_Y}{\alpha}=2 \qquad \wedge \qquad \lim_{x\to 0_+}\frac{\partial}{\partial x}\mathbf{D}_{X'}(x)>0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha}>\frac{2\cdot(3/2\delta)+2e^{-\xi/2\alpha}-e^{-\xi/\alpha}}{e^{-\xi/2\alpha}-e^{-\xi/\alpha}} \tag{40}$$

$$\geq 2\left(1+\frac{3}{\delta}+\sqrt[2]{\frac{3}{\delta}\left(1+\frac{3}{\delta}\right)}\right)=\frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{41}$$

$$\Longrightarrow \quad \delta \geq 3\alpha\cdot\frac{\theta_X-\alpha}{(\alpha-\theta_X/2)^2} \tag{42}$$

$$\wedge \quad e^{(x_{\mathrm{c},-}-\xi)/2\alpha}<e^{-\xi/2\alpha}<e^{(x_{\mathrm{c},+}-\xi)/2\alpha} \tag{43}$$

$$\Longrightarrow \quad x'_*=x_{\mathrm{c},+} \tag{44}$$

The author’s next focus of study is the cases where $\theta_Y = \alpha/3$ and $\theta_Y = 3\alpha$. The present cases will be resumed later. In the following analyses, assume that $\theta_X \neq \alpha/3$ and that $\theta_X \neq \alpha$. In addition, in the presence of multiple real roots, the author assumes no repeated roots for analytical convenience.

In the case where $\theta_Y = \alpha/3$, if $\lim_{x\to 0_+} \partial\mathbf{D}_{X'}(x)/\partial x > 0$, then Inequality (45) is implied. (45) mathematically implies (46); this can be proved by minimising the right-hand side of (45) while fixing the parametre δ. Inequality (46) implies (47), which implies more than one plausible real solution x_c. Inequalities (48) directly follow (45). $e^{(x_{\mathrm{c},3}-\xi)/3\alpha} < 0$, so $x_{\mathrm{c},3}$ is non-existent. From inequalities (48), it can be deduced that $x_{\mathrm{c},2} < 0$, so $x_{\mathrm{c},2}$ is invalid. By (46) and (48), $0 < x_{\mathrm{c},1} < \xi$, so $x_{\mathrm{c},1}$ is valid. By Equivalence (34), the author finds that $\partial^2 \mathbf{D}_{X'}(x)/(\partial x)^2 < 0$ at $x = x_{\mathrm{c},1}$, so $x'_* = x_{\mathrm{c},1}$:

$$\left(\frac{\theta_Y}{\alpha} = \frac{1}{3} \quad \wedge \quad \lim_{x\to 0_+} \frac{\partial}{\partial x}\mathbf{D}_{X'}(x) > 0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha} > \frac{(8/3\delta)/3 + e^{-\xi/\alpha} - e^{-3\xi/\alpha}/3}{e^{-\xi/\alpha} - e^{-3\xi/\alpha}} \tag{45}$$

$$\geq 1 + \frac{8}{3\delta}\cos\left(\frac{1}{3}\arccos\left(\frac{3\delta}{4}\right)\right) = \frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{46}$$

$$\Longrightarrow \quad \delta \geq \frac{4}{3}\alpha \cdot \sqrt[2]{\frac{\alpha - 3\theta_X}{(\alpha - \theta_X)^3}} \tag{47}$$

$$\wedge \quad e^{(x_{\mathrm{c},2}-\xi)/\alpha} < e^{-\xi/\alpha} < e^{(x_{\mathrm{c},1}-\xi)/\alpha} \tag{48}$$

$$\Longrightarrow \quad x'_* = x_{c,1} \tag{49}$$

In the case where $\theta_Y = 3\alpha$, the reasoning is analogous:

$$\left(\frac{\theta_Y}{\alpha} = 3 \quad \wedge \quad \lim_{x\to 0_+} \frac{\partial}{\partial x}\mathbf{D}_{X'}(x) > 0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha} > \frac{3\cdot(8/3\delta) + 3e^{-\xi/3\alpha} - e^{-\xi/\alpha}}{e^{-\xi/3\alpha} - e^{-3\xi/\alpha}} \tag{50}$$

$$\geq 3\left(1 + \frac{8}{\delta}\cos\left(\frac{1}{3}\arccos\left(\frac{\delta}{4}\right)\right)\right) = \frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{51}$$

$$\Longrightarrow \quad \delta \geq 4\alpha \cdot \sqrt[2]{\frac{\alpha - \theta_X}{(\alpha - \theta_X/3)^3}} \tag{52}$$

$$\wedge \quad e^{(x_{\mathrm{c},2}-\xi)/3\alpha} < e^{-\xi/3\alpha} < e^{(x_{\mathrm{c},1}-\xi)/3\alpha} \tag{53}$$

$$\Longrightarrow \quad x'_* = x_{\mathrm{c},1} \tag{54}$$

At this point, the author considers it to be appropriate to bring these four cases together and see, after a few minutes of calculations, that in all of these cases, as long as $\lim_{x\to 0_+} \partial \mathbf{D}_{X'}(x)/\partial x > 0$ in the event $I_{Y|X} = 1$, $\partial x'_*/\partial \xi > 0$, $\partial x'_*/\partial \theta_X > 0$, $\partial x'_*/\partial \alpha < 0$, and $\partial x'_*/\partial \delta > 0$; this is all as expected. Also, the critical focal object latent ambiguity $\theta_{X,\mathrm{c}}$ (which is given in (36), (41), (46), and (51)) is such that $\partial \theta_{X,\mathrm{c}}/\partial \delta < 0$.

Thus far the author has restricted the discussion to the focal object. Inquisitive readers may be curious about the posterior best representation of the contextual stimulus. That is where the reciprocity hypothesis concerns; see *Literature Review*. As mentioned in *Methods*, y'_* can be located simply by interchanging symbols; specifically, $\xi - x'_* \mapsto y'_*$ and $\theta_X \mapsto \theta_Y$. Accordingly, for $y'_* < \xi$, a sufficient condition is $\lim_{y\to \xi_-} \partial \mathbf{D}_{X'}(y)/\partial y > 0$; under that condition, the following can be derived:

$$\frac{\theta_X}{\alpha} = \frac{1}{2} \implies e^{-y'_*/\alpha} = \frac{\alpha - \theta_Y}{\alpha - 2\theta_Y} + \sqrt[2]{\left(\frac{\alpha - \theta_Y}{\alpha - 2\theta_Y}\right)^2 + \frac{3\alpha/2\delta}{\alpha - 2\theta_Y}} \tag{55}$$

$$\frac{\theta_X}{\alpha} = 2 \implies e^{-y'_*/2\alpha} = \frac{\alpha - \theta_Y/2}{\alpha - \theta_Y} + \sqrt[2]{\left(\frac{\alpha - \theta_Y/2}{\alpha - \theta_Y}\right)^2 + \frac{3\alpha/\delta}{\alpha - \theta_Y}} \tag{56}$$

$$\frac{\theta_X}{\alpha} = \frac{1}{3} \implies e^{-y'_*/\alpha} = 2 \cdot \sqrt[2]{\frac{\alpha - \theta_Y}{\alpha - 3\theta_Y}} \cos\left(\frac{1}{3}\arccos\left(\frac{4\alpha}{3\delta} \cdot \sqrt[2]{\frac{\alpha - 3\theta_Y}{(\alpha - \theta_Y)^3}}\right) - \frac{\pi}{3}\right) \tag{57}$$

$$\frac{\theta_X}{\alpha} = 3 \implies e^{-y'_*/3\alpha} = 2 \cdot \sqrt[2]{\frac{\alpha - \theta_Y/3}{\alpha - \theta_Y}} \cos\left(\frac{1}{3}\arccos\left(\frac{4\alpha}{\delta} \cdot \sqrt[2]{\frac{\alpha - \theta_Y}{(\alpha - \theta_Y/3)^3}}\right) - \frac{\pi}{3}\right) \tag{58}$$

These formulae derived, the author, however, is not studying the relationship between x'_* and y'_* in this subsubsection. To meaningfully study their relationship, i.e., to see whether it is always true, at least in some cases, that $x'_* < y'_*$ when $x'_* > 0$ and $y'_* < \xi$, with the results at hand, it would be necessary to derive an explicit formula for y'_* in some case where $\theta_X > 4\alpha$, but that is already proved to be intractable; refer to the Abel–Ruffini theorem. With the results at hand, i.e., cases where $\theta_X, \theta_Y \in \{\alpha/3,\ \alpha/2,\ 2\alpha,\ 3\alpha\}$, by Inequalities (36), (41), (46), and (51), $x'_* > 0$ implies $y'_* = \xi$, and $y'_* < \xi$ implies $x'_* = 0$.

The foregoing analyses are dedicated to perceptual assimilation. The author is now interested in perceptual contrast instead. To begin with, note that $x'_* \notin (0,\ \xi)$. This directly follows Equivalence (8). Next, the author notes

$$x > \xi \quad \Longrightarrow \quad \frac{\partial}{\partial x}\mathbf{D}_{X'}\left(x \middle| I_{Y|X}\right)$$
$$\equiv -\frac{\alpha + \theta_Y}{\alpha\delta} - I_{Y|X} \cdot \frac{(\alpha + \theta_X)e^{(\xi - x)/\alpha} - (\theta_Y + \theta_X)e^{(\xi - x)/\theta_Y}}{\alpha - \theta_Y} \tag{59}$$

Compare it with (24) and see the symmetry. The symmetry implies the possible presence of more than a single mode of the distribution, which echoes *General Properties*, unsurprisingly. Let $x_{\mathrm{c,L}}$ and $x_{\mathrm{c,R}}$ denote the plausible solutions to $\partial\mathbf{D}_{X'}(x)/\partial x = 0$ in the event $I_{Y|X} = -1$, with $x_{\mathrm{c,L}} < \xi$ and $x_{\mathrm{c,R}} > \xi$. Then by the symmetry between (24) and (59), it is decided that $x_{\mathrm{c,L}} + x_{\mathrm{c,R}} = 2\xi$; accordingly, by Equation (5), $\mathbf{M}_Y\left(x_{\mathrm{c,L}}\right) = \mathbf{M}_Y\left(x_{\mathrm{c,R}}\right)$, and therefore by Equation (1), $\mathbf{D}_{X'}\left(x_{\mathrm{c,L}}\right) > \mathbf{D}_{X'}\left(x_{\mathrm{c,R}}\right)$. If $x_{\mathrm{c,L}}$ is valid (and not just plausible), then $x'_* = x_{\mathrm{c,L}}$; if $x_{\mathrm{c,L}}$ is invalid, i.e., if $x_{\mathrm{c,L}}$ is non-existent or $x_{\mathrm{c,L}} \in (0,\ \xi)$, then $x'_* = 0$, for the reason which follows:

$$\frac{\mathbf{D}_{X'}\left(x_{\mathrm{c,R}}\right)}{\mathbf{D}_{X'}(0)} = e^{-x_{\mathrm{c,R}}/\theta_X} \cdot \frac{1 - \delta\,\mathbf{M}_Y\left(x_{\mathrm{c,R}}\right)}{1 - \delta\,\mathbf{M}_Y(0)}$$
$$= e^{-x_{\mathrm{c,R}}/\theta_X} \cdot \left(1 - \delta \cdot \frac{\mathbf{M}_Y\left(x_{\mathrm{c,R}}\right) - \mathbf{M}_Y(0)}{1 - \delta\,\mathbf{M}_Y(0)}\right)$$
$$< e^{-x_{\mathrm{c,R}}/\theta_X} \cdot \left(1 - \delta \cdot \frac{\mathbf{M}_Y(2\xi) - \mathbf{M}_Y(0)}{1 - \delta\,\mathbf{M}_Y(0)}\right)$$
$$= e^{-x_{\mathrm{c,R}}/\theta_X} \tag{60}$$

All in all, the focus of the following discussion shall be on $x < 0$. From Equivalence

(24), the author deduces that in the event $I_{Y|X} = -1$, for sufficiently ambiguous focal object, i.e., sufficiently large θ_X, and for sufficient volatility, i.e., sufficiently large δ, $\lim_{x\to 0_-} \partial \mathbf{D}_{X'}(x)/\partial x < 0$. By continuous differentiability of $\mathbf{D}_{X'}(x)$ for $x < 0$, $\lim_{x\to 0_-} \partial \mathbf{D}_{X'}(x)/\partial x < 0$ guarantees existence of a valid $x_{\mathrm{c,L}}$. This surprisingly agrees with the results obtained in the discussion on assimilation.

Presented below are four special cases. To lighten the notation, let $x_{\mathrm{c}} = x_{\mathrm{c,L}}$. In the event $I_{Y|X} = -1$, if $\partial \mathbf{D}_{X'}(x)/\partial x = 0$ for some $x < 0$, then

$$\frac{\theta_Y}{\alpha} = \frac{1}{2} \quad \Longrightarrow \quad e^{(x_{\mathrm{c},\pm} - \xi)/\alpha} = \frac{\alpha + \theta_X}{\alpha + 2\theta_X} \pm \sqrt[2]{\left(\frac{\alpha + \theta_X}{\alpha + 2\theta_X}\right)^2 - \frac{3\alpha/2\delta}{\alpha + 2\theta_X}} \tag{61}$$

$$\frac{\theta_Y}{\alpha} = 2 \quad \Longrightarrow \quad e^{(x_{\mathrm{c},\pm} - \xi)/2\alpha} = \frac{\alpha + \theta_X/2}{\alpha + \theta_X} \pm \sqrt[2]{\left(\frac{\alpha + \theta_X/2}{\alpha + \theta_X}\right)^2 - \frac{3\alpha/\delta}{\alpha + \theta_X}} \tag{62}$$

$$\left(\frac{\theta_Y}{\alpha} = \frac{1}{3} \quad \wedge \quad \frac{9\delta^2}{16\alpha^2} \cdot \frac{(\alpha + \theta_X)^3}{\alpha + 3\theta_X} < 1\right) \quad \Longrightarrow$$

$$\begin{aligned} e^{(x_{\mathrm{c},\#} - \xi)/\alpha} \; = \; &-\sqrt[3]{\frac{4\alpha/3\delta}{\alpha + 3\theta_X} + \sqrt[2]{\left(\frac{4\alpha/3\delta}{\alpha + 3\theta_X}\right)^2 - \left(\frac{\alpha + \theta_X}{\alpha + 3\theta_X}\right)^3}} \\ &-\sqrt[3]{\frac{4\alpha/3\delta}{\alpha + 3\theta_X} - \sqrt[2]{\left(\frac{4\alpha/3\delta}{\alpha + 3\theta_X}\right)^2 - \left(\frac{\alpha + \theta_X}{\alpha + 3\theta_X}\right)^3}} \end{aligned} \tag{63}$$

$$\left(\frac{\theta_Y}{\alpha} = 3 \quad \wedge \quad \frac{\delta^2}{16\alpha^2} \cdot \frac{(\alpha + \theta_X/3)^3}{\alpha + \theta_X} < 1\right) \quad \Longrightarrow$$

$$\begin{aligned} e^{(x_{\mathrm{c},\#} - \xi)/3\alpha} = &-\sqrt[3]{\frac{4\alpha/\delta}{\alpha + \theta_X} + \sqrt[2]{\left(\frac{4\alpha/\delta}{\alpha + \theta_X}\right)^2 - \left(\frac{\alpha + \theta_X/3}{\alpha + \theta_X}\right)^3}} \\ &-\sqrt[3]{\frac{4\alpha/\delta}{\alpha + \theta_X} - \sqrt[2]{\left(\frac{4\alpha/\delta}{\alpha + \theta_X}\right)^2 - \left(\frac{\alpha + \theta_X/3}{\alpha + \theta_X}\right)^3}} \end{aligned} \tag{64}$$

$$\left(\frac{\theta_Y}{\alpha}=\frac{1}{3} \quad \wedge \quad \frac{9\delta^2}{16\alpha^2}\cdot\frac{(\alpha+\theta_X)^3}{\alpha+3\theta_X}\geq 1\right) \qquad \Longrightarrow$$

$$e^{(x_{\mathrm{c},k}-\xi)/\alpha} = -2\cdot\sqrt[2]{\frac{\alpha+\theta_X}{\alpha+3\theta_X}}\cos\left(\frac{2k\pi}{3}+\frac{1}{3}\arccos\left(\frac{4\alpha}{3\delta}\cdot\sqrt[2]{\frac{\alpha+3\theta_X}{(\alpha+\theta_X)^3}}\right)\right) \tag{65}$$

$$\left(\frac{\theta_Y}{\alpha}=3 \quad \wedge \quad \frac{\delta^2}{16\alpha^2}\cdot\frac{(\alpha+\theta_X/3)^3}{\alpha+\theta_X}\geq 1\right) \qquad \Longrightarrow$$

$$e^{(x_{\mathrm{c},k}-\xi)/3\alpha} = -2\cdot\sqrt[2]{\frac{\alpha+\theta_X/3}{\alpha+\theta_X}}\cos\left(\frac{2k\pi}{3}+\frac{1}{3}\arccos\left(\frac{4\alpha}{\delta}\cdot\sqrt[2]{\frac{\alpha+\theta_X}{(\alpha+\theta_X/3)^3}}\right)\right) \tag{66}$$

The author is now studying these cases. Again, the author begins with the cases where $\theta_Y=\alpha/2$ and $\theta_Y=2\alpha$. In the event $I_{Y|X}=-1$, (68)–(77) can be deduced. With hindsight, the author derives from (24) the following result:

$$\frac{\partial^2}{(\partial x)^2}\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)\equiv\frac{I_{Y|X}}{\alpha-\theta_Y}\cdot\left(\left(1+\frac{\theta_X}{\alpha}\right)e^{(x-\xi)/\alpha}-\left(1+\frac{\theta_X}{\theta_Y}\right)e^{(x-\xi)/\theta_Y}\right) \tag{67}$$

In the case where $\theta_Y=\alpha/2$, if $\lim_{x\to 0_-}\partial\mathbf{D}_{X'}(x)/\partial x<0$, then Inequality (68) is implied. (68) mathematically implied (69); this can be proved by minimising the right-hand side of (67) while fixing the parametre δ. Inequality (69) implies (70), which guarantees that x_{c} is real. (71) directly follows (68). From (71), it can be deduced that $x_{\mathrm{c},\pm}\gtrless 0$; hence, $x_{\mathrm{c},-}$ is valid while $x_{\mathrm{c},+}$ is not. By Equivalence (67), $\partial^2\mathbf{D}_{X'}(x)/(\partial x)^2\gtrless 0$ at $x=x_{\mathrm{c},\pm}$, so in conclusion, $\lim_{x\to 0_-}\partial\mathbf{D}_{X'}(x)/\partial x<0$ implies $x'_*=x_{\mathrm{c},-}$.

$$\left(\frac{\theta_Y}{\alpha}=\frac{1}{2} \quad \wedge \quad \lim_{x\to 0_-}\frac{\partial}{\partial x}\mathbf{D}_{X'}(x)<0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha}>\frac{(3/2\delta)/2-e^{-\xi/\alpha}+e^{-2\xi/\alpha}/2}{e^{-\xi/\alpha}-e^{-2\xi/\alpha}} \tag{68}$$

$$\geq\frac{3}{2\delta}-1+\sqrt[2]{\frac{3}{2\delta}\left(\frac{3}{2\delta}-1\right)}=\frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{69}$$

$$\Longrightarrow \quad \delta \geq \frac{3}{2}\alpha \cdot \frac{2\theta_X + \alpha}{(\alpha + \theta_X)^2} \tag{70}$$

$$\wedge \quad e^{(x_{\mathrm{c},-} - \xi)/\alpha} < e^{-\xi/\alpha} < e^{(x_{\mathrm{c},+} - \xi)/\alpha} \tag{71}$$

$$\Longrightarrow \quad x'_* = x_{\mathrm{c},-} \tag{72}$$

In the case where $\theta_Y = 2\alpha$, the reasoning is analogous:

$$\left(\frac{\theta_Y}{\alpha} = 2 \qquad \wedge \qquad \lim_{x \to 0_-} \frac{\partial}{\partial x} \mathbf{D}_{X'}(x) < 0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha} > \frac{2 \cdot (3/2\delta) - 2e^{-\xi/2\alpha} + e^{-\xi/\alpha}}{e^{-\xi/\alpha} - e^{-\xi/2\alpha}} \tag{73}$$

$$\geq 2\left(\frac{3}{\delta} - 1 + \sqrt[2]{\frac{3}{\delta}\left(\frac{3}{\delta} - 1\right)}\right) = \frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{74}$$

$$\Longrightarrow \quad \delta \geq 3\alpha \cdot \frac{\theta_X + \alpha}{(\alpha + \theta_X/2)^2} \tag{75}$$

$$\wedge \quad e^{(x_{\mathrm{c},-} - \xi)/\alpha} < e^{-\xi/\alpha} < e^{(x_{\mathrm{c},+} - \xi)/\alpha} \tag{76}$$

$$\Longrightarrow \quad x'_* = x_{\mathrm{c},-} \tag{77}$$

The author's next focus of study is the cases where $\theta_Y = \alpha/3$ and $\theta_Y = 3\alpha$. Again, for convenience, assume no repeated roots.

In the case where $\theta_Y = \alpha/3$, if $\lim_{x \to 0_-} \partial \mathbf{D}_{X'}(x)/\partial x < 0$, then Inequality (78) is implied. (78) mathematically implies (79); this can be proved by minimising the right-hand side of (78) while fixing the parametre δ. Inequality (79) implies (80), which implies more than one plausible real solution x_c. Inequalities (81) directly follow (78). $e^{(x_{\mathrm{c},3} - \xi)/3\alpha} < 0$, so $x_{\mathrm{c},3}$ is non-existent. From Inequalities (81), it can be deduced that $x_{\mathrm{c},2} < 0$ while $x_{\mathrm{c},1} > 0$; accordingly, $x_{\mathrm{c},2}$ is valid while $x_{\mathrm{c},1}$ is not. By Equivalence (67), the author finds that $\partial^2 \mathbf{D}_{X'}(x)/(\partial x)^2 < 0$ at $x = x_{\mathrm{c},2}$, so in conclusion, if $\lim_{x \to 0_-} \partial \mathbf{D}_{X'}(x)/\partial x < 0$, then it is decided that $x'_* = x_{\mathrm{c},2}$:

$$\left(\frac{\theta_Y}{\alpha} = \frac{1}{3} \qquad \wedge \qquad \lim_{x \to 0_-} \frac{\partial}{\partial x} \mathbf{D}_{X'}(x) < 0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha} > \frac{(8/3\delta)/3 - e^{-\xi/\alpha} + e^{-3\xi/\alpha}/3}{e^{-\xi/\alpha} - e^{-3\xi/\alpha}} \tag{78}$$

$$\geq \frac{8}{3\delta} \cos\left(\frac{1}{3} \arccos\left(\frac{3\delta}{4}\right) - \frac{\pi}{3}\right) - 1 = \frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{79}$$

$$\Longrightarrow \quad \delta \geq \frac{4}{3} \alpha \cdot \sqrt[2]{\frac{\alpha + 3\theta_X}{(\alpha + \theta_X)^3}} \tag{80}$$

$$\wedge \quad e^{(x_{\mathrm{c},2} - \xi)/\alpha} < e^{-\xi/\alpha} < e^{(x_{\mathrm{c},1} - \xi)/\alpha} \tag{81}$$

$$\Longrightarrow \quad x'_* = x_{\mathrm{c},2} \tag{82}$$

In the case where $\theta_Y = 3\alpha$, the reasoning is analogous:

$$\left(\frac{\theta_Y}{\alpha} = 3 \qquad \wedge \qquad \lim_{x \to 0_-} \frac{\partial}{\partial x} \mathbf{D}_{X'}(x) < 0\right)$$

$$\Longrightarrow \quad \frac{\theta_X}{\alpha} > \frac{3 \cdot (8/3\delta) - 3e^{-\xi/3\alpha} + e^{-\xi/\alpha}}{e^{-\xi/3\alpha} - e^{-\xi/\alpha}} \tag{83}$$

$$\geq 3\left(\frac{8}{\delta} \cos\left(\frac{1}{3} \arccos\left(\frac{\delta}{4}\right) - \frac{\pi}{3}\right) - 1\right) = \frac{\theta_{X,\mathrm{c}}}{\alpha} \tag{84}$$

$$\Longrightarrow \quad \delta \geq 4\alpha \cdot \sqrt[2]{\frac{\alpha + \theta_X}{(\alpha + \theta_X/3)^3}} \tag{85}$$

$$\wedge \quad e^{(x_{\mathrm{c},2} - \xi)/3\alpha} < e^{-\xi/3\alpha} < e^{(x_{\mathrm{c},1} - \xi)/3\alpha} \tag{86}$$

$$\Longrightarrow \quad x'_* = x_{\mathrm{c},2} \tag{87}$$

At this point, the author is bringing these cases together and sees, after calculations, that in all of these cases, $\lim_{x \to 0_-} \partial \mathbf{D}_{X'}(x)/\partial x < 0$ in the event $I_{Y|X} = -1$ mathematically implies $\partial x'_*/\partial\alpha > 0$ and $\partial x'_*/\partial\delta < 0$; this is as expected. Nevertheless, meanwhile, it is implied that $\partial x'_*/\partial\xi > 0$ and that $\partial x'_*/\partial\theta_X < 0$; in addition, $\partial\theta_{X,\mathrm{c}}/\partial\delta < 0$. The latter three implications might appear to be perplexing; see *General Discussion*. The author is wondering whether or not these results hold for other distributions as well.

The next subsubsection will be dedicated to the logistic distribution, as previously announced. Before closing the present subsubsection, however, there is more to discuss.

Recall Equation (2); see *Methods*:

$$\int_{-\infty}^{\infty} \mathbf{D}_X(v) \ln\left(\frac{\mathbf{D}_Y(v)}{\mathbf{D}_X(v)}\right) \partial v$$

$$= \frac{1}{2\theta_X} \int_{-\infty}^{\infty} e^{-|v|/\theta_X} \ln\left(\frac{\theta_X}{\theta_Y} \cdot \frac{e^{-|v-\xi|/\theta_Y}}{e^{-|v|/\theta_X}}\right) \partial v$$

$$= \frac{1}{2\theta_X} \ln\left(\frac{\theta_X}{\theta_Y}\right) \int_{-\infty}^{\infty} e^{-|v|/\theta_X} \partial v + \frac{1}{2\theta_X} \int_{-\infty}^{\infty} \left(\frac{|v|}{\theta_X} - \frac{|v-\xi|}{\theta_Y}\right) e^{-|v|/\theta_X} \partial v$$

$$= \frac{1}{2} \ln\left(\frac{\theta_X}{\theta_Y}\right) \int_{-\infty}^{\infty} e^{-|v|} \partial v + \frac{1}{2} \int_{-\infty}^{\infty} |v| e^{-|v|} \partial v - \frac{\theta_X}{2\theta_Y} \int_{-\infty}^{\infty} \left|v - \frac{\xi}{\theta_X}\right| e^{-|v|} \partial v$$

$$= \frac{1}{2} \ln\left(\frac{\theta_X}{\theta_Y}\right) \left(\int_{0}^{\infty} e^{-v} \partial v + \int_{-\infty}^{0} e^{v} \partial v\right) + \frac{1}{2}\left(\int_{0}^{\infty} v e^{-v} \partial v - \int_{-\infty}^{0} v e^{v} \partial v\right)$$

$$- \frac{\theta_X}{2\theta_Y}\left(\int_{2\xi/\theta_X}^{\infty} \left(v - \frac{\xi}{\theta_X}\right) e^{-v} \partial v + \int_{\xi/\theta_X}^{2\xi/\theta_X} \left(v - \frac{\xi}{\theta_X}\right) e^{-v} \partial v\right.$$

$$\left. + \int_{0}^{\xi/\theta_X} \left(\frac{\xi}{\theta_X} - v\right) e^{-v} \partial v + \int_{-\infty}^{0} \left(\frac{\xi}{\theta_X} - v\right) e^{v} \partial v\right)$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) \int_{0}^{\infty} e^{-v} \partial v + \int_{0}^{\infty} v e^{-v} \partial v - \frac{\theta_X}{2\theta_Y}\left(\left(e^{\xi/\theta_X} + e^{-\xi/\theta_X}\right) \int_{\xi/\theta_X}^{\infty} v e^{-v} \partial v\right.$$

$$\left. + e^{-\xi/\theta_X} \int_{0}^{\xi/\theta_X} v(e^{v} + e^{-v}) \partial v\right)$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\theta_X}{\theta_Y} e^{-\xi/\theta_X} - \frac{\xi}{\theta_Y} + 1 \tag{88}$$

From Equation (88), the author derives

$$\frac{\partial}{\partial \theta_X} \mathbf{P}\left[I_{Y|X} = 1\right] \equiv \frac{\theta_Y}{\theta_X} - \left(1 + \frac{\xi}{\theta_X}\right) e^{-\xi/\theta_X} \tag{89}$$

$$\frac{\partial}{\partial \theta_Y} \mathbf{P}\left[I_{Y|X} = 1\right] \equiv \frac{\theta_X}{\theta_Y} e^{-\xi/\theta_X} + \frac{\xi}{\theta_Y} - 1 \tag{90}$$

Equation (89) suggests that $\partial\mathbf{P}\left[I_{Y|X}=1\right]/\partial\theta_X>0$ if and only if the contextual stimulus is sufficiently ambiguous and sufficiently extreme, and the focal object is sufficiently specific. Equation (90) suggests that $\partial\mathbf{P}\left[I_{Y|X}=1\right]/\partial\theta_Y>0$ if and only if the contextual stimulus is sufficiently specific and sufficiently extreme, and the focal object is sufficiently ambiguous. In addition, $\partial\mathbf{P}\left[I_{Y|X}=1\right]/\partial\xi<0$ unconditionally; that is to say, the extremity of the contextual stimulus hinders perceptual assimilation.

$$\frac{\partial}{\partial\xi}\mathbf{P}\left[I_{Y|X}=1\right]\equiv e^{-\xi/\theta_X}-1<0 \tag{91}$$

Logistic Distribution. In the previous subsubsection, the author assumed the Laplace distribution, which has a probability density curve indifferentiable at the mode. In this subsubsection, the logistic distribution is assumed instead:

$$\mathbf{D}_X(x)=\frac{1}{\theta_X}\cdot\frac{e^{x/\theta_X}}{(e^{x/\theta_X}+1)^2}\quad\wedge\quad\mathbf{D}_Y(y)=\frac{1}{\theta_Y}\cdot\frac{e^{(y-\xi)/\theta_Y}}{(e^{(y-\xi)/\theta_Y}+1)^2} \tag{92}$$

For computational convenience, let θ_Y be an integer multiple of α. Applying Equation (5), for $x>0$, the author derives

$$\mathbf{M}_Y(\xi+x)$$

$$=\frac{1}{\theta_Y}\int_x^\infty\frac{e^{v/\theta_Y+(x-v)/\alpha}}{(e^{v/\theta_Y}+1)^2}\partial v+\frac{1}{\theta_Y}\int_{-x}^\infty\frac{e^{v/\theta_Y-(v+x)/\alpha}}{(e^{v/\theta_Y}+1)^2}\partial v$$

$$=e^{x/\alpha}\int_{e^{x/\theta_Y}}^\infty\frac{1}{v^{\theta_Y/\alpha}(v+1)^2}\partial v+e^{-x/\alpha}\int_{e^{-x/\theta_Y}}^\infty\frac{1}{v^{\theta_Y/\alpha}(v+1)^2}\partial v$$

$$=e^{x/\alpha}\int_{e^{x/\theta_Y}}^\infty\left(\sum_{i=1}^{\theta_Y/\alpha}\frac{(-1)^{i-1}}{v^{\theta_Y/\alpha-i+1}(v+1)}+\frac{(-1)^{\theta_Y/\alpha}}{(v+1)^2}\right)\partial v$$

$$+e^{-x/\alpha}\int_{e^{-x/\theta_Y}}^\infty\left(\sum_{i=1}^{\theta_Y/\alpha}\frac{(-1)^{i-1}}{v^{\theta_Y/\alpha-i+1}(v+1)}+\frac{(-1)^{\theta_Y/\alpha}}{(v+1)^2}\right)\partial v$$

$$=e^{x/\alpha}\int_{e^{x/\theta_Y}}^\infty\left(\sum_{i=1}^{\theta_Y/\alpha}\sum_{j=1}^{\theta_Y/\alpha-i+1}\frac{(-1)^{i+j}}{v^{\theta_Y/\alpha-i-j+2}}+\frac{\theta_Y}{\alpha}\cdot\frac{(-1)^{\theta_Y/\alpha}}{v+1}+\frac{(-1)^{\theta_Y/\alpha}}{(v+1)^2}\right)\partial v$$

$$+e^{-x/\alpha}\int_{e^{-x/\theta_Y}}^{\infty}\left(\sum_{i=1}^{\theta_Y/\alpha}\sum_{j=1}^{\theta_Y/\alpha-i+1}\frac{(-1)^{i+j}}{v^{\theta_Y/\alpha-i-j+2}}+\frac{\theta_Y}{\alpha}\cdot\frac{(-1)^{\theta_Y/\alpha}}{v+1}+\frac{(-1)^{\theta_Y/\alpha}}{(v+1)^2}\right)\partial v$$

$$=e^{x/\alpha}\int_{e^{x/\theta_Y}}^{\infty}\left(\sum_{k=1}^{\theta_Y/\alpha}\frac{(-1)^{k-1}\cdot k}{v^{\theta_Y/\alpha-k+1}}+\frac{\theta_Y}{\alpha}\cdot\frac{(-1)^{\theta_Y/\alpha}}{v+1}+\frac{(-1)^{\theta_Y/\alpha}}{(v+1)^2}\right)\partial v$$

$$+e^{-x/\alpha}\int_{e^{-x/\theta_Y}}^{\infty}\left(\sum_{k=1}^{\theta_Y/\alpha}\frac{(-1)^{k-1}\cdot k}{v^{\theta_Y/\alpha-k+1}}+\frac{\theta_Y}{\alpha}\cdot\frac{(-1)^{\theta_Y/\alpha}}{v+1}+\frac{(-1)^{\theta_Y/\alpha}}{(v+1)^2}\right)\partial v$$

$$=\sum_{k=1}^{\theta_Y/\alpha-1}\frac{(-1)^{k-1}\cdot k}{\theta_Y/\alpha-k}\left(e^{kx/\theta_Y}+e^{-kx/\theta_Y}\right)-(-1)^{\theta_Y/\alpha}\left(\frac{\theta_Y}{\alpha}e^{x/\alpha}\ln\left(1+e^{-x/\theta_Y}\right)\right.$$

$$\left.+\frac{\theta_Y}{\alpha}e^{-x/\alpha}\ln\left(1+e^{x/\theta_Y}\right)-\frac{e^{x/\alpha}}{e^{x/\theta_Y}+1}-\frac{e^{-x/\alpha}}{1+e^{-x/\theta_Y}}\right) \tag{93}$$

Again, by symmetry, $x\mapsto|x-\xi|$ effectively equates to $\mathbf{M}_Y(\xi+x)\mapsto\mathbf{M}_Y(x)$. The magnitude of the influence of the contextual stimulus reaches its maximum at $x=\xi$:

$$\mathbf{M}_Y(x)\leq\sum_{k=1}^{\theta_Y/\alpha-1}\frac{(-1)^{k-1}\cdot 2k}{\theta_Y/\alpha-k}-(-1)^{\theta_Y/\alpha}\left(\frac{2\theta_Y}{\alpha}\ln(2)-1\right) \tag{94}$$

The regularising integral in Equation (1) in general is intractable; it is elementary when θ_X is an integer multiple of α, but even so the resultant expression is exceedingly complicated and unwieldy. Examples will be given on readers' request. Fortunately, it is actually irrelevant to the search for the best representation.

