An Illustrated Guide to Context Effects

Clintin P. Davis-Stober, A. A. J. Marley, William J. McCausland, and Brandon M. Turner

University of Missouri

University of Victoria

Université de Montréal

The Ohio State University

Author Note

Correspondence concerning this article should be addressed to William J. McCausland, Département de sciences économiques, Université de Montréal. E-mail: william.j.mccausland@umontreal.ca

Abstract

Three context effects pertaining to stochastic discrete choice have attracted a lot of attention in Psychology, Economics and Marketing: the similarity effect, the compromise effect and the asymmetric dominance effect. Despite this attention, the existing literature is rife with conflicting definitions and misconceptions. We provide theorems relating different variants of each of the three context effects, and theorems relating the context effects to conditions on discrete choice probabilities, such as random utility, regularity, the constant ratio rule, and simple scalability, that may or may not hold for any given discrete choice model. We show how context effects at the individual level can sometimes aggregate to context effects at the population level. Importantly, we offer this work as a field guide for researchers to sharpen empirical tests and aid future development of choice models.

Keywords: context effects, similarity effect, compromise effect, asymmetric dominance, random utility, regularity, common ratio, simple scalability

An Illustrated Guide to Context Effects

Introduction

Across the many disciplines that study decision making and choice, the term *context effect* broadly refers to how preferences between alternatives can systematically change when (typically non-preferred) alternatives are added to the choice set under consideration. Such behavior can call into question standard models of choice and suggest new directions for theoretic development. This article focuses on three well studied context effects relevant to the study of stochastic discrete choice: the similarity, compromise, and asymmetric dominance effects.

There is a vast empirical literature documenting various combinations of context effects across many different choice settings, including: consumer choice (e.g., Berkowitsch et al., 2014; Doyle et al., 1999), policy decisions (Herne, 1997), intertemporal choice (Gluth et al., 2017), and perceptual stimuli (Trueblood, Brown, Heathcote, & Busemeyer, 2013). Context effects have also been observed in the animal kingdom. For example, researchers have observed monkeys (Parrish et al., 2015), felines (Scarpi, 2011), and insects (Latty & Trueblood, 2020) exhibiting phantom decoy effects (where the decoy alternative is not available for selection).

Despite this interest, the empirical literature on context effects is largely disconnected and somewhat contradictory, with some context effects failing to be observed, or, in some cases, reversing their direction altogether (e.g., Spektor et al., 2018). Spektor et al. (2021) provide a comprehensive overview and synthesis of this contradictory empirical literature. To briefly summarize, they argue that whether or not a particular context effect manifests may depend upon three broad categories referring to the nature of the choice environment.

1. *The spatial arrangement of the choice stimuli.* As one example, Cataldo and Cohen (2018) demonstrated that the robustness and direction of some context effects depend

on whether the choice stimuli is presented in tabular formats that encourage either by-alternative or by-attribute comparisons.

2. *Whether the stimuli are abstract or concrete in nature.* There is good evidence that context effects can manifest when the choice alternatives are defined via concrete, clearly understood and easily comparable attributes (e.g., numerical monetary values). However, the evidence for context effects becomes murkier when the choice alternatives become more abstract in nature (e.g., rectangles of varying sizes; Spektor et al., 2018).

3. *How much time is available for individuals to make choices and how that time is controlled.* Generally speaking, context effects are more robust when decision makers have more time to deliberate (Spektor et al., 2021). Cataldo and Cohen (2021) provide a comprehensive analysis of time constraints on context effects (also see Molloy et al., 2019).

Given the many variables that impact the presence (or absences) of context effects, Spektor et al. (2021) conclude that a deeper understanding of how choice alternatives are cognitively represented is an important direction for future theory development.

Trueblood (2022) presents a theoretic synthesis of major competing theories of multi-alternative, multi-attribute choice that can account for various combinations of context effects. By focusing on dynamic theories of choice, Trueblood concludes that some of the disparate empirical observations in the study of context effects can be explained by changes in attention and attentional processes. Relatedly, there is growing work examining context effects from a neural perspective, focusing on value-based cognitive systems (Busemeyer et al., 2019; Gluth et al., 2020; Gluth et al., 2018).

In addition to the issues raised above, we argue that the literature on context effects is further complicated by inconsistencies in how context effects are defined, tested, and ultimately related to common decision models, such as random utility. To this end, we develop, under a universal notation, all relevant context effect definitions and known relations to a set of fundamental choice properties, which includes regularity, the common-ratio rule, and simple scalability. This section of the article can be considered as a precise tutorial review for readers interested in studying context effects. We also provide new theoretic results on how these context effects relate to one another and fundamental choice properties. We also explore how context effects at the individual level can sometimes aggregate to context effects at the population level. While we are certainly not the first to examine individual-population relationships among context effects, see Liew et al. (2016) and Katsimpokis et al. (2022), our results bring additional precision about which definitions of various context effects will aggregate. Altogether, we provide a comprehensive guide for investigating context effects, one that may help bring further clarity to the study of context effects.

Notation

We use the following notation throughout the paper. Let $U = \{x_1, \ldots, x_n\}$ be a finite universe or master set of choice objects. When faced with a non-empty choice set $A \subseteq U$, a decision maker (DM) chooses a single object from A. The probability that the DM chooses $x \in A$ is denoted $P_A(x)$. A *random choice structure* (RCS) on a master set U is the complete specification of the $P_A(x)$, $x \in A \subseteq U$, and is denoted P. For distinct $x, y \in U$, we use the shorthand notation p_{xy} to mean $P_{\{x,y\}}(x)$. We denote by $\Delta(U)$ the set of all RCSs on U; it is a Cartesian product of unit simplices of various dimensions.

Table 1 gives an example of an RCS on $U = \{x, y, z\}$ by specifying a complete list of choice probabilities on the non-empty subsets of U. The singleton choice probabilities in the first three rows are, of course, degenerate, but it is useful to include them in the definition of an RCS so that conditions like those in equation (1) below have concise expressions.

Figure 1 gives a graphical representation of the same RCS using a Barycentric coordinate system. Each point in the triangle xyz is a unique convex combination

$\lambda_x x + \lambda_y y + \lambda_z z$ of the vertices x, y and z, where $\lambda_x, \lambda_y, \lambda_z \geq 0$ and $\lambda_x + \lambda_y + \lambda_z = 1$. The vector $(\lambda_x, \lambda_y, \lambda_z)$ gives the Barycentric coordinates of the point. Thus, the vertices x, y and z have Barycentric coordinates $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$, respectively. Since xyz is equilateral, the distances of the point $(\lambda_x, \lambda_y, \lambda_z)$ to the sides yz, xz and xy of the triangle are fractions λ_x, λ_y and λ_z, respectively, of the height of the triangle.

We will be representing probability vectors on doubleton and tripleton choice sets as points in Barycentric coordinates. Here, for example, the ternary probability vector $P_U(\cdot)$ in the final row of Table 1 is represented by a solid dot in the interior of the triangle in Figure 1, with Barycentric coordinates $(P_U(x), P_U(y), P_U(z)) = (0.6, 0.1, 0.3)$. The lengths of the solid light grey line segments joining the point $(P_U(x), P_U(y), P_U(z))$ to the sides yz, xz and xy of triangle xyz are fractions 0.6, 0.1 and 0.3, respectively, of the height of xyz, as can easily be seen using the light grey dashed grid lines.

We represent the binary choice probabilities in rows four to six of Table 1 as points on the boundary of triangle xyz in Figure 1. The hollow dot on the left side of the triangle gives the choice probabilities $p_{xy} = 0.7$ and $p_{yx} = 1 - p_{xy} = 0.3$. It is the convex combination $p_{xy}x + p_{yx}y$ of vertices x and y, and has Barycentric coordinates $(p_{xy}, p_{yx}, 0)$. Similarly, the hollow dot on the right side gives the choice probabilities p_{yz} and p_{zy}; and the hollow dot on the base gives p_{xz} and p_{zx}. We adopt the convention that binary probabilities are indicated by hollow dots and ternary probability vectors by solid dots, so we can tell the difference between a binary probability and a ternary probability vector that happens to be on the boundary of the triangle.

We distinguish between two different interpretations of a RCS. An *individual* RCS governs the choices of a single individual; a *population* RCS, those of a random sample of individuals from a population. Note that if each individual in a population is governed by an individual RCS, which may be degenerate, then the population RCS will be a convex combination of individual RCSs: each population choice distribution $P_A(\cdot)$, $A \subseteq U$, is a mixture—with random sampling from the population being the common mixing

A	$P_A(x)$	$P_A(y)$	$P_A(z)$
$\{x\}$	1.0		
$\{y\}$		1.0	
$\{z\}$			1.0
$\{x, y\}$	0.7	0.3	
$\{y, z\}$		0.6	0.4
$\{x, z\}$	0.8		0.2
$\{x, y, z\}$	0.6	0.1	0.3

Table 1

A random choice structure on $U = \{x, y, z\}$.

distribution across choice sets A—of the various individual $P_A(\cdot)$.

Models for discrete choice and their properties

We will use the term *discrete choice model* to refer to any model specifying an RCS, directly or indirectly. Examples of indirect specifications include random utility models such as multinomial probit models as well as sequential elimination models such as Elimination by Aspects (EBA) models. In these examples, choice probabilities are not specified directly but are derived from the specification of a parametric random choice process and cannot be expressed as functions of the model parameters in closed form. Often in the literature, the term "discrete choice model" refers to the specification (direct or indirect) of an RCS up to a vector of parameters; that is, the specification of a set of models, with each model specifying an RCS. To make a clear distinction, we will call such a specification a *class* of discrete choice models. So for example, while many might refer to "*the* logit model" as "*a* discrete choice model", we will refer to logit as a class of models.

Over the last several decades, many conditions on choice probabilities, including various kinds of stochastic transitivity, regularity and simple scalability have been studied

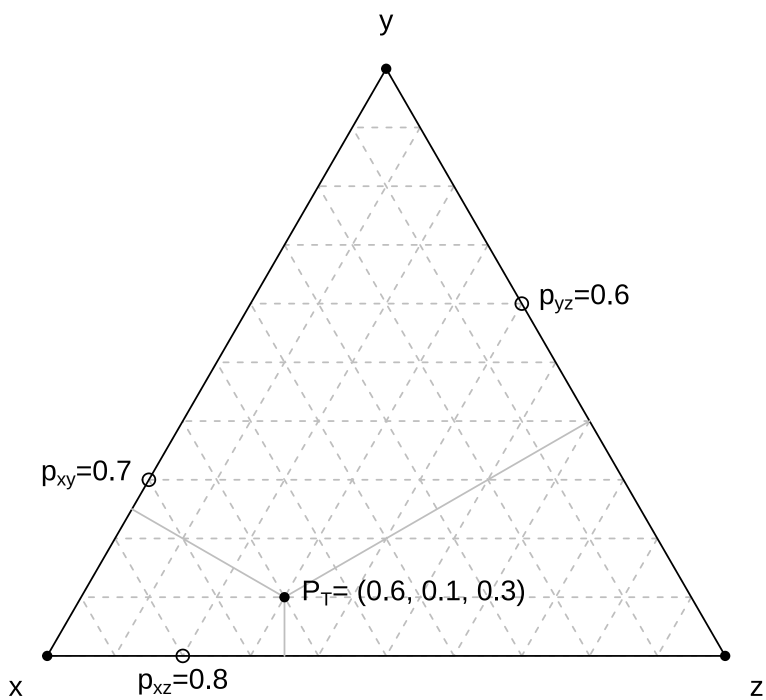

Figure 1

A random choice structure on $U = \{x, y, z\}$ as points in a Barycentric coordinate system.

(see Fishburn, 1999, for a thorough survey). For a given discrete choice model and a given condition, the model either satisfies the condition or it doesn't. When we say that a class of discrete choice models satisfies some condition, we mean that every model in the class satisfies that condition.

We will emphasize conditions that are targets of criticism (for being unrealistic) in the context effects literature. Thus, our discussion is organized around two properties: (i) regularity, a property of all random utility models that is inconsistent with the asymmetric dominance effect, and (ii) the constant ratio rule, a property of multinomial logit models that is inconsistent with the similarity and compromise effects. We discuss regularity and related properties in Section Regularity and random utility immediately below, then the constant ratio rule and related properties in Section The constant ratio rule and simple scalability.

Regularity and random utility

Most classes of discrete choice models in wide use in Economics and Marketing (and many, but fewer, in Psychology) satisfy a condition known as random utility. We will define this condition, but we first need to define a random utility model.

Definition 1 *A* random utility model *(RUM) for a master set U is a probability space $(\Omega, \mathcal{F}, \mu)$, where Ω is a sample space, $\mathcal{F}$ is an event space and μ is a probability measure on $\mathcal{F}$; and a measurable function $u\colon U \times \Omega \to \mathbb{R}$, where μ is* non-coincident*, meaning that for all distinct $x, y \in U$,*

$$\mu\left(\{\omega \in \Omega\colon u(x,\omega) = u(y,\omega)\}\right) = 0.$$

We call u a utility function; its maximization over the available options governs choice in state $\omega \in \Omega$. Non-coincidence rules out ties. Henceforth, we will denote a RUM by the utility function u, suppressing notation for $(\Omega, 2^\Omega, \mu)$. A RUM u for U *induces* the RCS $P^{(u)}$ through the construction

$$P_A^{(u)}(x) = \mu\left(\{\omega \in \Omega\colon u(x,\omega) = \max_{y\in A} u(y,\omega)\}\right), \quad x \in A \subseteq U.$$

We say that a RCS P satisfies random utility if it can be induced by a random utility model. This is a restrictive condition; Block and Marschak (1960) show that the following set of linear inequalities is necessary for random utility: for all non-empty $A \subseteq U$ and all $x \in A$,

$$\sum_{B\colon A\subseteq B\subseteq U} (-1)^{|B\setminus A|} P_B(x) \geq 0. \tag{1}$$

Provided P satisfies random utility, the left hand side of the inequality is, for given x and A, the probability of x having greater utility than each of the other elements of A and less utility than each of the elements of $U\setminus A$. Falmagne (1978) shows that this set of inequalities is also sufficient. McCausland and Marley (2013) summarize what is known about how random utility relates to other conditions on choice probabilities, and illustrate how strong a condition it is, in terms of its low prior probability as an event in $\Delta(U)$, for a large class of prior distributions on $\Delta(U)$.

