A New Foundation for Support Theory

Louis Narens*
Department of Cognitive Sciences
University of California, Irvine
email: lnarens@uci.edu

January 27, 2004

Abstract

A new foundation is presented for the theory of subjective judgments of probability generally known as "Support Theory." It is based on a new complementation operation that, unlike those of classical probability theory (set-theoretic complementation) and classical logic (negation), does not satisfy the principle of The Law Of The Excluded Middle. Interrelationships between the new complementation operation and the Kahneman and Tversky judgmental heuristics of availability and representativeness are described, and alternative explanations are provided for experimental studies involving subjective judgments of probability.

Subjective evaluations of degrees of belief are essential in human decision making. Numerous experimental studies have been conducted eliciting numerical judgments of probability, and many interesting phenomena have been uncovered. A cognitive theory was proposed by Amos Tversky and colleagues to explain some of the more prominent regularities revealed in these studies. This theory, known today as *Support Theory,* has its foundational basis in the articles of Tversky and Koehler (1994) and Rottenstreich and Tversky (1997), and incorporates Kahneman's and Tversky's seminal research on judgmental heuristics (e.g., Tversky and Kahneman, 1974).

Support Theory has an empirical base of results showing that different descriptions of the same event often produce different subjective probability estimates. It explains these results in terms of subjective evaluations of supporting evidence. It

*I would like to thank R. Duncan Luce for helpful comments on earlier drafts of this article.

assumes that events are evaluated in terms of subjective evidence invoked by their descriptions, and that the observed numerical probability judgments are the result of the combining of such evaluations of support in a manner that is consistent with a particular equation. The processes of evaluation are assumed to employ heuristics like those of Kahneman and Tversky, and because of this, are subject to the kinds of biases introduced by such heuristics.

This article provides a New Foundation for Support Theory. The New Foundation makes a sharp distinction between semantical representations of descriptions as part of natural language processing and cognitive representations of descriptions as part of a probabilistic judgment. In particular, judgments of probability employ a complementation operation that has no counterpart in the semantics. The complementation operation is used to construct cognitive events that are employed in the computation of the estimated probability. It has structural (= mathematical) features that differ significantly from the kind of complementation operations considered by cognitive psychologists in general, and support theorists in particular. That is, it has structural features different from the complementation operation used in the algebra of events (set theoretic complementation) or in classical logic (negation). The most important impact of the difference is that the Law of the Excluded Middle fails; that is, for event space, the union of some event A and its complement need not be the sure event, and for propositional logic, the disjunction of some proposition p and its negation need not be a tautology. The "middle" allowed by the new complementation operation is exploited by the New Foundation in a variety of ways, including modeling the impact of ambiguous or bad exemplars.

The new complementation operation has the formal mathematical properties of a negation operation employed in a non-classical logic called the *Intuitionistic Propositional Calculus.* This logic was invented by the mathematician Brouwer for his alternative foundation of mathematics, in which mathematical objects are considered to be constructions of the human mind. It was formalized by Heyting (1930), and since then has been a much studied subject by logicians. It has a variety of applications, including ones in artificial intelligence. Its relationship to the notion of "mathematical construction" (Kolmogorov, 1934) makes it a natural candidate for describing structural properties of "mental processing."

This article creates an event space corresponding to Intuitionistic Propositional Logic through the use basic topological concepts (e.g., open set, closed set, boundary). Open sets in the topology are interpreted as events consisting of cognitively clear exemplars, whereas the boundaries of open sets consisting of cognitive exemplars that are bad, ambiguous, or unrealized instances of the event under consideration.

In the New Foundation, a description α of a probabilistic event is assumed to give rise to a cognitive representation A. The probabilistic estimate for α is the result of weighing the evidence in favor of A against the evidence against A. It is assumed that the participant measures the evidence in favor of A by a positive real number, called *the support for* A, in symbols $S^+(A)$, and similarly the evidence against A by a positive real number, called *the support against* A, in symbols $S^-(A)$. The New Foundation assumes that the participant constructs a "complement" of A, $\ominus A$, that is a generalization of the complement operation used in set theory, and $S^-(A)$ is computed as the support for $\ominus A$, i.e., as $S^+(\ominus A)$. It should be noted that the generalized complement $\ominus A$ is created as part of the processing involved in making the probability judgment, and as such, need not correspond to a description.