To find the best representation, the author attempted to solve the Equation $\partial\mathbf{D}_{X'}\left(x\middle|I_{Y|X}\right)/\partial x=0$, which is transcendental and therefore in no way can be attacked analytically. In light of this, the author then tried indirect deduction by general properties (9)–(14), but the resultant inequalities are transcendental too, resisting the author's mathematical techniques.

In summary, there is little that can be done without simulations, within the author's capability. The author is now studying the probability of assimilation and contrast:

$$\int_{-\infty}^{\infty} \mathbf{D}_X(v) \ln\left(\frac{\mathbf{D}_Y(v)}{\mathbf{D}_X(v)}\right) \partial v$$

$$= \frac{1}{\theta_X} \int_{-\infty}^{\infty} \frac{e^{v/\theta_X}}{(e^{v/\theta_X}+1)^2} \ln\left(\frac{\theta_X}{\theta_Y} \cdot \frac{e^{(v-\xi)/\theta_Y}}{e^{v/\theta_X}} \cdot \frac{\left(e^{v/\theta_X}+1\right)^2}{(e^{(v-\xi)/\theta_Y}+1)^2}\right) \partial v$$

$$= \frac{1}{\theta_X} \ln\left(\frac{\theta_X}{\theta_Y}\right) \int_{-\infty}^{\infty} \frac{e^{v/\theta_X}}{(e^{v/\theta_X}+1)^2} \partial v - \frac{1}{\theta_X} \int_{-\infty}^{\infty} \left(\frac{v}{\theta_X} - \frac{v-\xi}{\theta_Y}\right) \frac{e^{v/\theta_X}}{(e^{v/\theta_X}+1)^2} \partial v$$

$$+ \frac{2}{\theta_X} \int_{-\infty}^{\infty} \frac{e^{v/\theta_X}}{(e^{v/\theta_X}+1)^2} \ln\left(\frac{e^{v/\theta_X}+1}{e^{(v-\xi)/\theta_Y}+1}\right) \partial v$$

$$= \frac{1}{\theta_X} \left(\ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi}{\theta_Y}\right) \int_{-\infty}^{\infty} \frac{e^{v/\theta_X}}{(e^{v/\theta_X}+1)^2} \partial v + \frac{2}{\theta_X} \int_{-\infty}^{\infty} \frac{e^{v/\theta_X}}{(e^{v/\theta_X}+1)^2} \ln\left(\frac{e^{v/\theta_X}+1}{e^{(v-\xi)/\theta_Y}+1}\right) \partial v$$

$$= \left(\ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi}{\theta_Y}\right) \int_{0}^{\infty} \frac{1}{(v+1)^2} \partial v + 2 \int_{0}^{\infty} \frac{1}{(v+1)^2} \ln\left(\frac{v+1}{1+v^{\theta_X/\theta_Y} e^{-\xi/\theta_Y}}\right) \partial v$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi}{\theta_Y} + 2 \int_{0}^{\infty} \frac{1}{(v+1)^2} \left(\ln(v+1) - \ln\left(1+v^{\theta_X/\theta_Y} e^{-\xi/\theta_Y}\right)\right) \partial v$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi}{\theta_Y} + 2 - 2 \int_{0}^{\infty} \frac{1}{(v+1)^2} \ln\left(1+v^{\theta_X/\theta_Y} e^{-\xi/\theta_Y}\right) \partial v$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi}{\theta_Y} + 2 - \frac{2\theta_X}{\theta_Y} \int_{0}^{\infty} \frac{1}{v(v+1)(v^{-\theta_X/\theta_Y} e^{\xi/\theta_Y}+1)} \partial v$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi}{\theta_Y} + 2 - \frac{2\theta_X}{\theta_Y} \int_{0}^{\infty} \frac{1}{(v+1)(v^{\theta_X/\theta_Y} e^{\xi/\theta_Y}+1)} \partial v \qquad (95)$$

The remaining integral is very unlikely to be tractable for irrational θ_X/θ_Y; it is elementary when θ_X/θ_Y is rational. From Equation (95), the author derives

$$\frac{\partial}{\partial \theta_X} \mathbf{P}\left[I_{Y|X} = 1\right] \equiv \frac{\theta_Y}{2\theta_X} - \int_{0}^{\infty} \frac{1 + (1 - (\theta_X/\theta_Y) \cdot \ln(v)) v^{\theta_X/\theta_Y} e^{\xi/\theta_Y}}{(v+1)(v^{\theta_X/\theta_Y} e^{\xi/\theta_Y}+1)^2} \partial v \qquad (96)$$

$$\frac{\partial}{\partial \theta_Y} \mathbf{P}\left[I_{Y|X} = 1\right] \equiv \frac{2\theta_X}{\theta_Y} \int_{0}^{\infty} \frac{1 + (1 - (\theta_X/\theta_Y) \cdot \ln(v) - \xi/\theta_Y) v^{\theta_X/\theta_Y} e^{\xi/\theta_Y}}{(v+1)(v^{\theta_X/\theta_Y} e^{\xi/\theta_Y}+1)^2} \partial v + \frac{\xi}{\theta_Y} - 1 \qquad (97)$$

From (96) and (97), the author arrives at the same conclusions as in the previous subsubsection. The underlying intuition is that in (96) and (97), the θ_X and θ_Y in the integrals exert minimal influence on $\partial\mathbf{P}\left[I_{Y|X}=1\right]/\partial\theta_X$ and $\partial\mathbf{P}\left[I_{Y|X}=1\right]/\partial\theta_Y$; that is, they are negligible. Finally, observe that $\partial\mathbf{P}\left[I_{Y|X}=1\right]/\partial\xi<0$ in any conditions; the proof follows:

$$\begin{aligned}\frac{\partial}{\partial\xi}\mathbf{P}\left[I_{Y|X}=1\right]&\equiv\frac{2\theta_X}{\theta_Y}\int_0^\infty\frac{v^{\theta_X/\theta_Y}e^{\xi/\theta_Y}}{(v+1)(v^{\theta_X/\theta_Y}e^{\xi/\theta_Y}+1)^2}\partial v-1\\&<\frac{2\theta_X}{\theta_Y}\int_0^\infty\frac{v^{\theta_X/\theta_Y}}{(v+1)(v^{\theta_X/\theta_Y}+1)^2}\partial v-1\\&=2\int_0^\infty\frac{v^{\theta_Y/\theta_X}}{(v+1)^2(v^{\theta_Y/\theta_X}+1)}\partial v-1\\&=2\int_0^\infty\frac{1}{(v+1)^2(v^{\theta_Y/\theta_X}+1)}\partial v-1\\&=\int_0^\infty\frac{1}{(v+1)^2}\partial v-1=0\end{aligned}\tag{98}$$

Gaussian Distribution. The Gaussian distribution is the author's last focus of study in *Analytical Approach*. In light of its apparently similar behaviour (i.e., a bell shape), agreement with *Logistic Distribution* is expected.

Consider the following functional form:

$$\mathbf{D}_X(x)=\frac{1}{\sqrt[2]{2\pi}\,\theta_X}e^{-x^2/2\theta_X^2}\qquad\wedge\qquad\mathbf{D}_Y(y)=\frac{1}{\sqrt[2]{2\pi}\,\theta_Y}e^{-(y-\xi)^2/2\theta_Y^2}\tag{99}$$

Review *General Properties*, applying Equation (5):

$$\begin{aligned}&\mathbf{M}_Y(\xi+x)\\&=\frac{1}{\sqrt[2]{2\pi}}e^{\theta_Y^2/2\alpha^2}\left(e^{x/\alpha}\int_{x/\theta_Y}^\infty e^{-(v+\theta_Y/\alpha)^2/2}\,\partial v+e^{-x/\alpha}\int_{-x/\theta_Y}^\infty e^{-(v+\theta_Y/\alpha)^2/2}\,\partial v\right)\\&=\frac{1}{2}e^{\theta_Y^2/2\alpha^2}\left(e^{x/\alpha}\left(1-\operatorname{erf}\left(\frac{\theta_Y/\alpha+x/\theta_Y}{\sqrt[2]{2}}\right)\right)+e^{-x/\alpha}\left(1-\operatorname{erf}\left(\frac{\theta_Y/\alpha-x/\theta_Y}{\sqrt[2]{2}}\right)\right)\right)\end{aligned}\tag{100}$$

Equation (100) holds for $x>0$. Letting $x\mapsto|x-\xi|$ completes the substitution

$\mathbf{M}_Y(\xi + x) \mapsto \mathbf{M}_Y(x)$. By symmetry, the magnitude is maximum at $x = \xi$:

$$\mathbf{M}_Y(x) \leq e^{\theta_Y^2/2\alpha^2}\left(1 - \operatorname{erf}\left(\frac{\theta_Y}{\sqrt[2]{2}\alpha}\right)\right) \tag{101}$$

For the reasons as given in Logistic Distribution, the author is skipping the regularising integral and skipping the search for the best representation to the probability of assimilation and contrast. Now,

$$\int_{-\infty}^{\infty} \mathbf{D}_X(v) \ln\left(\frac{\mathbf{D}_Y(v)}{\mathbf{D}_X(v)}\right) \partial v$$

$$= \frac{1}{\sqrt[2]{2\pi}\theta_X} \int_{-\infty}^{\infty} e^{-v^2/2\theta_X^2} \ln\left(\frac{\theta_X}{\theta_Y} \cdot \frac{e^{-(v-\xi)^2/2\theta_Y^2}}{e^{-v^2/2\theta_X^2}}\right) \partial v$$

$$= \frac{1}{\sqrt[2]{2\pi}\theta_X} \ln\left(\frac{\theta_X}{\theta_Y}\right) \int_{-\infty}^{\infty} e^{-v^2/2\theta_X^2} \partial v + \frac{1}{2\sqrt[2]{2\pi}\theta_X} \int_{-\infty}^{\infty} \left(\frac{v^2}{\theta_X^2} - \frac{(v-\xi)^2}{\theta_Y^2}\right) e^{-v^2/2\theta_X^2} \partial v$$

$$= \frac{1}{\sqrt[2]{2\pi}\theta_X} \left(\ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi^2}{2\theta_Y^2}\right) \int_{-\infty}^{\infty} e^{-v^2/2\theta_X^2} \partial v + \frac{1}{2\sqrt[2]{2\pi}\theta_X} \left(\frac{1}{\theta_X^2} - \frac{1}{\theta_Y^2}\right) \int_{-\infty}^{\infty} v^2 e^{-v^2/2\theta_X^2} \partial v$$

$$= \frac{1}{\sqrt[2]{\pi}} \left(\ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi^2}{2\theta_Y^2}\right) \int_{-\infty}^{\infty} e^{-v^2} \partial v + \frac{1}{\sqrt[2]{\pi}} \left(1 - \frac{\theta_X^2}{\theta_Y^2}\right) \int_{-\infty}^{\infty} v^2 e^{-v^2} \partial v$$

$$= \frac{1}{\sqrt[2]{\pi}} \left(\ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi^2}{2\theta_Y^2} - \frac{\theta_X^2}{2\theta_Y^2} + \frac{1}{2}\right) \int_{-\infty}^{\infty} e^{-v^2} \partial v$$

$$= \ln\left(\frac{\theta_X}{\theta_Y}\right) - \frac{\xi^2}{2\theta_Y^2} - \frac{\theta_X^2}{2\theta_Y^2} + \frac{1}{2} \tag{102}$$

Equation (102) implies

$$\frac{\partial}{\partial \theta_X} \mathbf{P}[I_{Y|X} = 1] \equiv \frac{\theta_Y^2}{\theta_X^2} - 1 \tag{103}$$

$$\frac{\partial}{\partial \theta_Y} \mathbf{P}[I_{Y|X} = 1] \equiv \frac{\theta_X^2}{\theta_Y^2} + \frac{\xi^2}{\theta_Y^2} - 1 \tag{104}$$

It is self-evident $\partial \mathbf{P}[I_{Y|X} = 1]/\partial \xi < 0$. Note that $\partial \mathbf{P}[I_{Y|X} = 1]/\partial \theta_X$ is irrespective of ξ. Otherwise, as expected, the findings in *Logistic Distribution* continue to hold.

Numerical Simulations

Thus far, the author has focused the discussion on unimodal, symmetric distributions. However, as having already been shown, even in simple cases such as the Laplace, logistic, and Gaussian cases, the integrals and equations involved in locating the posterior best representation are generally intractable. In light of the technical difficulties ahead, the author is closing *Analytical Approach* and starting *Numerical Simulations*. The author will be simulating the perceptual dynamics formulated in *Methods*, with Desmos, a graphing calculator which is free and accessible on the Web. On the following pages, the author will be assuming the skewed generalised Student's t distribution, which accommodates a vast variety of distributions, and the Weibull distribution, which will be made bimodal for the purpose of this study. For the former distribution, $\mathbb{V} = (-\infty, \infty)$; for the latter distribution, $\mathbb{V} = [0, \infty)$. In approximating the improper integrals involved, the author sets the bounds of integration to be $[-1000, 1000]$ for the former distribution, and $[0, 2000]$ for the latter distribution. The truncated probability in any of the reported cases is no greater than 1%. In producing the figures, the global mode is taken as the best representation.

Skewed Generalised Student's t Distribution

Assume the following functional form:

$$\mathbf{D}(v) = \frac{\kappa\zeta/2\theta}{\mathrm{B}(1/\kappa\,, \eta/\kappa)} \cdot \left(1 + \left(\zeta \cdot \frac{|v - \xi|/\theta}{1 + \rho\,\mathrm{sgn}(v - \xi)}\right)^{\kappa}\right)^{-(\eta+1)/\kappa} \tag{105}$$

This is Theodossiou's generalisation (1998) of Student's t distribution, as reviewed by Li and Nadarajah (2018). $\kappa > 0$; large κ gives a flat-top distribution. $\eta > 2$; large η mathematically implies fast convergence of the tail probability to zero. $|\rho| < 1$, which controls skewness; $\rho < 0$ gives a left-skewed distribution, and $\rho > 0$ gives a right-skewed distribution. The auxiliary parametre $\zeta > 0$:

$$\zeta^2 = \left(1 + \rho^2 \left(3 - 4 \cdot \frac{\mathrm{B}^2(2/\kappa\,, (\eta - 1)/\kappa)}{\mathrm{B}(1/\kappa\,, \eta/\kappa)\,\mathrm{B}(3/\kappa\,, (\eta - 2)/\kappa)}\right)\right) \cdot \frac{\mathrm{B}(3/\kappa\,, (\eta - 2)/\kappa)}{\mathrm{B}(1/\kappa\,,\ \eta/\kappa)}$$

$\mathrm{B}(\cdot\,,\ \cdot)$ is Euler's beta integral:

$$\mathrm{B}(x, y) = \int_0^1 v^{x-1}(1 - v)^{y-1}\partial v$$

To save paper space, the author is not including figures of the original and distorted distributions, which readers can easily produce with Desmos.

Hereafter in this subsection, the author is reporting and commenting on the simulation results. The figures being referred to are presented on pages 41–62. Note that $\xi_X = 0$.

Figures 1–8 consistently suggest that in the event of assimilation, the best representation of the focal object, in general, will assimilate to the context and that, in that event, context extremity magnifies the shift of the best representation when the context is moderate, but diminishes it when the context is extreme; the diminishment is unpredicted in *Laplace Distribution* (see *Analytical Approach*). Also, these figures consistently suggest that in the event of assimilation, the best representation of the focal object will unconditionally shift; this conflicts with the findings in *Laplace Distribution*.

See *General Discussion* for the cause of the conflict. In the event of contrast, the unconditional shift is suggested, too, but the direction is generally opposite; that is, in the event of contrast, the best representation of the focal object in general will contrast away from the context, as it should do, with only one exception: For contexts in the vicinity of the original best representation of the focal object, if the context is skewed to a symmetric focal object or if the focal object is skewed to a symmetric context, then surprisingly, the best representation will assimilate to the context; a Desmos graph will explain.

In the event of contrast, context extremity generally diminishes the shift of the best representation of the focal object, as already predicted in *Laplace Distribution*. Nevertheless, still, exceptions exist: For contexts in the vicinity of the original best representation of the focal object, the shift can be magnified by context extremity.

Figures 1–4 together suggest that in the event of assimilation, the shift of the best representation of the focal object, when the context is moderate, is greater when the context is skewed away from the focal object, latently specific, with a less flat top, and heavy-tailed; and when the context is extreme, is greater when the context is skewed to the focal object, latently ambiguous, with a flatter top, and light-tailed. The latter combination gives a greater shift in the event of contrast, generally.

Figures 5–8 together suggest that in the event of assimilation, the shift of the best representation of the focal object is greater when the focal object is skewed to the context, latently ambiguous, with a flatter top, and light-tailed; and in the event of contrast, the shift is greater when the focal object is skewed away from the context, latently specific, with a less flat top, and heavy-tailed.

Remarkably, in the event of assimilation, Figures 1–4 suggest that it is possible that the posterior best representation of the focal object is farther from its original than is the context if the context is skewed away from a symmetric focal object. Figures 5–8 suggest that when the context is symmetric, no shift will be predicted in the event of assimilation if the context coincides with the focal object.

Figures 9–16 consistently suggest that if the best representation of the focal object is expected to assimilate to the context, then the context is more likely to be moderate than extreme; and that if it is expected to contrast away, then the context is more likely to be extreme than moderate. The converse, however, is not implied. It is possible that even when the context is in the vicinity of the original best representation of the focal object, the best representation of the focal object is still expected to contrast away; this is, for example, when the context is seriously skewed away from the focal object, or the other way around. Skewness matters.

For now, assume that the converse argument holds, the cases right above being rare. As the context goes from moderate to extreme, the expectation will then go from weak, strong, and weak assimilation, to weak, strong, and weak contrast, in sequential order. Eventually, the magnitude of contrast will converge to zero, naturally; see *General Discussion*.

From Figures 9–12, it can be induced that if the focal object is symmetric, then for moderate contexts, the expected posterior best representation of the focal object is closer to the context generally when the context is skewed to the focal object, latently ambiguous, and with a less flat top; for extreme contexts, it is closer generally when the context is skewed away from the focal object, and latently specific. The author is expecting that for extreme contexts, it is closer generally when the context is with a flatter top, but sees no evidence in Figure 11. Figure 12 is inconclusive about the influence of the tail behaviour of the context distribution on the posterior best representation of the focal object when the context is moderate, yet it can be induced that for extreme contexts, it is closer generally when the context is light-tailed.

From Figures 13–16, it can be induced that if the context is symmetric, then the expected posterior best representation of the focal object is consistently closer to the context when the focal object is skewed to the context. When skewed to the context, the focal object is expected to shift to a position closer to the context for moderate contexts when the focal object is latently ambiguous, with a flatter top, and light-tailed; for extreme contexts, it is expected to shift to a position closer to the context when it is latently specific, with a less flat top, and heavy-tailed. When skewed away from the context, the focal object appears to be expected to shift to a position closer to the context, regardless of context extremity, when latently specific, with a less flat top, and heavy-tailed.

Figures 17–20 consistently suggest that context extremity generally reduces the probability of assimilation, or equivalently, enhances the probability of contrast. Nevertheless, exceptionally, for symmetric focal objects, when the context is seriously skewed to the focal object, and for symmetric contexts, when the focal object is seriously skewed to the context, if the context is in the vicinity of the original best representation of the focal object, then context extremity enhances the probability of assimilation, reducing the probability of contrast.

In addition to the foregoing, these figures consistently suggest that the probability of assimilation is higher when the context is skewed to a symmetric focal object than when it is skewed away from the focal object; and is higher when the focal object is skewed to a symmetric context than when it is skewed away from the context.

For contexts in the vicinity of the original best representation of the focal object, the probability of the focal distribution assimilating to the contextual distribution is higher when the context is similarly skewed, ambiguous, peaked, and tailed. For sufficiently extreme contexts, the probability of the focal distribution contrasting away from the contextual distribution is higher when the context is specific, with a flatter top, and light-tailed.

Figures 17–20, finally, consistently suggest that other things being equal, skewness and ambiguity varying, then higher probability of the focal distribution contrasting away from the contextual distribution generally implies the other way around. This, combined with Figures 1–8, echoes the reciprocity hypothesis; see *Literature Review*.

Figure 21 suggests that greater evaluative volatility, as its name literally suggests, allows for greater shift of the best representation in either event.

Figure 22 suggests that cognitive consumption diminishes the influence of context extremity on the magnitude of the shift of the best representation of the focal object. Given extremity, it can be solved for the critical cognitive consumption level that maximises context effects. As the context goes from moderate to extreme, the critical level steadily goes up.

Figure 1

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Context Skewness ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

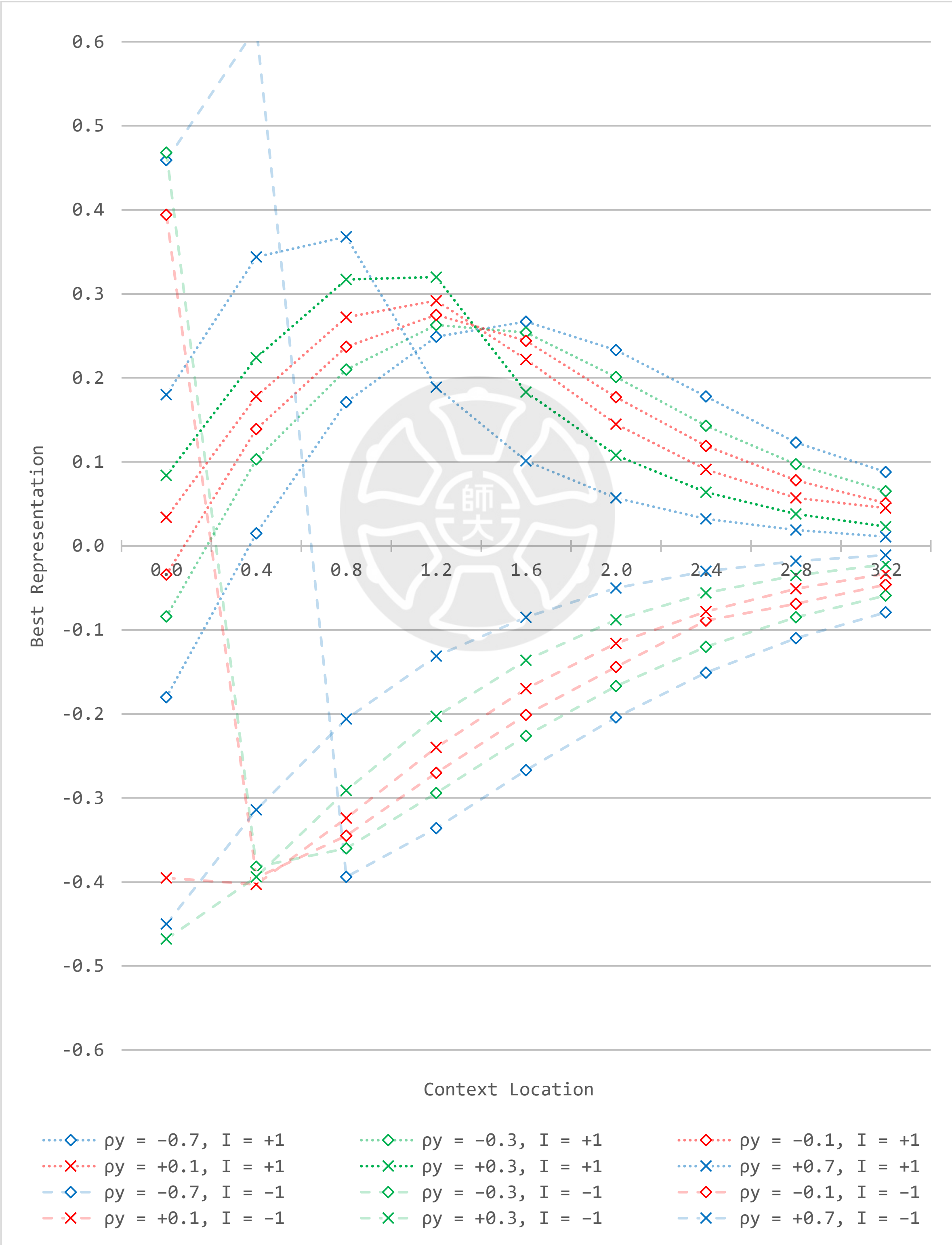


Figure 2

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Context Ambiguity ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

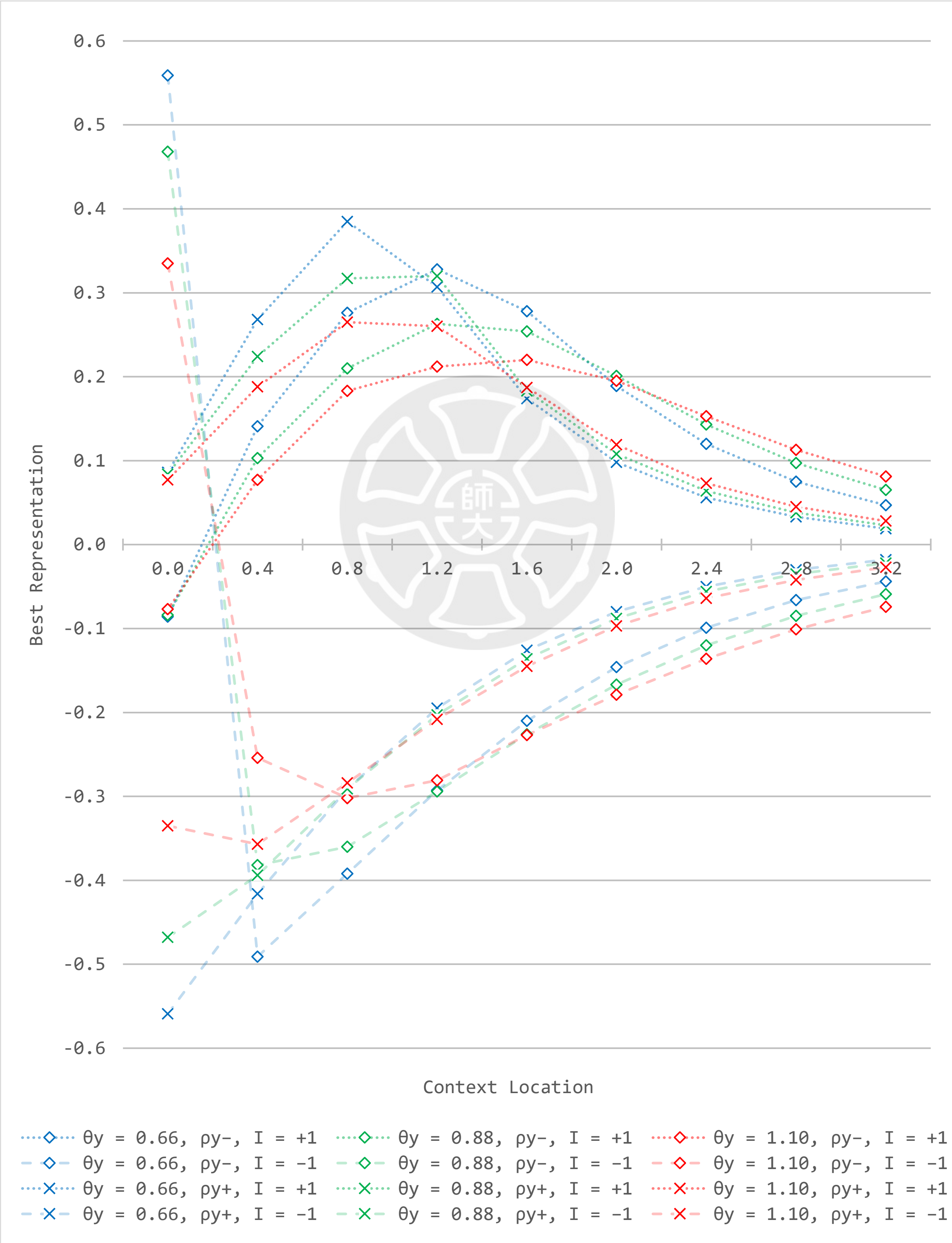


Figure 3

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Context Peak Behaviour ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = 2.5$, and $\eta_X = \eta_Y = 3.6$)