Examples of random utility models include multinomial logit, multinomial probit, mixed multinomial logit, and generalized extreme value models. In these examples and others, the specification of a discrete choice model proceeds by describing a random utility model and then computing—often numerically—the induced choice probabilities. However, there are other discrete choice models that are more naturally specified in some other way, but which satisfy random utility nonetheless; one can check the conditions in (1) without actually constructing a random utility model. For example, Tversky (1972a) shows that the class of Elimination By Aspects (EBA) models, introduced by Tversky (1972b), satisfies random utility, although EBA choice probabilities are usually derived as the result of a random sequential choice process.

Because the term "context effect" is open to interpretation, it will be useful to carefully examine the way in which random utility models—and by extension, their induced choice probabilities—are context invariant. In the definition of a random utility model, the utility distribution does not depend on the choice set offered. This is one kind of context invariance, and it is what gives content to the random utility condition. It is important to understand, however, that this does not imply that a so-called context effect is inconsistent with random utility. The similarity and compromise effects, as defined by Tversky (1972b) and Tversky and Simonson (1993), respectively, are both consistent with random utility, as we show below. In fact, we have just seen that EBA models, introduced by Tversky (1972b) specifically to account for the similarity effect, satisfy random utility.

Other classes of models generate choice probabilities in an explicitly context dependant manner. An example where individual models may or may not satisfy random utility, depending on the universe of choice objects, is the random regret minimization model of Chorus (2010). In simulations, not reported here but using R code available at 10.6084/m9.figshare.21791186, we found regions of a variable object's attribute space for which choice probabilities satisfy random utility and other regions for which they do not. Davis-Stober et al. (2017) provide a set of models which do or do not satisfy random utility depending upon their parametric specifications.

A weaker condition than random utility is *regularity*, the condition that adding objects to a choice set cannot increase choice probabilities of the objects already in the set.

Definition 2 *A RCS P satisfies* regularity *if for all* $x \in A \subset B \subseteq U$*,*

$$P_A(x) \geq P_B(x). \tag{2}$$

The regularity condition is relevant because it is both a consequence of random utility and (transparently) inconsistent with the original (i.e. the binary-ternary version, as defined below) asymmetric dominance effect.

Much of the evidence against random utility comes from the literature on the asymmetric dominance effect, where inequalities of the form (2), for particular choice sets A and B, are tested and often rejected: Rieskamp et al. (2006) survey empirical violations of five “consistency principles” in economics, regularity is the only one of these principles that is necessary for random utility, and the asymmetric dominance effect is the only empirical evidence they document against regularity.

There is also some literature on other kinds of tests, either direct tests of random utility or tests of other necessary conditions. Regenwetter et al. (2011) test a condition on binary choice probabilities called the triangle inequality, a condition that is itself implied by regularity. Using experimental binary choice data for 18 participants, they find strong evidence against the triangle inequality for one participant (also see Cavagnaro & Davis-Stober, 2014). McCausland and Marley (2014) and McCausland et al. (2020) jointly test the full set of necessary and sufficient conditions for random utility in (1). Using a subset of the data in Regenwetter et al. (2011), McCausland and Marley (2014) find moderate evidence against random utility for the same participant and mild evidence against for another participant. Using experimental choice data featuring choice from all binary and larger subsets of a master set McCausland et al. (2020) find strong evidence against random utility for four participants out of 141.

In some cases, it will be convenient to speak of random preference models, which are very similar to random utility models for finite master sets. In a random preference model, a random strict preference (or ranking) replaces the random vector of utility values in a random utility model. A random preference model induces choice probabilities when a decision maker chooses the highest ranked object (rather than the object with the highest utility) from a given choice set. For a finite master set U, the discrete choice models that can be induced by a random preference are the same ones that can be induced by a random utility model (see Marley & Regenwetter, 2018). Analogous to the definition of a random utility model and the construction of a RCS from one, we have the following:

Definition 3 *A* random preference model *(RPM) on a master set U is a probability space $(R, 2^R, \pi)$ where R is the set of strict linear orders on U.*

We use the notation $\succ \in R$ to denote an outcome, so that, for example, $\pi(\{\succ \in R \colon x \succ y\}) = \sum_{\{\succ \in R \colon x \succ y\}} \pi(\{\succ\})$ is the probability that the fixed object $x \in U$ is ranked above (i.e., is preferred to) the fixed object $y \in U$, $y \neq x$.

A random preference model with random binary relation $\succ$ *induces* the RCS $P^{(\succ)}$ through the construction

$$P_A^{(\succ)}(x) = \pi\left(\{\succ \in R \colon x \succ y \text{ for all } y \in A \backslash \{x\}\}\right), \quad x \in A \subseteq U.$$

The constant ratio rule and simple scalability

A second pair of conditions targeted by the context effects literature is the constant ratio rule and simple scalability. The constant ratio rule underlies the multinomial logit model, a model that has been widely used since the onset of the context effects literature and is inconsistent with all three context effects, as they are defined in the seminal context effect articles outlined in Section The Seminal Papers below.

Definition 4 (Tversky, 1972b) *P satisfies the* constant ratio rule *if for all $x, y \in A$,*

$$\frac{p_{xy}}{p_{yx}} = \frac{P_A(x)}{P_A(y)},$$

whenever the denominators do not vanish.

The term *independence of irrelevant alternatives* (or IIA) is commonly used, especially in economics, to mean the same thing, although the term has other meanings: for example, Ray (1973) documents three different conditions with the IIA name and clarifies the relationships among them. To avoid ambiguity or confusion, we will use the term constant ratio rule.

If, in addition to the constant ratio rule, we also require all choice probabilities to be positive, we get the class of multinomial logit models. The constant ratio rule is a consequence of Luce's choice axiom, introduced in Luce (1959), but because this axiom does not rule out choice probabilities equal to zero, there are models satisfying Luce's choice axiom that are not multinomial logit models.

Many theoretical and empirical problems with the constant ratio rule were raised before the seminal papers in the context effect literature. Some of the theoretical arguments and empirical evidence against the constant ratio rule also apply to a weaker condition, called simple scalability.

Definition 5 (Krantz, 1964) *P satisfies* simple scalability *if there exists a function $u\colon U \to \mathbb{R}$ and functions $F_2\colon \mathbb{R}^2 \to \mathbb{R}, F_3\colon \mathbb{R}^3 \to \mathbb{R}, \ldots, F_{|U|}\colon \mathbb{R}^{|U|} \to \mathbb{R}$ such that for all $A = \{x, y, \ldots, z\} \subseteq U$,*

$$P_A(x) = F_{|A|}(u(x), u(y), \ldots, u(z)),$$

where each F_i is increasing in its first argument, and strictly so if $P_A(x) < 1$; and decreasing in its other arguments, and strictly so if $P_A(x) > 0$.

Tversky (1972b) shows that simple scalability is equivalent to the following condition:

Definition 6 (Tversky, 1972b) *P satisfies* order independence *if for all $A, B \subseteq U$, $x, y \in A - B$ and $z \in B$*

$$P_A(x) \geq P_A(y) \Leftrightarrow P_{B\cup\{x\}}(z) \leq P_{B\cup\{y\}}(z)$$

provided the choice probabilities on the two sides of either inequality are not both 0 or 1.

The order independence condition can be seen as an ordinal version of the constant ratio rule. Since it is a condition on choice probabilities, it allows for statistical testing of simple scalability and the exploration of the logical relationship between simple scalability and various context effects.

One of the first theoretical counterexamples[1] to simple scalability was provided by Debreu (1960) in a review of Luce (1959). Another well-known counterexample is the Bicycle/Pony example provided in Luce and Suppes (1965), attributed there to L.J. Savage. The Red bus/Blue bus example of McFadden (1974) appears shortly after Tversky (1972b), but it is a particularly well known one.

Tversky (1972b) cites the following papers as providing empirical evidence against simple scalability: Becker et al. (1963), Chipman (1960), Coombs (1958), Krantz (1967), and Tversky and Russo (1969). Luce (1977) provides an even more extensive list and some details.

Illustrations

Figure 2 illustrates how regularity, the constant ratio rule, and simple scalability constrain random choice structures. Again, we consider a RCS on the set $U = \{x, y, z\}$. We use a Barycentric coordinate system to plot choice probability vectors and sets of them. We fix the binary probabilities $p_{xz} = 0.45$ (point A) and $p_{yz} = 0.35$ (point B) and show the values of p_{xy} and $P_U(\cdot)$ that are consistent with the fixed p_{xz} and p_{yz}, under each of regularity, the constant ratio rule and simple scalability.

[1] A *theoretical counterexample* to a given condition is a specification of choice probabilities, claimed to be plausible, that is inconsistent with the condition.

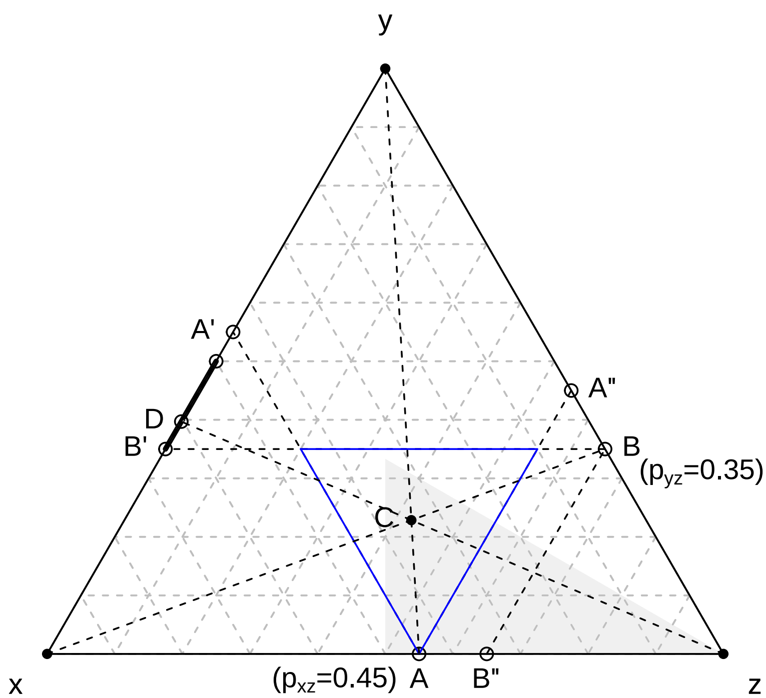

Figure 2

Regularity, the constant ratio rule and simple scalability

First, suppose that regularity holds. Then p_{xz} is an upper bound for $P_U(x)$, and its complementary probability $p_{zx} = 1 - p_{xz}$ is an upper bound for $P_U(z)$. The set of ternary probability vectors $P_U(\cdot)$ satisfying $P_U(x) \leq p_{xz}$ and $P_U(z) \leq p_{zx}$ is the parallelogram $A'AA''y$. Likewise, p_{yz} and p_{zy} give upper bounds on $P_U(y)$ and $P_U(z)$, and the set of ternary probability vectors satisfying $P_U(y) \leq p_{yz}$ and $P_U(z) \leq p_{zy}$ is the parallelogram $B'BB''x$. The intersection of the two parallelograms, the inverted equilateral triangle in solid lines, is the boundary of the set of ternary probability vectors $P_U(\cdot)$ consistent with $p_{xz} = 0.45$ and $p_{yz} = 0.35$ under regularity.

With p_{xz} and p_{yz} still fixed, we now relax regularity and suppose that the constant ratio rule holds. Then given $p_{xz} = 0.45$, the ternary probability $P_U(\cdot)$ must fall on the line segment Ay, the locus of ternary probabilities satisfying $P_U(x)/P_U(z) = p_{xz}/p_{zx} = 0.45/0.55 = \frac{9}{11}$; given $p_{yz} = 0.35$, it must fall on Bx, the locus of ternary probabilities satisfying $P_U(y)/P_U(z) = p_{yz}/p_{zy} = 0.35/0.65 = \frac{7}{13}$. The intersection of these two line segments, the point C, is the only value of the ternary choice probability vector that is consistent with both $p_{xz} = 0.45$ and $p_{yz} = 0.35$ under the constant ratio rule. The extension of line segment Cz intersects side xy at D, which gives the only value of the binary probability p_{xy} consistent with $p_{xz} = 0.45$ and $p_{yz} = 0.35$ under the constant ratio rule.

Finally, we relax the constant ratio rule and suppose that simple scalability holds. Here a range of possible values of $P_U(\cdot)$ and p_{xy} are possible. Rieskamp et al. (2006) show that simple scalability implies strong stochastic transitivity (SST) of binary choice probabilities, the condition that for all distinct $a, b, c \in U$, if $p_{ab} \geq \frac{1}{2}$ and $p_{bc} \geq \frac{1}{2}$, then $p_{ac} \geq \max(p_{ab}, p_{bc})$. For the given values of p_{xz} and p_{yz}, SST constrains p_{xy} to be in the interval $[1/2, \min(1 - p_{xz}, 1 - p_{yz})]$, shown as a thick line segment in xy. The interior of the gray shaded region gives the set of ternary choice probabilities consistent with simple scalability, provided that p_{xy} is selected to satisfy SST: the set satisfying $P_U(z) > P_U(x) > P_U(y)$.