The New Foundation presents ways that the generalized complement interacts with the heuristics of availability and representativeness. Representativeness is given a new formulation based on sets of properties. (The Kahneman and Tversky formulation is based on prototypes.) Features of the generalized complement are then used to explain the empirical results observed in Support Theory as well as some new empirical results that challenge Support Theory's basic tenets. The New Foundation provides alternative accounts for both kinds of results.

Section 1 presents a brief outline of Support Theory. First, the judgmental heuristics of Kahneman and Tversky are defined and illustrated. These heuristics are part of the foundational basis of Support Theory and appear in many explanations as phenomena generated by judgments of probability. Then a short description is given of the theoretical foundation for Support Theory, as described in Tversky and Koehler (1994) and Rottenstreich and Tversky (1997). Later sections of the article provide empirical counter-examples to some of basic tenets of this theoretical foundation. However, independent of these empirical counter-examples, other weaknesses of Support Theory's foundation are discussed in later sections of the article.

Section 2 presents the New Foundation and several empirical violations due to other researchers of some of its major tenets. The New Foundation's explanation for the observed phenomena often differ from these researchers, who often try to generalize Support Theory in manner that incorporates their phenomena. The New Foundation is not an extension or generalization of Support Theory' Foundation, but is a different approach to Support Theory's phenomena, based on an entirely different modeling of subjective interpretations of descriptions of events.

Section 3 is a discussion. Part of it consists of a critical evaluation of some of Support Theory's concepts, particularly its foundational use of its concepts of "extensional" and "nonextensional," and another part discusses relationship of the New Foundation to related theories in the literature.

Section 4 provides an algebraic characterization of the New Foundation's event space in terms of lattice theory. (Classical) probabilistic event space and classical propositional calculus have a common characterization as a particular kind of lattice, known as a *boolean lattice.* The complementation operation of a boolean lattice corresponds to set-theoretic complementation in probabilistic event space, and to negation in classical propositional calculus. A property of boolean complementation is deleted from the definition of boolean lattice, yielding a more general form of complementation. Lattices with this generalized form of complementation, when interpreted as a calculus of propositions, produces a logic that is similar to (but not identical with) Intuitionistic Propositional Logic. When interpreted as a lattice of events, the generalized complementation permits structure to be added to the event space. It is shown that the additional structure can be interpreted topologically with elements of the lattice corresponding to open sets in some topology. This allows the event space of New Foundation to be interpreted as the event space corresponding to intuitionistic logic. The lattice formulation provides for a clear description of how the event spaces of (classical) probability theory, the New Foundation, and quantum mechanics are related. It also provides the intuitive basis for the author's contention that the non-classical event space of the New Foundation is the best kind of event space that one is likely to find for the modeling of judgments of probability.

1 Support Theory

1.1 Psychological Background: Judgmental Heuristics

The heuristics of availability, representativeness, adjustment and anchoring, as described in Tversky and Kahneman (1974), play a major role in this article in explaining empirical phenomena and in developing theory.

Tversky and Kahneman introduce *availability* as follows:

> The are situations in which people assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind. For example, one may assess the risk of heart attack among middle-aged people by recalling occurrences among one's acquaintances. Similarly, one may evaluate the probability that a given business venture will fail by imagining various difficulties it could encounter. This judgmental heuristic is called availability. Availability is a useful clue for assessing frequency or probability, because instances of large classes are usually recalled better and faster than instances of less

> frequent classes. However, availability is affected by factors other than frequency and probability. *(pg. 1127)*

They illustrate and described *representativeness* as follows:

> For an illustration of judgment by representativeness, consider an individual who has been described by a former neighbor as follows: "Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail." How do people assess the probability that Steve is engaged in a particular occupation from a list of possibilities (for example, farmer, salesman, airline pilot, librarian, or physician)? How do people order these occupations from most to least likely? In the representativeness heuristic, the probability that Steve is a librarian, for example, is assessed by the degree to which he is representative of, or similar to, the stereotype of a librarian. Indeed, research with problems of this type has shown that people order the occupations in exactly the same way (Kahneman and Tversky, 1973). This approach to judgment of probability leads to serious errors, because similarity, or representativeness, is not influenced by several factors that should affect judgments of probability. *(pg. 1124)*