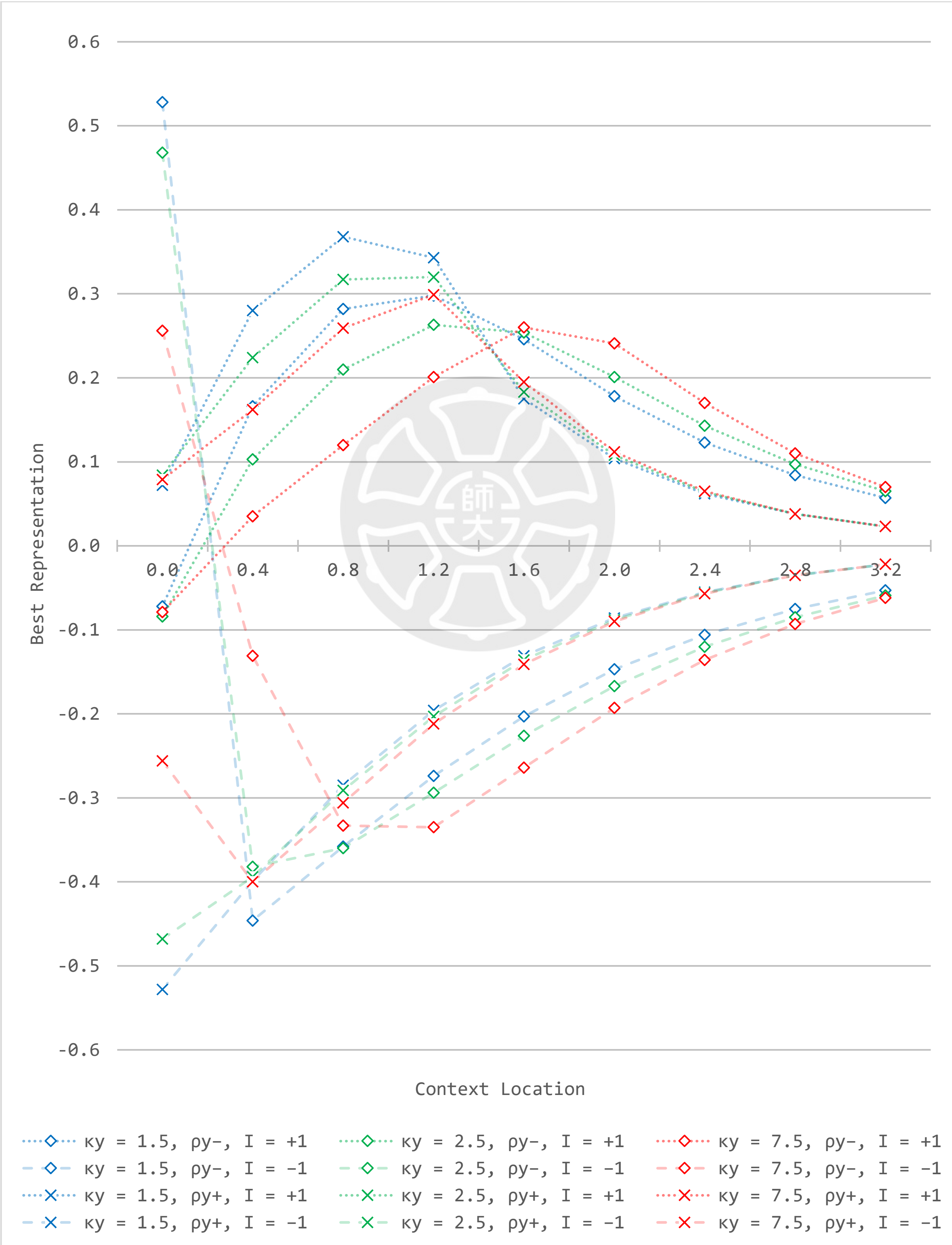


Figure 4

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Context Tail Behaviour ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = 3.6$)

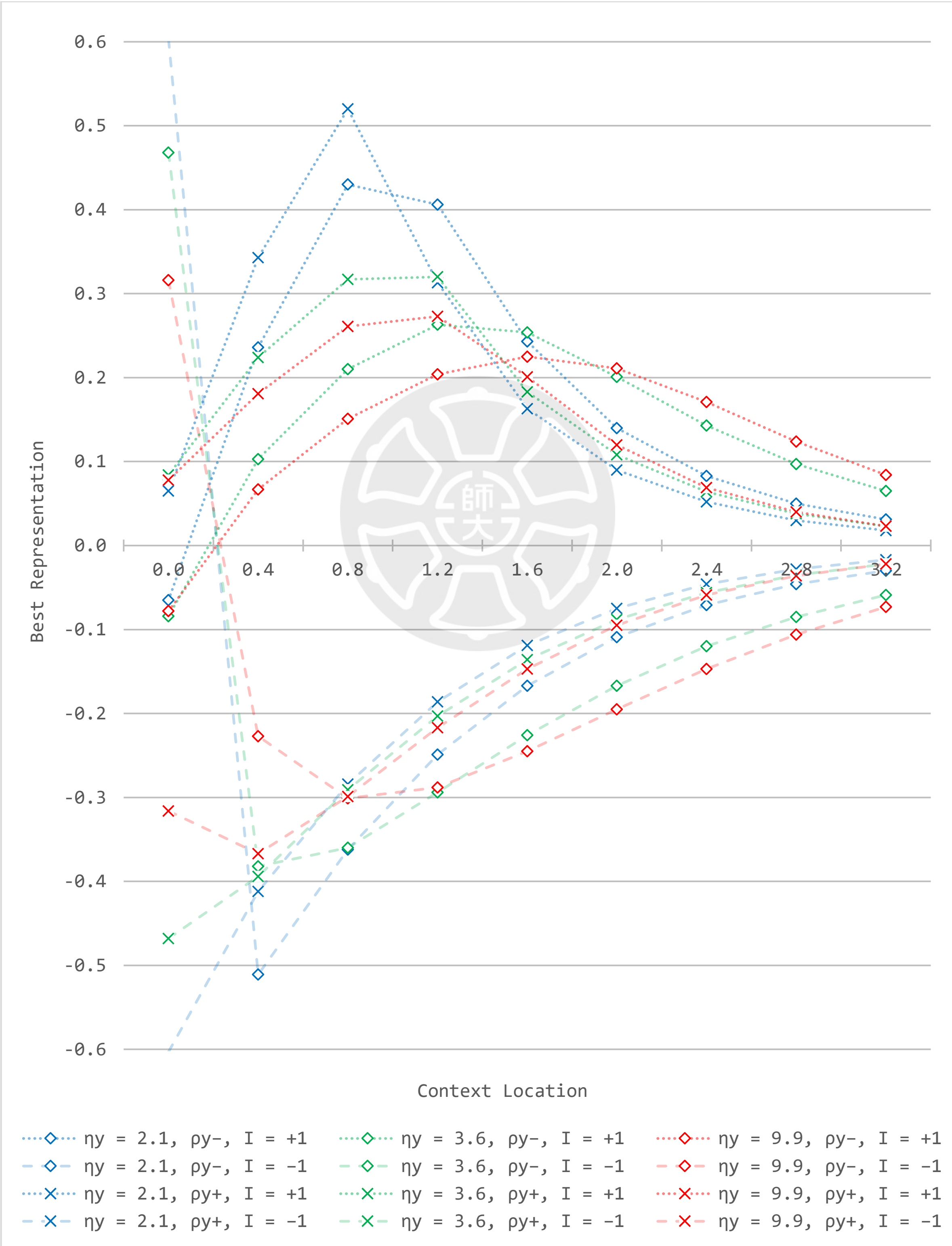


Figure 5

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Focal Object Skewness ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_Y = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

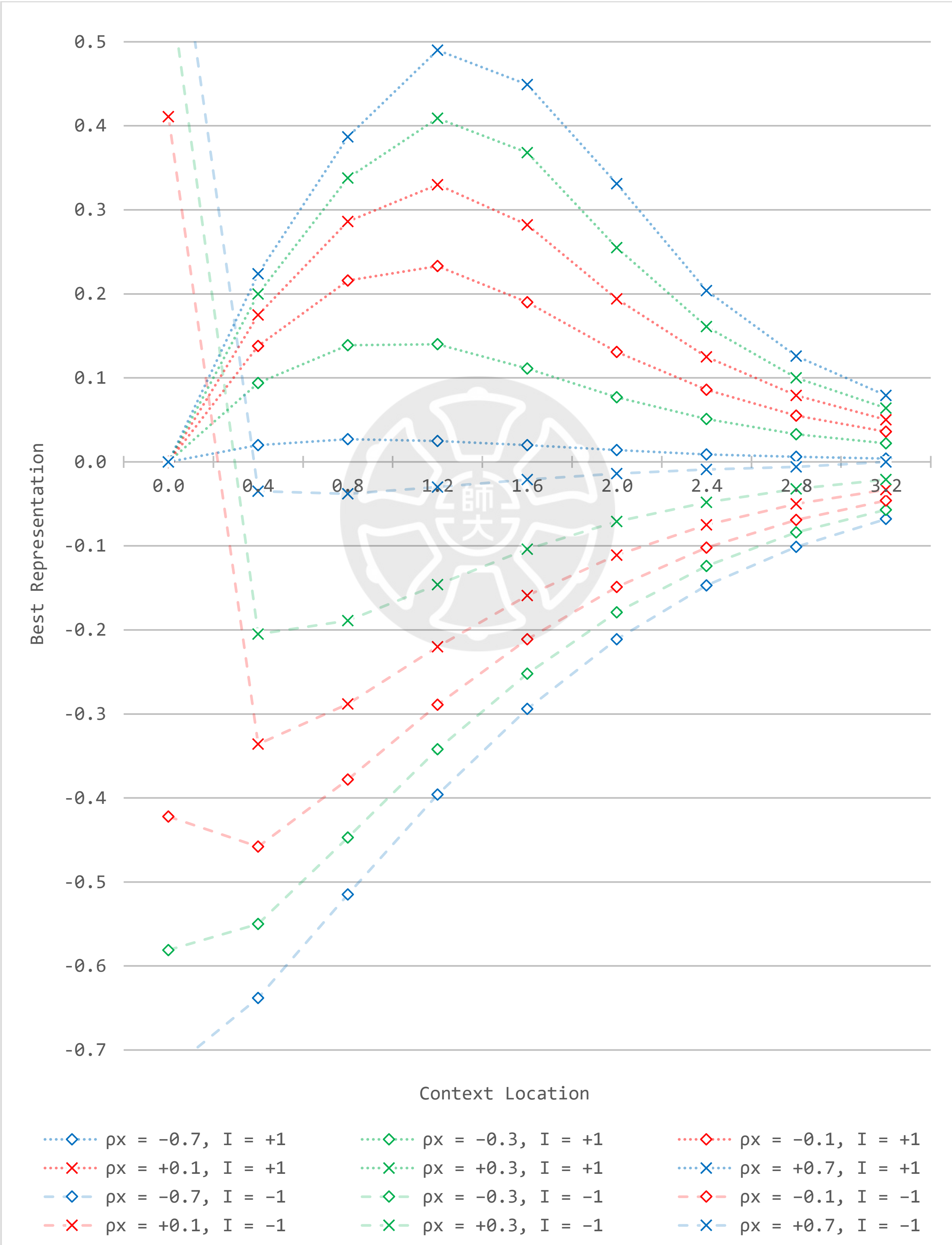


Figure 6

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Focal Object Ambiguity ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_Y = 0.88$, $|\rho_X| = 0.3$, $\rho_Y = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

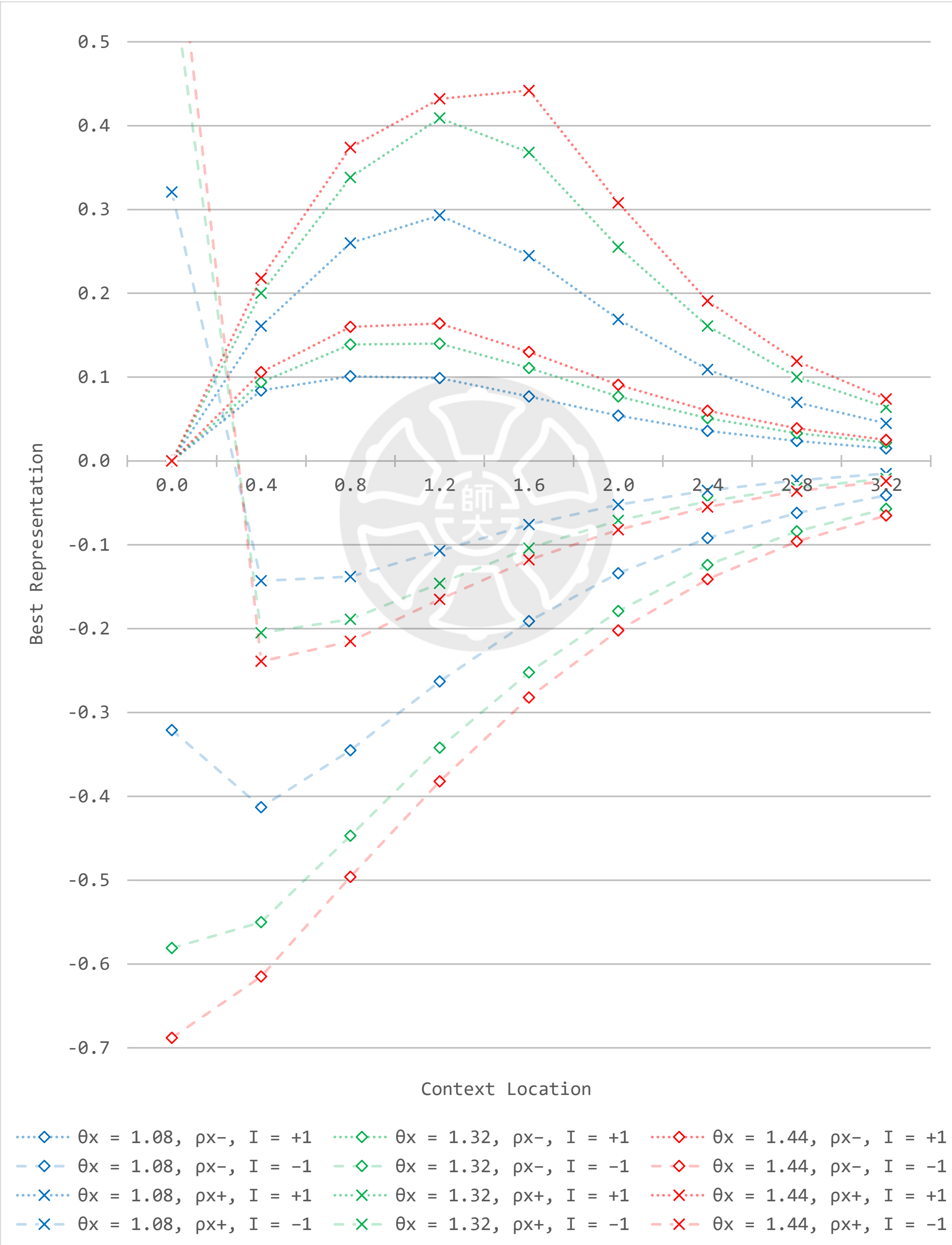


Figure 7

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Focal Object Peak Behaviour ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $|\rho_X| = 0.3$, $\rho_Y = 0.0$, $\kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

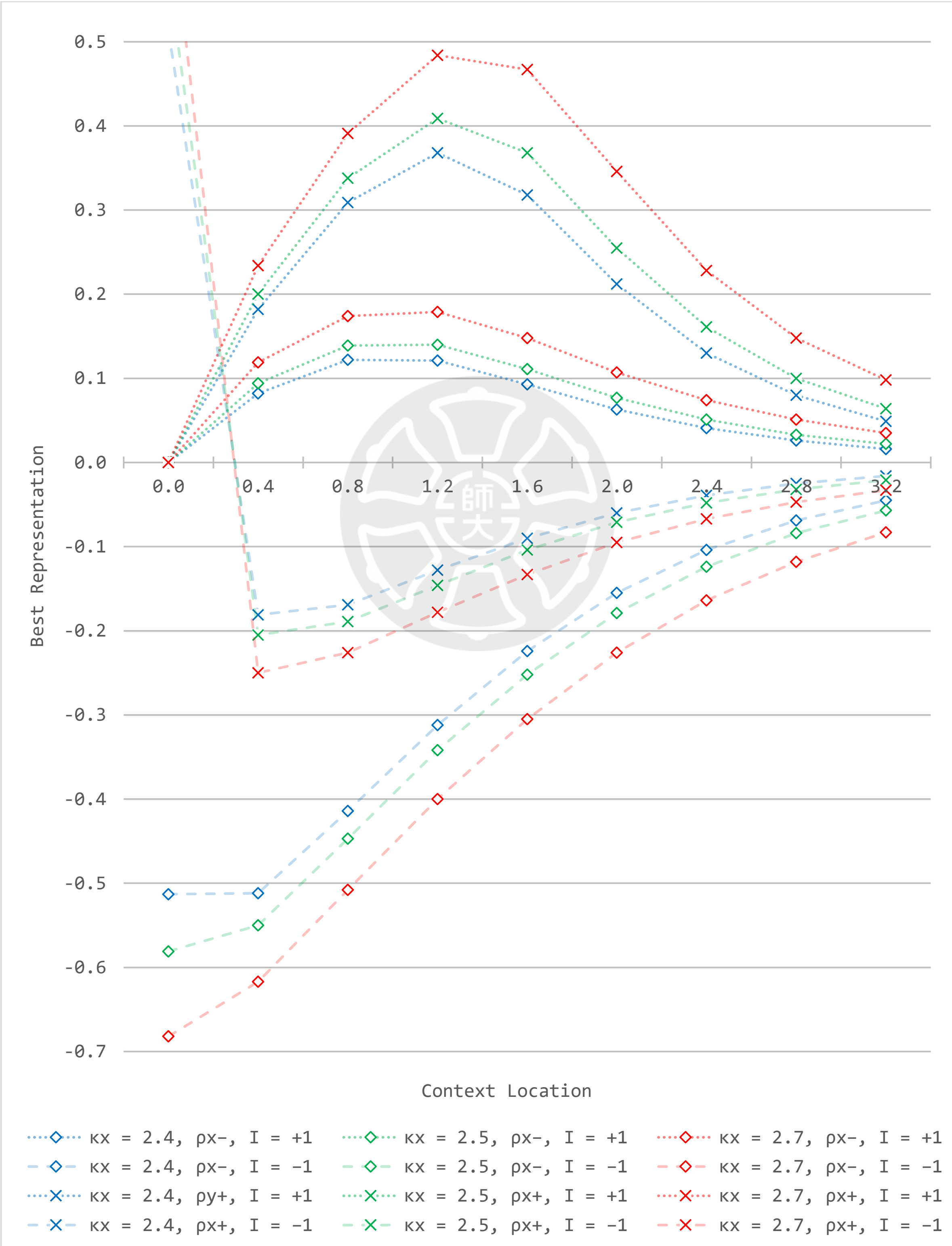


Figure 8

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Focal Object Tail Behaviour ($\alpha = 0.5$, $\delta = 0.9$, $\xi_X = 0.0$, $|\rho_X| = 0.3$, $\rho_Y = 0.0$, $\theta_X = 5.0$, $\theta_Y = 1.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_Y = 3.6$)

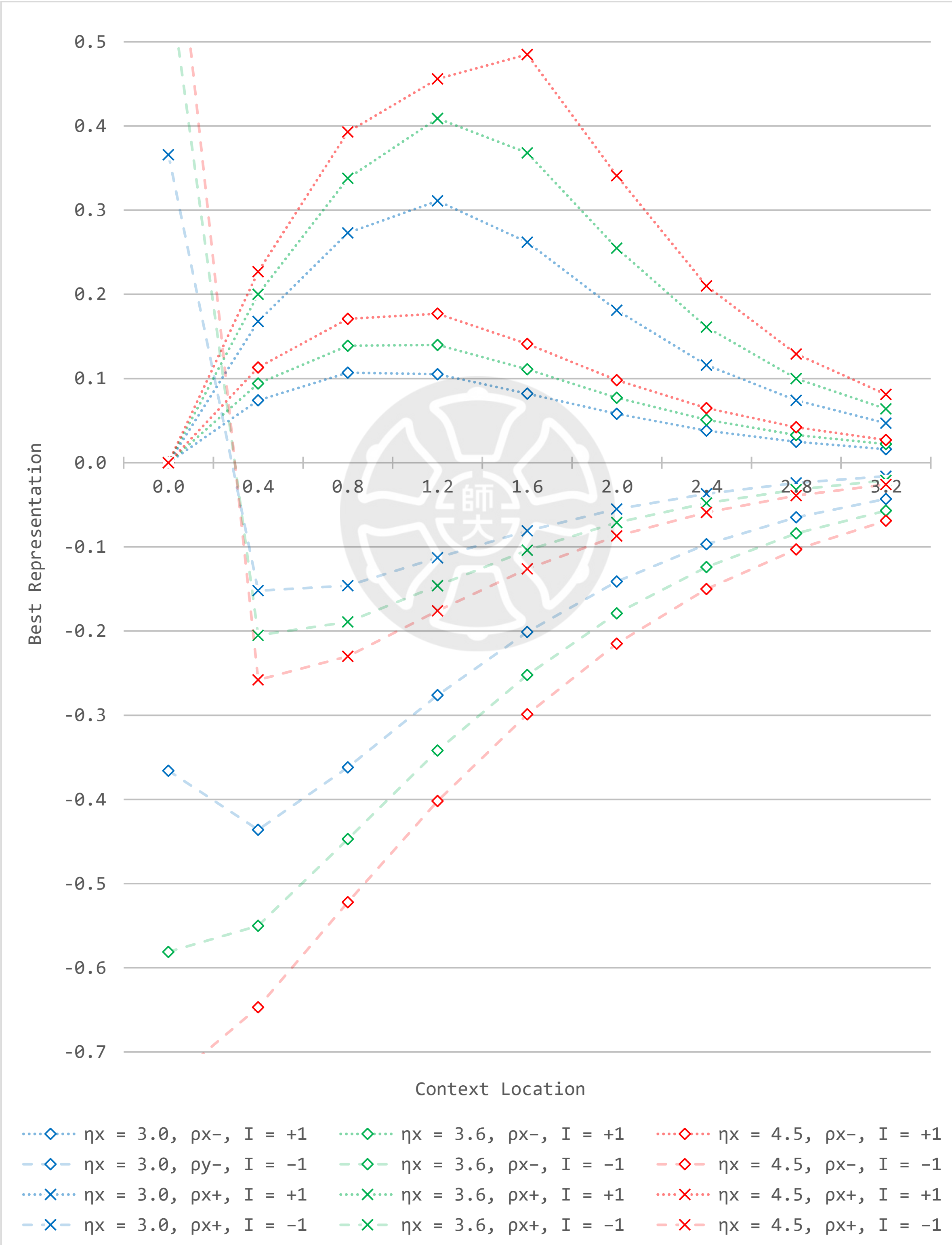


Figure 9

Skewed Generalised Student's t: The Expected Influence of Context Extremity on the Best Representation, Moderated by Context Skewness ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

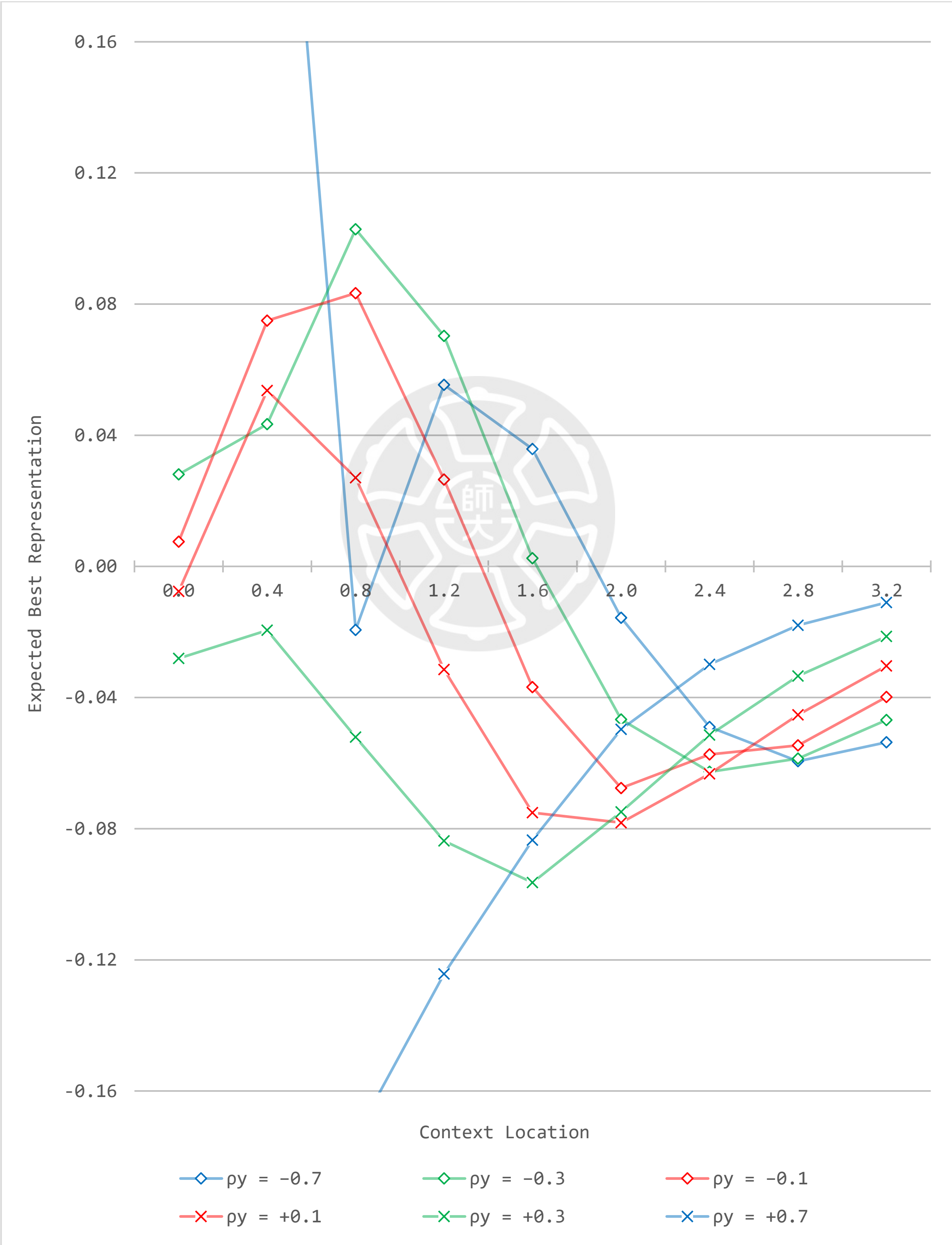


Figure 10

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Context Ambiguity ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

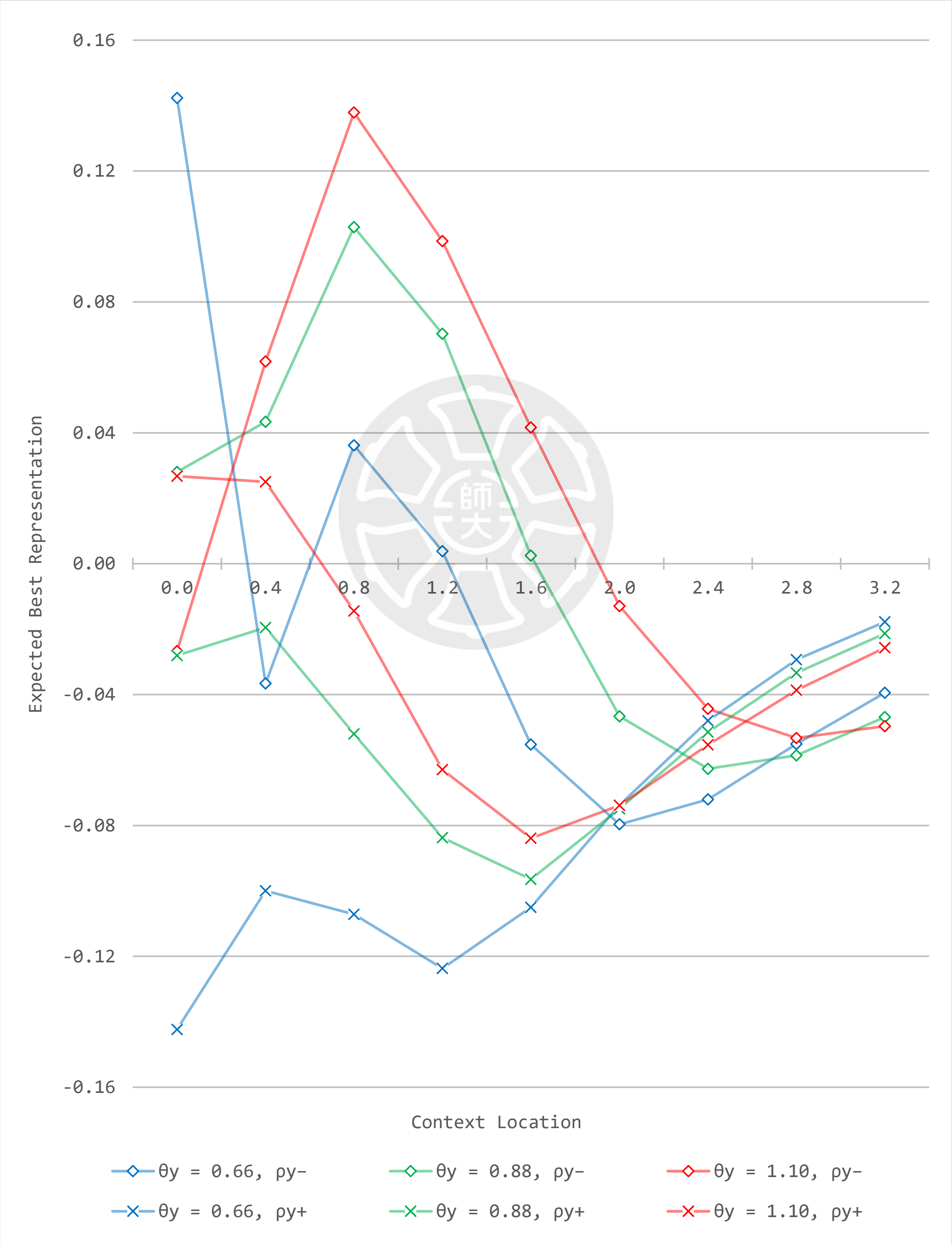


Figure 11

Skewed Generalised Student's t: The Expected Influence of Context Extremity on the Best Representation, Moderated by Context Peak Behaviour ($\alpha = 0.5$, $\lambda = 1.8$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = 2.5$, and $\eta_X = \eta_Y = 3.6$)