The Seminal Papers

The similarity effect is commonly attributed to Tversky (1972b), who describes a "similarity hypothesis" relating to a choice environment in which choice objects x and y are similar and object z is dissimilar to both x and y. The hypothesis is that the addition of y to choice set $\{x, z\}$ will reduce the share of x of the total probability of choosing x or z and symmetrically, that the addition of x to $\{y, z\}$ will reduce the share of y of the total probability of choosing y or z.

Similarity Effect. Tversky (1972b) and Tversky (1972a) motivate the similarity hypothesis by providing several theoretical counterexamples to simple scalability, and by extension, to the constant ratio rule. Tversky (1972a) (page 292) claims that the hypothesis is "incorporated into the [EBA] model". The two papers make it clear that the similarity hypothesis is a prediction about individual and not population choice. The evidence of this is three-fold: first, the theoretical counterexamples in these papers pertain to individual choice; second, EBA models, introduced in Tversky (1972b) to accommodate similarity effects, is described as a random process giving rise to individual choice probabilities; third, the empirical results in Tversky (1972b) are based on observations of the repeated choices of individuals.

Compromise Effect. Simonson (1989) and Tversky and Simonson (1993) each introduce a version of the compromise effect; both effects pertain to a situation where an object y is "between" two other objects, x and z. Simonson (1989) describes a strong version of the effect, where adding x to the choice set $\{y, z\}$ increases the probability of choosing y, and adding z to the choice set $\{x, y\}$ also increases the probability of choosing y. Tversky and Simonson (1993) describe a weak version, where adding x to $\{y, z\}$ increases the *relative* probability of y over z *and* where adding z to the choice set $\{x, y\}$ increases the *relative* probability of y over x. Tversky and Simonson (1993) interpret probabilities exclusively as population probabilities. However, the theoretical result they provide to argue that the compromise effect gives indirect evidence against random preference applies equally well to individual and population random preferences: they show that a condition on random preferences (the ranking condition, defined below) rules out the compromise effect. The plausibility of the ranking condition is subjective, and one may regard its plausibility differently for individual and population random preferences, but the theoretical result applies regardless.

Asymmetric Dominance Effect. Huber et al. (1982) introduce the asymmetric dominance effect. They consider the addition of a third object, called a decoy, to a binary choice set in which neither object "dominates" the other. The decoy is dominated by one object in the pair but not the other. They claim that adding the decoy "*can* [emphasis added] increase the probability of choosing the item that dominates it." When it does, the asymmetric dominance effect is said to occur. Their discussion of choice probabilities does not distinguish between individual and population choice probabilities, and the empirical evidence provided in the paper comes from both within- and between-subject experimental designs. There is a tighter connection between individual and population effects for the asymmetric dominance effect than for the other two effects: aggregation properties of the former are much more solid than those of the latter, as we will see in Section Aggregation of context effects.

Huber et al. (1982) describe an "attraction" effect, similar to the asymmetric dominance effect, but where the decoy is not necessarily dominated by its "target". Many authors have since treated the terms "attraction effect" and "asymmetric dominance effect" as synonymous and requiring dominance, including the original authors of both papers, in Huber et al. (2014). We adopt the latter term, to avoid any ambiguity.

Already, three interpretive issues arise in these descriptions: the distinction between population and individual RCSs, the distinction between a weak and a strong version of an effect, and the question of whether a claim about a "context effect" is a prediction that a phenomenon *will* occur or a claim that a phenomenon *can* occur. Before we provide a more complete list of the issues in Section An overview of the issues, we break down the descriptions of context effects into three components.

Structure, preconditions and effect regions

In this section we introduce some terminology that will allow us to decompose the description of a context effect into three components:

1. the structure of the universe U of choice objects that allows us to talk about similarity, between-ness or dominance;

2. a precondition on choice objects that can be asserted or verified before observing choice probabilities, such as two objects being similar to each other and each dissimilar to a third object; and

3. the effect region, a subset of $\Delta(U)$ defining a restriction on choice probabilities

associated with the precondition.

To illustrate, we will look more closely at the similarity hypothesis, as described by Tversky (1972b), which will serve as an example. In defining the similarity hypothesis, Tversky (1972b) describes a tripleton set $T \equiv \{x, y, z\}$ of choice objects where x and y are "similar" and z is "dissimilar" to each of x and y. We will take $T \subseteq U$, with T not necessarily the same set as the universe U. He defines the "similarity hypothesis" as follows, on page 292. Mathematical notation and equation numbers have been modified to be consistent with the current paper.

> ...the similarity hypothesis ...predicts that the addition of alternative y to the set $\{x, z\}$ will reduce $P_T(x)$ proportionally more than $P_T(z)$. That is, the similar alternative, x, will lose relatively more than the dissimilar alternative, z, by the addition of y. Likewise, y is expected to lose relatively more than z by the introduction of x. ...the similarity hypothesis implies
>
> $$p_{xz} > \frac{P_T(x)}{P_T(x) + P_T(z)} \tag{3}$$
>
> and
>
> $$p_{yz} > \frac{P_T(y)}{P_T(y) + P_T(z)}. \tag{4}$$

In Section Definitions of context effects, we develop notation allowing us to express the similarity hypothesis above as follows: let $\sim$ be a similarity relation on U and suppose that $T = \{x, y, z\} \subseteq U$ is such that $x \sim y$, $x \not\sim z$ and $y \not\sim z$; then $P \in S_{xy}$. Here the structure of the universe U is the similarity relation $\sim$, the precondition is the triple $T = \{x, y, z\} \subseteq U$ satisfying $x \sim y$, $x \not\sim z$ and $y \not\sim z$; and the effect region S_{xy} is the subset of $\Delta(U)$ where both—note the "and" between the two inequalities—inequalities (3) and (4) hold.

An overview of the issues

We will argue for more precision in the descriptions of context effects, along the following dimensions.

1. *Whether an effect pertains to individual or population probabilities.* We have seen that in the seminal papers, the similarity effect applies to individual choice probabilities; the compromise effect, to population choice probabilities; and the asymmetric dominance effect, to both. In the subsequent literature, the similarity and compromise effects have been more flexibly defined. Compounding problems of imprecision on this point is the fact that not all effects aggregate: we will show below that it is possible for the RCS of each individual in a population to satisfy a context effect and not the RCS for the population, and vice versa.

2. *Whether an effect is* **weak***, such as the compromise effect described in Tversky and Simonson (1993) or* **strong***, such as the compromose effect described in Simonson (1989) and the asymetric dominance effect described in Huber et al. (1982).* We will see that the weak and strong effects have very different implications about whether choice probabilities satisfy random utility. Also, the aggregation properties of the weak and strong effects are very different, and the degree of empirical evidence supporting the weak effect is much more solid than that supporting the strong effect.

3. *In the case of the similarity and compromise effects, whether the claimed effect is one- or two-sided.* Tversky (1972b) makes the prediction that both (3) and (4) will occur when the precondition is satisfied, and we will call this a two-sided effect. However, much of the evidence presented in that paper provides more support for one-sided effects, in which one of those equations holds but not the other. In the subsequent literature, both kinds of effects are reported and the distinction between the two is not always made explicit.

4. *Whether the claim one makes about a context effect is that it is predicted to occur in a given context, as in the case of the seminal papers on the similarity and compromise effects, or that the context effect is possible, as in the seminal paper on the asymmetric dominance effect.* The former is a much stronger claim, although the latter is still important, as the effect may be inconsistent with properties of commonly used discrete choice models.

5. *Whether the precondition is asserted, as the similarity relation in the seminal similarity effect paper, or can be derived from a more basic and objective criterion, such as the between-ness relation in Tversky and Simonson (1993), one of the seminal compromise effects paper.*

6. *Whether preconditions and effect regions are binary-ternary or ternary-ternary.* The context effects introduced in the seminal papers are binary-ternary, as they relate choice probabilities on a ternary choice set and those on one or more binary subsets. Since then, articles have tested ternary-ternary versions of these context effects, which relate choice probabilities on two ternary choice sets having two elements in common. While the ternary-ternary and binary-ternary versions of an effect are related, they are distinct effects.

Without clarity in these dimensions, claims that theories "capture" a context effect, claims about evidence for context effects, and claims about what this evidence implies about conditions such as regularity, random utility, the constant ratio rule, and simple scalability are hopelessly vague.

Outline

The paper is organized as follows. We introduce definitions for various versions of the three context effects in Section Definitions of context effects. We survey some known theoretical results and identify some open theoretical questions in Section Theoretical

Results and conclude in Section Conclusions.

Definitions of context effects

In this section we elaborate on the terminology set forth in the introduction that allow us to decompose the description of a context effect into three components: a structure on the master set (such as a similarity relation $\sim$), preconditions (such as $x \sim y$, $x \not\sim z$, $y \not\sim z$) and an effect region (such as that defined by equations (3) and (4)). We first describe similarity, compromise and asymmetric dominance effects where choice probabilities on binary and ternary choice sets are compared. The seminal papers all describe effects of this kind, and we will call them binary-ternary (or 2-3) effects. Then we will describe ternary-ternary (or 3-3) variants of the three effects where choice probabilities on two different ternary sets are compared.

The structure of the master set

Each of the three context effects is described in terms of a relation defined on the master set of objects. As we have seen, for the similarity effect, there is a binary similarity relation $\sim$. For the compromise effect, there is a ternary between-ness relation $\cdot|\cdot|\cdot$, where $x|y|z$ means that y is "between" x and z. For the asymmetric dominance effect, there is a dominance relation, which we will denote $\triangleright$. We will assume that the similarity condition is symmetric, that the between-ness relation is symmetric in the sense that $x|y|z \Leftrightarrow z|y|x$, and that the dominance relation is transitive and irreflexive. This allows us to specify relations more succinctly, as the assumptions "fill in" the missing information. We consider the assumptions reasonable and are not aware of examples where they do not hold.

In some cases, the relation is asserted, with an appeal to the reader's intuition. Most people would accept the characterization, in Tversky (1972b), of two different recordings of the same Beethoven symphony as being similar to each other and a recording of a Debussy suite as being dissimilar to each of the Beethoven recordings. Luce and Suppes (1965) describe an example they attribute to L. J. Savage in his correspondence with Luce, where dominance is asserted. The objects in the example are a pony, a bicycle, and "another bicycle...which, although basically the same as [the first] is better in minor ways, such as having a speedometer". McCausland et al. (2021) describe an experiment where participants are asked to choose the event they "think is most likely to happen in the next twenty years". They propose that the disjunction of two events ("Either Scotland or Quebec become independent countries") dominates any one of the two events ("Scotland becomes an independent country"), which dominate the conjunction of the two events ("Scotland and Quebec become independent countries.") based on coherent probability judgements.

Often, however, the similarity, between-ness and dominance relations are derived relations, objectively defined in terms of a deeper structure on the master set. Very often, this deeper structure consists of an *attribute* space. Choice object are identified with vectors, points in a Euclidean space of finite dimension I. Each dimension is an attribute (such as price) of a single indivisible choice object and the value of the corresponding element of the vector is an attribute level (such as 10 dollars).

Two relations defined on attribute spaces are relevant to context effects, vector dominance and between-ness, defined below. We use the term vector dominance, rather than dominance, to emphasize that it is an objective property of objects, rather than a property of preferences. However, it is worth noting that this kind of dominance is usually considered relevant only when some kind of monotonicity of preferences is anticipated: in most cases, if x vector dominates y, x is clearly the better choice and we expect people to choose x over y when $\{x, y\}$ is presented. We can always transform attributes like price, where we might expect preferences to be monotonically decreasing, so that preferences are increasing in the transformed attribute levels.

Definition 7 *Let $x = (x_1, \ldots, x_I)$ and $y = (y_1, \ldots, x_I)$ be two objects in an attribute space. We say that $x = (x_1, \ldots, x_I)$* vector dominates *$(y_1, \ldots, y_I)$, which we denote $x \gneqq y$, if $x_i \geq y_i$ for all $i \in \{1, \ldots, I\}$ and $x \neq y$.*

Definition 8 (Tversky and Simonson, 1993) *Let x, y and z be three objects in an attribute space. We say that y is* between *x and z, denoted $x|y|z$, if neither x nor z vector dominates the other, and for all $i \in \{1, \ldots, I\}$, either $x_i \leq y_i \leq z_i$ or $z_i \leq y_i \leq x_i$.*

Note the distinction between the dominance relation $\triangleright$, which does not rely on any attribute structure and in some cases may be primitive, and the vector dominance relation $\gneqq$, which is an objectively defined relation on attribute spaces. In many cases, the $\triangleright$ relation is derived from the $\gneqq$ relation. We are not aware of any cases in which the compromise effect is based on a between-ness relation other than that just defined, so there is no need to make a similar distinction here.

We discuss below the relationship among the three context effects, and so we emphasize here that a master set can be endowed with more than one of the three relations. For example, we might specify both a similarity relation $\sim$ and a between-ness relation $\cdot|\cdot|\cdot$ on the same master set. If we have a tripleton set $\{x, y, z\} \subseteq U$ where $x \sim y$, $x \not\sim z$, $y \not\sim z$ and $x|y|z$, we simultaneously satisfy the preconditions for the similarity and compromise effects.