Tversky and Kahneman describe adjustment and anchoring as follows:

> In many situations, people make estimates by starting from an initial value that is adjusted to yield the final answer. The initial value, or starting point, may be suggested by the formation of the problem, or it may be the result of a partial computation. In either case, adjustments are typically insufficient. That is, different starting points yield different estimates, which are biased toward the initial value. *(pg. 1128)*

In the New Foundation presented in this article, availability is taken as fundamental. When representativeness is employed as a basis for the probabilistic evaluation of items, the availability of the properties employed in determining an item's representativeness is used for judging its probability. In general, a judged probability of an item is theorized to be determined by the availability of instances of the item or the availability of features of the item. Whether it is instances or features that are acted upon will, in general, vary with the type of item and task. Adjustment and anchoring play only a minor role in the New Foundation.

1.2 Empirical Basis of Support Theory

"Support Theory" is an approach to understanding and modeling human judgments of probability that was initiated in Tversky and Koehler (1994), expanded upon by Rottenstreich and Tversky (1997), and extended by a number of investigators. Support Theory was originally designed to explain the following two empirical results, which together appear to be puzzling:

1. *Binary complementarity:* For a binary partition of an event, the sum of the judged probabilities of elements of the partition is 1.

2. *Subadditivity:* for partitions consisting of three or more elements, the sum of the judged probabilities of the elements is ≥ 1, with > 1 often being observed.

The empirical basis for Support Theory was founded on numerous experimental results showing binary complementarity for partitions consisting of two elements, and subadditivity for partitions consisting of three or more elements. The following example (Redelmeier, Koehler, Liberman, and Tversky, 1995) shows such subadditivity in a between-subjects design. (Other studies show similar effects for within-subject designs.)

Redelmeier et al. presented the the following scenario to a group of 52 expert physicians a Tel Aviv University.

> B. G. is a 67-year-old retired farmer who presents to the emergency department with chest pain of four hours' duration. The diagnosis is acute myocardial infarction. Physical examination shows no evidence of pulmonar edema, hypothension, or mental status changes. His EKG shows ST-ssegment elevation in the anterior leads, but no dysrythmia or heart block. His past medical history is unremarkable. He is admitted to the hospital and treated in the usual manner. Consider the possible outcomes.

Each physician was randomly assigned to evaluate on the the following four prognoses for this patient:

- dying during this admission
- surviving this admission but dying within one year
- living for more than one year but less than ten years
- surviving for more than ten years.

Redelmeier, et al. write,

> The average probabilities probabilities assigned to to these prognoses were 14%, 26%, 55%, and 69%, repsectively. According to standard theory, the probabilities assigned to these outcomes should sum to 100%. In contrast, the average judgments added to 164% (95% confidence interval: 134% to 194%).

What is striking is that the participants in this example are all expert and well experienced, judging a situation like ones they encounter routinely and have to make weighty decisions about.

Another example is an experiment of Fox, Rogers, and Tversky (1996). They asked professional option traders to judge the probability that the closing price of Microsoft stock would fall within a particular interval on a specific future date. They found subadditivity: E.g., when four disjoint intervals that spanned the set of possible prices were presented for evaluation, the sums of the assigned probabilities were typically about 1.50. And they also found binary complementarity: When binary partitions were presented the sums of the assigned probabilities were very close to 1, e.g., .98.

Other examples include Tversky and Fox (2000) who examined probability judgments involving future temperature in San Francisco, the point-spread of selected NBA and NFL professional sports games, and many other quantities. They found sums of assigned probabilities greater than 1 for partitions consisting of more than two elements and sums nearly equal 1 for binary partitions. Other researchers, e.g., Wallsten, Budescu, and Zwick (1992) have also observed Binary Complementarity in experimental settings.