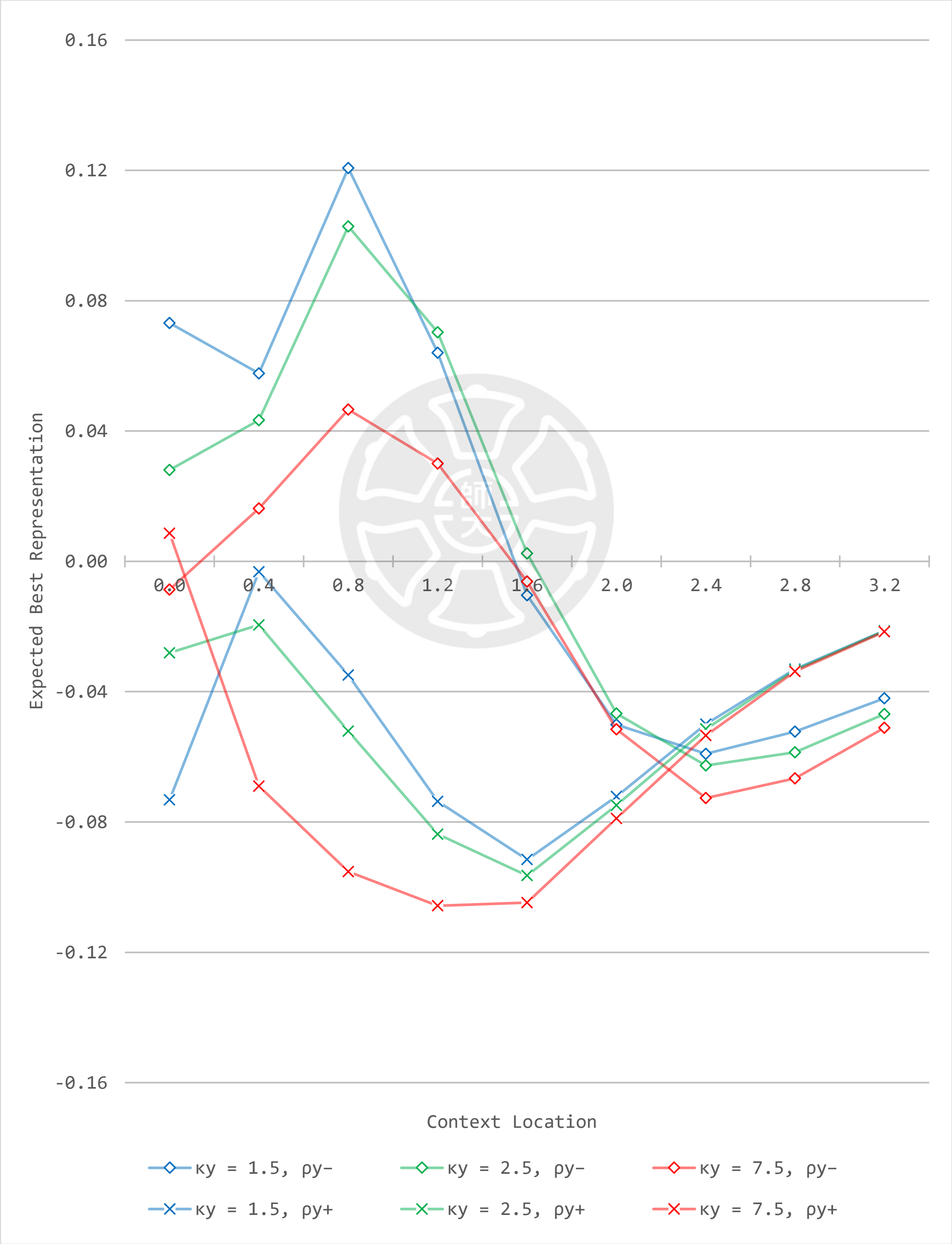


Figure 12

Skewed Generalised Student's t: The Expected Influence of Context Extremity on the Best Representation, Moderated by Context Tail Behaviour ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = 3.6$)

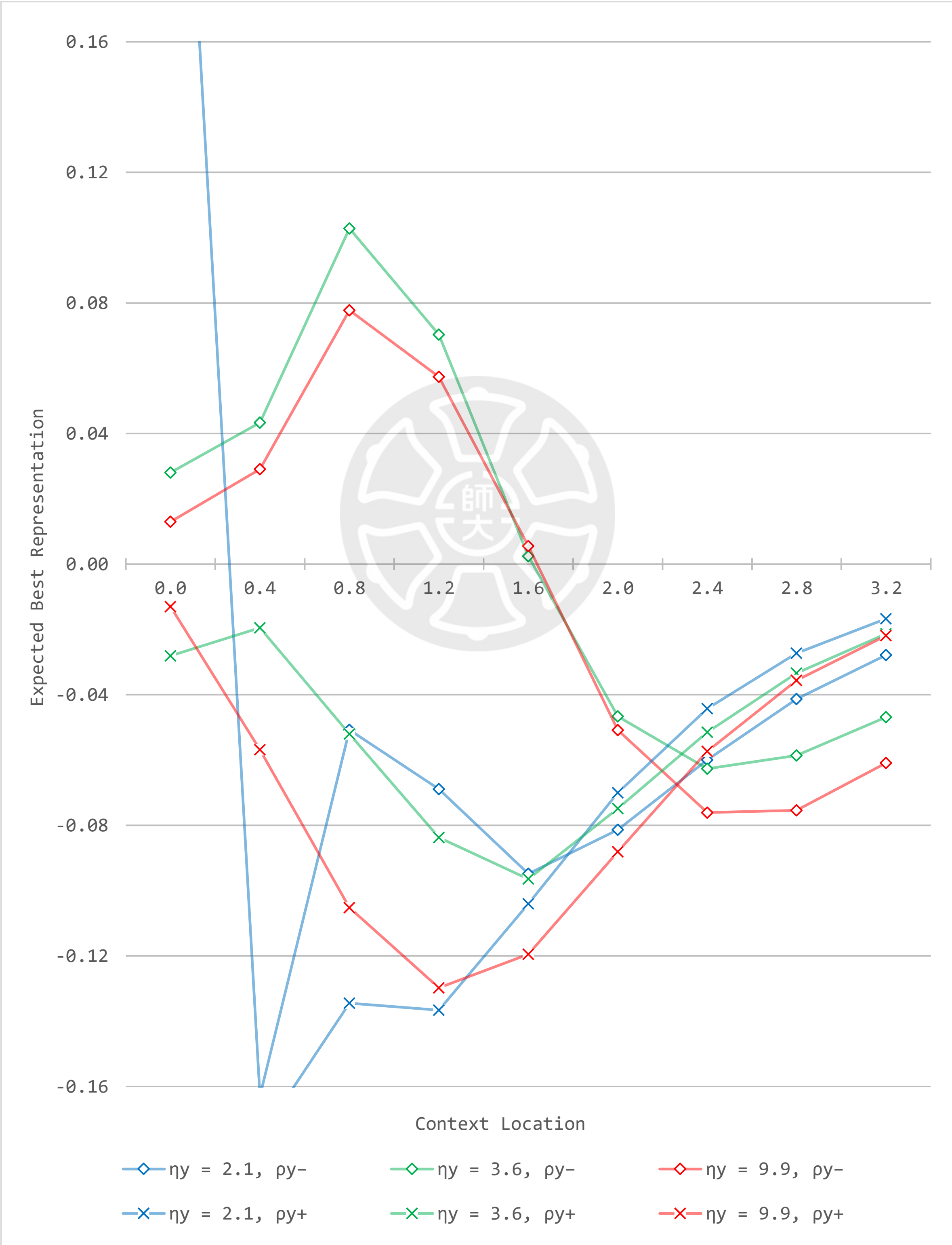


Figure 13

Skewed Generalised Student's t: The Expected Influence of Context Extremity on the Best Representation, Moderated by Focal Object Skewness ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_Y = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

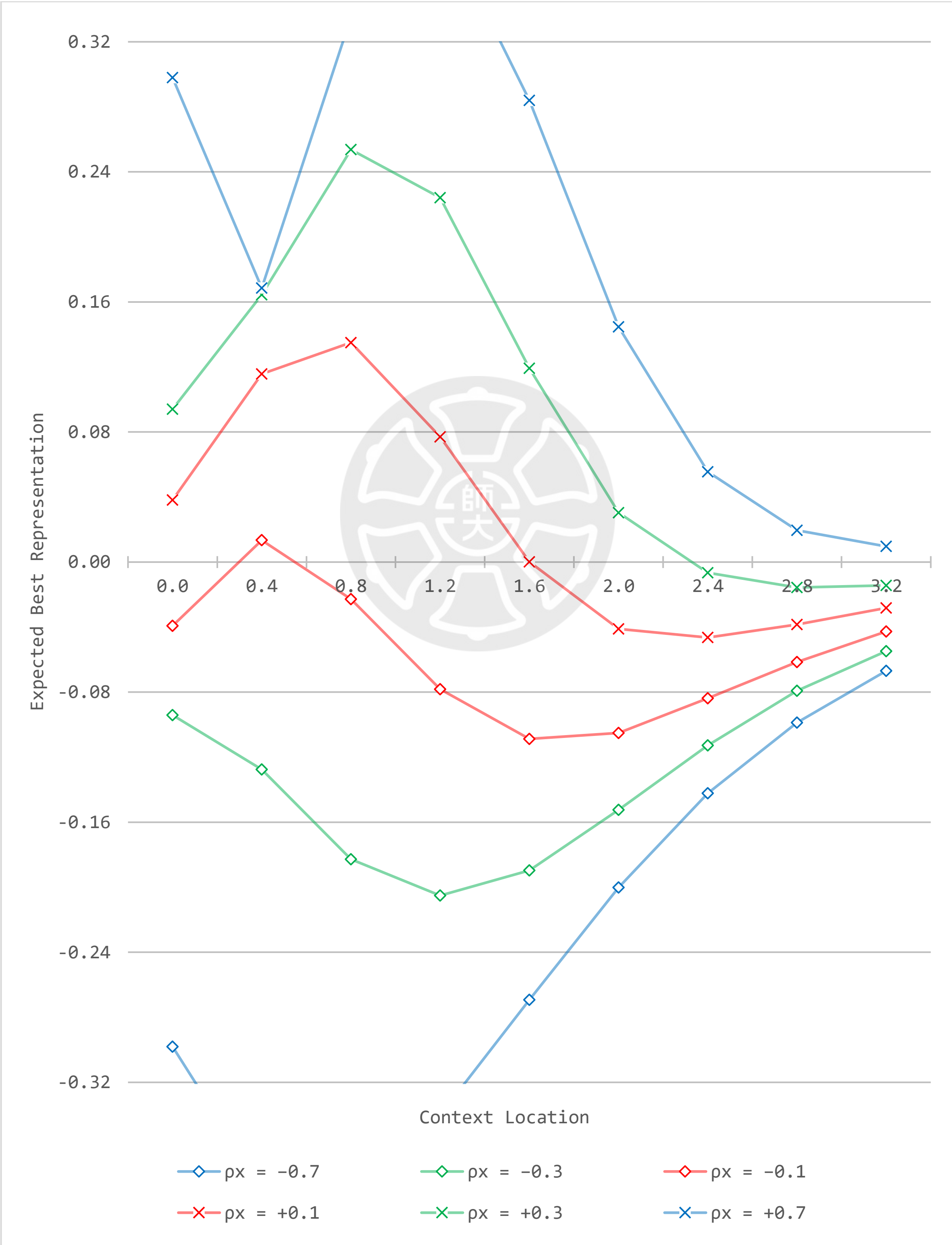


Figure 14

Skewed Generalised Student's t: The Expected Influence of Context Extremity on the Best Representation, Moderated by Focal Object Ambiguity ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_Y = 0.88$, $|\rho_X| = 0.3$, $\rho_Y = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

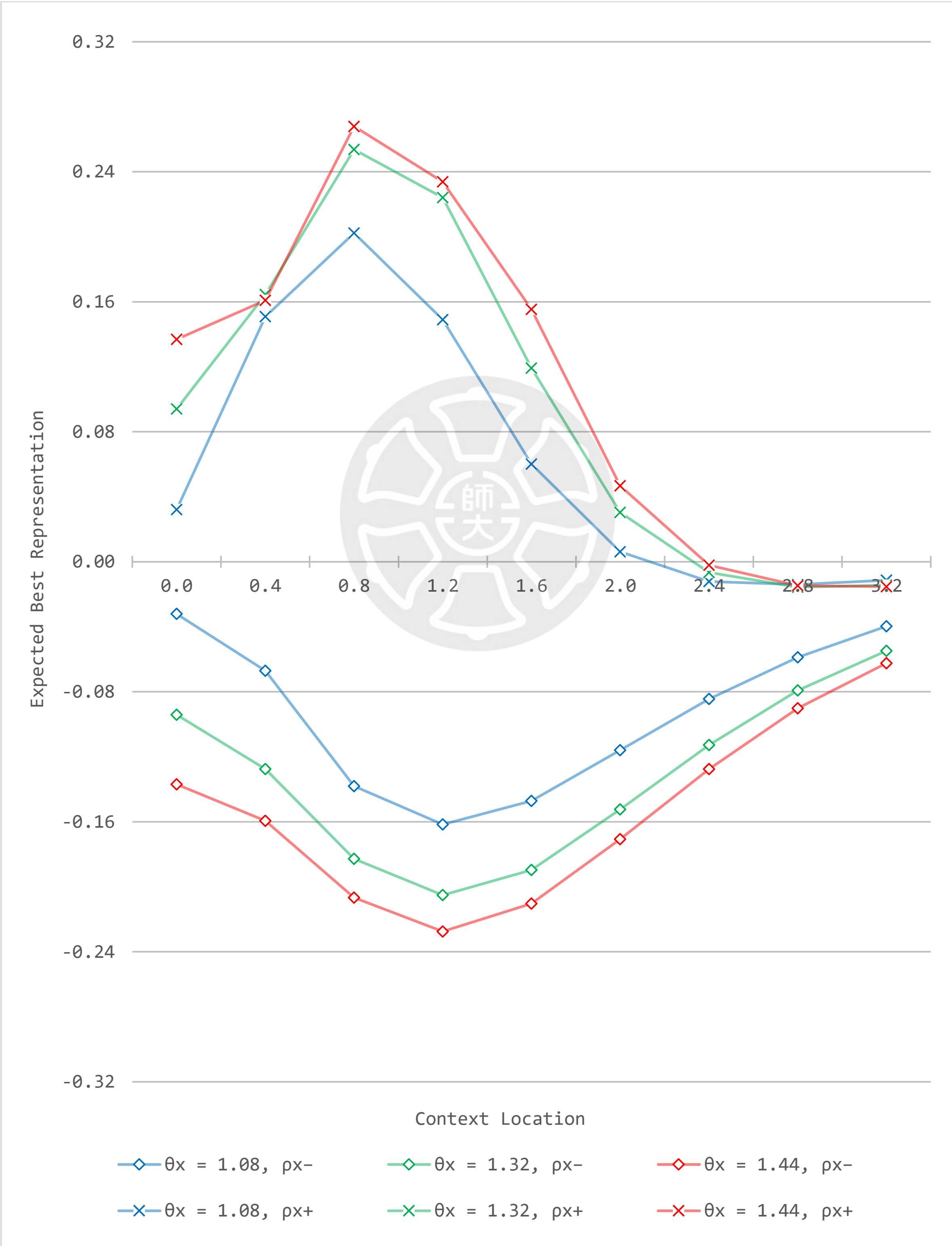


Figure 15

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Focal Object Peak Behaviour ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $|\rho_X| = 0.3$, $\rho_Y = 0.0$, $\kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

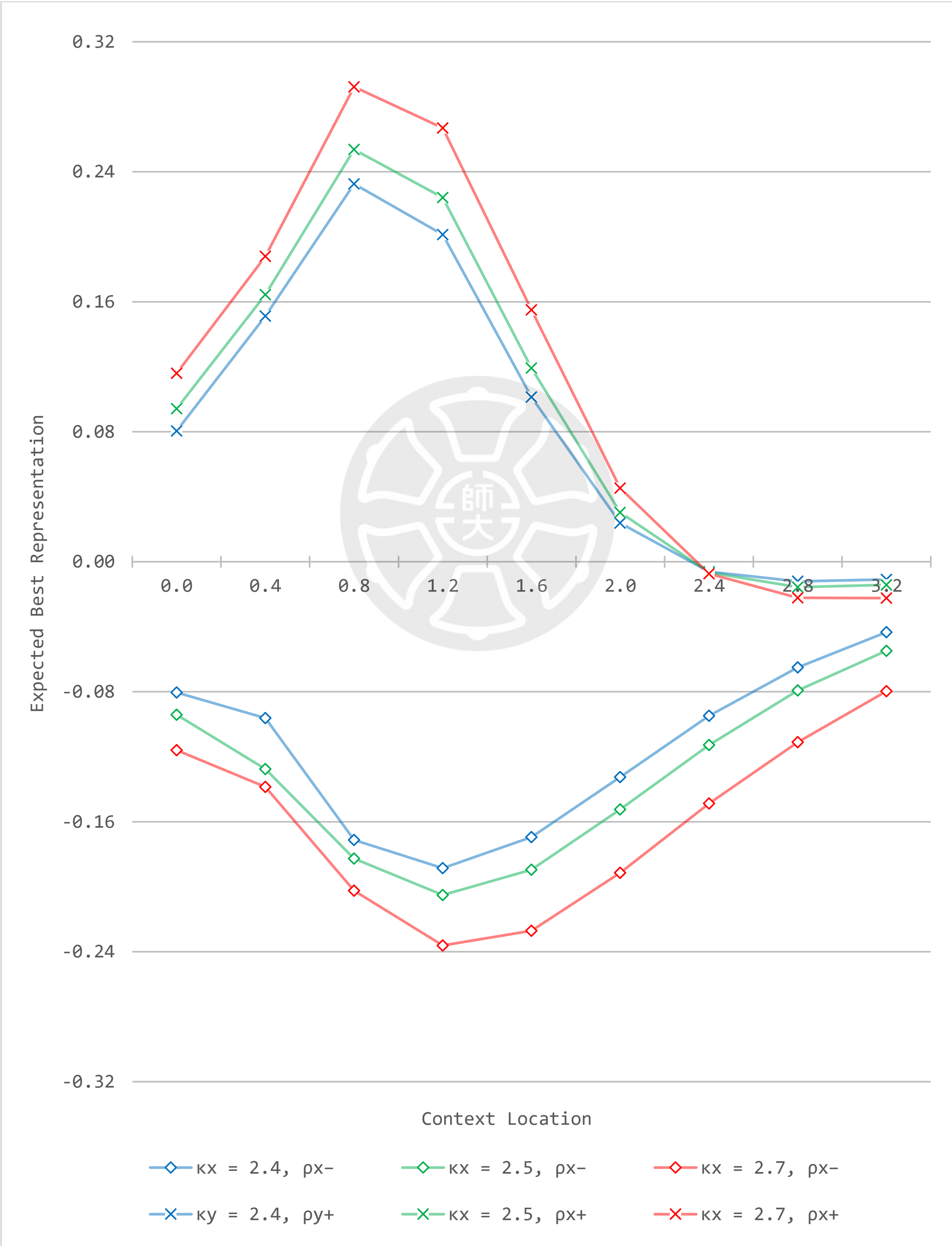


Figure 16

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Focal Object Tail Behaviour ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\xi_X = 0.0$, $|\rho_X| = 0.3$, $\rho_Y = 0.0$, $\theta_X = 5.0$, $\theta_Y = 1.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_Y = 3.6$)

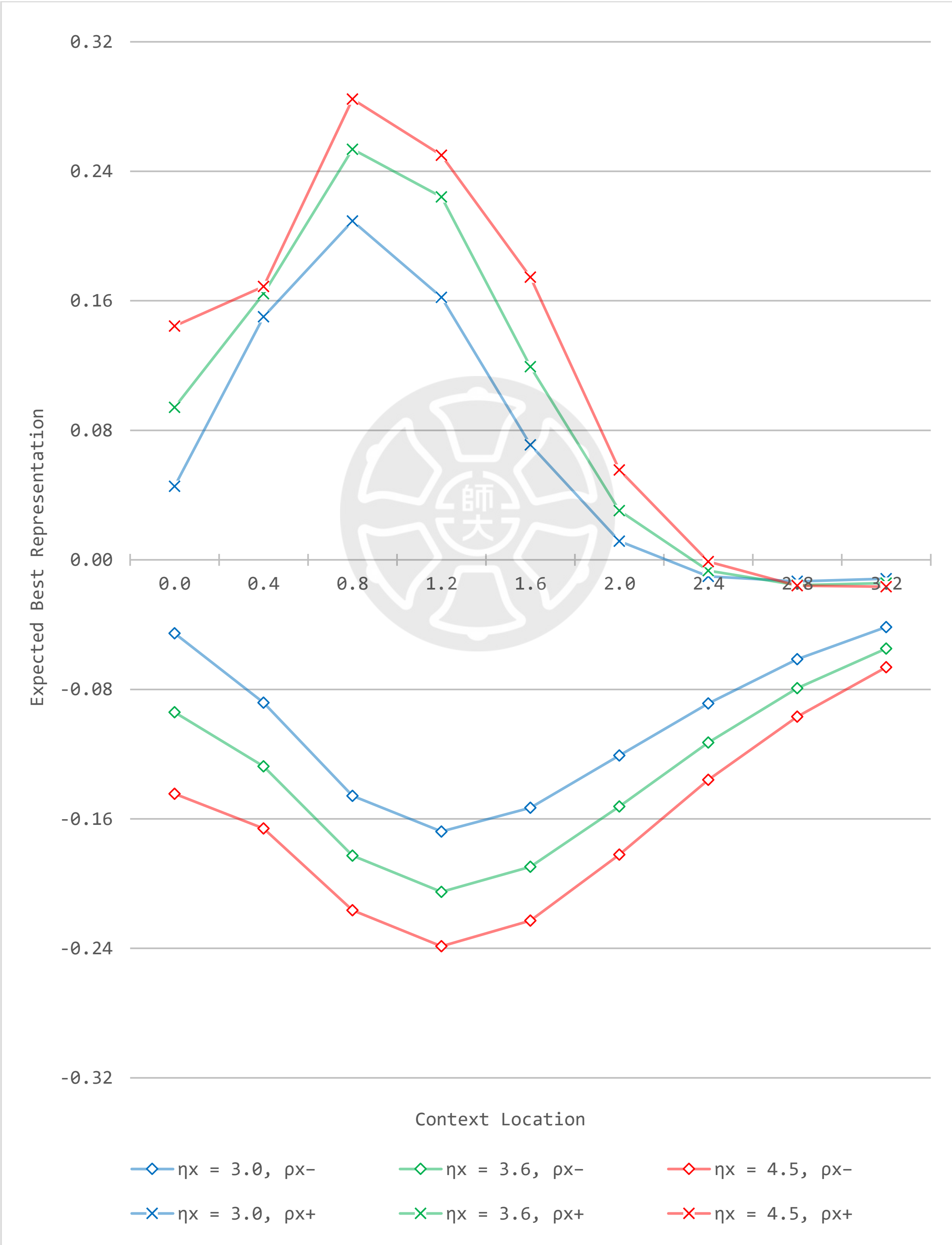


Figure 17

Skewed Generalised Student's t: Skewness on the Probability of Assimilation ($\lambda = 1.8$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

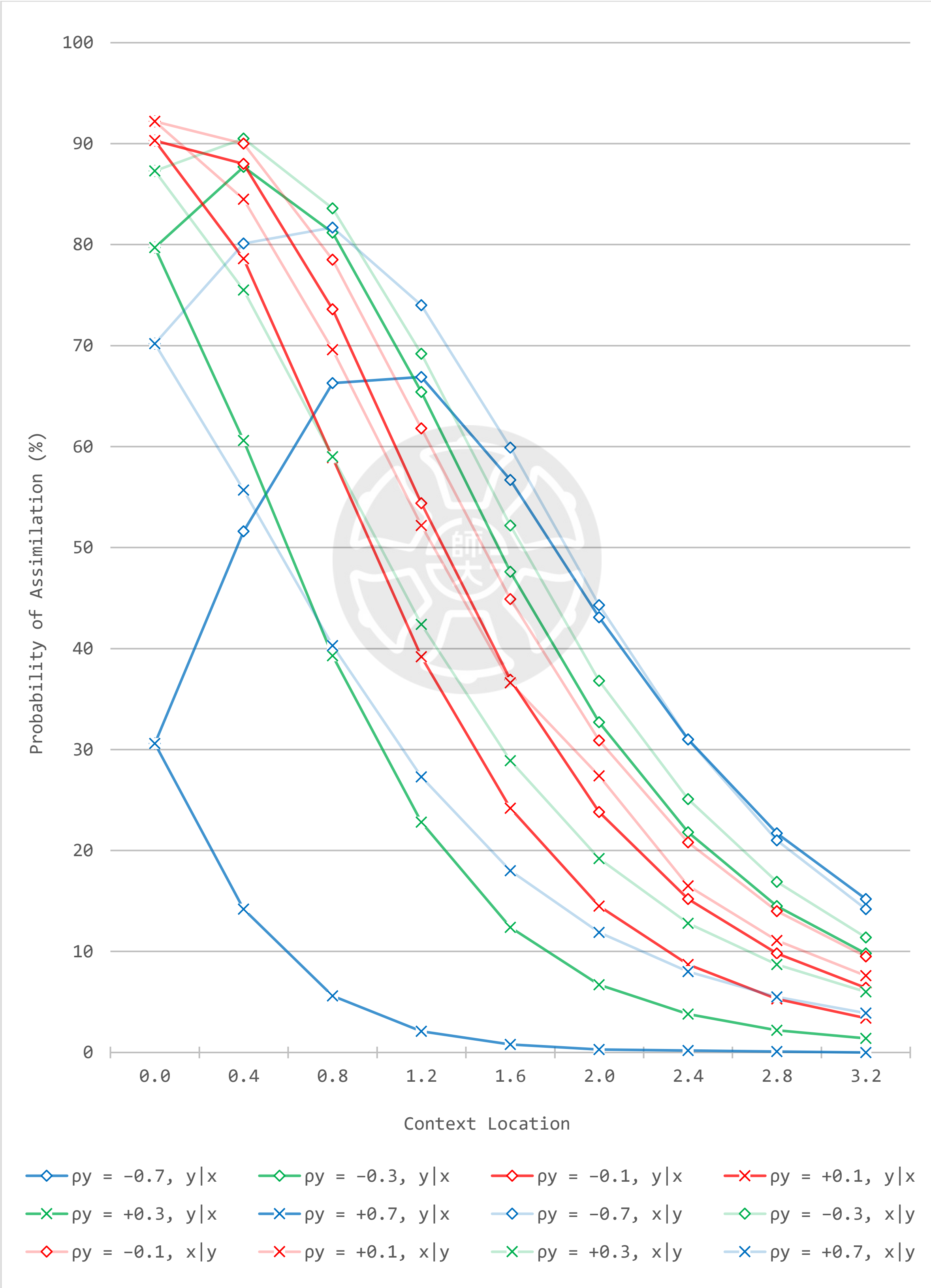


Figure 18

Skewed Generalised Student's t: Ambiguity on the Probability of Assimilation ($\lambda = 1.8$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

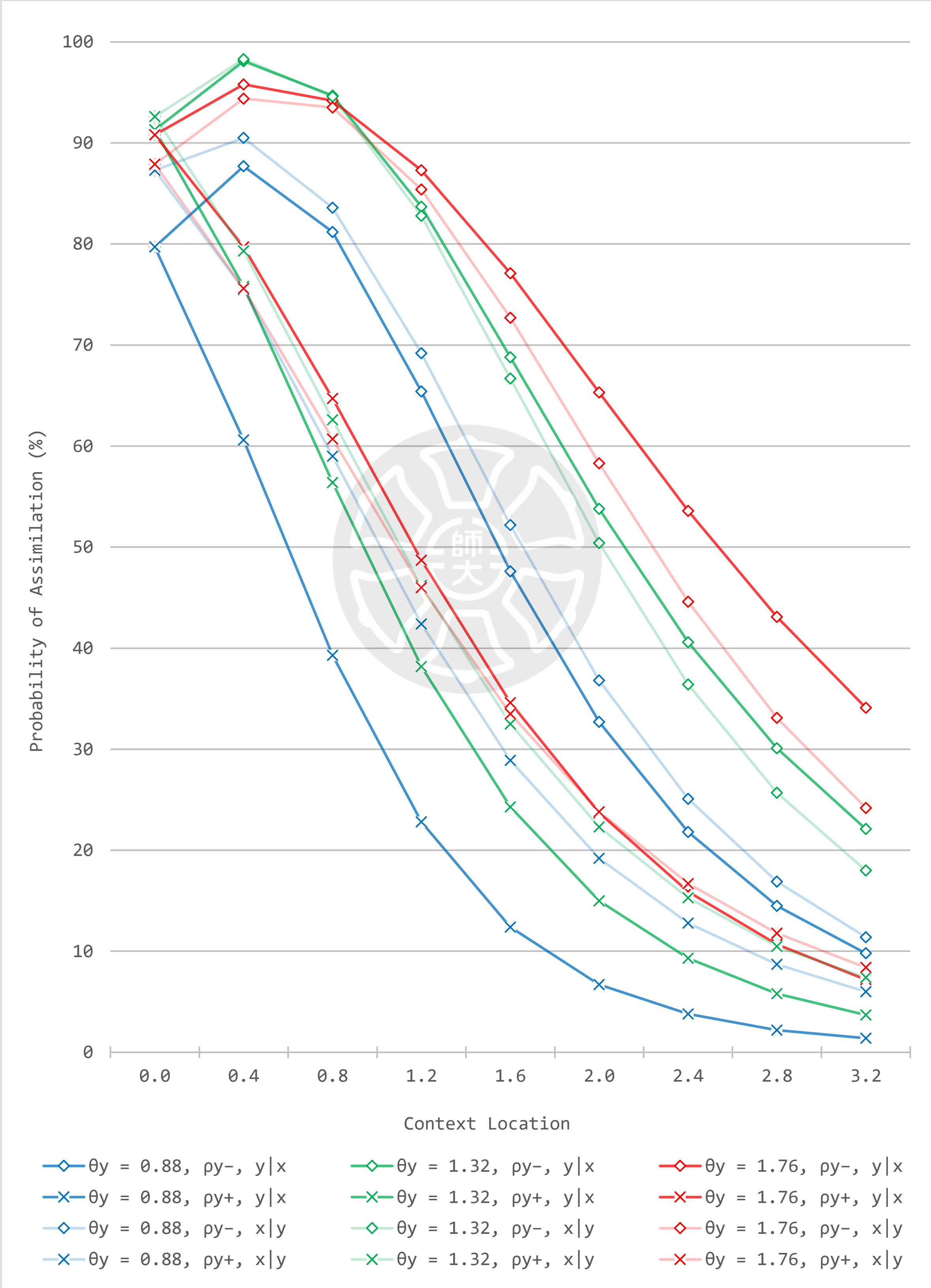


Figure 19

Skewed Generalised Student's t: Peak on the Probability of Assimilation ($\lambda = 1.8$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = 2.5$, and $\eta_X = \eta_Y = 3.6$)