Preconditions

The precondition for the similarity effect is the existence of a set $T = \{x, y, z\} \subseteq U$ satisfying $x \sim y$, $x \not\sim z$ and $y \not\sim z$; that for the compromise effect is the existence of a set $T = \{x, y, z\} \subseteq U$ such that $x|y|z$; that for the asymmetric dominance effect is the existence of a set $T = \{x, y, z\} \subseteq U$ satisfying $x \triangleright y$, $z \not\triangleright y$ and $x \not\triangleright z$.

Preconditions of different context effects can be satisfied simultaneously, and in some reasonable cases, associated effects may be inconsistent with each other. For example, an object y may be between objects x and z (i.e. $x|y|z$) but closer to x than y. Suppose we posit $x \sim y$, $x \not\sim z$, $y \not\sim z$. One of the inequalities that the weak similarity effect associates with this precondition is $P_T(z)/P_T(y) > p_{zy}/p_{yz}$. However, one of the inequalities the weak compromise effect associates with the precondition $x|y|z$ is $P_T(z)/P_T(y) < p_{zy}/p_{yz}$. Clearly the two inequalities cannot both be satisfied. For another example, consider a triple $T = \{x, y, z\}$ satisfying $x \triangleright y$, $z \not\triangleright y$, $x \not\triangleright z$. The asymmetric dominance effect associates this precondition with the inequality $P_T(x) > p_{xz}$. If x is close to y, we might reasonably posit $x \sim y$, $x \not\sim z$, $y \not\sim z$. However, the weak similarity effect associates this additional precondition with the inequality $P_T(x)/P_T(z) < p_{xz}/p_{zx}$. Again, we have two inequalities that cannot both hold.

These examples illustrate the importance of clarifying whether one is predicting that an effect will occur or claiming that it is possible. They also illustrate the importance of distinguishing between one-sided effects and the stricter two-sided effects, since the latter may be in conflict with another type of context effect even when the former is not. Finally, they encourage precise descriptions of the structure of the master set, to avoid conflicting predictions.

Effect regions

Part of the description of a context effect is a condition on choice probabilities associated with a precondition. We represent these conditions as *effect regions*, subsets of $\Delta(U)$, the space of all RCSs on the master set U. We can interpret effect regions and their complements as hypotheses amenable to statistical testing; detecting the presence of a context effect, contingent on the precondition being satisfied, can be framed as a test of such a hypothesis.

We introduce some notation that helps reveal the resemblance among the various context effects. For a given master set U and distinct objects $a, b, b' \in U$, we define the *weak generic effect region* $E_{abb'}$ as

$$E_{abb'} \equiv \left\{ P \in \Delta(U) \colon \frac{P_{\{a,b,b'\}}(b)}{P_{\{a,b,b'\}}(a) + P_{\{a,b,b'\}}(b)} > p_{ba} \right\},$$

where adding the object b' to the set $\{a, b\}$ increases the probability *share* of b, out of the two elements a and b. We define the *strong generic effect region* $\mathbf{E}_{abb'}$ as

$$\mathbf{E}_{abb'} \equiv \{P \in \Delta(U) \colon P_{\{a,b,b'\}}(b) > p_{ba}\},$$

where adding the object b' to the set $\{a, b\}$ increases the probability of b. Note that $\mathbf{E}_{abb'} \subset E_{abb'}$: the strong generic effect is indeed more restrictive than the weak generic effect. The weak generic effect region can also be written as

$$E_{abb'} = \left\{ P \in \Delta(U) \colon p_{ab} > \frac{P_{\{a,b,b'\}}(a)}{P_{\{a,b,b'\}}(a) + P_{\{a,b,b'\}}(b)} \right\},$$

where the defining inequality resembles equations (3) and (4) more transparently. However, this form obscures the relationship between the weak and strong generic effect regions that is more obvious in the previous definition. We will define $E^c_{abb'}$ and $\mathbf{E}^c_{abb'}$ as the complements of $E_{abb'}$ and $\mathbf{E}_{abb'}$, respectively, in $\Delta(U)$. We now express the effect regions for various versions of the three types of context effects, in terms of the generic regions.

Similarity effects

Let $T \equiv \{x, y, z\} \subseteq U$. The similarity effects we describe here are associated with a similarity relation $\sim$ on the master set U, together with the precondition that $x \sim y$, $x \not\sim z$ and $y \not\sim z$.

We call the similarity effects described by Tversky (1972b), restated in equations (3) and (4), weak similarity effects. In terms of weak generic effect regions, the region where (3) holds is E_{xzy} and the region where (4) holds is E_{yzx}. Later, we will use the following shorthand notation for one- and two-sided weak similarity effects. We define $S_{xy} \equiv E_{xzy} \cap E_{yzx}$, the region in which (3) and (4) both hold, and call the associated effect a two-sided weak similarity effect. Similarly, we define $S_x \equiv E_{xzy} \cap E^c_{yzx}$ and $S_y \equiv E^c_{xzy} \cap E_{yzx}$, and call the associated effects one-sided weak similarity effects. Finally, we let $\not{S} \equiv E^c_{xzy} \cap E^c_{yzx}$, the region where neither (3) nor (4) holds.

Figure 3 shows cross sections of the sets S_x, S_y and S_{xy}. For the same fixed values $p_{xz} = 0.45$ and $p_{yz} = 0.35$ as in Figure 2, Figure 3 shows the set of values of the ternary probability vector $P_T(\cdot)$ in each of these regions, in Barycentric coordinates. Also indicated on the diagram, to help illustrate results in Section Theoretical Results, are the regularity and simple scalability regions from Figure 2. The boundary of those regions are indicated here by dashed and dotted lines, respectively. Note that for the given binary choice

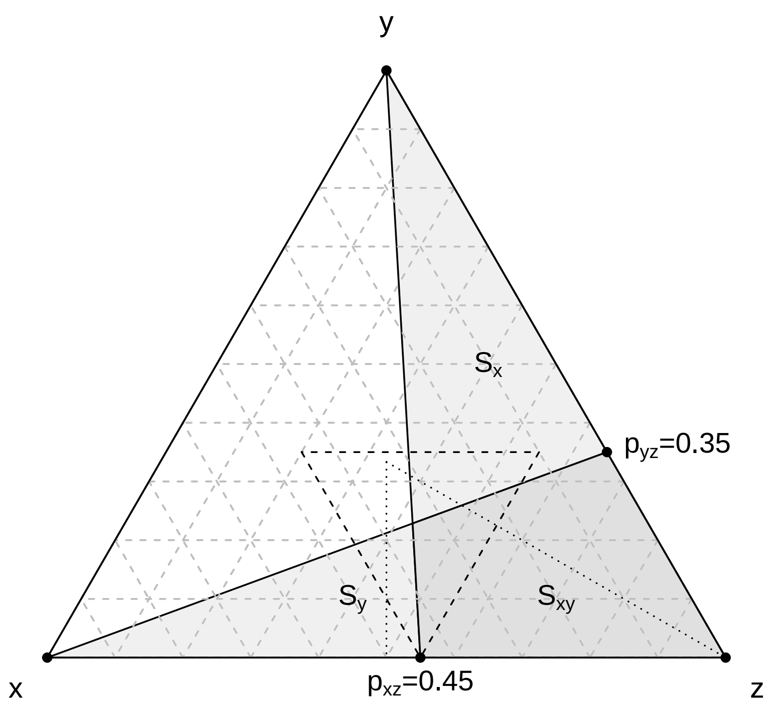

Figure 3

One- and two-sided weak similarity effects: given $p_{xz} = 0.45$ and $p_{yz} = 0.35$, S_{xy} is the region where (3) and (4) holds, S_x is the region where (3) holds and (4) does not and S_y is the region where (4) holds and (3) does not. $\not{S}$, the region where there is no weak similarity effect is unshaded.

probabilities, all of the regions S_x, S_y, S_{xy} and $\not{S}$ intersect with the region where simple scalability holds, illustrating that the type of weak similarity effect is logically independent of simple scalability. Similarly, the four regions all intersect with the region where regularity holds, showing that the type of weak similarity effect is logically independent of regularity.

We now define notation for what we will call strong similarity effects. Corresponding to the weak similarity regions E_{xzy} and E_{yzx}, described by equations (3) and (4), we have the strong similarity regions $\mathbf{E}_{xzy}$ and $\mathbf{E}_{yzx}$. We define, analogously with the weak similarity effects, the two-sided strong similarity effect $\mathbf{S}_{xy}$, the two one-sided strong similarity effects $\mathbf{S}_x$ and $\mathbf{S}_y$, and the absence of a strong similarity effect $\not{\mathbf{S}}$. Clearly $\mathbf{S}_{xyz} \subseteq S_{xyz}$, justifying the qualifiers weak and strong.

Compromise effects

Again, let $T \equiv \{x, y, z\} \subseteq U$, but now we describe compromise effects associated with the precondition that x, y and z satisfy the between-ness relation $x|y|z$.

The compromise effects described by Tversky and Simonson (1993) are weak effects. The effect region associated with adding z to $\{x, y\}$, making y a "compromise" option, is the weak effect region E_{xyz}; the region associated with adding x to $\{y, z\}$, also making y a compromise option, is E_{zyx}.

In a similar manner to the weak similarity effects, we use the shorthand notation C_{xz} to denote $E_{xyz} \cap E_{zyx}$ and call the associated effect a two-sided weak compromise effect. We use the notation $C_x \equiv E_{xyz} \cap E^c_{zyx}$ and $C_z \equiv E^c_{xyz} \cap E_{zyx}$ for the two one-sided weak compromise effect regions and $\not{C} \equiv E^c_{xyz} \cap E_{zyx}$ for the region where there is no weak compromise effect.

The compromise effects described by Simonson (1989) are strong effects, where the probability of choosing the between object y goes up when it becomes a compromise object. We use the notation $\mathbf{C}_x \equiv \mathbf{E}_{xyz} \cap \mathbf{E}^c_{zyx}$ and $\mathbf{C}_z \equiv \mathbf{E}^c_{xyz} \cap \mathbf{E}_{zyx}$ for the two one-sided strong compromise effect regions and $\not{\mathbf{C}} \equiv \mathbf{E}^c_{xyz} \cap \mathbf{E}_{zyx}$ for the region where there is no strong compromise effect.

The notation for the weak and strong generic effect regions helps reveal the common mathematical structure of the weak similarity and compromise effects and that of the strong similarity and compromise effects. In the similarity effect, there is a symmetry between the two similar choice objects, and the dissimilar object is in the same relation to both. In the compromise effect, there is a symmetry between the two extreme objects, and the between object is in a symmetric relation to the two others, in the sense that $x|y|z$ if and only if $z|y|x$. Furthermore, in both weak effects, the distinct object (dissimilar or between) is the one that is relatively favoured in ternary choice, compared to binary choice. In both strong effects, the distinct object is the one whose probability increases in ternary choice. Note that the definitions of C_{xz}, C_x, C_z, $\not{C}$, $\mathbf{C}_x$, $\mathbf{C}_z$ and $\not{\mathbf{C}}$ can be obtained by exchanging the elements y and z in the definitions of S_{xy}, S_x, S_y, $\not{S}$, $\mathbf{S}_x$, $\mathbf{S}_y$ and $\not{\mathbf{S}}$, respectively. We will exploit this common structure in the proofs of Section Theoretical Results.

Asymmetric dominance effects

For the same $T = \{x, y, z\} \subseteq U$, we now consider the effect region associated with the precondition that $x \rhd y$, $z \not\rhd y$ and $x \not\rhd z$, where $\rhd$ is a dominance relation on U. As far as we know, the asymmetric dominance effect is consistently described as a strong effect. The decoy object y in the asymmetric dominance effect, unlike the decoy in the similarity and compromise effects, is expected to have a very low choice probability. As a consequence, there is little difference in practice between the strong asymmetric dominance effect and a weak version of it. The effect region for the asymmetric dominance effect is $\mathbf{E}_{zxy}$, where the probability of choosing x increases when the decoy y, dominated by x but not by z, is added to the choice set $\{x, z\}$. For the sake of completeness of the results below, we define a weak asymmetric dominance effect as the effect with the same precondition and effect region E_{zxy}.

The asymmetric dominance effect differs in another way: there is no symmetry in the precondition, as there is between the two similar objects of the similarity effect and the two extreme objects of the compromise effect. Thus, there is only ever a one-sided effect. However, when we consider the ternary-ternary effects, we will see that in this case, the asymmetric dominance effect is more similar to the other two effects.

Ternary-ternary effects

Trueblood et al. (2014) give clear definitions of ternary-ternary versions of the similarity, compromise, and asymmetric dominance effects, although the idea goes back to at least Wedell (1991). All three ternary-ternary effects have the following features in common: there are two main choice objects, a and b, which appear in both relevant choice sets, and two other objects, a' and b'. Two ternary choice sets are presented to decision makers, $\{a, b, a'\}$ and $\{a, b, b'\}$. Relative to the inclusion of b', the inclusion of a' in the choice set favours the choice of a relative to b. For a given master set U and distinct elements a, b, a' and b', the generic ternary-ternary effect region is

$$E_{a'abb'} \equiv \left\{P \in \Delta(U) \colon P_{\{a,b,a'\}}(a) > P_{\{a,b,b'\}}(a) \text{ and } P_{\{a,b,a'\}}(b) < P_{\{a,b,b'\}}(b)\right\},$$

so that when we compare choice probabilities on the ternary set where the third element is a' with those where the third element is b', the probability of choosing a is higher and the probability of choosing b is lower.

In the ternary-ternary similarity effect, there is a similarity relation $\sim$ on U and the precondition is $a' \sim b$, $b' \sim a$ and $a \not\sim b$. In the ternary-ternary compromise effect, there is a between-ness relation $\cdot|\cdot|\cdot$ and the precondition is $a'|a|b$ and $a|b|b'$. In the ternary-ternary asymmetric dominance effect, there is a dominance relation $\triangleright$ and the precondition is $a \triangleright a'$, $b \triangleright b'$, $a \not\triangleright b'$, $b \not\triangleright a'$.