1.2.1 Support Theory's Foundation

The theoretical foundation for Support Theory is given in Tversky and Koehler (1994) and Rottenstreich and Tversky (1997). The basic units in Support Theory are descriptions of events (often called "hypotheses" by support theorists). Descriptions are presented to participants for probabilistic evaluation. It is assume that participants evaluate the descriptions in terms of a "support function," s, which is a ratio scale into the positive reals. Support Theory assumes that the value of $s(\alpha)$ for a description α is determined by a participant through the use of judgmental heuristics. Most Support Theory models employ judged (conditional) probability of descriptions of the form, "α occurring rather than γ occurring," in symbols, $P(\alpha, \gamma)$, where the logical conjunction of α and γ is understood by the participant to be impossible. It is additionally assumed that $P(\alpha, \gamma)$ is determined by the equation,

$$P(\alpha, \gamma) = \frac{s(\alpha)}{s(\alpha) + s(\gamma)} \,.$$

(A notable exception to this is Idson, Krantz, Osherson, and Bonini, 2001, which presents an extension of Support Theory that employs a different equation.)

Support Theory makes a distinction between "implicit" and "explicit" disjunctions. A description is said to be *null* if and only if it describes the null event. Descriptions of the form "α or γ," where α and γ are nonnull and the description "α and γ" is null, are called *explicit disjunctions.* Explicit disjunctions are often written as $\alpha \vee \gamma$, with the understanding that the description "α and γ" is null. A description is called *implicit* (or an *implicit disjunction*) if it is nonnull and and not an explicit disjunction. An explicit disjunction δ and an implicit disjunction γ may describe the same event—that is, have the same *extension,* in symbols, $\delta' = \gamma'$—but have different support assign to them. Tversky and Koehler (1994) provides the following illustration:

> For example, suppose A is "Ann majors in a natural science," B is "Ann majors in biological science," and C is "Ann majors in a physical science." The explicit disjunction, $B \vee C$ ("Ann majors in in either a biological or physical science"), has the same extension as A (i.e., $A' = (B \vee C)' = (B' \cup C')$), but A is an implicit disjunction because it is not an explicit disjunction. *(*pg. 548)

In their generalization of the Support Theory of Tversky and Koehler (1994), Rottenstreich and Tversky (1997) distinguishes two ways in which support and explicit disjunction relate. Suppose α is implicit, $\delta \vee \gamma$ is explicit, and α and $\delta \vee \gamma$ describe the same event. Rottenstreich and Tversky assume the following two conditions linking support to implicit descriptions and explicit disjunctions:

(1) *implicit subadditivity:* $s(\alpha) \leq s(\delta \vee \gamma)$.

(2) *explicit subadditivity:* $s(\delta \vee \gamma) \leq s(\delta) + s(\gamma)$.

A direct consequence of (1) and (2) is

(3) $s(\alpha) \leq s(\delta) + s(\gamma)$.

Instead of (2), Tversky and Koehler assumed *additivity,* $s(\delta \vee \gamma) = s(\delta) + s(\gamma)$, which along with (1) yields (3). Rottenstreich and Tversky observed cases where additivity failed but explicit subadditivity (2) held.

Using the above notation, $\delta \vee \gamma$ is said to be an *unpacking* of α. Of course, α may have many unpackings. The following has been much observed in the Support Theory literature.

Subadditive unpacking: A partition $\mathcal{P}_1 = (\kappa_1, \ldots, \alpha, \ldots, \kappa_n)$ with $n \geq 2$ elements is judged to have probability p_1, and when α is replaced by its unpacking $\delta \vee \gamma$ to yield the partition $\mathcal{P}_2 = (\kappa_1, \ldots, \delta \vee \gamma, \ldots, \kappa_n)$ and a judged probability p_2 for $\mathcal{P}_2$, then $p_1 < p_2$.