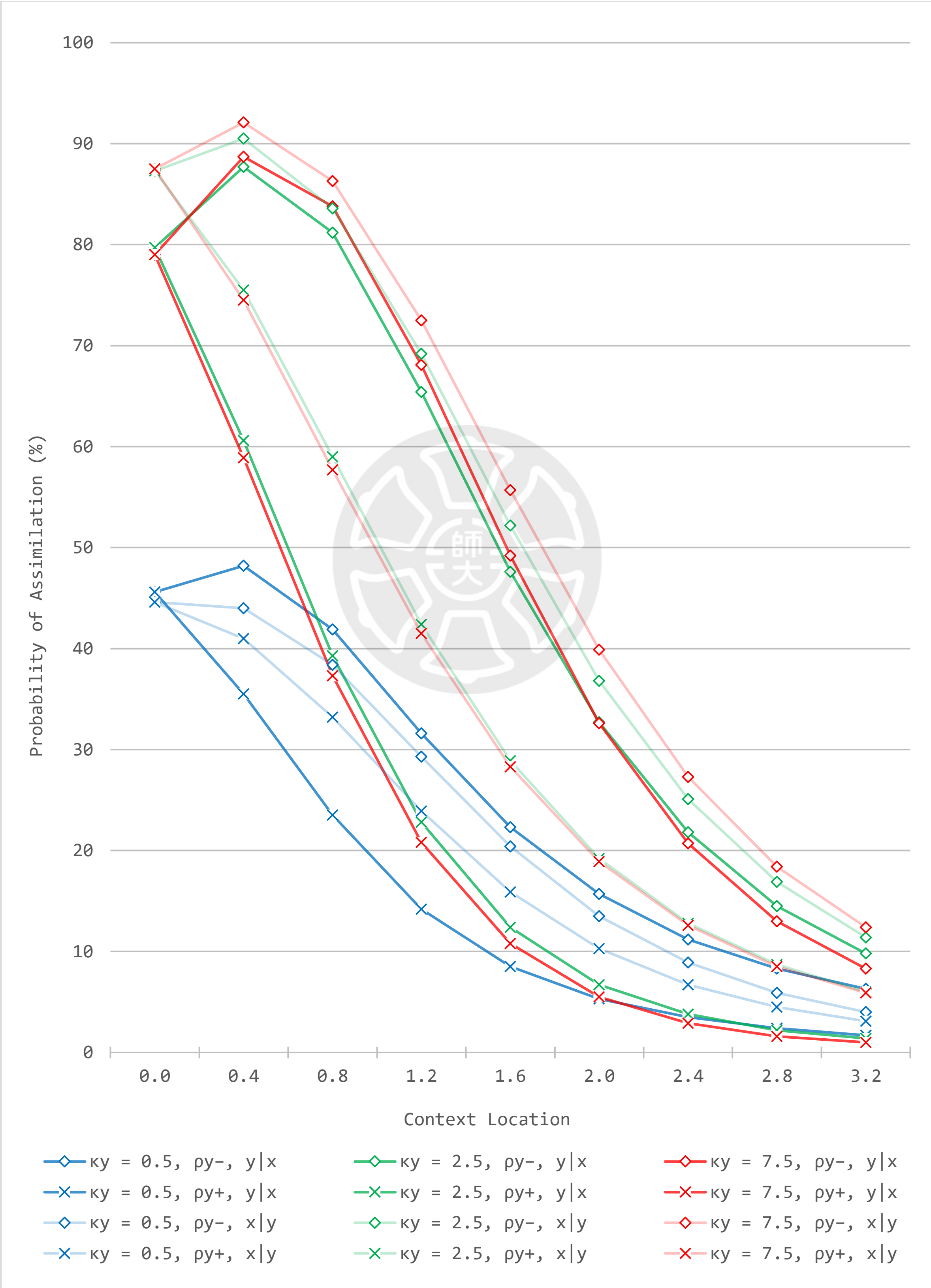


Figure 20

Skewed Generalised Student's t: Tails on the Probability of Assimilation ($\lambda = 1.8$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $|\rho_Y| = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = 3.6$)

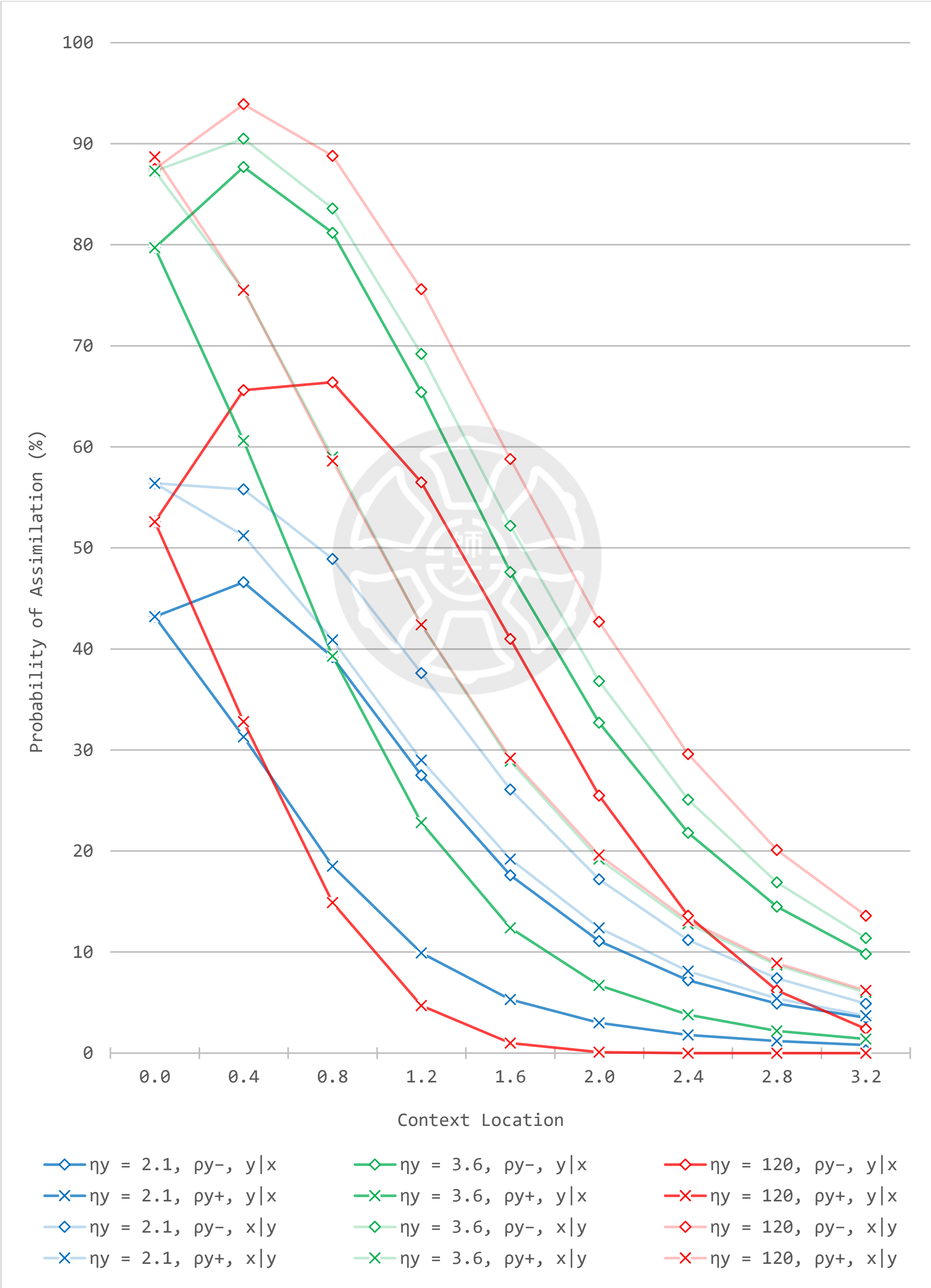


Figure 21

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Evaluative Volatility ($\alpha = 0.5$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\rho_Y = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

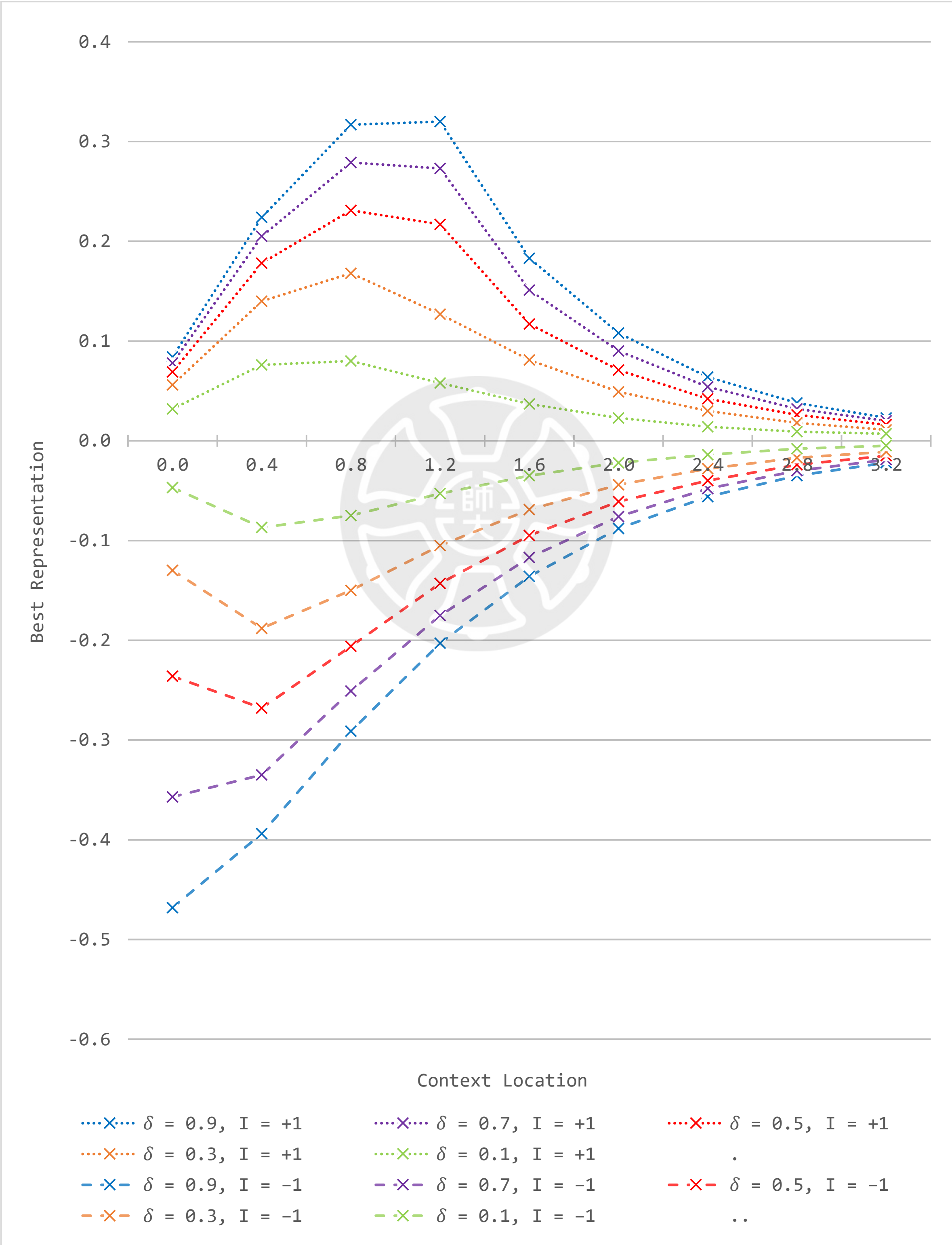
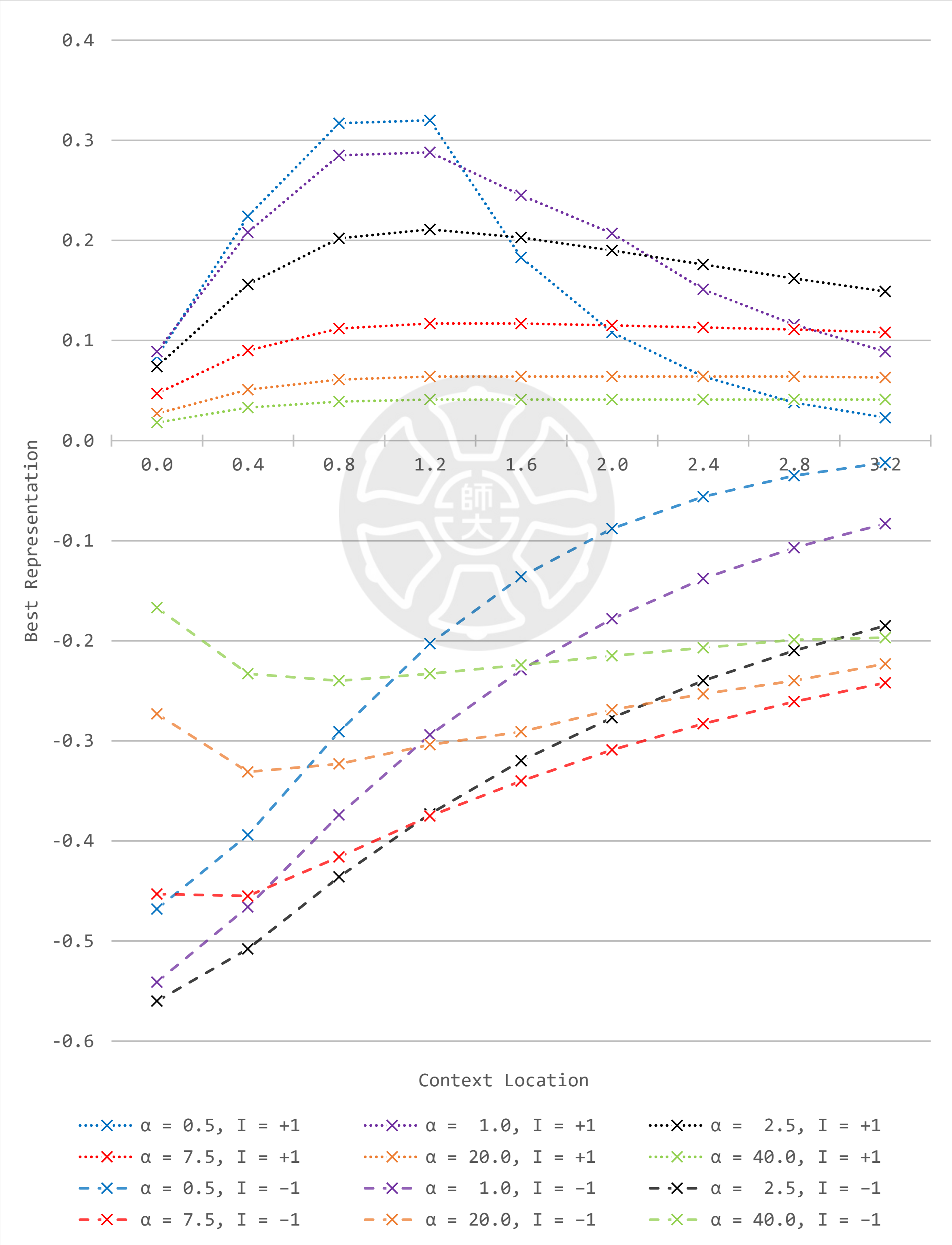


Figure 22

Skewed Generalised Student's t: The Realised Influence of Context Extremity on the Best Representation, Moderated by Cognitive Consumption ($\delta = 0.9$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\rho_Y = 0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)



This is the end of this subsection. The inclusion of skewness, peak and tail behaviour, on the one hand, extends the results in *Analytical Approach*; on the other hand, it gives conflicting findings which call for careful discussion. In the following subsection, the author attempts further generalisation by introducing bimodality and considering a differently supported distribution.

Two Mixed Weibull Distributions

The Weibull distribution is supported on a half-bounded interval $[0, \infty)$, rather than $(-\infty, \infty)$. The distribution has the following probability density:

$$\mathbf{D}_k(v) = \frac{\tau_k}{\theta_k}\left(\frac{v}{\theta_k}\right)^{\tau_k - 1} e^{-(v/\theta_k)^{\tau_k}} \tag{106}$$

$\tau_k > 0$. $\mathbf{D}_k(v)$ is maximum at $v = 0$ if $\tau_k \leq 1$, and $v = \xi_k$ if $\tau_k > 1$, where

$$\frac{\xi_k}{\theta_k} = \left(1 - \frac{1}{\tau_k}\right)^{1/\tau_k} \tag{107}$$

The derivation is left to readers. An expedient way of constructing a bimodal distribution is to let $\mathbf{D}(v) = \omega\, \mathbf{D}_0(v) + (1 - \omega)\, \mathbf{D}_1(v)$ where $1/2 < \omega < 1$ and $k \in \{0, 1\}$. $\mathbf{D}_0(\cdot)$ herein is referred to as the dominant distribution, $\mathbf{D}_1(\cdot)$ the recessive. The author is now interpreting the results. See pages 64–66 for the figures. The author is planning a light study.

Figures 23 and 24 consistently suggest that in general, the recessive distribution is more influential in shifting the best representation of the focal object when it is latently specific than ambiguous. The quirks in both figures are attributed to the fact that the recessive mode becomes more salient than the dominant mode, which is possible.

Take either of the blue distributions in Figure 25. Finally, recall the parametre β; see *Methods*. Here is an illustrative example: When β is small, one would take the recessive mode, despite its smaller contribution to the distribution; when β is large, one would instead take the dominant mode, despite it being clearly less salient.

Figure 23

Mixed Weibull: The Realised Influence of Recessive Context Extremity on the Best Representation, Moderated by Recessive Context Ambiguity ($\alpha = 4.0$, $\delta = 0.9$, $\xi_{X,0} = 12.5$, $\xi_{X,1} = 8.5$, $\xi_{Y,0} = 16.5$, $\theta_{X,0} = 12.75$, $\theta_{X,1} = 8.56$, $\theta_{Y,0} = 16.65$, $\omega_X = \omega_Y = 0.7$)

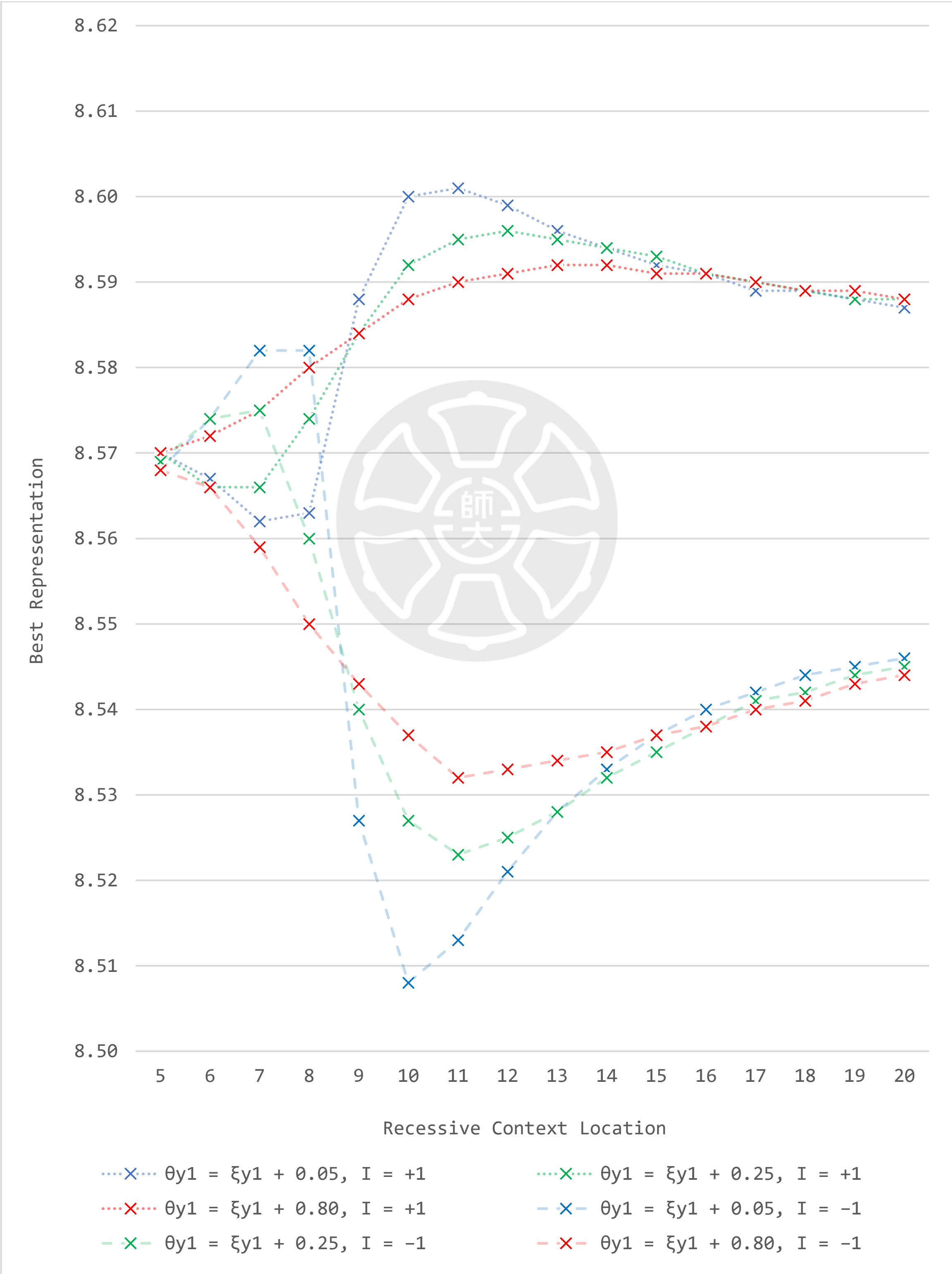


Figure 24

Mixed Weibull: The Realised Influence of Recessive Focal Object Extremity on the Best Representation, Moderated by Recessive Focal Object Ambiguity ($\alpha = 4.0$, $\delta = 0.9$, $\xi_{X,0} = 12.5$, $\xi_{Y,0} = 16.5$, $\xi_{Y,1} = 20.5$, $\theta_{X,0} = 12.75$, $\theta_{Y,0} = 16.65$, $\theta_{Y,1} = 20.52$, $\omega_X = \omega_Y = 0.7$)

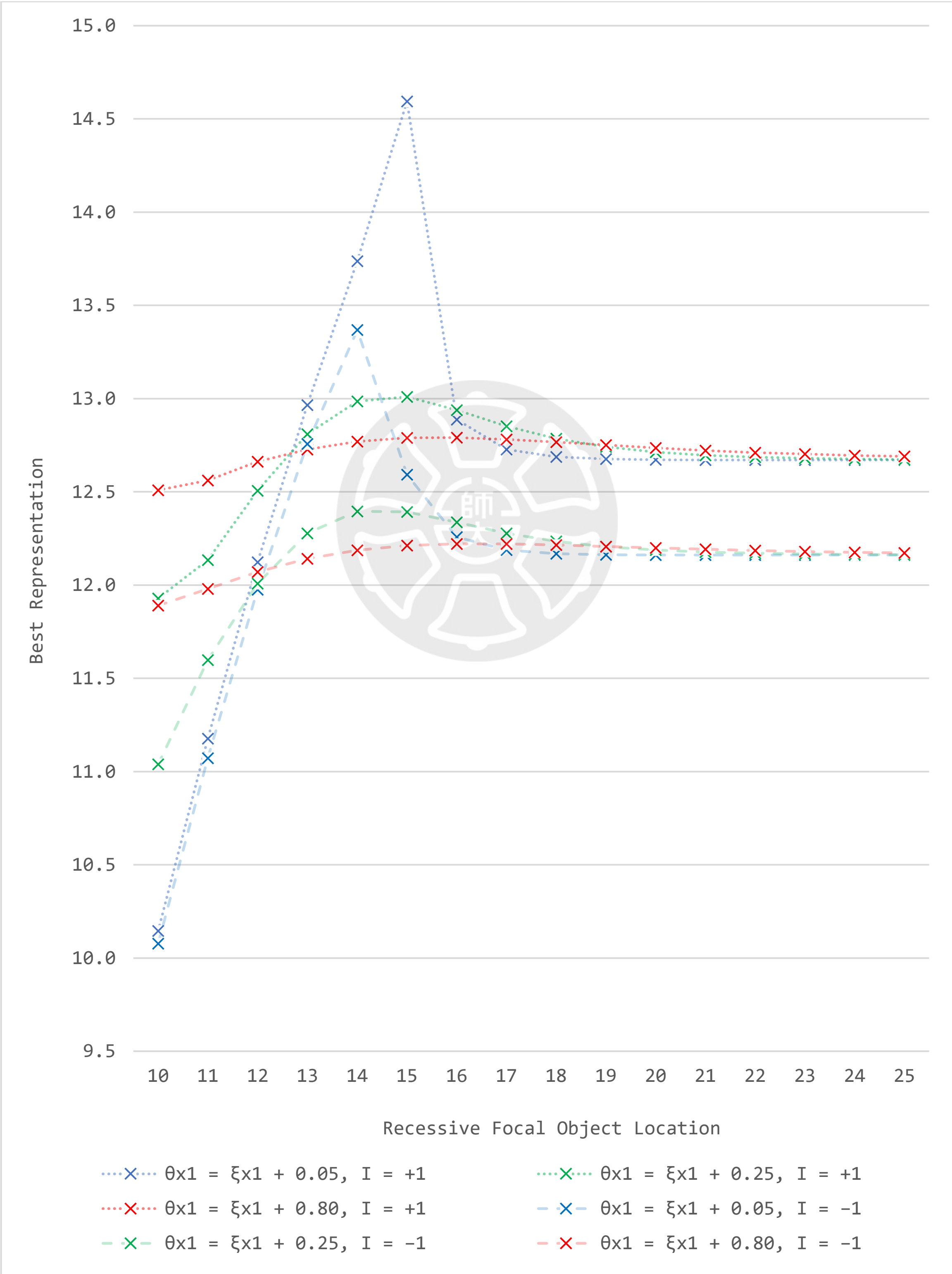
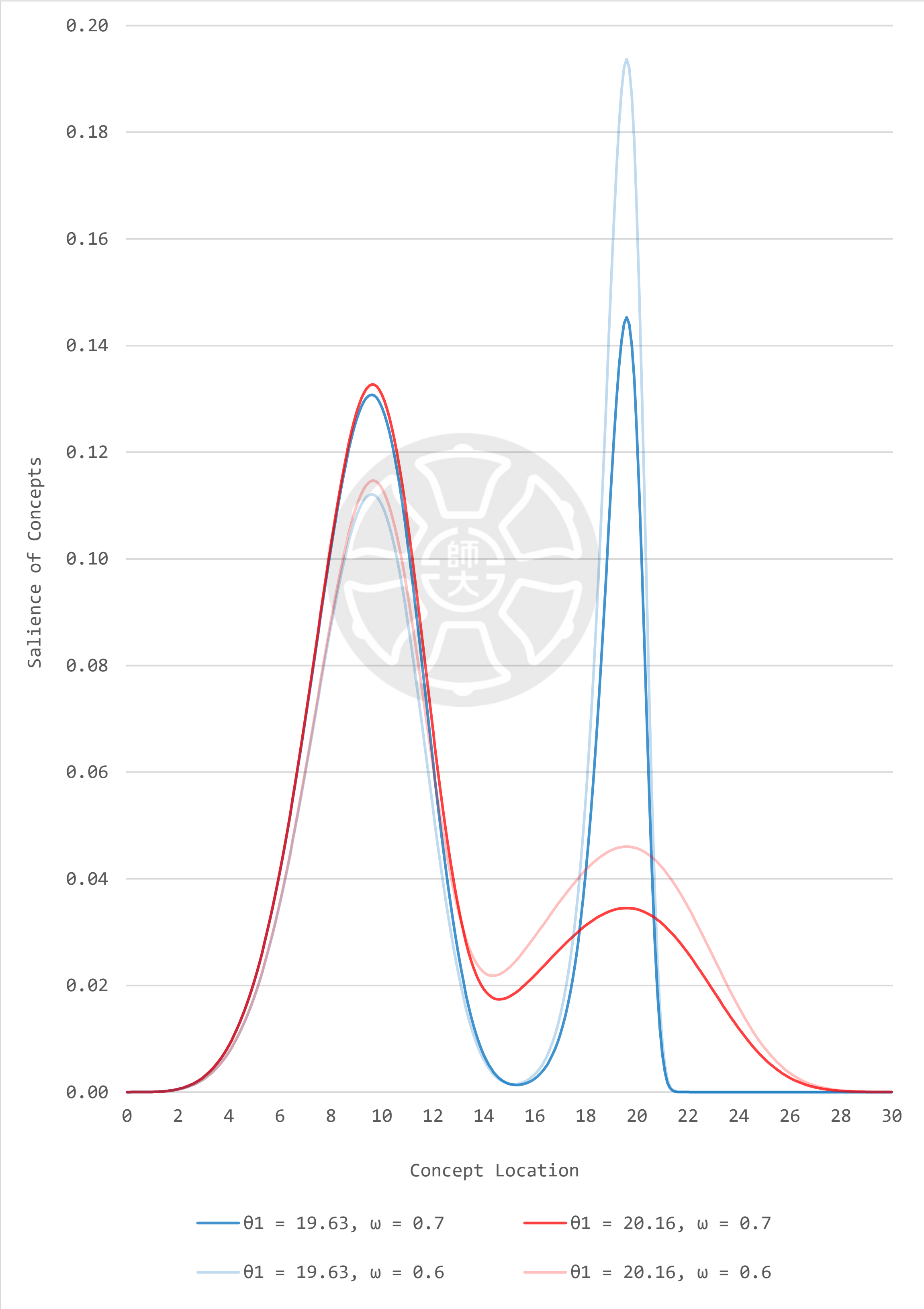


Figure 25

Mixed Weibull: Selecting the Best Representation in the Presence of Two Different Modes

($\xi_0 = 9.6$, $\xi_1 = 19.6$, $\theta_0 = 10.04$)



General Discussion

This chapter is the very end of this article, and will begin by reflecting on the model of dimensional range overlap and the equations proposed in this article. Also, the reciprocity hypothesis will be recalled. The author will then propose a new research hypothesis: the diminishment hypothesis, which underlies the proposed equations. A number of research recommendations will follow. Some of the details and apparently perplexing findings left unexplained in the foregoing chapters will be addressed in this chapter.

Advances in the Dimensional Range Overlap

Distribution Characteristics

In the present article, the author proposes Equations (1), (2) and (3), and others, which are based on Hsiāo's mathematical formulation (2002) of the distribution of concepts. The distribution underlies the model of dimensional range overlap, first proposed by Chiĕn and Hsiāo (2001), and later refined by Chiĕn et al. (2010). Nevertheless, the distribution, despite its role which is essential to the model, has scarcely been studied; rather, it was the dimensional range (see *Methods*) that received academic attention. Such focus of study ignored the influence of the behaviour of the distribution within and outside the dimensional range; one can easily sketch distributions with identical dimensional ranges but strikingly different skewness, peak and tail behaviours. In fact, the threshold of salience and dimensional ranges are mathematically completely irrelevant to the location of the best representation.

This is not to say that Chiĕn and Hsiāo's focus on dimensional ranges was senseless; see *Psychological Measurement* in *Research Limitations and Recommendations*. The point is that what actually causes people to give a different evaluation after their exposure to contextual stimuli is the stimuli's distorting the distribution describing the focal object, making some concepts temporarily more salient and therefore more representative of the focal object, even in the event of perceptual contrast, and some concepts temporarily less salient and therefore less representative, even in the event of perceptual assimilation. At this point, it may be worth clarifying that herein, the event of assimilation and the event of contrast refer respectively to the event $I = 1$ and the event $I = -1$. Inquisitive readers may be wondering why the salience of some concepts, in a theoretical view, will be enhanced even in the event of contrast, and will be suppressed even in the event of assimilation. A superficial answer is regularisation, which is necessary since probability is a unit measure, and the author, following Hsião's mathematical formulation, assumes Kolmogorov's axioms, too; see *Research Framework*. A psychological answer is that different concepts on a dimension conflict and therefore compete for salience. The competition also accounts for the dual role of latent ambiguity, which changes salience while scaling distributions. The dual role explains why it is possible that latent ambiguity, as suggested by Equations (16) and (17), diminishes perceived ambiguity.