Although the same name is used to describe both ternary-ternary and binary-ternary effects, sometimes in the same paper, they are distinct effects. They are related, however, and we explore this relation below.

Theoretical Results

In this section we provide some results relevant to the issues raised in Section An overview of the issues. Some are previously known results and others are new. We begin in Section Compatibility (or not) of context effects with regularity, random utility, the constant ratio rule and simple scalability with results on the compatibility, or not, of the three context effects with regularity, random utility, the constant ratio rule and simple scalability. We then discuss results that relate different kinds of context effect. Section Aggregation of context effects relates individual and population versions of the same effects; some effects aggregate and others do not. Section Theorems relating binary-ternary and ternary-ternary context effects relates binary-ternary and ternary-ternary versions of the same effect. Section Theorems related to the compromise effect summarizes known results about the compromise effect.

Compatibility (or not) of context effects with regularity, random utility, the constant ratio rule and simple scalability

The compatibility results of this section are summarized in Table 2 and can be expressed briefly as follows. All context effects, weak or strong, one- or two-sided, binary-ternary or ternary-ternary, are incompatible with the constant ratio rule ($\times_1$). All ternary-ternary effects are incompatible with simple scalability ($\times_2$). All strong binary-ternary context effects, whether one- or two-sided, are incompatible with regularity and, by extension, with random utility ($\times_3$). However, all weak binary-ternary context effects, whether one- or two-sided, and all ternary-ternary context effects, are compatible with regularity, in the sense that for each type of effect, there are examples where regularity holds ($\checkmark_1$) Also, all binary-ternary context effects, whether weak or strong, whether one- or two-sided, are compatible with simple scalability, in the sense that for each type of effect, there are examples where simple scalability holds ($\checkmark_2$).

Effect type	Constant ratio rule	Simple scalability	Regularity
Weak 2-3	$\times_1$	$\checkmark_2$	$\checkmark_1$
Strong 2-3	$\times_1$	$\checkmark_2$	$\times_3$
3-3	$\times_1$	$\times_2$	$\checkmark_1$

Table 2

Summary of compatibility results

The incompatibility results are mostly well known and fairly obvious. Suppose we have a universe that includes objects a, b, a' and b': $\{a, a', b, b'\} \subseteq U$. The condition in the definition of the binary-ternary weak generic context effect $E_{abb'}$ is clearly inconsistent with $P_{\{a,b,b'\}}(a)/P_{\{a,b,b'\}}(b) = p(a,b)/p(b,a)$, a necessary condition for the constant ratio rule. By extension, all binary-ternary weak and strong context effects, one-sided and two-sided, are inconsistent with the constant ratio rule. If the two conditions in the definition of the ternary-ternary generic effect region $E_{a'abb'}$ hold, then $P_{\{a,b,a'\}}(a)/P_{\{a,b,a'\}}(b) = p(a,b)/p(b,a)$ and $P_{\{a,b,b'\}}(a)/P_{\{a,b,b'\}}(b) = p(a,b)/p(b,a)$ cannot both hold. Therefore all ternary-ternary effects are incompatible with the constant ratio rule. The stronger condition that ternary-ternary effects are incompatible with simple scalability is easily shown directly without appeal to its equivalence to the order independence condition. Suppose that simple scalability holds and let $F_2(\cdot,\cdot)$, $F_3(\cdot,\cdot,\cdot)$, $F_4(\cdot,\cdot,\cdot,\cdot)$ and $u(\cdot)$ be functions satisfying the conditions of Definition 5. The first defining condition of the generic ternary-ternary effect gives

$$F_3(u(a), u(b), u(a')) = P_{\{a,b,a'\}}(a) > P_{\{a,b,b'\}}(a) = F_3(u(a), u(b), u(b')),$$

which, together with the monotonicity property of F_3 implies $u(a') < u(b')$. The second defining condition gives

$$F_3(u(b), u(a), u(a')) = P_{\{a,b,a'\}}(b) < P_{\{a,b,b'\}}(b) = F_3(u(b), u(a), u(b'))$$

which, with monotonicity, implies $u(a') > u(b')$. Clearly the two conditions cannot both hold if simple scalability does. Finally, the condition in the definition of the strong generic context effect $\mathbf{E}_{abb'}$ is clearly inconsistent with $P_{\{a,b,b'\}}(b) \leq p_{ba}$, a necessary condition for regularity. By extension, all strong context effects, one-sided and two-sided, are inconsistent with regularity. Since regularity is itself necessary for random utility, they are inconsistent with random utility.

We now turn to compatibility results for binary-ternary effects. Since these results consist of examples of complete RCSs satisfying various combinations of conditions, we provide examples with the simplest structure possible. Thus, we restrict our attention to the universe $U = \{x, y, z\}$. We begin with weak similarity effects and show that whether we have a two-sided effect, a one-sided effect or no effect is logically independent of regularity and logically independent of simple scalability. We do this by providing examples of RCSs satisfying both simple scalability and regularity in each of the regions S_{xy}, S_x, S_y and $\not S$; and other examples of RCSs satisfying neither regularity nor simple scalability in each of those same regions. Figure 3 illustrates these facts graphically. In all our examples, $p_{xz} = 0.45$, $p_{yz} = 0.35$ and $p_{xy} = 117/194 \approx 0.603$. The values of p_{xz} and p_{yz} are the same as those in Figures 2 and 3. The value $p_{xy} = 117/194$ corresponds to point D in Figure 2, and is the only value of p_{xy} that makes the three binary choice probabilities consistent with the constant ratio rule. Given these three binary choice probabilities, the set of ternary choice probability vectors completing a RCS on $U = \{x, y, z\}$ satisfying simple scalability is the region where $P_U(z) > P_U(x) > P_U(y)$, indicated in grey in Figure 2 and in dotted lines in Figure 3. We turn now to the regularity condition—and indirectly, random utility, since they are equivalent in universes of size three. Given $p_{xz} = 0.45$ and $p_{yz} = 0.35$, the value $p_{xy} = 117/194$ imposes no further regularity restrictions on ternary choice probabilities: since $p_{xy} > p_{xz}$ and $p_{yx} > p_{yz}$, the constraint $P_U(x) \leq p_{xz}$ on $P_U(x)$ and the constraint $P_U(y) \leq p_{yz}$ make the constraints $P_U(x) \leq p_{xy}$ and $P_U(y) \leq p_{yx}$ redundant. Thus the set of ternary choice probability vectors consistent with the three binary choice probabilities we have just specified is the same set of vectors consistent with $p_{xz} = 0.45$ and $p_{yz} = 0.35$ alone, namely the inverted triangle in Figure 2, also indicated in Figure 3 with dashed lines.

Recall that the regions in Figure 3 labelled S_{xy}, S_x, S_y and $\not S$ are cross sections of the regions S_{xy}, S_x, S_y and $\not S$ defined in Section Definitions of context effects. Now it just a matter of pointing out that each of these cross sections intersects with the regularity region, the simple scalability region and each of their complements. Table 3 provides particular numerical examples, which can be verified directly. The first four rows show examples of ternary probability vectors, which, together with the three fixed binary probabilities, form a RCS satisfying both simple scalability and regularity, in the regions S_{xy}, S_x, S_y and $\not S$, respectively. The next four rows show examples giving RCSs satisfying neither simple scalability nor regularity, in the same four effect regions.

Effect	$P_T(x)$	$P_T(y)$	$P_T(z)$	$\frac{P_T(x)}{P_T(x)+P_T(z)}$	$\frac{P_T(y)}{P_T(y)+P_T(z)}$
S_{xy}	0.3	0.2	0.5	$3/8 = 0.375$	$2/7 \approx 0.286$
S_x	0.3	0.27	0.43	$30/73 \approx 0.411$	$27/70 \approx 0.386$
S_y	0.4	0.17	0.43	$40/83 \approx 0.482$	$17/60 \approx 0.283$
$\not{S}$	0.33	0.3	0.37	$33/70 \approx 0.471$	$30/67 \approx 0.448$
S_{xy}	0.0	0.2	0.8	0.0	0.2
S_x	0.0	0.8	0.2	0.0	0.8
S_y	0.8	0.0	0.2	0.8	0.0
$\not{S}$	0.5	0.5	0.0	1.0	1.0

Table 3

Compatibility of two-sided (S_{xy}), one-sided (S_x and S_y) and no ($\not{S}$) weak similarity effects with simple scalability and regularity (first four rows) and compatibility with violations of simple scalability and regularity (last four rows). In all eight cases, $p_{xy} = 117/194 \approx 0.603$, $p_{xz} = 0.45$ and $p_{yz} = 0.35$.

Given the common mathematical structure of the various similarity effects and their compromise effect counterparts, outlined in Section Compromise effects, we obtain similar compatibility results for weak compromise effects by exchanging the objects y and z. Table 4 is the resulting counterpart to Table 3. The first four rows give examples of ternary probability vectors, which, together with the binary choice probabilities $p_{xy} = 0.45$, $p_{xz} = 117/194 \approx 0.603$ and $p_{yz} = 0.65$, form a RCS satisfying both simple scalability and regularity, in the regions C_{xz}, C_x, C_z and $\not{C}$, respectively. The next four rows give RCSs satisfying neither simple scalability nor regularity, in the same effect regions.

Table 5 shows examples of RCSs explicitly induced by random preferences, in the regions S_{xy}, S_x, $\not{S}$, C_{xz}, C_x and $\not{C}$. The first column identifies which subset of the partition defined in Section Similarity effects that the example falls into. Columns π_{xyz} through π_{zyx}

Effect	$P_T(x)$	$P_T(y)$	$P_T(z)$	$\frac{P_T(x)}{P_T(x)+P_T(y)}$	$\frac{P_T(z)}{P_T(z)+P_T(y)}$
C_{xz}	0.3	0.5	0.2	$3/8 = 0.375$	$2/7 \approx 0.286$
C_x	0.3	0.43	0.27	$30/73 \approx 0.411$	$27/70 \approx 0.386$
C_z	0.4	0.43	0.17	$40/83 \approx 0.482$	$17/60 \approx 0.283$
$\not{C}$	0.33	0.37	0.3	$33/70 \approx 0.471$	$30/67 \approx 0.448$
C_{xz}	0.0	0.8	0.2	0.0	0.2
C_x	0.0	0.2	0.8	0.0	0.8
C_z	0.8	0.2	0.0	0.8	0.0
$\not{C}$	0.5	0.0	0.5	1.0	1.0

Table 4

Compatibility of two-sided (C_{xz}), one-sided (C_x and C_z) and no ($\not{C}$) weak compromise effects with simple scalability and regularity (first four rows) and compatibility with violations of simple scalability and regularity (last four rows). In all eight cases, $p_{xy} = 0.45$, $p_{xz} = 117/194 \approx 0.603$ *and* $p_{zy} = 0.35$.

describe a random preference and columns p_{xy} through $P_T(z)$ describe the choice probabilities induced by that random preference. The effects S_y and C_z are omitted from the table because of symmetry between x and y in the weak similarity effect and the symmetry between x and z in the weak compromise effect.

We now show that strong binary-ternary effects are compatible with simple scalability. We begin with two examples showing the compatibility of one- and two-sided strong similarity effects. In both examples, the binary choice probabilities are $p_{xy} = 117/194 \approx 0.603$, $p_{xz} = 0.45$ and $p_{yz} = 0.35$, as they were in previous examples. We have seen that SST holds for these probabilities. In the first example, we complete the RCS on $U = \{x, y, z\}$ with the ternary choice probability vector $(P_U(x), P_U(y), P_U((z)) = (0.3, 0.1, 0.6)$. Since $0.6 = P_U(z) > p_{zx} = 0.55$ but $0.6 = P_U(z) \not> p_{zy} = 0.65$, we have a one-sided, but not two-sided strong effect. Since

Effect	π_{xyz}	π_{xzy}	π_{yxz}	π_{yzx}	π_{zxy}	π_{zyx}	p_{xy}	p_{yz}	p_{xz}	$P_T(x)$	$P_T(y)$	$P_T(z)$
S_{xy}	1/4	0	1/4	0	1/4	1/4	1/2	1/2	1/2	1/4	1/4	1/2
S_x	0	1/8	3/8	1/8	3/8	0	1/2	1/2	1/2	1/8	1/2	3/8
$\not S$	0	1/2	0	1/2	0	0	1/2	1/2	1/2	1/2	1/2	0
C_{xz}	0	1/4	1/4	1/4	1/4	0	1/2	1/2	1/2	1/4	1/2	1/4
C_x	1/8	0	3/8	0	3/8	1/8	1/2	1/2	1/2	1/8	3/8	1/2
$\not C$	1/2	0	0	0	0	1/2	1/2	1/2	1/2	1/2	0	1/2

Table 5

Compatibility of weak similarity and weak compromise effects with random preferences. Each line gives the specification of a single random preference, in columns π_{xyz} through π_{zyx}; and the induced RCS, in columns p_{xy} through $P_T(z)$.

$P_U(z) > P_U(y) > P_U(x)$, simple scalability holds. In the second example, $(P_U(x), P_U(y), P_U((z)) = (0.2, 0.1, 0.7)$, and since $P_U(z) = 0.7$ is greater than both p_{zx} and p_{zy}, we have a two-sided strong effect. Because of common mathematical structure, the strong one- and two-sided binary-ternary compromise effects, and the strong binary-ternary asymmetric dominance effect—which is only ever one-sided—are also compatible with simple scalability.