In Support Theory there are several different cognitive processes that determine the support of a description:

> The support associated with a given [description] is interpreted as a measure of the strength of evidence in favor of this [description] to the judge. The support may be based on objective data (e.g., frequency of homicide in the relevant data) or on a subjective impression mediated by judgmental heuristics, such as representativeness, availability, or anchoring and adjustment (Kahneman, Solvic, and Tversky, 1982). For example, the hypothesis "Bill is an accountant" may be evaluated by the degree to which Bill's personality matches the stereotype of an accountant, and the prediction "An oil spill along the eastern coast before the end of next year" be be assessed by the ease with which similar accidents come to mind. Support may also reflect reasons or arguments recruited by the judge in favor of the hypothesis in question (e.g., if the defendant were guilty, he would not have reported the crime). *(Tversky and Koehler, 1994, pg. 549)*

How particular heuristically based processes differentially affect probability judgments has been the focus of much recent research by support theorists.

In Support Theory, participants are presented with a description β that establishes the context for the probabilistic judgment of the description α. α is called the *focal* description. Support Theory studies are almost always designed so that the context β contains a description γ, called the *alternative* description, such that it is clear to the participant that β implies that the intersection of the extensions of α and γ is null and that either α or γ must occur. Throughout this article, this situation is described as, *the participant is asked to judge* $\alpha|\beta$*, "the probability of* α *given* β*."* In addition, in within-subject designs the participant is asked to judge $\gamma|\beta$, and in between-subjects designs other participants are asked to judge $\gamma|\beta$.

For a binary partition (α, γ), both $\alpha|\beta$ and $\gamma|\beta$ are presented. For a ternary partition (α, δ, γ), all three of $\alpha|\beta$, $\delta|\beta$, and $\gamma|\beta$ are presented, where, of course,

$$(\text{the extension of } \alpha, \text{ the extension of } \delta, \text{ the extension of } \gamma)$$

is a partition of the extension of β. (Similarly for n-ary partitions for $n > 3$).

2 New Foundation

2.1 Notation

Throughout this article, $\cup$, $\cap$, $-$, and $\in$ stand for, respectively, set-theoretic intersection, union, complementation, and membership. $\subseteq$ is the subset relation, and $\subset$ is the proper subset relation. $\notin$ stands for "is not a member of" and $\not\subseteq$ for "is not a subset of." For nonempty sets $\mathcal{E}$, $\bigcup\mathcal{E}$ and $\bigcap\mathcal{E}$ have the following definitions:

$$\bigcup\mathcal{E} = \{x|x \in E \text{ for some } E \text{ in } \mathcal{E}\} \text{ and } \bigcap\mathcal{E} = \{x|x \in E \text{ for all } E \text{ in } \mathcal{E}\}\,.$$

"iff" stands for "if and only if." □

2.2 Basic Concepts

This article presents a different foundation for the empirical phenomena of Support Theory. Three kinds of structures are assumed: A set of descriptions, a natural language semantics for the descriptions, and a structure of cognitive representations for interpreting descriptions for judgments of probability.

2.2.1 Semantical Representations

The descriptions are natural language formulations, say English, of individuals, categories, and propositions. Only descriptions of propositions are presented to the participant for probabilistic evaluation. *The natural language semantics* provides an abstract representation that associates with each propositional description α, a *semantical representation* of α, $\mathbf{SR}(\alpha)$, which is assumed to be an event. As is customary in probability theory, events are considered to be sets which forms a dl:

Definition 1 $\langle\mathfrak{X} = \langle\mathcal{X}, \subseteq, \cup, \cap, -, X, \varnothing\rangle$ is said to be a *boolean algebra of sets* if and only if X is a nonempty set and $\mathcal{X}$ is a set of subsets of X such that the following two conditions hold:

(i) X and $\varnothing$ are in $\mathcal{X}$, and

(ii) for all E and F in $\mathcal{X}$, $E \cup F$, $E \cap F$, and $-E$ are in $\mathcal{X}$.

In probabilistic situations, boolean algebras of sets are often called *boolean algebras of events,* and the sets in the domain of a boolean algebra of sets are often called *events.* □

(Theorem 6 of Section 4 provides a partial justification for the common practice in science and mathematics of representing propositions as events in a boolean algebra of events .)