Hsião, for some reason, was silent about the possible diminishment in formulating the distribution. The author makes it explicit here. The diminishment cannot be accounted for solely by the model of dimensional range overlap, since the model per se, as aforesaid, focuses on the dimensional range rather than the underlying distribution of concepts. The difference in focus allows the author to discover the roles of skewness and other distribution characteristics in shifting people's evaluations; see *Numerical Simulations*.

Additional Parametres

In addition to distribution characteristics, the author introduces a number of personal, situational parametres, i.e., evaluative volatility (denoted by δ), cognitive consumption (α and β), and attention to dissimilarities (denoted by λ). In particular, by evaluative volatility and cognitive consumption, the author incorporates the elaboration likelihood model (Petty & Wegener, 1999) into the prediction of people's perceptions. By the elaboration likelihood model, the author expects low evaluative volatility if and only if one's evaluation is formed only after effortful elaboration, and high cognitive consumption if and only if one is faced with personally consequential decisions, has the relevant cognitive resources (i.e., knowledge) and is not distracted or under time pressure.

Accordingly, without empirical evidence, the author expects that more often than not, the parametre α is sufficiently small that the diminishment hypothesis (see the next page) is practically testable, assuming availability of the necessary psychometric techniques; and that, by the same token, the parametre β is small enough for the calculation and comparison of $\mathbf{P}[W \in \mathbb{T}_w]$ to be effectively unneeded and for the best representation to be simply the most salient concept. (Refer back to *Methods* for the definitions of W and $\mathbb{T}_w$.)

Reciprocity Hypothesis

In summary, the author advances the model of dimensional range overlap by considering the characteristics of the distribution of concepts and distribution-independent factors. This, however, is not the sole contribution of this article. In *Numerical Simulations*, the author implicitly addresses a question stemming from Hsiāo's reciprocity hypothesis (2002): Is it possible that after exposure to a contextual stimulus, people's evaluation of the focal object becomes farther away from their original than their posterior evaluation of the contextual stimulus, in the event of mutual perceptual assimilation? Mathematically put, the question asks about the possibility that in the event $I_{Y|X} = I_{X|Y} = 1$, $(y_*' - x_*')(y_* - x_*) < 0$. According to Figures 1–8, the author affirms such a possibility. In those figures, $y_* = \xi_Y$ and $x_* = \xi_X$ (symbols defined by Equation (105)). Figures 1–4 consistently suggest that $\rho_X = 0$ and $\rho_Y(y_* - x_*) > 0$ together imply that in the event $I_{Y|X} = 1$, $(x_*' - y_*)(x_* - y_*) \leq 0$ if y_* is the vicinity of x_*. Figures 5–8 consistently suggest (after interchanging symbols) that $\rho_X = 0$ implies that $(y_*' - y_*)(x_* - y_*) \geq 0$ in the event $I_{X|Y} = 1$. Now, observe

$$(y_*' - x_*')(y_* - x_*) = (x_*' - y_*)(x_* - y_*) - (y_*' - y_*)(x_* - y_*) \quad (108)$$

The inequality in question immediately follows.

In addition to affirming that possibility, the author echoes the reciprocity hypothesis; see the paragraphs summarising Figures 17–20 in *Numerical Simulations*. In particular, recall that Figures 1–8 consistently suggest that in general, distributionwise assimilation and contrast (defined by the random indicator I) implies pointwise assimilation and contrast (defined by shifts of the best representation). Without this implication, the author would not be able to echo the hypothesis with those figures.

Diminishment Hypothesis

Finally, before the ending section *Research Limitations and Recommendations*, the author, as already announced, would like to propose a new research hypothesis: the diminishment hypothesis, which predicts diminishing context effects as the context slides (rather than shifts) away from the focal object along the focal dimension, from a position far away from the focal object to infinity (or the end of the dimension, if the dimension is bounded). See Figures 1–8 for visualisations.

Readers with a good memory may be reminded of the results in *Laplace Distribution*; see *Analytical Approach*. There, the diminishment hypothesis is supported in the event of contrast in all of the four special cases considered and analysed; refer back to Equations (61), (62), (65), and (66): $\partial x_*'/\partial \xi > 0$ (symbols defined there). This, in fact, is decided at the beginning, since the diminishment hypothesis, at a granular level, is the conceptual basis for Equation (3). The distribution-distorting integral $\mathbf{M}(\cdot)$, as defined by Equation (3), penalises the relative influence of contextual concepts for their difference from a given focal concept (which is the argument of the integral function).

In the event of assimilation, in *Laplace Distribution*, the diminishment hypothesis is also supported, although only in a rather loose sense. Recall Inequalities (35), (40), (45), and

(50), which consistently suggest that other things being equal, as the context slides away, then at some point, the inequality $\lim_{x \to 0_+} \partial \mathbf{D}_{X'}(x)/\partial x > 0$ will be violated, and the posterior focal distribution will become bimodal. As the context continues to slide away, then conceivably, the global mode of the distribution will suddenly skip (rather than shift) back to its original at some point. (For those interested, here is an example: Let $\alpha = 0.5$, $\delta = 0.9$, $\theta_X = 7.29$, and $\theta_Y = 0.13$. Try $\xi = 0.95$, $\xi = 3.43$, and $\xi = 8.76$; graph and see.)

The author is now closing this section. In doing so, the author is expecting a question: Why is it that context effects are hypothesised to diminish rather than be magnified? The answer, admittedly, is not psychologically grounded. The diminishment hypothesis as it is is a mere analogy with physical laws governing gravitational and electromagnetic forces, and such an analogy is indeed liable to be faulty. A justification by simulated results will be provided, but still, empirical tests are needed.

In the following section, a list of research recommendations will be offered. Of them, *Psychological Measurement* is perhaps the first and foremost priority.

Research Limitations and Recommendations

Psychological Measurement

In the present research, no laboratory or field experiments are conducted. This is because of the psychometric difficulties involved. In the following subsubsections, the author will propose a possible mental map from the latent dimension to a bounded measurement scale, then a tentative, partial psychological measurement approach, and finally, an explanation of the difficulties aforementioned.

Mapping to a Bounded Measurement Scale. As given in *Methods*, the minimum requirement for a latent dimension-to-measurement scale mental map to be possible is order preservation; that is to say, it is required that $W \mapsto \widehat{W}$ be monotonic everywhere and strictly monotonic almost everywhere on the dimension $\mathbb{V}$. Typically, the mapped dimension as displayed on a measurement scale $\widehat{\mathbb{V}}$ is bounded, whereas $\mathbb{V}$ can be bounded, half-bounded, or unbounded. In certain cases, an unbounded dimension is unnatural. For example, wealth is naturally non-negative, so a lower bound is expected. In other cases, an attribute may be coded differently by different people. For example, unattractiveness may be coded zero, which indicates void of attractiveness, or it may be coded negative infinity, as opposed to attractiveness, which is positive infinity.

It is very unlikely, in the author's view, to know how people actually code. Assumptions are at researchers' discretion. With limited loss of generality, the author focuses the following discussion on unbounded dimensions. The author proposes

$$\frac{\widehat{w}/A}{2/\pi} = \arctan\left(\frac{w}{\sigma}\right) - \frac{1}{4\mu A}\sin\left(4\mu A \arctan\left(\frac{w}{\sigma}\right)\right) \tag{109}$$

For the sake of convenience, the author assumes a continuous measurement scale for now. Also, the anchors are equally-spaced on the scale. The neutral point is anchored at zero; otherwise, recalibrate the scale before applying the map proposed above.

$\sigma > 0$; the parametre σ may be interpreted as insensitivity of measured evaluations to latent change. The parametre A is a positive integer for the bound of a measurement scale. For example, a measurement scale with bounds labelled -5 and 5 has $A = 5$. $\mu > 0$ and is such that μA is an integer. Assume a balanced scale. While $2A + 1$ indicates the face evaluative granularity, the parametre μ adjusts it to $2\mu A + 1$, which is the true evaluative granularity. Evaluative granularity herein is defined as the maximum number of anchors possibly existing in the evaluator's conscious mind, which may or may not really be displayed on the measurement scale. The anchors aforementioned are equally-spaced. For example, if $A = 5$ and $\mu = 2.0$, then the measurement scale in the evaluator's conscious mind has a total of 21 anchors evenly distributed in units of 0.5, between -5 and 5, inclusive.

Figure 26 (see the next page) visualises the map being proposed. The first thing to note is that the map is strictly monotonic almost everywhere, and monotonic everywhere, as suggested by Equivalence (110), which follows Equation (109) directly:

$$\frac{\partial \widehat{w}}{\partial w} \equiv \frac{2w\sigma}{w^2 + \sigma^2} \sin^2\left(2\mu A \arctan\left(\frac{w}{\sigma}\right)\right) \geq 0 \tag{110}$$

Equivalence (110) implies

$$\left.\frac{\partial \widehat{w}}{\partial w}\right|_{w=w_g} = 0 \quad \Leftrightarrow \quad \frac{w_g}{\sigma} = \tan\left(\frac{g\pi}{2\mu A}\right) \quad \Leftrightarrow \quad \widehat{w_g} = \frac{g}{\mu} \tag{111}$$

g is a non-negative integer such that $g < \mu A$. Notice that $\widehat{w_g}$ equals the kth anchor from the neutral point in the evaluator's conscious mind. This is not a mathematical coincidence, but is a result of the author's careful construction. The author posits that people, in general, have a tendency to map latent concepts to the vicinity of anchors, displayed or imaginary, on the measurement scale in their conscious minds. The tendency is implied by the fact that $\partial\widehat{w}/\partial w = 0$ at $w = w_g$, and is responsible for the stair-like behaviour of the mental map as visualised. The underlying intuition is that in mapping a latent concept to a measurement scale, one may begin by attempting to choose from some very loosely distributed anchors (e.g., -5, 0, and 5 in the previous example), next from some relatively densely distributed anchors (e.g., 0, 1, 2, 3, 4, and 5 in that example), and then from consecutive anchors in units of $1/\mu$ (e.g., 3.0, 3.5, and 4.0). If any of those consecutive anchors appears to be accurate, one then stops, and marks a position that he believes to be that anchor. Otherwise, one bases his final mark on the most accurate of those consecutive anchors, his final mark being a result of adjustments made to the most accurate anchor. This procedure should minimise his cognitive consumption. People, in general, are thrifty in the use of cognitive resources; see *Additional Parametres*.

Without loss of generality, let $\sigma = 1$, for now. Below are two Maclaurin series approximations of the map. Equation (112) is meant for minuscule w, whereas Equation (113) is for large w. For the sake of compactness, the author has used the big O notation, which readers are expected to know:

$$\frac{\widehat{w}}{A} = \frac{16}{3\pi}\mu^2 A^2 w^3 + \mathbf{O}(w^5) \tag{112}$$

$$\frac{\widehat{w}}{A} = 1 - \frac{16}{3\pi}\mu^2 A^2 w^{-3} + \mathbf{O}(w^{-5}) \tag{113}$$

Applying these series approximations, the author deduces that evaluative granularity facilitates detection of minuscule latent deviations from the neutral concept, and detection of latent differences between extreme objects, even when the bounds of the measurement scale are regularised. When the bounds are not regularised, i.e., are in their original, evaluative granularity facilitates detection of latent differences, in general. This all together seems to suggest that anchoring helps even in the case of a continuous scale.

Finally, the author is applying the proposed map. Take Figures 1–4 and 17–20, extract data therefrom, and produce Figures 27 and 28, accordingly. *c* is for *continuous* and *d* is for *discrete*. By *discrete*, the author means a discrete scale. Presented with a discrete scale, people find a continuous score and then round it to the specified digit.

$$t_{\widehat{X'}} = \frac{n_{X_*,+}\widehat{x'_{*,+}} + n_{X_*,-}\widehat{x'_{*,-}}}{\left|\widehat{x'_{*,+}} - \widehat{x'_{*,-}}\right| \cdot \sqrt[2]{\frac{n_{X_*,+}n_{X_*,-}}{n_{X_*,+} + n_{X_*,-} - 1}}} \tag{114}$$

In calculating the p values in Figure 28, Student’s t is necessary. Equation (114) gives that. $\widehat{x'_{*,+}}$ and $\widehat{x'_{*,-}}$ are the predicted posterior best representations of the focal object mapped to the measurement scale as described in Figure 28 in the event of assimilation and contrast, respectively. $n_{X_*,+}$ and $n_{X_*,-}$ are the expected numbers of observations of $\widehat{x'_{*,+}}$ and $\widehat{x'_{*,-}}$, respectively: $n_{X_*,+} = \lfloor n\,\mathbf{P}[I_{Y|X} = 1] + 1/2 \rfloor$ and $n_{X_*,-} = n - n_{X_*,+}$. In Figure 28, $n = 33$.

Figure 26

Mapping from the Latent Dimension to a Bounded, Continuous Measurement Scale ($\sigma = 0.51$)

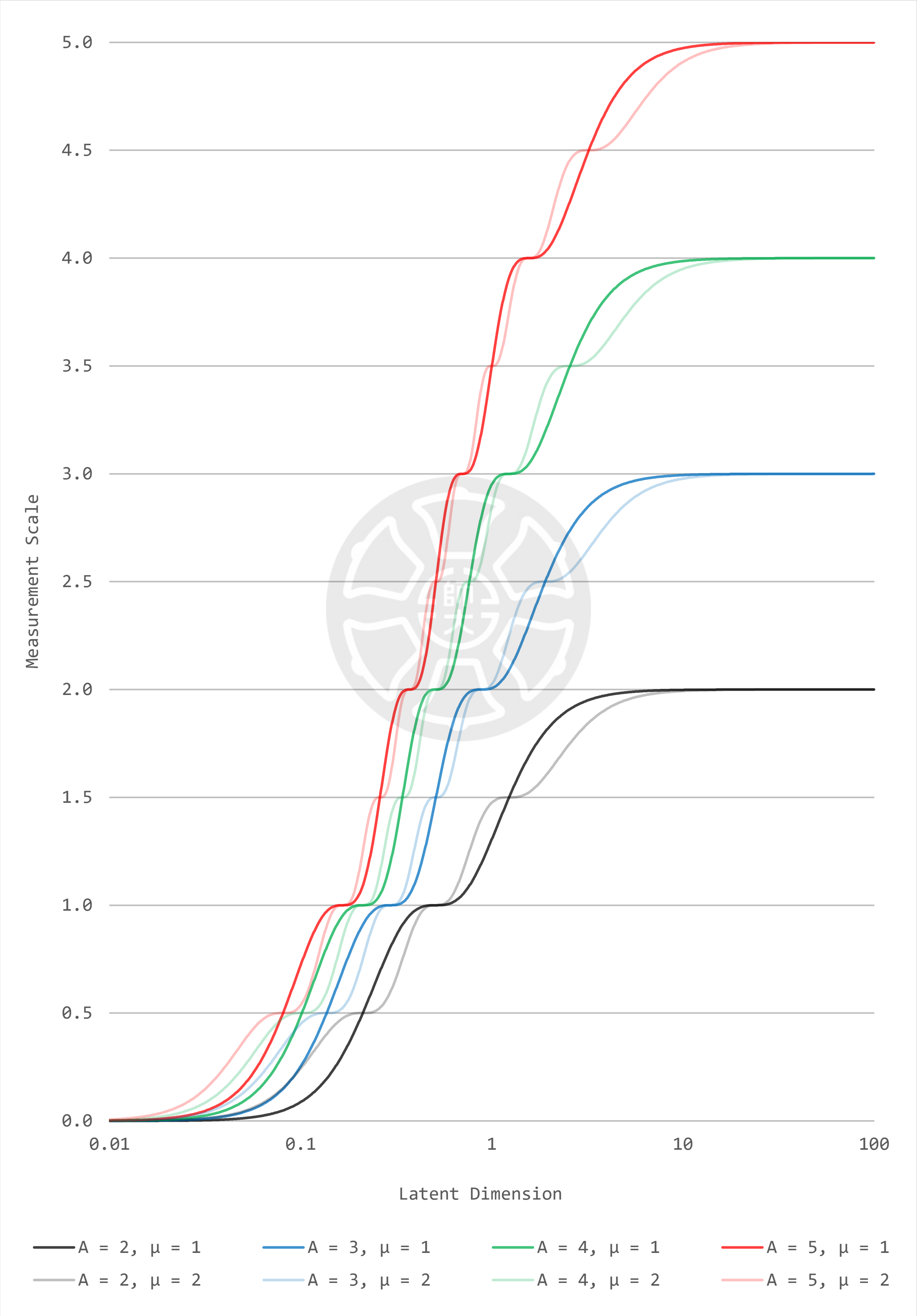


Figure 27

Skewed Generalised Student's t: The Measured Best Representation, by Context Extremity ($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\sigma = 0.51$, $A = 5$, $n = 33$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\rho_Y = -0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)

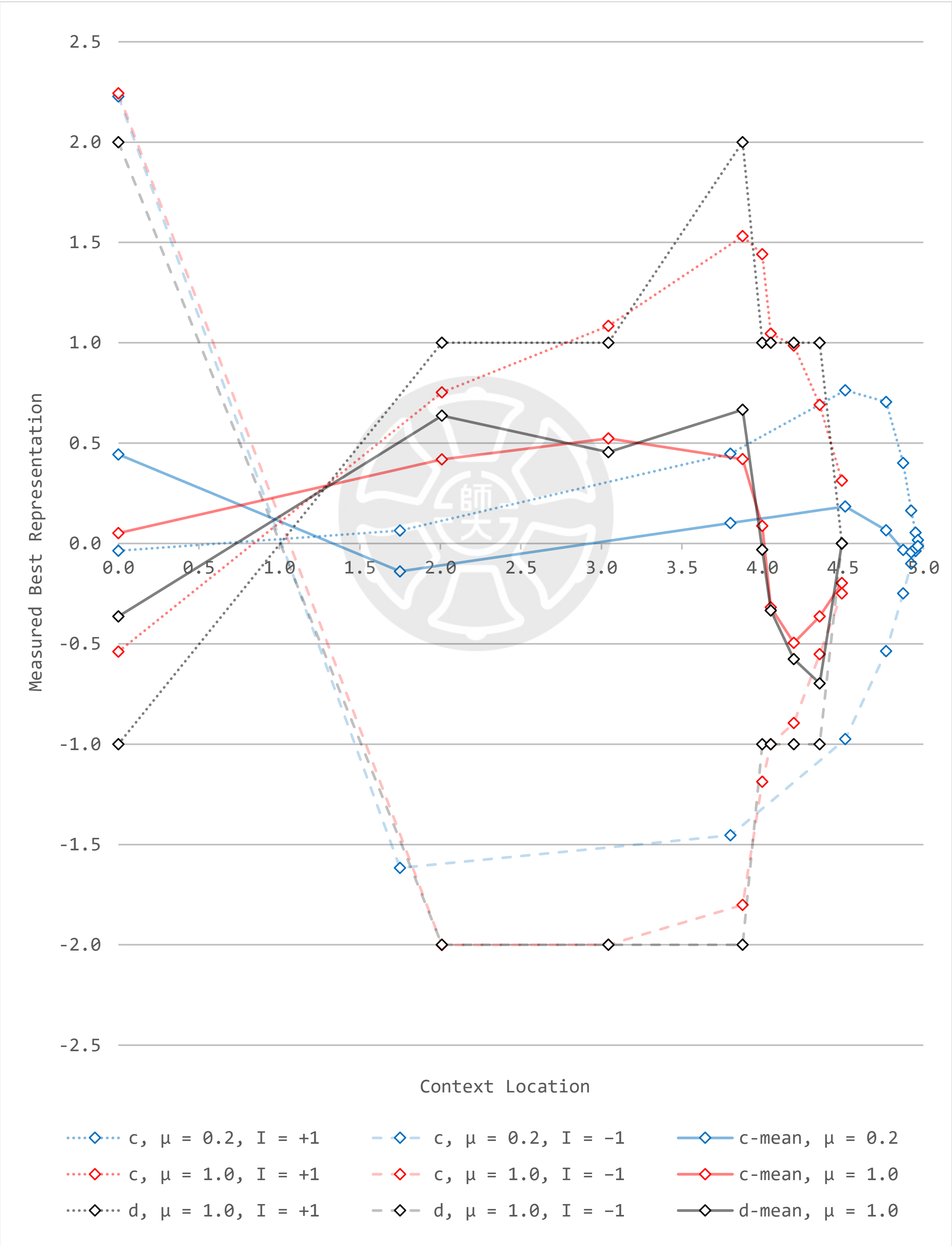
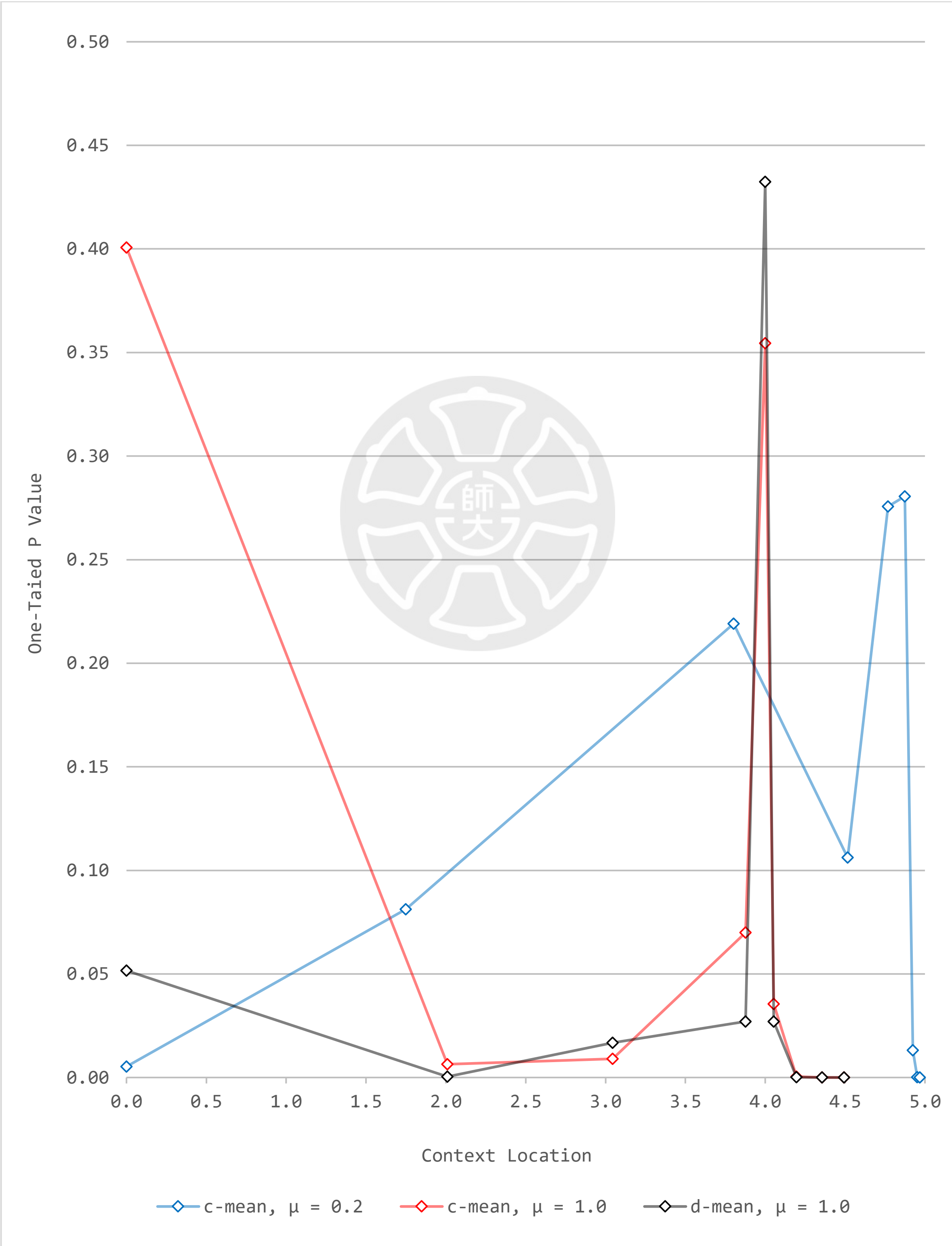


Figure 28

Skewed Generalised Student's t: Student's t Test for the Measured Effect, of Context Extremity

($\alpha = 0.5$, $\lambda = 1.8$, $\delta = 0.9$, $\sigma = 0.51$, $A = 5$, $n = 33$, $\xi_X = 0.00$, $\theta_X = 1.32$, $\theta_Y = 0.88$, $\rho_X = 0.0$, $\rho_Y = -0.3$, $\kappa_X = \kappa_Y = 2.5$, and $\eta_X = \eta_Y = 3.6$)



Inferring the Latent and Estimating the Parametres. In the previous subsubsection, the author simulated an experiment with 9 different conditions, conditions differentiated by context extremity, each involving 33 participants. The participants are perfectly psychologically homogeneous with respect to their perceptions, evaluative volatility, evaluative granularity, attention to dissimilarities, and so forth. The results are visualised in Figures 27 and 28. Figure 27 tautologically echoes Figures 1–4; the echo is tautological since the map being applied, as minimally required, preserves order. But its echoing Figures 9–12 is not tautological: Figure 27 is not, even partially, a mapping result of any of Figures 9–12. Figures 9–12 and 27 consistently suggest that on average, as Herr (1989) and Chiěn et al. (2010) observed, the best representation of the focal object generally will shift towards the context when the context is moderate and away from the context when the context is extreme. Figure 28 adds that the assimilation and contrast will be more easily statistically detected when the context is sufficiently extreme or when the context is moderate, but is not in the vicinity of the original best representation of the focal object.

All in all, the simulated experiment appears to provide evidence of convergent validity of Equations (1), (2), and (3). In particular, recall the diminishment hypothesis (see the so-named subsection in this chapter), which underlies Equation (3). By causally justifying Equation (3), the experiment supports the hypothesis, too.

In the present subsubsection, the author, after seeing the foregoing evidence of validity, is interested to know whether or not the evidence is replicable in real-world experiments. Real-world experiments, as opposed to computer-simulated experiments, are vital in that, ultimately, the equations being proposed are supposed to predict real human perceptions. In the following paragraphs, the author will be documenting some preliminary thoughts regarding real-world experiments testing the equations. The author does not run such an experiment for the reason aforesaid: the presence of two major psychometric challenges, i.e., inferring the latent distribution of concepts and estimating the parametres involved.

The author is now taking the first challenge, i.e., inferring the latent. The author's main idea is simple: estimating the cumulative probability. To this end, the experimenter asks participants to indicate their confidence c_k, which is bounded by 0 and 1, in the following statement: The object being considered cannot be represented by any scores higher than $\widehat{w_k}$ on the measurement scale. Without loss of generality, let $\widehat{w_k} < \widehat{w_{k+1}}$. For the sake of simplicity, assume a continuous measurement scale. Also, assume that $c_k = 0$ for $k < 1$, and that $c_k = 1$ for $k > d$; for $k \in [1, d)$, $0 < c_k < c_{k+1} < 1$.

Now, the author assumes

$$c_k = \mathbf{P}\left[\widehat{W} < \widehat{w_k} \,\middle|\, \widehat{W} \in \widehat{\mathbb{R}_W}\right] \tag{115}$$

Recall the definition of a dimensional range: 'the range of accessible concepts relating to a [given] object'; see *Literature Review*. The key word is *accessible*, which underlies the assumption and explains Equation (16). With this in mind, it can be easily deduced that $w_0 \le \mathbf{D}_{W,-}^{[-1]}(\varepsilon)$ and that $w_{d+1} \ge \mathbf{D}_{W,+}^{[-1]}(\varepsilon)$.

From Equation (115), the author derives

$$c_k = \mathbf{P}[W < w_k \mid W \in \mathbb{R}_W] \tag{116}$$

Next, the author defines

$$w_k = \mathbf{Q}_W(q_k) = \mathbf{Q}_W(\mathbf{P}[W < w_k]) \tag{117}$$

$\mathbf{Q}(\cdot)$ takes quantiles. Then, the author rewrites Equation (116), and plugs it into (117) to see (118):

$$w_k = \mathbf{Q}_W\left(c_k\, \mathbf{P}\left[W < \mathbf{D}_{W,+}^{[-1]}(\varepsilon)\right] + (1 - c_k)\, \mathbf{P}\left[W < \mathbf{D}_{W,-}^{[-1]}(\varepsilon)\right]\right) \tag{118}$$

Accordingly, the parametres needing estimation are ε and those characterising the distribution of W. (It may be tempting to estimate $\mathbf{P}[W < \mathbf{D}_{W,+}^{[-1]}(\varepsilon)]$ and $\mathbf{P}[W < \mathbf{D}_{W,-}^{[-1]}(\varepsilon)]$ and those characterising W as if they were independent, but that is problematic: It would introduce an additional, superfluous degree of freedom, which would allow the threshold of salience defining dimensional ranges to vary, which is illegitimate; refer back to *Research Framework* for the assumptions in this research.)

In order to more specifically show the mathematical workability of the idea, the author is now illustrating the inference procedure with a logistic example:

$$q = \frac{1}{e^{(\xi - w)/\theta} + 1} \tag{119}$$

Equation (119) directly follows (92) after change of symbols, and integration. To begin with, one solves $\mathbf{D}_W(w) = \varepsilon$:

$$\mathbf{D}_{W,\pm}^{[-1]}(\varepsilon) = \xi \pm \theta \ln\left(\frac{1 + \sqrt[2]{1 - 4\theta\varepsilon}}{1 - \sqrt[2]{1 - 4\theta\varepsilon}}\right) \tag{120}$$

$\mathbf{D}_{W,+}^{[-1]}(\varepsilon) = \sup(\mathbb{R}_W)$, and $\mathbf{D}_{W,-}^{[-1]}(\varepsilon) = \inf(\mathbb{R}_W)$, so

$$\mathbf{P}[W \in \mathbb{R}_W] = \sqrt[2]{1 - 4\theta\varepsilon} \tag{121}$$

Plug Equation (121) into (116) and then (117) to see

$$q_k = \frac{1}{2} - \left(\frac{1}{2} - c_k\right)\sqrt[2]{1 - 4\theta\varepsilon} \tag{122}$$

Finally, derive the quantile function and apply Equation (122), to arrive at the following:

$$w_k = \xi + \theta \ln\left(\frac{q_k}{1 - q_k}\right) \tag{123}$$

$$= \xi + \theta \ln\left(\frac{1 - (1 - 2c_k)\sqrt[2]{1 - 4\theta\varepsilon}}{1 + (1 - 2c_k)\sqrt[2]{1 - 4\theta\varepsilon}}\right) \tag{124}$$

Equation (124), which is the logistic case of (118), provides the regression model for estimating the parametres ξ, θ, and ε. Since the model is non-linear and cannot be linearised, advanced techniques are needed in estimation. Fortunately, this is easily doable with the `nls` function in the R language.