Our final example in this section shows that the ternary-ternary effects are compatible with random preference, and by extension, random utility and regularity. The random preference that assigns probability $\epsilon \in (0, \frac{1}{2})$ to the rankings $b' \succ a \succ b \succ a'$ and $a' \succ b \succ a \succ b'$ and probability $\frac{1}{2} - \epsilon$ to the rankings $a \succ b \succ a' \succ b'$ and $b \succ a \succ a' \succ b'$ induces probabilities satisfying the two inequalities in the definition of the generic ternary-ternary effect: $\frac{1}{2} = P_{\{a,b,a'\}}(a) > P_{\{a,b,b'\}}(a) = \frac{1}{2} - \epsilon$ and

$\frac{1}{2} = P_{\{a,b,b'\}}(b) > P_{\{a,b,a'\}}(b) = \frac{1}{2} - \epsilon$. There are simpler examples, such as the random preference assigning probability $\frac{1}{2}$ to the first two rankings above, but the example provided is a more realistic outcome of an asymmetric dominance experiment: the probabilities of $a \succ a'$ and $b \succ b'$ are both $1 - \epsilon$ and can be set arbitrarily close to one.

Because only some versions of the context effects conflict with simple scalability, it is important to be precise about effect regions. We take Trueblood, Brown, and Heathcote (2013) (TBH) as a cautionary example. TBH assert (page 179)

> Beyond their practical implications for areas such as consumer choice, the three effects are important for theories of preference because they violate the property of simple scalability (Krantz, 1964; Tversky, 1972).

Krantz (1964) defines simple scalability, and the paper predates the similarity effect. Tversky (1972b) provides many theoretical arguments and much empirical evidence against simple scalability, introduces the EBA model specifically to "resolve this problem" and claims that the (weak two-sided binary-ternary) similarity effect is "incorporated into the [EBA] model". Then in the empirical section, he tests the stronger hypothesis of the constant-ratio rule, not simple scalability, against the alternative hypothesis of the similarity effect. He never claims, however, that the effect violates simple scalability, and we have seen that it does not. TBH provide three different descriptions of the "similarity effect", none of which corresponds to the two-sided weak effect introduced by Tversky (1972b) or to the one-sided version of it. In the context of binary-ternary choice, they give the following descriptions: (page 179 and page 181, respectively)

> The similarity effect arises from the introduction of an option that is similar to, and competitive with, one of the original alternatives, and causes a reduction in the probability of choosing the similar alternative.

and

> The similarity effect occurs when a competitive option ... that is similar to one of the existing alternatives is added to the choice set and the probability of selecting the dissimilar option increases.

The first of these descriptions is not much more than what regularity predicts. The second is a strong similarity effect, which does not necessarily violate simple scalability either. TBH also use the term "similarity effect" to describe (page 181) the ternary-ternary similarity effect, which does in fact violate simple scalability. Fortunately, this is the definition to which the empirical results of their paper pertain.

Aggregation of context effects

Suppose we have a population of decision makers, each of whose choices is governed by an RCS. For any given context effect, an individual RCS is either inside or outside the associated effect region, and the same is true for the population RCS. We discuss here the logical relationships between the occurrence or not of a given effect for individual RCSs and its occurrence or not for the population RCS.

It turns out that the aggregation properties of the weak similarity and weak compromise effects are very different from those of the strong versions of those effects. We can say much more about the aggregation of strong effects. This is because the effect regions for the strong effects are defined by one or more linear inequalities in choice probabilities and are therefore convex in $\Delta(U)$. Because a population RCS is a convex combination of individual RCSs, any strong effect that is satisfied by each individual RCS in the population must be satisfied by the population RCS. For the weak effects, the occurrence or not of an effect in the various individual RCSs implies very little about its occurrence or not in the population RCS; although we have seen convex cross-sections of the effect regions of weak context effects, the effect regions are not in fact **convex** and neither are their complements.

We support our claim about the (lack of) aggregation of weak effects by providing a long list of counterexamples. We start with the weak similarity effect and suppose we have (x, y, z) with $x \sim y$, $x \not\sim z$ and $y \not\sim z$. Each counterexample consists of the specification of two individual RCSs on $\{x, y, z\}$, $P^{(1)}$ and $P^{(2)}$; in each case, we compute an aggregate RCS P that is the equally weighted mixture of $P^{(1)}$ and $P^{(2)}$. We can think of a random draw from a population of the two individuals or, alternatively, a random draw from a population with two equally likely types.

Each of the three RCSs ($P^{(1)}$, $P^{(2)}$ and P) is in exactly one of S_{xy}, S_x, S_y or $\not S$; in each counterexample, we indicate, for each RCS, which is the case. There are $4^3 = 64$ combinations in total, but some are equivalent to others, because of symmetry between x and y, and are excluded. There is one counterexample for each row of Tables 6, 7 and 8, each one demonstrating the possibility of a given combination. It turns out that all combinations are possible.

We start with Table 6, which illustrates some of the more intuitively obvious possibilities. The binary probabilities do not vary from case to case within this table. They are $p^{(1)}_{xz} = p^{(2)}_{xz} = p_{xz} = 0.5$ and $p^{(1)}_{yz} = p^{(2)}_{yz} = p_{yz} = 0.5$. The binary probabilities $p^{(1)}_{xy}$, $p^{(2)}_{xy}$ and p_{xy} are not specified as they have no bearing on the occurence or not of the various effects. Each row gives gives the ternary probabilities for the individual RCSs $P^{(1)}$ and $P^{(2)}$ and for the aggregate RCS P. Each of the three RCSs is identified as satisfying one of S_{xy}, S_x, S_y or $\not S$, for convenience. All can be verified fairly easily by hand using the information provided in the table.

Take as an example the first row of Table 6. The two individual RCSs $P^{(1)}$ and $P^{(2)}$ are the same, and thus the aggregate RCS P is the same as $P^{(1)}$ and $P^{(2)}$. The first column indicates that $P^{(1)}$ satisfies S_{xy}, the two-sided weak similarity effect. The next three columns indicate that the ternary probabilities for $P^{(1)}$ in this case are $P^{(1)}(x) = P^{(1)}(y) = 0.2$ and $P^{(1)}(z) = 0.6$. The next four columns give the same information for $P^{(2)}$; the last four columns, for the aggregate RCS P. The relevant binary probabilities are given above the table and do not vary from case to case within the table.

Of course, it is obvious that two RCSs exhibiting the two-sided weak similarity effect can aggregate to an RCS satisfying the two-sided weak similarity effect. Some other cases are just as obvious but we include them all for completeness. We will see there are also cases that are not obvious, and even counter-intuitive.

Table 7 is similar to Table 6, except for the specification of the binary choice probabilities. As before, binary choice probabilities do not vary within the table, but now they are $p^{(1)}_{xz} = 0.2$, $p^{(2)}_{xz} = 0.8$ and $p_{xz} = 0.5$; and $p^{(1)}_{yz} = 0.8$, $p^{(2)}_{yz} = 0.2$ and $p_{yz} = 0.5$. Again, $p^{(1)}_{xy}$, $p^{(2)}_{xy}$ and p_{xy} need not be specified. The cases in this table are more counterintuitive. For example, in the second case of that table, two RCSs exhibiting weak two-sided similarity effects aggregate to an RCS satisfying no similarity effects. This case is illustrated in Figure 4, where the left, middle and right panels show the RCSs $P^{(1)}$, $P^{(2)}$ and P, respectively, in the same way as Figure 1 shows a particular RCS. The shaded regions of the three panels indicate cross sections of S_{xy}, S_x, S_y and $\not S$ in the same way that Figure 3 does.

Table 8 gives another set of examples. As in the two previous tables, binary choice probabilities do not vary within the table; here, they are $p^{(1)}_{xz} = 0.8$, $p^{(2)}_{xz} = 0.2$ and $p_{xz} = 0.5$; and $p^{(1)}_{yz} = 0.8$, $p^{(2)}_{yz} = 0.2$ and $p_{yz} = 0.5$. Again, $p^{(1)}_{xy}$, $p^{(2)}_{xy}$ and p_{xy} need not be specified. Here, too, there are counterintuitive cases. For example, in the fourth row of that table, two RCSs, neither exhibiting any kind of weak similarity effect aggregate to an RCS that exhibits a two-sided weak similarity effect. This case is illustrated in Figure 5 in the same manner as Figure 4 illustrated another case; again, the left, middle and right panels show the RCSs $P^{(1)}$, $P^{(2)}$ and P, respectively.

Table 9 gives a single example, the remaining case where $P^{(1)}$ and $P^{(2)}$ both satisfy S_x and P satisfies S_y. The case is illustrated in Figure 6, where the left middle and right panels show the RCSs $P^{(1)}$, $P^{(2)}$ and P, respectively.

Taking the four tables together, we have several remarks. First, several cases are omitted because of symmetry; for example, all cases where $P^{(1)}$ satisfies S_y are omitted

$p^{(1)}_{xz} = 0.5$, $p^{(1)}_{yz} = 0.5$				$p^{(2)}_{xz} = 0.5$, $p^{(2)}_{yz} = 0.5$				$p_{xz} = 0.5$, $p_{yz} = 0.5$			
	$P^{(1)}_T(x)$	$P^{(1)}_T(y)$	$P^{(1)}_T(z)$		$P^{(2)}_T(x)$	$P^{(2)}_T(y)$	$P^{(2)}_T(z)$		$P_T(x)$	$P_T(y)$	$P_T(z)$
S_{xy}	0.2	0.2	0.6	S_{xy}	0.2	0.2	0.6	S_{xy}	0.2	0.2	0.6
S_{xy}	0.2	0.2	0.6	S_x	0.1	0.5	0.4	S_{xy}	0.15	0.35	0.50
S_{xy}	0.2	0.2	0.6	S_x		0.8	0.2	S_x	0.1	0.5	0.4
S_{xy}	0.2	0.2	0.6	$\not{S}$	0.4	0.4	0.2	S_{xy}	0.3	0.3	0.4
S_{xy}	0.2	0.2	0.6	$\not{S}$	0.3	0.7		S_x	0.25	0.45	0.30
S_{xy}	0.3	0.3	0.4	$\not{S}$	0.5	0.5		$\not{S}$	0.4	0.4	0.2
S_x	0.1	0.5	0.4	S_x	0.1	0.5	0.4	S_x	0.1	0.5	0.4
S_x	0.1	0.5	0.4	S_y	0.5	0.1	0.4	S_{xy}	0.3	0.3	0.4
S_x		0.8	0.2	S_y	0.5	0.1	0.4	S_x	0.25	0.45	0.30
S_x		0.8	0.2	S_y	0.8		0.2	$\not{S}$	0.4	0.4	0.2
S_x	0.1	0.5	0.4	$\not{S}$	0.4	0.4	0.2	S_x	0.25	0.45	0.30
S_x	0.1	0.5	0.4	$\not{S}$	0.5	0.5		$\not{S}$	0.3	0.5	0.2
$\not{S}$	0.5	0.5		$\not{S}$	0.5	0.5		$\not{S}$	0.5	0.5	

Table 6

Patterns of occurrence of weak similarity effects in two individual RCSs and their population RCS.

because each is equivalent to another case where $P^{(1)}$ satisfies S_x. However, taking into account this symmetry, the list of cases is exhaustive: any pattern is possible. Second, within each table we have listed cases in the order S_{xy}, S_x, S_y, $\not{S}$, lexicographically with respect to $P^{(1)}$, $P^{(2)}$ and P. This is to help the reader find particular cases in the three tables.

Again, we can exploit the common mathematical structure of the various similarity effects and their compromise effect counterparts to obtain a full set of examples showing that any aggregation pattern is also possible for weak compromise effects. We do not

$p^{(1)}_{xz} = 0.2$, $p^{(1)}_{yz} = 0.8$				$p^{(2)}_{xz} = 0.8$, $p^{(2)}_{yz} = 0.2$				$p_{xz} = 0.5$, $p_{yz} = 0.5$			
	$P^{(1)}_T(x)$	$P^{(1)}_T(y)$	$P^{(1)}_T(z)$		$P^{(2)}_T(x)$	$P^{(2)}_T(y)$	$P^{(2)}_T(z)$		$P_T(x)$	$P_T(y)$	$P_T(z)$
S_{xy}		0.66	0.34	S_{xy}	0.58	0.08	0.34	S_x	0.29	0.37	0.34
S_{xy}		0.70	0.30	S_{xy}	0.70		0.30	$\not S$	0.35	0.35	0.30
S_{xy}	0.10	0.40	0.50	S_x	0.70	0.10	0.20	S_y	0.40	0.25	0.35
S_{xy}	0.05	0.70	0.25	S_x	0.55	0.20	0.25	$\not S$	0.30	0.45	0.25
S_x		0.85	0.15	S_x	0.60	0.15	0.25	$\not S$	0.30	0.50	0.20

Table 7

Patterns of occurrence of weak similarity effects in two individual RCSs and their population RCS.

tabulate these explicitly; the reader can construct individual examples as desired by exchanging objects y and z. For example, to construct an example where two RCSs in C_x aggregate to an RCS in C_{xz}, take the first example in Table 8. Exchanging objects y and z gives one RCS type with $p_{xy} = 0.8$, $p_{yz} = 0.2$, $P_T(x) = 0.42$, $P_T(y) = 0.11$ and $P_T(z) = 0.47$; a second RCS type with $p_{xy} = 0.2$, $p_{yz} = 0.8$, $P_T(x) = 0.12$, $P_T(y) = 0.69$ and $P_T(z) = 0.19$; and an aggregate RCD with $p_{xy} = 0.5$, $p_{yz} = 0.5$, $P_T(x) = 0.27$, $P_T(y) = 0.40$ and $P_T(z) = 0.33$.

Theorems relating binary-ternary and ternary-ternary context effects

Here we explore the theoretical relationship between the ternary-ternary (3-3) context effects described in Section Ternary-ternary effects, and related binary-ternary context effects. We will take a generic approach since the results here apply equally to the cases of similarity, compromise and asymmetric dominance effects.

We start with the choice environment outlined in Section Ternary-ternary effects, where U is a universe of objects including a, b, a' and b'. The relevant choice sets are $\{a, b\}$, $\{a, b, a'\}$ and $\{a, b, b'\}$. The presence of object a' is supposed to favour the choice of

$p^{(1)}_{xz} = 0.8$, $p^{(1)}_{yz} = 0.8$				$p^{(2)}_{xz} = 0.2$, $p^{(2)}_{yz} = 0.2$				$p_{xz} = 0.5$, $p_{yz} = 0.5$			
	$P^{(1)}_T(x)$	$P^{(1)}_T(y)$	$P^{(1)}_T(z)$		$P^{(2)}_T(x)$	$P^{(2)}_T(y)$	$P^{(2)}_T(z)$		$P_T(x)$	$P_T(y)$	$P_T(z)$
S_x	0.42	0.47	0.11	S_x	0.12	0.19	0.69	S_{xy}	0.27	0.33	0.40
S_x	0.42	0.47	0.11	$\not S$	0.18	0.17	0.65	S_{xy}	0.30	0.32	0.38
S_x	0.42	0.47	0.11	$\not S$	0.32	0.15	0.53	S_y	0.37	0.31	0.32
$\not S$	0.45	0.45	0.10	$\not S$	0.17	0.17	0.66	S_{xy}	0.31	0.31	0.38
$\not S$	0.45	0.45	0.10	$\not S$	0.15	0.31	0.54	S_x	0.30	0.38	0.32

Table 8

Patterns of occurrence of weak similarity effects in two individual RCS and their population RCS.

a; the presence of b', the choice of b.

The 3-3 effect pertains to the two ternary choice probability vectors $P_{\{a,b,a'\}}(\cdot)$ and $P_{\{a,b,b'\}}(\cdot)$. The effect region is defined by the following inequalities, both of which must hold for the effect to occur:

$$P_{\{a,b,a'\}}(a) > P_{\{a,b,b'\}}(a), \qquad (5) \qquad\qquad P_{\{a,b,a'\}}(b) < P_{\{a,b,b'\}}(b). \qquad (6)$$

Related to this 3-3 effect are two distinct 2-3 effects, one of them a constraint on $P_{\{a,b,a'\}}(\cdot)$ and p_{ab} and the other, a constraint on $P_{\{a,b,b'\}}(\cdot)$ and p_{ab}. Each of these effects can be strong, weak or absent. The strong versions of the two effects are

$$P_{\{a,b,a'\}}(a) > p_{ab}, \qquad (7) \qquad\qquad P_{\{a,b,b'\}}(b) > p_{ba}, \qquad (8)$$

and the weak versions can be expressed as

	type (1), S_x	type (2), S_x	population, S_y
p_{xz}	0.95	0.05	0.50
p_{yz}	0.80	0.50	0.65
$P_T(x)$	0.71	0.01	0.36
$P_T(y)$	0.24	0.52	0.38
$P_T(z)$	0.05	0.47	0.26

Table 9

Example showing two one-sided weak similarity effects (same side S_x) in two equi-probable types, (1) *and* (2)*, aggregating to a population RCS with a one-sided weak similarity effect on the opposite side S_y*

$$\frac{P_{\{a,b,a'\}}(a)}{P_{\{a,b,a'\}}(b)} > \frac{p_{ab}}{p_{ba}}, \quad (9) \qquad \frac{P_{\{a,b,b'\}}(b)}{P_{\{a,b,b'\}}(a)} > \frac{p_{ba}}{p_{ab}}. \quad (10)$$

We now establish some results that relate the occurrence of the 3-3 effect to the occurrence and strength of the two 2-3 effects. These can be summarized as follows. If the 3-3 effect occurs, then at least one weak 2-3 effect must occur. If both strong 2-3 effects occur, then the 3-3 effect must occur. All patterns of occurrence of the 3-3 and 2-3 effects that are not ruled out by these two results are possible.

First we show that the 3-3 effect implies that there must be a weak 2-3 effect. If the 3-3 similarity effect occurs, then both (5) and (6) hold and therefore

$$\frac{P_{\{a,b,a'\}}(a)}{P_{\{a,b,a'\}}(b)} > \frac{P_{\{a,b,b'\}}(a)}{P_{\{a,b,b'\}}(b)}.$$

The ratio p_{ab}/p_{ba} must either be less than the left-hand side of this equation, in which case (9) holds; or greater than the right-hand side, in which case (10) holds, or both. Therefore there must be at least one weak similarity effect.

We now show that if both strong 2-3 effects occur, then the 3-3 effect must occur. If

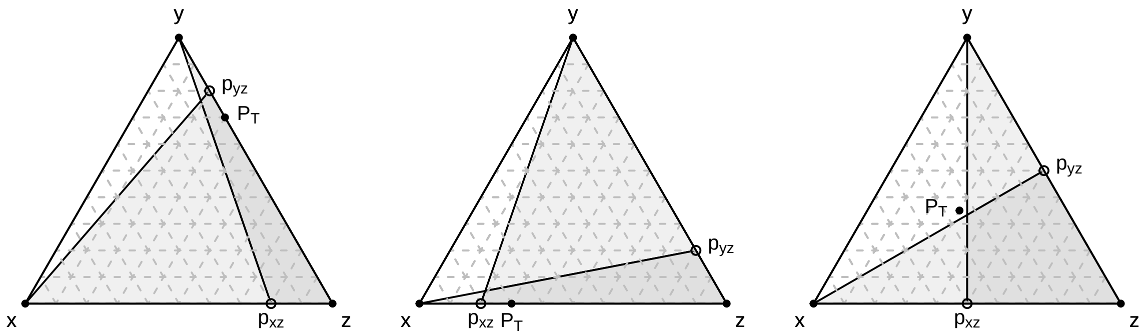

Figure 4

Two individuals satisfying a two-sided weak similarity effect aggregating to a population that fails to satisfy any two-sided weak similarity effect.

both strong 2-3 similarity effects occur, then (7) and (8) hold. Summing these two equations gives $P_{\{a,b,a'\}}(a) + P_{\{a,b,b'\}}(b) > 1$. Using the fact that choice probabilities on $\{a, b, b'\}$ sum to one, in order to eliminate $P_{\{a,b,b'\}}(b)$, gives $P_{\{a,b,a'\}}(a) > P_{\{a,b,b'\}}(a) + P_{\{a,b,b'\}}(b')$, which implies (5). Similarly, eliminating $P_{\{a,b,a'\}}(a)$ gives $P_{\{a,b,b'\}}(b) > P_{\{a,b,a'\}}(b) + P_{\{a,b,a'\}}(a')$, which implies (6).

Table 10 shows examples of all patterns of ternary-ternary (3-3) and binary-ternary (2-3) similarity effects that we have not just ruled out. Each row gives one example. In all examples, the binary choice probabilities are $p_{ab} = p_{ba} = \frac{1}{2}$. The first column indicates whether the 3-3 effect occurs or not. The second column gives the status of the two 2-3 effects; in each case, there is either no effect, a weak effect (understood here to mean only a weak effect and not a strong effect) and a strong effect. The third and fourth columns specify the probability vectors $P_{\{a,b,a'\}}(\cdot)$ for the various examples: the third column gives numerical values of the choice probabilities of a, b and a', respectively, as a triple; the fourth column gives the label identifying the point $P_{\{a,b,a'\}}(\cdot)$ in Barycentric coordinates in Figure 7, with the vertex on the right identified with the point a'. Similarly, the fourth and fifth columns specify $P_{\{a,b,b'\}}(\cdot)$. Here, the Barycentric coordinates are for $P_{\{a,b,b'\}}(\cdot)$ and the vertex on the right is identified with the point b'.

In addition to illustrating the values of $P_{\{a,b,a'\}}(\cdot)$ and $P_{\{a,b,b'\}}(\cdot)$ in the various

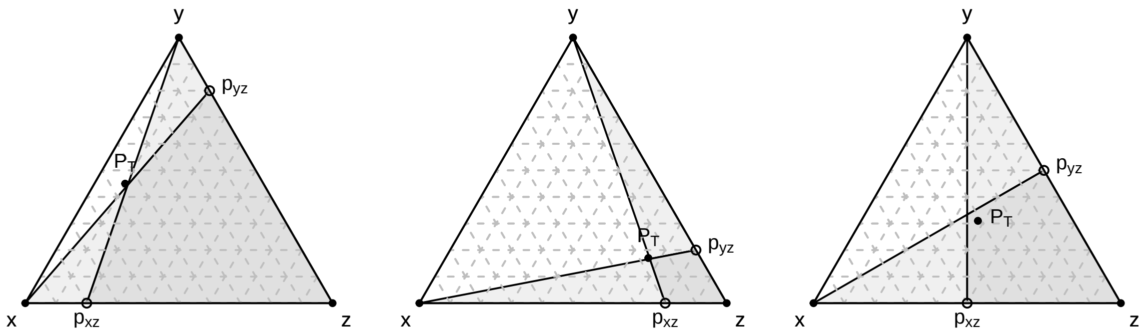

Figure 5

Two individuals failing to satisfy any weak similarity effect aggregating to a population satisfying a two-sided weak similarity effect.

examples, Figure 7 also illustrates cross-sections of the various binary-ternary effect regions, for the fixed binary probability $p_{ab} = \frac{1}{2}$. A strong effect occurs for the a' decoy when $P_{\{a,b,a'\}}(\cdot)$ is in the dark triangle containing S_a and S'_a; a strong effect occurs for the b' decoy when $P_{\{a,b,b'\}}(\cdot)$ is in the dark triangle containing S_b. A weak effect occurs for the a' decoy when $P_{\{a,b,a'\}}(\cdot)$ is in the light triangle containing W_a and W'_a; a weak effect occurs for the b' decoy when $P_{\{a,b,b'\}}(\cdot)$ is in the light triangle containing W_b.

Theorems related to the compromise effect

We have seen that random utility is compatible with the weak one- and two-sided compromise effects. However, Tversky and Simonson (1993) show that a random preference satisfying a theoretically compelling *ranking* condition is incompatible with the weak one- and two-sided compromise effects.

In this discussion, we take $T \equiv \{x, y, z\}$ as a fixed subset of objects in the master set U. The ranking condition is as follows:

Definition 9 (Tversky and Simonson, 1993) *The random preference model $\succ$ on U satisfies the ranking condition if $x|y|z$ implies*

$$\frac{\Pr[z \succ y \succ x]}{\Pr[z \succ x \succ y]} \geq \frac{P_T^{(\succ)}(y)}{P_T^{(\succ)}(x)} = \frac{\Pr[y \succ x \succ z] + \Pr[y \succ z \succ x]}{\Pr[x \succ y \succ z] + \Pr[x \succ z \succ y]}, \tag{11}$$

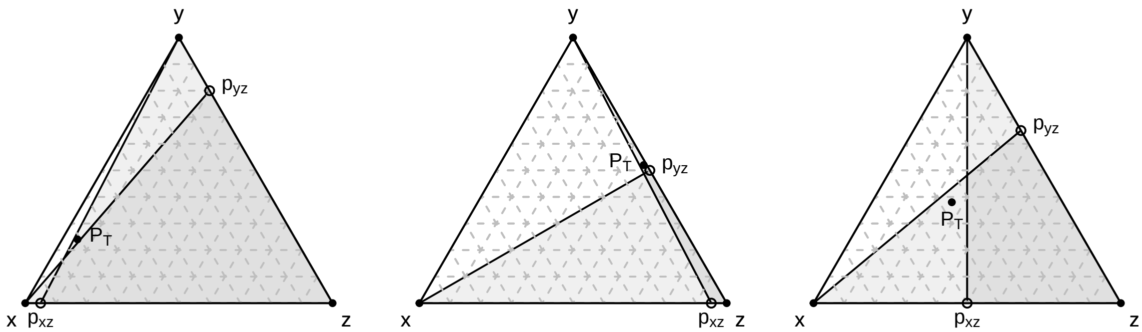

Figure 6

Two individuals satisfying the one-sided weak similarity effect S_x aggregating to a population satisfying the opposite weak similarity effect S_y

whenever the two denominators are both non-zero. The inequality in (11) is part of the definition of the ranking condition; the equality is true by the definition of $P_T^{(\succ)}(\cdot)$, whether or not the ranking condition holds.

The ranking condition is equivalent to the condition that when y is between x and z, the conditional probability of the event $y \succ x$ given that z is ranked first is at least as great as the conditional probability of $y \succ x$ given that z is not ranked first. To see this, apply the transformation $f(x) = x/(1+x)$ on $[0, \infty)$, which is strictly increasing and maps a/b to $a/(a+b)$, to the extreme left and right hand sides of (11). The transformed left-hand side is the first conditional probability and the transformed right-hand side is the second.

The plausibility of the ranking condition, for y between x and z, is clearest when y either is or vector-dominates a convex combination of x and z in attribute space. We conjecture that in this case, many readers would consider counter-intuitive any preference order where y is ranked last in $\{x, y, z\}$; no preference over the attribute space satisfying both monotonicity and strict convexity can rank y last in $\{x, y, z\}$.

Random preference and the ranking condition together imply the following between-ness inequality.