In the illustrative example, the author takes the logistic distribution solely for simplicity. In practice, one may want a more flexible, accommodating distribution that potentially fits the observed data better. For this purpose, the author recommends the meta-logistic distribution, which was introduced by Keelin (2016) and is constructed as follows:

$$\xi = \sum_{h=1}^{m} \psi_{2h-1} \left(q - \frac{1}{2}\right)^{h-1} \tag{125}$$

$$\theta = \sum_{h=1}^{m} \psi_{2h} \quad \left(q - \frac{1}{2}\right)^{h-1} \tag{126}$$

Refer to Equation (123) for ξ and θ. The substitutions above are said to be legitimate if and only if both of the following requirements are met: First, $\partial \mathbf{Q}_W(q)/\partial q > 0$, as a quantile function mathematically should be. Second, in the present settings, the author imposes that $w_0 \leq \mathbf{D}_{W,-}^{[-1]}(\varepsilon)$ and that $w_{d+1} \geq \mathbf{D}_{W,+}^{[-1]}(\varepsilon)$; this is exactly as aforementioned. With $m = 2$, skewness and kurtosis are allowed to vary; with $m \geq 3$, bimodality is possible. For psychological data, $m = 3$ should normally suffice.

The meta-logistic distribution has the following probability density:

$$\begin{aligned}
\mathbf{D}_W(w) &= \left(\frac{\partial}{\partial q}\mathbf{Q}_W(q)\right)^{-1}\Bigg|_{q=\mathbf{P}[W<w]} \\
&= \left(\frac{\partial \xi}{\partial q} + \frac{1-2q}{q(1-q)} \cdot \theta + \ln\left(\frac{q}{1-q}\right) \cdot \frac{\partial \theta}{\partial q}\right)^{-1}\Bigg|_{q=\mathbf{P}[W<w]} \\
&= \Bigg(\psi_2 \frac{1-2q}{q(1-q)} + \sum_{h=1}^{m-1} \psi_{2h+1} h \left(q - \frac{1}{2}\right)^{h-1} \\
&\qquad + \sum_{h=1}^{m-1} \psi_{2h+2} \left(q - \frac{1}{2}\right)^{h-1} \left(h \ln\left(\frac{q}{1-q}\right) - \frac{(1-2q)^2}{2q(1-q)}\right)\Bigg)^{-1}\Bigg|_{q=\mathbf{P}[W<w]}
\end{aligned} \tag{127}$$

The equation $\mathbf{D}_W(w) = \varepsilon$ is transcendental for all $m \geq 2$, so the regression model is practically almost always not analytically expressible, and therefore, the `nls` function in R, which requires an explicit equation, is not applicable. The author, admittedly, does not know any existing statistical solutions. Below is a tentative, iterative procedure that may work, provided convergence:

1. Begin by assuming $q_k = c_k$.
2. Assume an arbitrary but possible σ.
3. Regress the inferred w_k on q_k.
4. Assume an arbitrary but possible ε.
5. Plug the estimated ψ_h into $\mathbf{D}_W(w) = \varepsilon$. Solve for $\mathbf{P}[W < \mathbf{D}_{W,\pm}^{[-1]}(\varepsilon)]$ and recalculate q_k.
6. Repeat Steps 3 and 5 until all the ψ_h appear to converge.
7. Repeat Steps 4–6 and select an ε.

The regression model is linear, given q_k:

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ \vdots \\ w_d \end{bmatrix} = \mathcal{M} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \vdots \\ \psi_{2m} \end{bmatrix} \tag{128}$$

$$\mathcal{M} = \begin{bmatrix} 1 & \ln\left(\frac{q_1}{1-q_1}\right) & q_1 - \frac{1}{2} & \left(q_1 - \frac{1}{2}\right)\ln\left(\frac{q_1}{1-q_1}\right) & \cdots & \left(q_1 - \frac{1}{2}\right)^{m-1}\ln\left(\frac{q_1}{1-q_1}\right) \\ 1 & \ln\left(\frac{q_2}{1-q_2}\right) & q_2 - \frac{1}{2} & \left(q_2 - \frac{1}{2}\right)\ln\left(\frac{q_2}{1-q_2}\right) & \cdots & \left(q_2 - \frac{1}{2}\right)^{m-1}\ln\left(\frac{q_2}{1-q_2}\right) \\ 1 & \ln\left(\frac{q_3}{1-q_3}\right) & q_3 - \frac{1}{2} & \left(q_3 - \frac{1}{2}\right)\ln\left(\frac{q_3}{1-q_3}\right) & \cdots & \left(q_3 - \frac{1}{2}\right)^{m-1}\ln\left(\frac{q_3}{1-q_3}\right) \\ 1 & \ln\left(\frac{q_4}{1-q_4}\right) & q_4 - \frac{1}{2} & \left(q_4 - \frac{1}{2}\right)\ln\left(\frac{q_4}{1-q_4}\right) & \cdots & \left(q_4 - \frac{1}{2}\right)^{m-1}\ln\left(\frac{q_4}{1-q_4}\right) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \ln\left(\frac{q_d}{1-q_d}\right) & q_d - \frac{1}{2} & \left(q_d - \frac{1}{2}\right)\ln\left(\frac{q_d}{1-q_d}\right) & \cdots & \left(q_d - \frac{1}{2}\right)^{m-1}\ln\left(\frac{q_d}{1-q_d}\right) \end{bmatrix}$$

Above is the regressor matrix $\mathcal{M}$. The author advocates use of the meta-logistic distribution not only for its immense flexibility, but for its unparalleled ease of use: Elementary techniques such as the ordinary linear least-square normally work. In addition, it can be made bounded or half-bounded. Also, it is possible to construct meta-Laplace, Gaussian, and other distributions by analogy. See the reference paper for more on the distribution.

Now, recall the heading of the present subsubsection. The second challenge is estimating the parametres involved, e.g., evaluative volatility. Considering its relatively simple nature, the author is deferring the discussion to the end of the present subsubsection. In the following paragraphs preceding that, the author is addressing a number of questions that readers may have regarding the inference of the latent. The author is more than happy to answer or discourse if readers have more in their minds.

First, thus far, the author has assumed a continuous measurement scale for simplicity. Readers may be wondering whether or not the idea being proposed is also applicable to discrete versions. The answer is yes, but only after the necessary adjustments. The adjustments should address rounding and be made in accordance with one's assumed mental map.

Second, readers may feel like assuming that $\mathbb{R}_W = (w_0, w_{d+1})$; this is very intuitive. In fact, the assumption is justifiable if the intervals (w_0, w_1) and (w_d, w_{d+1}) are narrow. However, it can undermine goodness of fit, partially by reducing free parametres. While one should trade accuracy off for efficiency, the decision is not always easy.

Within the scope of this article, the author prefers to stay silent about the best practice. Below the author is simply resuming the logistic example and working on the assumption:

If $w_0 = \mathbf{D}_{W,-}^{[-1]}(\varepsilon)$ and $w_{d+1} = \mathbf{D}_{W,+}^{[-1]}(\varepsilon)$, then one can derive the following equations:

$$\xi = \frac{w_0 + w_{d+1}}{2} \tag{129}$$

$$\varepsilon = \frac{1}{\theta} \cdot \frac{1}{(r_\theta + 1/r_\theta)^2} \tag{130}$$

For the sake of compactness, the author has let $r_\theta = e^{(w_{d+1} - w_0)/4\theta}$. The only free parametre is latent ambiguity, i.e., θ. This is as expected: two degrees of freedom lost as a result of two additional constraints in estimation. Plugging the equations displayed above into (124) immediately yields

$$w_k = \frac{w_0 + w_{d+1}}{2} + \theta \ln\left(\frac{1/r_\theta + c_k(r_\theta - 1/r_\theta)}{r_\theta - c_k(r_\theta - 1/r_\theta)}\right) \tag{131}$$

This, as its original, is obviously non-linear, so in R, one may want to use the `nls` function, which was mentioned earlier. Diligent readers may also try the meta-logistic example: One should find, again, two degrees of freedom lost; specifically, the regressor matrix $\mathcal{M}$ will reduce to a size of d by $2m - 2$.

The author hopes that readers feel satisfied. The author is now answering a third possible question: Instead of inferring the dimensional range, is it possibly appropriate to ask participants directly for a range estimate and then take it as the dimensional range? Actually, this is exactly what Hsiāo (2002) did in his experiments; Chiĕn et al. (2010) did the same. The author, while being greatly inspired by Hsiāo and Chiĕn et al., feels sorry to say that such a measurement approach, in the author's opinion, is not perfectly aligned with Hsiāo's own definition of dimension ranges. Recall the definition, which emphasises concept accessibility, which is defined by the latent salience of a concept and is not by participants' confidence in their estimates. When asked to indicate a range estimate, the author posits that people will generally indicate a range, denoted by $\widehat{\mathbb{C}_W^*}$, with the minimum latent coverage, denoted by $\mathbb{C}_W^*$, in which their confidence is at a particular level, denoted by γ. Mathematically formulated, $|\mathbb{C}_W^*| = \min(|\mathbb{C}_W|)$, $w_* \in \mathbb{C}_W$, and $\mathbf{P}[W \in \mathbb{C}_W \mid W \in \mathbb{R}_W] = \gamma$.

$0 < \gamma < 1$, so $\mathbb{C}_W$ is indefinite. On a continuous scale, $\widehat{\mathbb{C}_W^*} \subset \widehat{\mathbb{R}_W}$; on a discrete scale,

$\widehat{\mathbb{C}_W^*} \subseteq \widehat{\mathbb{R}_W}$. The possible equality is due to rounding.

At this point, curious readers may be pondering over a question: Why is there a need for an additional parametre γ? The answer lies in the first postulate of the elaboration likelihood model (Petty & Wegener, 1999). As reviewed at the beginning of this article, people generally want their estimates to be accurate, and when the estimates are a range, people then want their estimates to be sharp. While an all-encompassing range is collectively always accurate, when taken individually, most of its elements may be misleading. A range as such is far from informative. People want their estimates to be considered to be useful.

In summary, people's range estimates can be a proxy for their dimensional ranges, but ultimately, it is a proxy: It is confounded with the additional parametre γ. Investigating what factors may influence γ may be interesting, but the author is leaving that to future scholars. Herein, only point estimates are of interest.

Here comes a fourth possible question: According to Figures 10 and 14, one would induce that when the context is extreme, latent ambiguity is expected to magnify pointwise contrast, i.e., the shift of the best representation of the focal object away from that of the contextual stimulus. This appears to conflict with the past findings. In response to this, the author would argue that it is not really meaningful to compare those findings in respect of ambiguity directly with the predictions being made in this article. The reasoning follows: First, in the past, ambiguity was usually qualitatively described (e.g., in Herr's experiments (1989), a realistic, yet unfamiliar, fictitious car); when it was quantified, it was only simplistically, crudely measured (e.g., in Chiěn et al.'s experiments (2010), the size of the range describing a given object). All in all, the language being used and the measurement in the past research fell below the level of precision that the author aims at in building the model presented in this article. Second, in the existing academic literature, it seems that ambiguity always refers to perceived ambiguity, rather than latent ambiguity, which appears to be new. For their difference, see Equations (16) and (17). Third and last, in all of the past empirical studies, the skewness, kurtosis (which influences the peak and tail), and modality (i.e., the number of peaks) of the latent distribution of concepts were completely ignored, and consequently not controlled in the experiments or the subsequent analyses. Take Figure 10 and compare the sequences: (a) $\rho_Y = 0.3$ and $\theta_Y = 1.10$; and (b) $\rho_Y = -0.3$ and $\theta_Y = 0.66$. At $\xi_Y = 3.2$, the sequence (a) has $x'_* \approx -0.0257$, whereas (b) has $x'_* \approx -0.0395$. In this example, the apparent consistency between the model predictions and the past findings is a confounded result.

Summarising the foregoing, the author would suggest, in order to empirically test the proposed mathematical model, one need to focus on the latent and not on the perceived, and that he need to account for skewness, at least. There is no short cut. The measurement and inference are not easy, indeed, but such an academic attempt can be fruitful.

Up till the present, four possible questions have already been addressed. The rest is another four, which are no less important despite their order of presentation.

Here is the fifth one: What is the recommended number of testing anchors $\widehat{w_k}$? Without actually running an experiment, the author would say: a number such that it is expected that $d \geq 7$, which is a very humble estimate. Ideally, the author would like as many as possible, but in reality, obtaining $d \geq 10$ may be difficult enough. Remember that participants are human and may feel fatigued, bored, or impatient when asked a long list of almost identical questions demanding their cognitive resources. Also, naturally, researchers usually need at least two more testing anchors ($\widehat{w_0}$ and $\widehat{w_{d+1}}$) than their actual input in the subsequent statistical analyses. For example, for $d = 7$, researchers may need 9 testing anchors on the measurement scale for a presumably ambiguous object, and 14 or more for a presumably specific object. For $d = 10$, if the object to be evaluated is presumably ambiguous, 13 may suffice, but if the object is presumably highly specific, then 20 or more may be necessary.

The number of testing anchors may also have an effect on people's evaluative granularity. This is one of the concerns of the next question. For now, the author is clarifying that, by presumably specific or ambiguous objects, the author means objects expected to be perceived as specific or ambiguous. The author opines that perceived ambiguity is usually a rather good proxy for the size of the dimensional range, at least when the latent distribution of concept is logistic. Presented below is the mathematical reasoning:

Recall Equations (16) and (120). The logistic perceived ambiguity, relative to the logistic latent ambiguity, is given by the following integral:

$$
\begin{aligned}
\frac{\varphi_W}{\theta} &= \frac{1}{\theta}\int_{\mathbf{D}_{W,-}^{[-1]}(\varepsilon)}^{\mathbf{D}_{W,+}^{[-1]}(\varepsilon)} \frac{|w-\xi|}{\theta}\cdot\frac{e^{(w-\xi)/\theta}}{\left(e^{(w-\xi)/\theta}+1\right)^2}\,\partial w \\
&= \frac{2}{\theta}\int_{0}^{\mathbf{D}_{W,+}^{[-1]}(\varepsilon)} \frac{w-\xi}{\theta}\cdot\frac{e^{(w-\xi)/\theta}}{\left(e^{(w-\xi)/\theta}+1\right)^2}\,\partial w \\
&= 2\int_{0}^{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta} \frac{we^{w}}{(e^{w}+1)^2}\,\partial w \\
&= 2\left(\int_{0}^{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta} \frac{1}{e^{w}+1}\,\partial w - \frac{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}{e^{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}+1}\right) \\
&= 2\left(\frac{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}{1+e^{-\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}} - \int_{0}^{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta} \frac{e^{w}}{e^{w}+1}\,\partial w\right) \\
&= 2\left(\frac{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}{1+e^{-\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}} - \ln\left(e^{\left(\mathbf{D}_{W,+}^{[-1]}(\varepsilon)-\xi\right)/\theta}+1\right)+\ln(2)\right) \\
&= 2\ln\left(1-\sqrt[2]{1-4\theta\varepsilon}\right)+\left(1+\sqrt[2]{1-4\theta\varepsilon}\right)\ln\left(\frac{1+\sqrt[2]{1-4\theta\varepsilon}}{1-\sqrt[2]{1-4\theta\varepsilon}}\right) \\
&= \ln(4\theta\varepsilon)+\sqrt[2]{1-4\theta\varepsilon}\,\ln\left(\frac{1+\sqrt[2]{1-4\theta\varepsilon}}{1-\sqrt[2]{1-4\theta\varepsilon}}\right) \qquad (132)
\end{aligned}
$$

From this, the author immediately derives

$$\frac{\partial \varphi_W}{\partial \theta} = \frac{\varphi_W}{\theta} - \frac{|\mathbb{R}_W| \cdot \varepsilon}{\sqrt[2]{1 - 4\theta\varepsilon}} \tag{133}$$

$|\mathbb{R}_W| = \mathbf{D}_{W,+}^{[-1]}(\varepsilon) - \mathbf{D}_{W,-}^{[-1]}(\varepsilon)$. Also, note

$$\frac{\partial |\mathbb{R}_W|}{\partial \theta} = \frac{|\mathbb{R}_W|}{\theta} - \frac{2}{\sqrt[2]{1 - 4\theta\varepsilon}} \tag{134}$$

Neither $\partial\varphi/\partial\theta = 0$ nor $\partial|\mathbb{R}_W|/\partial\theta = 0$ is analytically tractable, to the author's knowledge. Numerically solved with Desmos (which is heavily used in *Numerical Simulations*), under the assumption that $\varepsilon = 1/4$ so that it is implied that $\theta \in (0, 1)$, $\partial\varphi_W/\partial\theta \lesseqgtr 0$ holds approximately for $\theta \gtreqless 0.47$, and $\partial|\mathbb{R}_W|/\partial\theta \lesseqgtr 0$ holds approximately for $\theta \gtreqless 0.58$, so $\partial|\mathbb{R}_W|/\partial\varphi_W > 0$ except for $\theta \in (0.47, 0.58)$. (Acute readers might have guessed that the author has implicitly assumed the uniform distribution, which maximises information entropy in the absence of prior knowledge, for θ.)

Now, the author is proceeding to the next possible question, which is a triplet: Is there an effect (a) of the geometric length of the measurement scale, (b) of the number of anchors displayed on the scale, or (c) of the numeric labels attached to the bounds of the scale, on evaluative granularity? Again, the author is assuming a continuous measurement scale, bounded and balanced. Without a diligent search in the literature, the author hazards the following answers, based on the author's self-observation:

(a) Yes, (b) yes, and (c) yes. Specifically, (a) a geometrically longer measurement scale with (b) more displayed anchors and (c) a larger numerical range may signal a greater need for precision, and people may therefore try to evaluate with higher granularity. This is consistent with the first postulate of the elaboration likelihood model, earlier reviewed.

For the effect of evaluative granularity, see the previous subsubsection; in particular, the author would like to direct readers back to Equation (112). Equation (112) suggests that lower evaluative granularity implies a stronger tendency to trivialise latent conceptual differences from neutrality, and to indicate a neutral or nearly neutral position on a measurement scale, accordingly, which hinders statistical detection of context effects, and is therefore undesirable.

In summary, higher evaluative granularity facilitates psychological measurement in general. How, then, can the experimenter improve participants' evaluative granularity? The answers to the triplet (a), (b), and (c) in the present question (which is the sixth one) exactly address that. For example, rather than a mere 3-inch continuous scale with 5 anchors, bounded by the labels 1 and 5, one may want to use a 12-inch version with 11 anchors, its bounds labelled 0 and 100. (Such an approach is cosmetic. One may also gamify the experiment in an attempt to enhance the participants' level of engagement, motivating them to evaluate with higher granularity.)

Following the sixth, which is on the measurement scale, here comes the seventh question, which is on framing. Consider the following four statements: The object as described (or as presented) (a) cannot be represented by any scores higher than $\widehat{w_k}$; (b) cannot be represented by any scores lower than $\widehat{w_k}$; (c) is possibly represented by some score higher than $\widehat{w_k}$; (d) is possibly represented by some score lower than $\widehat{w_k}$. Of these four, which one is the most suitable for the confidence-indicating task in the experiment?

The author does not have a definite answer. While in mathematical language, (a) is concerning $1 - \mathbf{P}\left[\widehat{W} > \widehat{w_k} \mid \widehat{W} \in \widehat{\mathbb{R}_W}\right]$, (b) is concerning $1 - \mathbf{P}\left[\widehat{W} < \widehat{w_k} \mid \widehat{W} \in \widehat{\mathbb{R}_W}\right]$, (c) is concerning $\mathbf{P}\left[\widehat{W} > \widehat{w_k} \mid \widehat{W} \in \widehat{\mathbb{R}_W}\right]$, and (d) is concerning $\mathbf{P}\left[\widehat{W} < \widehat{w_k} \mid \widehat{W} \in \widehat{\mathbb{R}_W}\right]$, so logically speaking, (a) and (d) are almost equivalent, provided a continuous scale, so are (b) and (c); pragmatically speaking, the author does not expect such equivalence to be observed. Anecdotal evidence agrees with the author and suggests that (a) and (b), owing to the mere presence of *any*, which is a strong word encompassing all of the negated possibilities, may lead to a lower level of confidence; and that, on the other hand, (c) and (d), owing to the heuristic, ‘There is always a possibility’, may lead to a higher level of confidence.

A simple solution mitigating the possible bias is to average (a) and (d), (b) and (c), but it would double the task. The author is unfortunately unaware of better solutions.

Finally, in this paragraph is the eighth question: How can one estimate the latent distribution the mode of which is given? This is the case where the experimenter asks the participants to indicate the best representation of the object as described (or as presented) before the confidence-indicating task, assuming their latent distribution to be unimodal. In answer to this, the author would suggest that one simply impose that $\partial \mathbf{D}_W(w)/\partial w = 0$, and that $\partial^2 \mathbf{D}_W(w)/(\partial w)^2 < 0$, at $w = w_*$. Some degrees of freedom will be lost, consequently.

This is near the end of the present subsubsection. Careful readers should ask more: For example, what if the dimensional range does not appear to be a single interval? What if instead, it appears to be the union of two or more intervals? What if the estimated meta-logistic distribution is illegitimate? The author is work-shy and is leaving all of this to future scholars. The whole topic of inferring the latent distribution of concepts well deserves a separate research.

What remains herein is estimating the distribution-independent parametres involved. As previously said, this is relatively simple: Some parametres are not supposed to be estimated, e.g., β, σ, μ, and A. β is to be qualitatively inferred. For any given observed data, the estimated θ is proportional to σ (since both are scaling parametres), and hence, σ may be set to unity. μ is at researchers’ discretion; A is decided by the measurement scale. For the other parametres, i.e., α, δ, λ, and ε, regression is the way. Computers are widely available and affordable, so what is the challenge?

From the author's perspective, the challenge lies in two areas: The first one is technical, and the second one is practical. Below the author is explaining:

In estimating the parametres involved, one can opt for either the sequential approach or the simultaneous approach. The sequential approach begins by estimating $\mathbf{D}_X(\cdot)$ and $\mathbf{D}_Y(\cdot)$, $\mathbf{D}_{X'}(\cdot)$ and $\mathbf{D}_{Y'}(\cdot)$, and then plugs the estimation results into Equations (1), (2), and (3) to estimate α, δ, and λ. The simultaneous approach, on the other hand, simultaneously estimates $\mathbf{D}_X(\cdot)$ and $\mathbf{D}_Y(\cdot)$, $\mathbf{D}_{X'}(\cdot)$ and $\mathbf{D}_{Y'}(\cdot)$, and Equations (1), (2), and (3). The choice between these two needs deliberation. While the sequential approach is potentially problematic in that the residuals across the equations are not necessarily mutually independent, it appears to be much easier than the simultaneous approach. For the latter, the author admittedly does not know any existing statistical techniques that work (the author's knowledge is limited, though).

This is the technical challenge that the author sees. The author is hoping for modelling experts' input. This challenge, however, is not as big and serious as what follows:

Recall the probabilistic nature of the proposed model, which predicts that for any given contextual stimulus, be it moderate or extreme, specific or ambiguous, perceptual assimilation and contrast both can possibly occur to the latent distributions of concepts describing the focal object and the contextual stimulus. Accordingly, to empirically test the model, one theoretically needs to repeatedly ask the participants to indicate the best representation of the object being considered. This is easier said than done, however. First, the repetition is distinctly unnatural from the participants' perspective to the extent that in the second or some later repeated session, they may be guessing the experimenter's purpose of repetition, and the guessing can lead them to consciously bias their own evaluations, for example, in the hope of satisfying the expectation that they think or feel that the experimenter is holding. Second, the participants may want their evaluations to be consistent or they may feel fatigued, or bored by the repeated experimental tasks, so in the later sessions, they may simply copy their previous evaluations, marking a position on the measurement scale mindlessly. Third, perceptual change is irreversible.

A usual way to solve this is to use group data, assuming that participants are psychologically homogeneous. This, however, as Luce (1997) pointed out, is rarely the case in reality. Perhaps this is a limitation common to psychological research, or in a broader sense, to social science. The author is rather pessimistic about a better solution.

This is the end of the present subsubsection. Much work is needed in order to empirically test the equations being proposed, and if the equations turn out to be valid, to know the distribution of the parametres involved. For example, the author is particularly interested to know the distribution of θ (or partially equivalently, ψ_2 in Equation (126)). As mentioned in *Distribution Characteristics*, latent ambiguity can not only magnify, but also diminish perceived ambiguity. Which of these two is more likely, then?

The author is leaving the question to interested future scholars, and is now moving to the next subsection *Mathematical Analysis*, which, as the five subsections following it, is supposed to be short and less calculation-heavy, hence a relatively light read. The five subsections following it are *Alternative Views*, *Subliminal Stimuli*, *Multiple Objects*, *Multiple Dimensions*, and finally, *Theory-Based Bias Correction*.

Mathematical Analysis

In *Numerical Simulations*, the author reports a goodly number of non-trivial results. The results are so rich, especially when compared with *Analytical Approach*, that one may be wondering, what is *Analytical Approach* there for? Why does the author bother to write the section? *Numerical Simulations* alone suffices. What is the point of all the mathematics?

Rigour, or mathematical rigour. The point is that, as Luce (1997) emphasised, simulations do not substitute for mathematical analyses; they do not provide the rigour that mathematical analyses do. In *Analytical Approach*, each and every result is proved (some proofs are omitted for brevity, though), and a proved result cannot possibly be wrong, for certain. In contrast, in *Numerical Simulations*, all of the results are based solely on naïve induction, which can overgeneralise.

Why, then, does the author not prove the results in *Numerical Simulations*? This is a natural question, but the answer is mundane: It is simply because the author is incapable. As explained at the beginning of *Numerical Simulations*, even in the case of the Laplace distribution, which is the simplest case in *Analytical Approach*, locating the posterior best representation is analytically intractable, in general, not to mention the skewed generalised Student's t distribution, which is more complex. Perhaps there is some devilish clever way to prove the results without locating the posterior best representation; if so, the author would love to be enlightened. In the foreseeable future, however, the author is not really hopeful of that. Instead, the author would recommend that future scholars try replicating the reported results by simulating with different distributions, e.g., as reviewed by Li and Nadarajah (2018), in the Student's t distribution family, Mittnik and Paolella's generalisation (2000), Jones and Faddy's generalisation (2003), Aas and Haf's generalisation (2006), and so forth. (The meta-logistic distribution by Keelin (2016) would be inconvenient.)

Finally, the author is answering a question earlier raised: In *Skewed Generalised Student's t Distribution*, the best representation of the focal object, as shown in Figures 1–8, shifts without any conditions, but in *Laplace Distribution*, the shift is found to be conditional. How is it that? A simple mathematical analysis reveals the key factor deciding whether or not the shift is conditional: differentiability of the focal latent probability density (i.e., salience) $\mathbf{D}_X(\cdot)$ at the best representation. Undifferentiability at a mode implies a pointed top, which hinders pointwise effects. Take Equation (105). Differentiability at $x = \xi_X$ implies $\kappa_X \geq 2$; the proof is left to diligent readers. This is the case of Figures 1–8. On the other hand, the Laplace distribution is a special case where $\rho_X = 0$, $\kappa_X = 1$, and $\eta_X \to \infty$. The conflict is explained.

Alternative Views

Throughout this article, the author assumes uncountably many concepts on a dimension. However, as admitted in *Research Framework*, this assumption is actually questionable. For example, consider a queue of people. When asked to estimate the number of people in queue, one will naturally think in terms of integers, which are countable.

Being countable implies being discrete. That being said, the author does not develop a discrete version of the proposed equations. Why? The answer is twofold. First, in the author's own experience, discrete mathematics is typically less elegant than continuous mathematics, and mathematical elegance translates to convenience. Second, in the author's belief, for most applications, a continuous model as proposed in this article should suffice. To see this, resume the queuing example. When the queue is short, of only three people, the number of people is indisputable; the focal distribution is degenerate, so no context effects are expected. Nevertheless, when the queue is long and extends over three major city blocks, one may be unable to count the number of people at a glance, and even if he counts it, he may still be uncertain of the result, fearing miscounts. In such scenarios, context effects can be expected; since the number is large, the discrete nature may be neglected.

The author is expecting sceptical readers: Why does a large number necessarily make the queue approximately continuous? To these attentive, scrupulous readers, the author would recommend a commentary by Luce (1997), who openly expressed the doubt; in fact, he extended it to sensory dimensions, e.g., visual temporal continuity. His doubt is reasoned; see the commentary. However, he also acknowledged the fact that continuous mathematics seemed to work well in describing the world in many different sciences.

In conclusion, the author is comfortable with the assumption of uncountably many concepts. The author is now moving forwards, reviewing another modelling assumption:

Throughout this article, the author assumes that people normally take the most salient concept as the best representation; mathematically put, the global mode is taken. This should be natural: The most salient concept is the most easily accessed. However, some scholars held different views. Chiĕn et al. (2010) suggested use of the median as the best representation. The author acknowledges that using the median has an advantage over using the global mode: Median is unique, and when conditional on accessibility, is unique as long as the dimensional range is a single interval, whereas a distribution can be multimodal. The possible plurality of the global mode can cause inconvenience in giving a point estimate, in forming a perception of ambiguity in accordance with Equation (16), and in judging whether a shift is assimilation or contrast by Equation (4). Nevertheless, that is purely academic. In reality, the author doubts that people's distribution of concepts is ever as pathological as described. Moreover, in the case of two global modes, assuming that β is small enough for the testing interval $\mathbb{T}_w$ to be effectively irrelevant, one may do as suggested in *Methods*. Regarding Equation (4), defining pointwise assimilation and contrast, the author would say, is actually for the mere purpose of easier communication. The greatest interest, as said in *Introduction*, is clearer (not necessarily easier) communication and predicting the posterior observed evaluation.

All in all, use of the global mode is not any less appropriate. Is it, then, more appropriate? The author, of course, would say yes; otherwise, the median would have been used. This, however, is not to say that using the median is unreasonable. If the median were taken, one would have $\mathbf{P}[W < w_* \mid W \in \mathbb{R}_W] = \mathbf{P}[W > w_* \mid W \in \mathbb{R}_W]$. The best representation should then be interpreted as the equilibrium concept bisecting the dimensional range into two equally representative halves. If one were to be more representative, the evaluator would think or feel that his evaluation is biased towards the less representative. People, as Petty and Wegener

(1999) postulated, are averse to biases, in general. For this reason, the author considers use of the median as the best representation to be possible.

Now the author accepts the median, how about the mean? The mean was considered in Chiěn et al.'s discussion (2010), too. Chiěn et al. considered the mean to be inferior to the median as the best representation, especially in the case of a skewed distribution. But why? In the reference discussion, they were silent. The author's guess is that they were under the impression that the median, as often depicted in introductory statistics textbooks, would be on the right of the mean if the distribution were skewed to the left, and on the left of the mean if the distribution were skewed to the right. This picture, while accurate for many widely known distributions (e.g., the log-Gaussian distribution), can be inaccurate for many others (e.g., bimodal and trimodal distributions). The author does concur with Chiěn et al., but would validate their assertion by reminding readers of the fact that the dimensional range need not be a single interval and that, mathematically, it is possible that $\mathbf{E}(W \mid W \in \mathbb{R}_W) \notin \mathbb{R}_W$. Psychologically interpreted, this means that, were the mean taken as the best representation, it would be possible that one took an inaccessible concept and marked it on a measurement scale. This is self-contradictory; thus, the author rejects the mean as the best representation.

For comparison and contrast, the author is now digressing a little and reviewing another modelling work, by Givon and Shapira (1984), which was on the topic of the optimal number of anchors on a discrete measurement scale. That is an interesting topic but is irrelevant. What is relevant to the current review is the model that Givon and Shapira developed. Similar to the model that the author constructed, Givon and Shapira's model was based on the assumption of a latent distribution of concepts possibly describing a given object. Yet, remarkably, they took the mean. Why? There is a fundamental difference distinguishing their model from the one that the author proposes herein: the definition of the distribution.