Definition 10 (Tversky and Simonson, 1993) *A RCS P on U satisfies the*

3-3	2-3	$P_{\{a,b,a'\}}(\cdot)$	label	$P_{\{a,b,b'\}}(\cdot)$	label
yes	2 strong	$(0.55, 0.10, 0.35)$	S_a	$(0.10, 0.55, 0.35)$	S_b
yes	1 strong, 1 weak	$(0.35, 0.30, 0.35)$	W_a	$(0.10, 0.55, 0.35)$	S_b
yes	1 strong	$(0.30, 0.35, 0.35)$	N_a	$(0.10, 0.55, 0.35)$	S_b
yes	2 weak	$(0.35, 0.30, 0.35)$	W_a	$(0.20, 0.45, 0.35)$	W_b
yes	1 weak	$(0.30, 0.35, 0.35)$	N_a	$(0.20, 0.45, 0.35)$	W_b
no	1 strong, 1 weak	$(0.35, 0.30, 0.35)$	W_a	$(0.40, 0.55, 0.05)$	S'_b
no	1 strong	$(0.30, 0.35, 0.35)$	N_a	$(0.40, 0.55, 0.05)$	S'_b
no	2 weak	$(0.35, 0.30, 0.35)$	W_a	$(0.15, 0.25, 0.60)$	W'_b
no	1 weak	$(0.30, 0.35, 0.35)$	N_a	$(0.15, 0.25, 0.60)$	W'_b
no	no effect	$(0.30, 0.35, 0.35)$	N_a	$(0.40, 0.25, 0.35)$	N_b

Table 10

Examples of all possible combinations of ternary-ternary (3-3) and binary-ternary (2-3) similarity effects. In all examples, $p_{ab} = p_{ba} = \frac{1}{2}$. The 3-3 effect is either present (yes in first column) or not (no). The two 2-3 effects can be both strong, one strong and one weak, one strong, two weak, one weak or both absent. Label refer to the labels of points in Figure 7 specifying $P_{\{a,b,a'\}}(\cdot)$ (column four) or $P_{\{a,b,a'\}}(\cdot)$ (column six).

between-ness inequality *if*

$$x|y|z \Rightarrow p_{yx} \geq \frac{P_T(y)}{P_T(y) + P_T(x)}. \tag{12}$$

Theorem 1 (Tversky and Simonson, 1993) *Let $\succ$ be a random preference on U. Then if the ranking condition holds, $P^{(\succ)}$ satisfies the between-ness inequality.*

Note that the set of RCSs satisfying (12) is the complement of the effect region of the two-sided weak compromise effect. Recall also the symmetry of the between-ness relation: $x|y|z$ implies $z|y|x$. Thus if $x|y|z$ and the random preference $\succ$ satisfies the ranking condition and then we have not only (12) but also, through $z|y|x$,

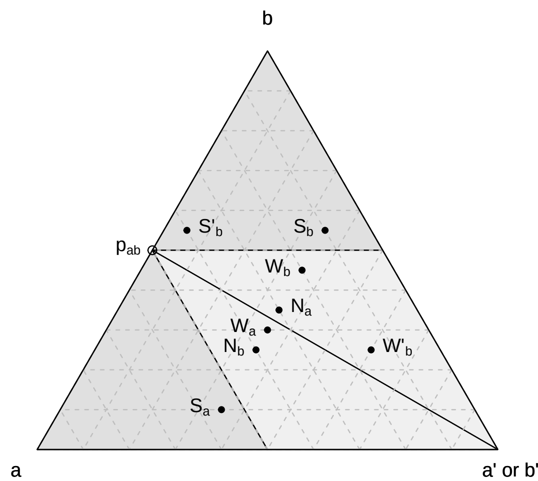

Figure 7

Points representing $P_{\{a,b,a'\}}(\cdot)$ or $P_{\{a,b,b'\}}(\cdot)$ in the various examples of Section Theorems relating binary-ternary and ternary-ternary context effects. The labels of points in this figure correspond to the labels in Table 10.

$p_{yz} \geq P_T(y)/(P_T(y) + P_T(z))$. Together, these two conditions are precisely the condition that there is no one- or two-sided weak compromise effect.

Conclusions

We presented a comprehensive 'field guide' for evaluating three common context effects: similarity, compromise, and asymmetric dominance. We presented precise definitions for these effects and how they relate to standard properties of choice, such as regularity. Our field guide is 'illustrated' in the sense that it offers graphical comparisons of how these context effects, and related choice properties, constrain choice probabilities. This offers a concrete way to think about parsimony when developing theories and empirical tests. We also summarized existing theorems on these context effects as well as offering a series of new theoretic results.

Our consideration of context effects focused exclusively on systems of choice probabilities. In other words, the probability of selecting one alternative from a set was treated as a primitive, without requiring the choice probabilities to be generated by a model of multi-alternative, multi-attribute choice. Future work could explore how the definitions and results we've provided interact with and/or constrain prominent models of multi-alternative, multi-attribute choice. Please see Turner et al. (2018) for a comprehensive analysis of competing models of multi-alternative, multi-attribute choice under context effects (also see Trueblood, 2022, and the references therein).

All of the context effects we've considered in the current work are defined via order constraints on systems of choice probabilities. Recent advances in order-constrained statistical analysis and related software now allow researchers to test these definitions directly, without requiring an explicit model of multi-alternative, multi-attribute choice. For example, Heck and Davis-Stober (2019) provide a general Bayesian approach for evaluating choice under general linear and ordinal constraints, which includes an R (Team et al., 2013) package (also see Regenwetter et al., 2014; Zwilling et al., 2019). We also direct the reader to Katsimpokis et al. (2022) for a novel Bayesian approach for testing context effects. Future work could consider extending the approach of Katsimpokis et al. to the additional formulations we consider here.

In synthesizing the disparities in the context effects literature, Spektor et al. (2021) drew a distinction between 'deliberative models' and 'representation models.' The idea being that the observed fragility of some context effects may likely depend upon how choice alternatives are cognitively represented, perhaps independently of the process in which the representations are evaluated. Our theoretic results, and those related to random utility in particular, may be useful in further exploring how deliberative formulations may or may not lead to context effects, depending upon the nature of the stimuli representation.

In the present work, we identified multiple discrepant context effect definitions, such as weak versus strong effects, 2-3 versus 3-3 comparisons, and the implications thereof. Future work could consider a broader meta-analysis of the empirical literature on context effects, examining which definitions were used, whether the analyses were individual or population-based, and the resulting consequences for prominent multi-alternative, multi-attribute models of choice.

References

Becker, G. M., DeGroot, M. H., & Marschak, J. (1963). Probabilities of choices among very similar objects. *Behavioral Science*, *8*, 306–311.

Berkowitsch, N. A., Scheibehenne, B., & Rieskamp, J. (2014). Rigorously testing multialternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, *143*(3), 1331.

Block, H. D., & Marschak, J. (1960). Random orderings and stochastic theories of responses. In I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, & H. B. Mann (Eds.), *Contributions to probability and statistics: Essays in honor of harold hotelling* (pp. 97–132). Stanford University Press.

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in cognitive sciences*, *23*(3), 251–263.

Cataldo, A. M., & Cohen, A. L. (2018). Reversing the similarity effect: The effect of presentation format. *Cognition*, *175*, 141–156.

Cataldo, A. M., & Cohen, A. L. (2021). Modeling preference reversals in context effects over time. *Computational Brain & Behavior*, *4*(1), 101–123.

Cavagnaro, D. R., & Davis-Stober, C. P. (2014). Transitive in our preferences, but transitive in different ways: An analysis of choice variability. *Decision*, *1*, 102–122.

Chipman, J. S. (1960). Stochastic choice and subjective probability. In D. Willner (Ed.), *Decisions, values, and groups*. Pergamon Press.

Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport and Infrastructure Research*, *10*(2), 181–196.

Coombs, C. H. (1958). On the use of inconsistency of preferences in psychological measurement. *Journal of Experimental Psychology*, *55*, 1–7.

Davis-Stober, C. P., Brown, N., Park, S., & Regenwetter, M. (2017). Recasting a biologically motivated computational model within a fechnerian and random utility framework. *Journal of mathematical psychology*, *77*, 156–164.

Debreu, G. (1960). Review of R. D. Luce, Individual choice behavior: A theoretical analysis. *American Economic Review*, *50*, 186–188.

Doyle, J. R., O'Connor, D. J., Reynolds, G. M., & Bottomley, P. A. (1999). The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases. *Psychology & Marketing*, *16*(3), 225–243.

Falmagne, J. C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, *18*, 52–72.

Fishburn, P. C. (1999). Stochastic utility. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of utility theory: Volume 1, principles* (pp. 273–319). Kluwer Academic Publishers.

Gluth, S., Hotaling, J. M., & Rieskamp, J. (2017). The attraction effect modulates reward prediction errors and intertemporal choices. *Journal of Neuroscience*, *37*(2), 371–382.

Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*, *4*(6), 634–645.

Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision making. *Elife*, *7*, e39659.

Heck, D. W., & Davis-Stober, C. P. (2019). Multinomial models with linear inequality constraints: Overview and improvements of computational methods for bayesian inference. *Journal of mathematical psychology*, *91*, 70–87.

Herne, K. (1997). Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy*, *13*(3), 575–589.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*, 90–98.

Huber, J., Payne, J. W., & Puto, C. P. (2014). Let's be honest about the attraction effect. *Journal of Marketing Research*, *51*.

Katsimpokis, D., Fontanesi, L., & Rieskamp, J. (2022). A robust bayesian test for identifying context effects in multiattribute decision-making. *Psychonomic Bulletin & Review*, 1–18.

Krantz, D. H. (1964). *The scaling of small and large color differences.* (Doctoral dissertation). University of Pennsylvania.

Krantz, D. H. (1967). Rational distance function for multidimensional scaling. *Journal of Mathematical Psychology*, *4*, 226–245.

Latty, T., & Trueblood, J. S. (2020). How do insects choose flowers? a review of multi-attribute flower choice and decoy effects in flower-visiting insects. *Journal of Animal Ecology*, *89*(12), 2750–2762.

Liew, S. X., Howe, P. D., & Little, D. R. (2016). The appropriacy of averaging in the study of context effects. *Psychonomic bulletin & review*, *23*(5), 1639–1646.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* John Wiley & Sons.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, *15*, 215–233.

Luce, R. D., & Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 249–410). John Wiley & Sons.

Marley, A. A. J., & Regenwetter, M. (2018). Choice, preference, and utility: Probabilistic and deterministic representations. In W. Batchelder, H. Colonius, E. Dzhafarov, & J. Myung (Eds.), *New handbook of mathematical psychology, volume 1: Measurement and methodology.* Cambridge University Press.

McCausland, W. J., Davis-Stober, C., Marley, A. A., Park, S., & Brown, N. (2020). Testing the random utility hypothesis directly. *The Economic Journal*, *130*(625), 183–207.

McCausland, W. J., & Marley, A. A. J. (2013). Prior distributions for random choice structures. *Journal of Mathematical Psychology*, *57*, 78–93.

McCausland, W. J., & Marley, A. A. J. (2014). Bayesian inference and model comparison for random choice structures. *Journal of Mathematical Psychology*, *62-63*, 33–46.

McCausland, W. J., Marley, A. A. J., & Davis-Stober, C. (2021). *Testing axioms of stochastic discrete choice using population choice probabilities* [manuscript].

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*, 303–328.

Molloy, M. F., Galdo, M., Bahg, G., Liu, Q., & Turner, B. M. (2019). What's in a response time?: On the importance of response time measures in constraining models of context effects. *Decision*, *6*(2), 171.

Parrish, A. E., Evans, T. A., & Beran, M. J. (2015). Rhesus macaques (macaca mulatta) exhibit the decoy effect in a perceptual discrimination task. *Attention, Perception, & Psychophysics*, *77*(5), 1715–1725.

Ray, P. (1973). Independence of irrelevant alternatives. *Econometrica*, *41*.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, *118*, 42–56.

Regenwetter, M., Davis-Stober, C. P., Lim, S. H., Guo, Y., Popova, A., Zwilling, C., Cha, Y.-S., & Messner, W. (2014). Qtest: Quantitative testing of theories of binary choice. *Decision*, *1*(1), 2.

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, *44*, 631–661.

Scarpi, D. (2011). The impact of phantom decoys on choices in cats. *Animal Cognition*, *14*(1), 127–136.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*.

Spektor, M. S., Bhatia, S., & Gluth, S. (2021). The elusiveness of context effects in decision making. *Trends in Cognitive Sciences*, *25*(10), 843–854.

Spektor, M. S., Kellen, D., & Hotaling, J. M. (2018). When the good looks bad: An experimental exploration of the repulsion effect. *Psychological science*, *29*(8), 1309–1320.

Team, R. C. et al. (2013). R: A language and environment for statistical computing.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2013). *The multiattribute linear ballistic accumulator model* [Manuscript, Department of Cognitive Sciences, University of California Irvine].

Trueblood, J. S. (2022). Theories of context effects in multialternative, multiattribute choice. *Current Directions in Psychological Science*, *31*(5), 428–435.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological review*, *121*(2), 179.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological science*, *24*(6), 901–908.

Turner, B. M., Schley, D. R., Muller, C., & Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological review*, *125*(3), 329.

Tversky, A. (1972a). Choice by elimination. *Journal of Mathematical Psychology*, *9*, 341–367.

Tversky, A., & Russo, J. E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology*, *6*, 1–12.

Tversky, A. (1972b). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*, 1179–1189.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learn- ing, Memory, and Cognition*, *17*, 767–778.

Zwilling, C. E., Cavagnaro, D. R., Regenwetter, M., Lim, S. H., Fields, B., & Zhang, Y. (2019). Qtest 2.1: Quantitative testing of theories of binary choice using bayesian inference. *Journal of Mathematical Psychology*, *91*, 176–194.