Let $W_{X;*}$ be a random variable such that in formal, measure-theoretic language,

$W_{X;\,*}\colon \mathfrak{S}_{(\mathcal{P}_Y,\boldsymbol{\Sigma})} \mapsto \mathbb{W}_{X;\,*}$. The sample space $\mathfrak{S}_{(\mathcal{P}_Y,\boldsymbol{\Sigma})}$ is defined as the set of such duplets $(\wp_Y, \boldsymbol{\varsigma})$ where $\wp_Y \in \mathcal{P}_Y$. $\mathcal{P}_Y$ is defined as the power set of the set of all the random variables Y, Y as already defined in *Methods*. $\boldsymbol{\varsigma} \in \boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma} = \prod \Sigma$. The set product is Cartesian and the multiplicand Σ is the support of some salience-independent psychological parametres involved. These parametres per se are random variables. The image $\mathbb{W}_{X;\,*}$ is defined as the set of x_* and all the possible expected posterior best representation of the focal object $\mathbf{E}\big(x'_* \,\big|\, (\wp_Y, \boldsymbol{\varsigma})\big)$.

Now, with these symbols defined, the author is explaining the fundamental difference: In the present article, by *latent distribution*, the author is consistently concerning $W_X \in \{X, X'\}$. Givon and Shapira, in subtle contrast, by conceptually following Thurstone's law of comparative judgment (1927a and 1927b), were concerning the asterisked $W_{X;\,*}$ conditional on $\mathbb{W}_{X;\,*} \subseteq \mathbb{R}_{W_X}$. Accordingly, they held a frequentist view, and argued that the latent dimension was bounded; $\left|\mathbb{W}_{X;\,*} \cap \mathbb{R}_{W_X}\right| \leq \left|\mathbb{R}_{W_X}\right| < \infty$. The randomness, as formulated in the previous paragraph, is attributed to contextual stimuli and the salience-independent parametres involved. Neither of these two sources of randomness was Givon and Shapira's modelling focus. Rather, they seemingly assumed that these two, in the language of signal processing, were noises. This was evidenced by their wording: They referred to the mean $\mathbf{E}(W_{X;\,*} \,|\, \mathbb{W}_{X;\,*} \subseteq \mathbb{R}_{W_X})$ as the true and not the best representation. As the author read their paper, Givon and Shapira, with a focus on measurement scale development, were actually interested in $\mathbf{E}(W_{X;\,*} \,|\, (\mathbb{W}_{X;\,*} \subseteq \mathbb{R}_{W_X}) \wedge (\wp_Y = \emptyset))$, which is free of contextual stimuli. What they underplayed is dedicated to herein.

This is the end of the current review. The purpose of reviewing Givon and Shapira's modelling work is demonstrating that, while apparently similar, their model is different in essence from the model proposed in this article, and that the difference becomes evident on mathematical formulation. This shows the first of the two benefits of use of mathematics mentioned in *Introduction*: clearer and consequently more efficient communication. Clarity is in the sense of no ambiguities, and efficiency is in the sense of no misunderstandings, and no clarifications; the saved energy may be devoted to research.

Before proceeding to the next discussion, the author goes back from the digression. Summarising the previous comparison of the mean, median, and mode, the author considers the mode to be the most likely best representation, generally speaking, and rejects the mean for the possible contradiction. Conceivably, readers may not agree. In the following paragraphs, the author will continue comparing the mode with the other two competing measures of distribution central tendency, i.e., the mean and median.

Begin with the mean. The author is expecting readers to contend that in reality, the dimensional range, as Chiĕn et al. (2010) implicitly assumed, should be a single interval. That is fair enough; the author actually does believe that is the usual case. But still, there are real-world exceptions: Consider a Dalmatian and rate its whiteness. Whiteness is indisputably opposed to blackness. The midpoint on the scale represents grey. The black dots or patches of the Dalmatian under consideration are visually easily distinguished from the rest of the dog's coat, which is white. Those dots or patches visually cover 45% of its coat. Then, realistically, the dimensional range describing the Dalmatian cannot be a single interval; the Dalmatian is visually a discrete mix of black and white, and can never be perceived as grey.

The Dalmatian example should convince readers. But still, one may argue that there are occasions on which people consciously consider the mean. For example, when asked, 'How, on average, was your sex experience in the past year', people are supposed to consider the mean. In reply to such arguments, the author would invite readers to answer a basic question: What exactly is being evaluated? It is the mean that is being evaluated. People can be uncertain of their average sex experience; see *Multiple Objects*.

Thus far, the author has refuted the mean as a possible best representation. How about the median? As aforesaid, the author considers the median to be possible. Then, what makes the author more inclined to the mode as the best representation, than the median? The answer is twofold. The psychological answer is presented as follows:

Recall the Dalmatian. Now, imagine that the author paints a number of visually easily distinguished grey dots or patches on the Dalmatian’s white coat. The grey dots or patches visually cover 10% of the dog’s entire coat. Rerate the whiteness of the dog. Still, the scale is bipolar and balanced. With those artificial grey dots or patches, grey is perceived, therefore in the dimensional range which is defined by accessibility. Is it possible, then, that people indicate grey, which is the midpoint, as the best representation? Without actually carrying out a survey, the author is not absolutely sure, but considers it to be unlikely. Black and white are much more salient. The author personally would indicate either.

Empirically testing whether or not people would do as the author predicts in the second Dalmatian example is left to future scholars. In the testing, be cautious: In the Dalmatian examples, the midpoint, allow the author to repeat, represents grey. The colour of grey may or may not be produced by mixing 50% black and 50% white; the mix can be discrete. When explained as a mix of black and white, rather than grey, the midpoint is perfectly likely to be marked even when the respondent does not see any grey on the coat. Semantic ambiguity can yield the unexpected.

In summary, the author believes that the mode is more natural, but is open to counterevidence. Below the author is presenting the other answer, which is less psychological than the first:

Recall the discussion on the countability of concepts on a latent dimension. The author opines that continuous distributions should work in most cases where context effects can be expected. That being said, the author acknowledges that there are cases where the latent distribution must actually be discrete. Now, consider a discrete distribution. If the median were taken as the best representation, the equation justifying its use could be easily violated; that is to say, it would be perfectly possible that $\mathbf{P}[W < w_* \mid W \in \mathbb{R}_W] \neq \mathbf{P}[W > w_* \mid W \in \mathbb{R}_W]$. In fact, the contradiction $w_* \notin \mathbb{R}_W$ also. The even more disturbing contradiction $w_* \notin \mathbb{V}$ is nowhere close to being pathological.

This should explain the author's general preference for the mode, but still, the author conceptually accepts the median; these issues with the median can actually be fixed quite easily. For example, define w_* such that $w_* \in \mathbb{R}_W$ and that the difference $|\mathbf{P}[W < w| \, W \in \mathbb{R}_W] - \mathbf{P}[W > w \mid W \in \mathbb{R}_W]|$ is minimum. This modification preserves the notion that people are generally disposed to give subjectively unbiased evaluations.

In closing, the author would recommend future scholars to consider the median in comparison with the mode. One may continue the theoretical comparison or empirically compare their explanatory power. Before opening the next discussion, allow the author to add a parenthetical note to the discussion on countability. The author is fully aware that a random variable with an uncountable support (i.e., a support with uncountable elements) is not necessarily entirely continuous, but can be partially continuous and partially discrete, or nowhere continuous and nowhere discrete; it can be singular. Nevertheless, the author considers all of these cases to be pathological, especially in everyday psychological settings. The author would be shocked if proved to be wrong.

The rest of the present subsection will be dedicated to the next discussion, focused on Equations (1) and (2), both central to the model being proposed in this article.

Throughout this article, the author assumes that for any given contextual stimulus, people's posterior best representation x'_* is a binary random variable. The randomness originates from the indicator $I_{Y|X}$, which is present in Equation (1). Presented below is a deterministic alternative to Equation (1):

$$\mathbf{D}_{X'}(x) = \frac{\mathbf{D}_X(x)\left(1 + \delta\, \mathbf{M}_Y(x)\left(\mathbf{P}\left[I_{Y|X} = 1\right] - \mathbf{P}\left[I_{Y|X} = -1\right]\right)\right)}{1 + \delta\left(\mathbf{P}\left[I_{Y|X} = 1\right] - \mathbf{P}\left[I_{Y|X} = -1\right]\right)\int_{\mathbb{V}} \mathbf{D}_X(v)\, \mathbf{M}_Y(v) \partial v} \tag{135}$$

Equation (135) suggests that perceptual assimilation and contrast, rather than mutually exclusive, are concurrent; the random indicator $I_{Y|X}$ is asbsent. $\mathbf{P}\left[I_{Y|X} = 1\right]$ and $\mathbf{P}\left[I_{Y|X} = -1\right]$ are the weights of perceptual assimilation and contrast. Comparing the x'_* predicted by (135) and the $\mathbf{E}(x'_*)$ predicted by (1), the author expects overall agreement in pattern since both are summaries of perceptual assimilation and contrast. This is merely an educated guess, though.

Mathematical analysis, numerical simulations, and experiments await future scholars. The author is now reviewing Equation (2), which measures distribution dissimilarity:

Recall the relative information entropy (which is formally referred to as the Kullback–Leibler divergence) (1951), which is used throughout this article. As mentioned in *Methods*, it is not symmetric. The asymmetry is natural in the present settings: If a concept is highly salient when one is evaluating the focal object but is not accessible at all when one evaluates the contextual stimulus, then hypothesising (in unconsciousness) that the focal object were an instance or occurrence of the sort where the contextual stimulus belongs should result in greater surprise when the surprise is weighted by the salience of the given concept in describing the tested object (i.e., in this case, the focal object), than should hypothesising the reverse. The weighting, in plain language, is playing up or down, and explains the asymmetry.

While the asymmetry is reasonable, it is admittedly not very intuitive, so one may want a symmetric measure. Jeffrey (1946), as cited by Kullback and Leibler (1951), proposed such a measure, which is the base for Equation (136):

$$\ln(\mathbf{P}[I = 1]) = \lambda \int_{\mathbb{V}} \mathbf{D}_X(v) \ln\left(\frac{\mathbf{D}_Y(v)}{\mathbf{D}_X(v)}\right) \partial v + \lambda \int_{\mathbb{V}} \mathbf{D}_Y(v) \ln\left(\frac{\mathbf{D}_X(v)}{\mathbf{D}_Y(v)}\right) \partial v \tag{136}$$

Equation (136) achieves symmetry by reciprocal testing. The degree of dissimilarity is given by the summary of the two hypothesis tests. Note that the tests are equally weighted. Lin (1991) introduced a weighting parametre, which herein is denoted by ω. In addition, he modified the null hypotheses: A mixture distribution of the focal object and the context stimulus was used. The equation right below applies Lin's work:

$$(\mathbf{P}[I=-1])^{\lambda} = \frac{1}{\omega\ln(\omega)+(1-\omega)\ln(1-\omega)} \left(\omega\int_{\mathbb{V}} \mathbf{D}_X(v)\ln\left(\omega+(1-\omega)\cdot\frac{\mathbf{D}_Y(v)}{\mathbf{D}_X(v)}\right)\partial v + (1-\omega)\int_{\mathbb{V}} \mathbf{D}_Y(v)\ln\left((1-\omega)+\omega\cdot\frac{\mathbf{D}_X(v)}{\mathbf{D}_Y(v)}\right)\partial v\right) \quad (137)$$

Thus far, all the measures are information-theoretic. Careful readers might have noticed that the surprise can be positive or negative. The sign is justified by information theory. When unweighted, the surprise in Equation (2) actually, precisely represents the coding or decoding effort needed to adjust the contextual stimulus information to the focal object information. More effort demanded implies greater distribution dissimilarity.

Other alternative measures abound. The author is not enumerating all. Rather, the author is just mentioning one more measure: the Wasserstein distance. The Wasserstein distance, as reviewed by Panaretos and Zemel (2019), originally arose from the optimal transport problem. Being a distance, it is symmetric and satisfies the triangle inequality. In the following equation, it is reduced to an analytically explicitly expressible case; specifically, it is the expression following the parametre λ (including the exponent):

$$\ln(\mathbf{P}[I=1]) = -\lambda\left(\int_0^1 |\mathbf{Q}_X(q)-\mathbf{Q}_Y(q)|^{\varrho}\,\partial q\right)^{1/\varrho} \quad (138)$$

$\varrho \geq 1$; the parametre ϱ may be interpreted as the penalty for divergence. In fact, as readers might have guessed, the Wasserstein distance is the solution to the optimal transport problem aforesaid in the sense of the ϱ-Lebesgue norm (or the $\mathbb{L}_{\varrho}$-norm in symbols).

This is the end of the discussion on Equation (2) and the end of the present subsection. Inquisitive readers may be wondering which one of these four is the most recommendable. The author must confess that in the absence of empirical data, not a single one appears to be absolutely superior to the others. For example, while the Wasserstein distance, as noted in the reference paper, takes the geometry of the spaces supporting the distributions being compared into account, Lin's proposal (formally referred to as the Jensen–Shannon divergence) allows for $\mathbf{P}[I = 1] = 0$ even when the latent dimension is bounded. Each has its appeal.

Subliminal Stimuli

This is a short subsection. Recall that the author imposes no accessibility constraints on the occurrence of context effects, which implies that context effects can occur even when the contextual stimulus is sensed but is not perceived at all. Admittedly, this may seem like a hoax, but several studies have confirmed the possibility. For example, refer to Draine and Greenwald's (1998) and Karremans, Stroebe, and Claus's (2006) experiments.

In brief, the author believes in subliminal context effects. Nevertheless, the author is seriously doubtful about the possibility of empirically testing the model being proposed using subliminal contextual stimuli. The approach presented in *Inferring the Latent and Estimating the Parametres* does not work. In conclusion, the author would recommend future scholars to put this topic aside, for efficiency. Supraliminal contextual stimuli should be more critical than the subliminal.

Multiple Objects

In *Methods* and *Results*, the author restricts the work to a system exclusively comprising a focal object and a contextual stimulus. Obviously, this is for the sake of simplicity. In reality, contextual stimuli are everywhere, at any time, for anyone, with very few exceptions. (Infants with anophthalmia are exceptions with respect to vision, for example.) Nevertheless, one should also acknowledge the unnecessity and impossibility of building a model accounting for all the contextual stimuli; such a model, if really built, would be unwieldy.

The real issue is to identify the critical contextual stimuli and include them in the model. What, then, is a critical contextual stimulus? Following intuition, the author would say that a contextual stimulus is critical if it is sensorially, temporally, or conceptually easily related to the focal evaluation. Note that criticality at this step is in the mind of the researcher. Researchers' asking respondents to rate a contextual stimulus's criticality per se will enhance the contextual stimulus' criticality, at least in the temporal sense. The author, unfortunately, knows no other ways and is not genuinely hopeful of any other ways. Researchers' knowledge about the respondents is valuable at this step.

Now, having already identified the critical contextual stimuli, a model can be built. Recall Equation (1), which is generalised as follows:

$$\mathbf{D}_{X'}\left(x \middle| \wedge I_{Y_k|X}\right) = \frac{\mathbf{D}_X(x)\left(1 + \sum \sigma_{Y_k} I_{Y_k|X}\ \mathbf{M}_{Y_k}(x)\right)}{1 + \sum \sigma_{Y_k} I_{Y_k|X} \int_{\mathbb{V}} \mathbf{D}_X(v)\ \mathbf{M}_{Y_k}(v) \partial v} \tag{139}$$

$0 < \sigma_{Y_k} = \delta\, \varkappa_{Y_k} < 1$. The parametre σ_{Y_k} is a mix of evaluative volatility (δ) and criticality ($\varkappa_{Y_k}$). Criticality following the first step is decided by data and is estimated in a stepwise manner. Researchers may add a previously unidentified critical contextual stimulus to the model or remove a previously misidentified contextual stimulus.

Conceivably, the generalised model is more difficult to analyse and empirically test. The author is leaving all those challenges to ambitious, talented future scholars. In the following two paragraphs, the author will resume the sex experience example in the subsection *Alternative Views*. There, what is finally evaluated is the mean, as clarified.

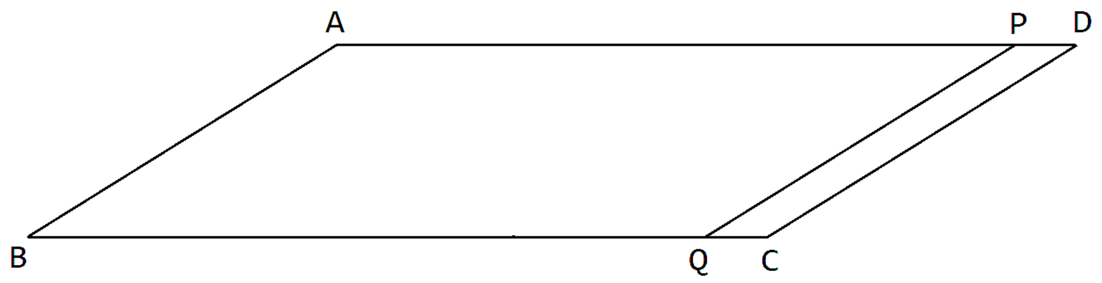
Let X be a random variable describing the evaluator's sex experience, the only one or one of a few, or one of many, in the past year that the evaluator recalls; and X_ε be a random variable defined by the equation: $\mathbf{D}_{X_\varepsilon}(x) = \mathbf{D}_X(x \mid X \in \mathbb{R}_X)$. The author is curious, which of the following, $\bar{X}$ or $\overline{X_\varepsilon}$, will be taken? $\bar{X} = (\sum \varkappa_X X)/(\sum \varkappa_X)$ and $\overline{X_\varepsilon} = \left(\sum \varkappa_{X_\varepsilon} X_\varepsilon\right)/\left(\sum \varkappa_{X_\varepsilon}\right)$. Taking $\overline{X_\varepsilon}$ implies that subliminal concepts have no effects on the final mean distribution; people consciously take the mean exclusively and precisely of what they access. The author, while having no empirical evidence, is more inclined to the other possibility, i.e., $\bar{X}$ taken as the mean random variable.

The choice between $\bar{X}$ and $\overline{X_\varepsilon}$ deserves academic deliberation. Again, the author is closing the discussion and looking forward to future scholars. What follows is a mere side note: In the sex experience example, people may actually be motivated to bias their own evaluations. For example, a man may overrate his average sex experience because he imagines that people will judge a man with a happy sex life as a man good at sex, and that is the image that he wants. The first postulate of the elaboration likelihood model (Petty & Wegener, 1999) is meant to be general and not universal.

Readers are now in the midst of this subsection. What remains is a number of four well-studied, but still rather interesting illusions. The first one is the Sander illusion:

Figure 29

Sander Illusion



Which point is farther from Q, A or D? Research reveals that people, in general, will consider A to be farther than D, while in fact, they are equidistant. Cooper et al. (1972) identified two factors contributing to this illusion: the area ratio of the left parallelogramme ABQP to the right parallelogramme PQCD, and visibility of AQ and DQ. A larger area ratio of the parallelogrammes and invisible AQ and DQ generally will yield a greater effect of the illusion. In the language of context effects research, it seems that visibility translates to ambiguity, and that the parallelogrammes are contextual stimuli.

See *Multiple Dimensions* for why the parallelogrammes are contextual stimuli even though they are not line segments. Now, refer to Figure 29 again. In addition to the two parallelogrammes, ABCD and all the line segments in sight, visible or invisible, except for AQ and DQ, are arguably contextual stimuli. All the visual elements in the figure that Cooper et al. took into account in their experiment were what they initially considered to be critical. And the two parallelogrammes, ABQP and PQCD, were empirically proved to be very likely to be truly critical to people's evaluations of AQ and DQ.

Presented below is the second illusion:

Figure 30

Delboeuf–Ebbinghaus Illusion

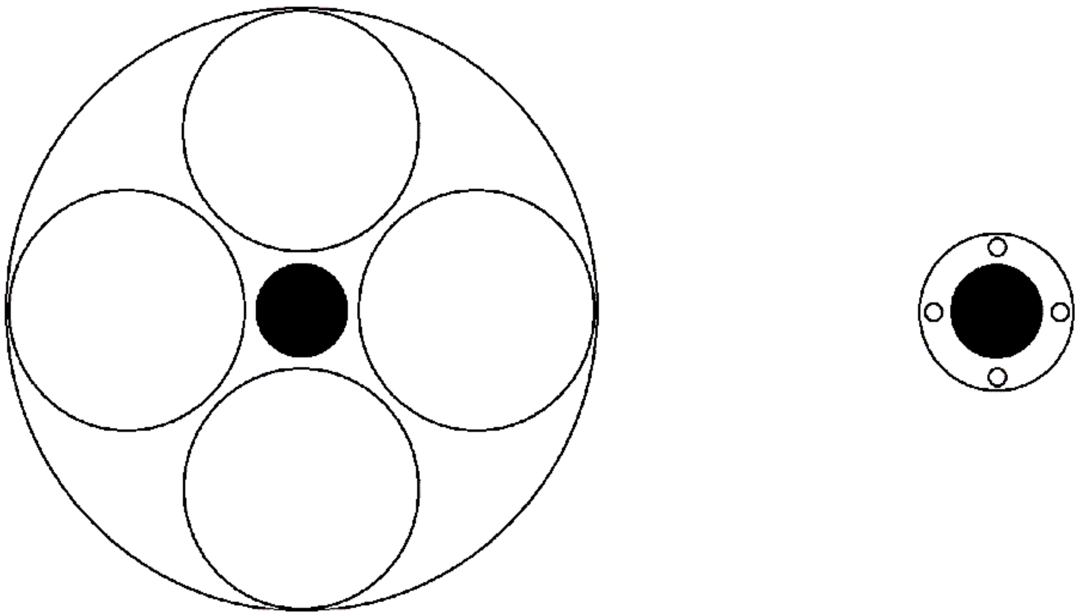
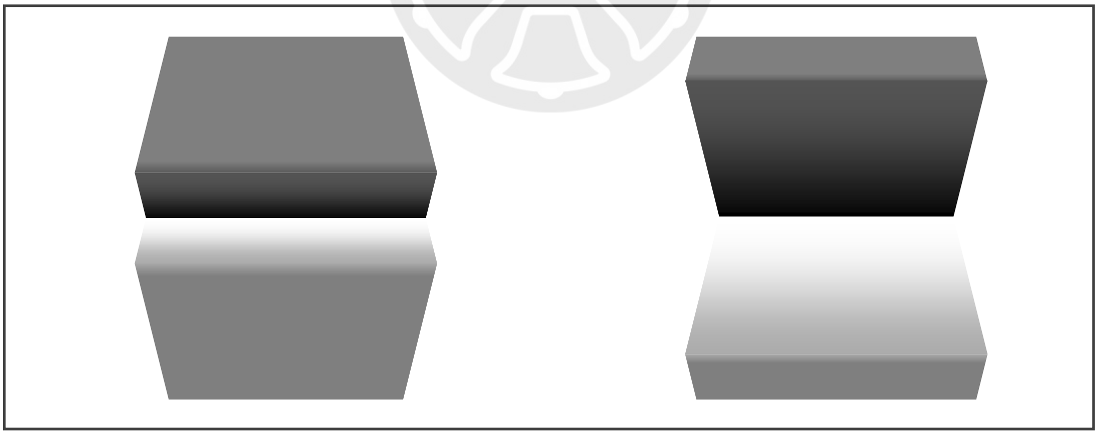


Figure 30 is a combination of the Delboeuf illusion and the Ebbinghaus illusion, and would be the Ebbinghaus illusion in itself if the largest circles on both sides were removed. Compare the sizes of the two black circles. Research reveals that people, in general, will indicate that the left black circle is smaller while in fact, they are of equal size. In the language of context effects research, the surrounding circles (including the largest) are the contextual stimuli and the black circles are the focal objects. Massaro and Anderson (1971) identified four factors deciding the magnitude of the context effects: on either side, the number and sizes of the contextual circles, the size of the focal circles, and the distances between the focal and contextual circles. When both sides are displayed together, as the case in Figure 30, the two systems merge into one, getting complicated.

Now, here comes the third illusion:

Figure 31

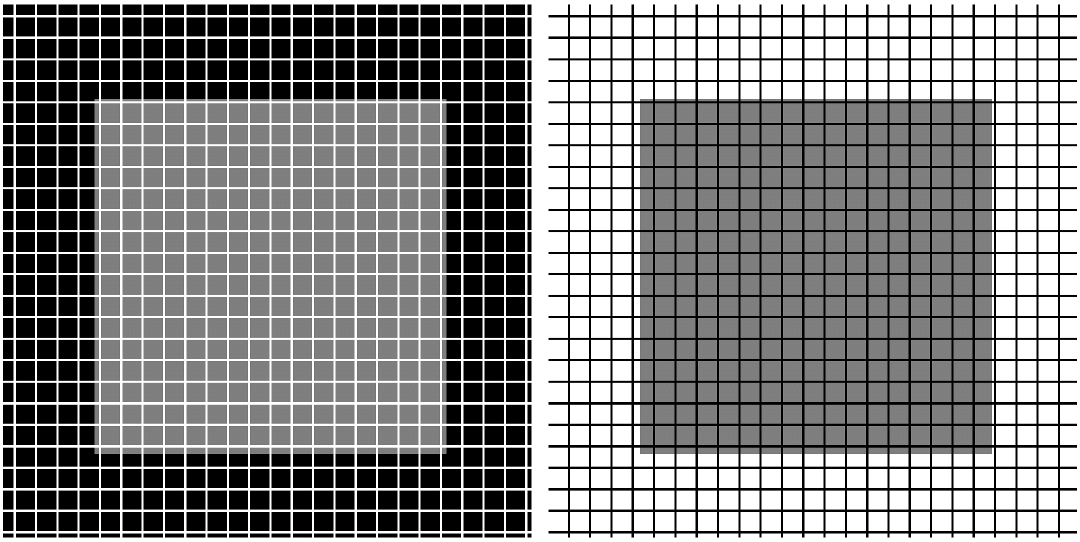
Cornsweet Illusion



Rate the brightness of the uppermost and lowermost grey trapezia in the figure on either side. People, in general, will find the lowermost trapezium to be brighter than the uppermost trapezium, while in fact, they are equally bright. See Purves et al.'s explanation (2002). Obviously, assimilation and contrast are concurrent.

Figure 32

White Illusion



Finally, this is the fourth illusion. Rate the brightness of the tiny squares in the two medium squares at the centres of the largest squares. People, in general, will find the tiny squares on the left side to be brighter than those on the right side, while in fact, they are equally bright. Again, this is a result of simultaneous assimilation and contrast.

This is the end of the present subsection. The author hopes that readers enjoy these illusions, which the author means to serve as examples of multiple contextual stimuli effecting perceptual change. The fourth illusion will be resumed in *Theory-Based Bias Correction*, which is the very end of this very last chapter.

Multiple Dimensions

Thus far, the author has focused the discussion on cases where only a single latent dimension of concepts is concerned. Naturally, one may be curious about cases where two or more dimensions are involved. For example, Givon and Shapira (1984), in developing their model, were addressing a composite scaling problem. A composite scale, by definition, involves multiple items, i.e., multiple dimensions. Accordingly, Givon and Shapira's model was multidimensional.

Is there, then, a need to develop a multidimensional version of the present model? The author is not sure. Take Figure 29, for example. If it is subjectively possible to evaluate the length of a parallelogramme, then there is no need. If, however, people consider it to be senseless to evaluate a parallelogramme by its length, since it is undefined in people's minds what exactly is the length of a parallelogram, then a model with two different dimensions will be necessary; their relationship will then be an interesting topic.

Theory-Based Bias Correction

Finally, here comes the ending subsection. Although the author makes it clear that 'Herein, context effects are the sole interest' (see *Literature Review*), the author considers the following finding to be possibly worthy of note:

Review Figure 32. Now, readers already possess the knowledge that the tiny squares in the medium squares on both sides are equally bright. Presumably, readers want to give an accurate rating, and are aware that the largest squares are producing contrast, and that the grid lines in the medium squares are producing assimilation. Now, rerate the brightness of the two collections of tiny squares. Will readers' ratings change? Anecdotal evidence suggests that some people will not, and the author is among those people. (Interested readers may refer to the flexible correction model (Wegener & Petty, 1997).)

References

Chiěn, Y.-W. & Hsiāo, C.-C. (2001). Dimensional range overlap model for explanation of contextual priming effects. In M. C. Gilly & J. Meyers-Levy (Eds.), *Advances in Consumer Research* (Vol. 28, p. 136). Valdosta, GA: Association for Consumer Research.

Chiěn, Y.-W., Wegener, D. T., Hsiāo, C.-C., & Petty, R. E. (2010). Dimensional range overlap and context effects in consumer judgments. *Journal of Consumer Research*, 37(3), 530–542. https://doi.org/10.1086/652415.

Cooper, M. R., Runyon, R. P., Tatz, S. J., & Heimer, W. I. (1972). The Sander illusion as a function of relative space and component lines. *Perception & Psychophysics*, 11(1), 102–104. https://doi.org/10.3758/bf03212695.

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, 127(3), 286–303. https://doi.org/10.1037/0096-3445.127.3.286.

Givon, M. M., & Shapira, Z. (1984). Response to rating scales: A theoretical model and its application to the number of categories problem. *Journal of Marketing Research*, 21(4), 410–419. https://doi.org/10.2307/3151467.

Herr, P. M. (1989). Priming price: Prior knowledge and context effects. *Journal of Consumer Research*, 16(1), 67–75. https://doi.org/10.1086/209194.

Hsiāo, C.-C. (2002). *The Reciprocity Hypothesis as an Explanation of Perception Shifts in Consumer Judgments* (AAI3099155) [doctoral dissertation]. Purdue University.

Karremans, J. C., Stroebe, W., & Claus, J. (2006). Beyond Vicary’s fantasies: The impact of subliminal priming and brand choice. *Journal of Experimental Social Psychology*, 42(6), 792–798. https://doi.org/10.1016/j.jesp.2005.12.002.

Keelin, T. W. (2016). The metalog distributions. *Decision Analysis*, 13(4), 243–277. https://doi.org/10.1287/deca.2016.0338.

Kullback, S. & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86. https://doi.org/10.1214/aoms/1177729694.

Li, R. & Nadarajah, S. (2018). A review of Student's t distribution and its generalisations. *Empirical Economics*, 58(3), 1461–1490. https://doi.org/10.1007/s00181-018-1570-0.

Lin, J. (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1), 145–151. https://doi.org/10.1109/18.61115.

Luce, R. D. (1997). Several unresolved conceptual problems of mathematical psychology. *Journal of Mathematical Psychology*, 41(1), 79–87. https://doi.org/10.1006/jmps.1997.1150.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, 51(3), 493–504. https://doi.org/10.1037/0022-3514.51.3.493.

Massaro, D. W. & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *Journal of Experimental Psychology*, 89(1), 147–151. https://doi.org/10.1037/h0031158.

Panaretos, V. M., & Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annual Review of Statistics and Its Application*, 6(1), 405–431. https://doi.org/10.1146/annurev-statistics-030718-104938.

Petty, R. E. & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology* (pp. 41–72), New York: Guildford Press.

Purves, D., Lotto B. R., & Nundy S. (2002) Why we see what we do: A probabilistic strategy based on past experience explains the remarkable difference between what we see and physical reality. *American Scientist*, 90(3), 236–243.

Thurstone, L. L. (1927a). Psychophysical analysis. The American Journal of Psychology, 38(3), 368–389. https://doi.org/10.2307/1415006.

Thurstone, L. L. (1927b). A law of comparative judgment. *Psychological Review*, 34(4), 273–286. https://doi.org/10.1037/h0070288.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. https://doi.org/10.1007/bf00122574.

Wegener, D. T. & Petty, R. E. (1997). The flexible correction model: The role of naïve theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 29, pp. 141c208). San Diego: Academic Press.