For our purposes only a bare minimum about the semantical representations needs to be known. It is assumed that **SR** is a function from propositional descriptions into the set of subsets of some nonempty set, and that for propositional descriptions α and γ,

- $\mathbf{SR}(\alpha \text{ and } \gamma) \;=\; \mathbf{SR}(\alpha) \cap \mathbf{SR}(\gamma)$,
- $\mathbf{SR}(\alpha \text{ or } \gamma) \;=\; \mathbf{SR}(\alpha) \cup \mathbf{SR}(\gamma)$, and
- $\mathbf{SR}(\text{not } \alpha) \;=\; -\mathbf{SR}(\alpha)$.

The semantics provide a means for classifying and relating propositional descriptions in terms of their semantical representation. Let α, δ, γ, and θ be arbitrary descriptions. Then,

- α and δ are *disjoint* if and only if $\mathbf{SR}(\alpha) \cap \mathbf{SR}(\delta) = \varnothing$;
- α is *null* if and only if $\mathbf{SR}(\alpha) = \varnothing$;
- $\gamma = \alpha \vee \delta$ is an *explicit disjunction* if and only if α and δ are disjoint and nonnull and γ is the description "α or δ".
- $\alpha \vee \delta$ is said to be an *unpacking of* θ if and only if $\alpha \vee \delta$ is an explicit disjunction and $\mathbf{SR}(\alpha \vee \delta) = \mathbf{SR}(\theta)$.

Within the New Foundation, $\mathbf{SR}(\alpha)$ corresponds to Support Theory's concept of "extension of α."

2.2.2 Cognitive Representations

The natural language semantics plays a relatively minor role in this article. The above properties are mostly what is needed. A different form of "representation" plays the major role. Descriptions of propositions are interpreted again as events, but events that are part of a more structured environment consisting of open sets from a topology.

Definition 2 A collection $\mathcal{U}$ is said to be a *topology with universe* X if and only if X is a nonempty set, $X \in \mathcal{U}$, $\varnothing \in \mathcal{U}$, for all A and B in $\mathcal{U}$, $A \cap B$ is in $\mathcal{U}$, and for all nonempty $\mathcal{W}$ such that $\mathcal{W} \subseteq \mathcal{U}$,

$$\bigcup \mathcal{W} \text{ is in } \mathcal{U}\,. \tag{1}$$

Note that it is immediate from Equation 1 that for all A and B in $\mathcal{U}$, $A \cup B$ is in $\mathcal{U}$.

Let E be an arbitrary subset of X and $\mathcal{U}$ be a topology with universe X. Then the following definitions hold:

- E is said to *open (in the topology* $\mathcal{U}$*)* if and only if $E \in \mathcal{U}$.
- E is said to be *closed (in the topology* $\mathcal{U}$*)* if and only if $X - E$ is open.

It immediately follows that X and $\varnothing$ are closed as well as open. In some cases $\mathcal{U}$ may have X and $\varnothing$ as the only elements that are both open and closed, while in other cases $\mathcal{U}$ may have additional elements that are both open and closed, or even have all of its elements both open and closed. The following definitions hold for all $E \subseteq X$:

- The *closure* of E, $\mathsf{cl}(E)$, is, the smallest closed set C such that $E \subseteq C$; that is,
$$\mathsf{cl}(E) = \bigcap \{B | B \text{ is closed and } E \subseteq B\}\,.$$
- The *boundary* of E, $\mathsf{bd}(E)$, is $\mathsf{cl}(E) - E$.
- The *interior* of E, $\mathsf{int}(E)$, is the largest open set D such that $D \subseteq E$; that is,
$$\mathsf{int}(E) = \bigcup \{F | F \text{ is open and } F \subseteq E\}\,.$$

It easily follows that the definition of "topology" implies the existence of the closure, interior, and boundary of E for all $E \subseteq X$, and that for all closed F, $\mathsf{cl}(F) = F$. □

A boolean algebra of events can be viewed as a structure of open sets from a topology where each open set is also a closed set:

Theorem 1 *Suppose* $\mathfrak{X} = \langle \mathcal{X}, \cup, \cap, {}^-, X, \varnothing \rangle$ *is a boolean algebra of events. Let* $\mathcal{U}$ = *the set of all subsets of* X*. Then* $\mathcal{U}$ *is a topology in which each element of* $\mathcal{U}$ *is both open and closed. Thus in particular each element of* $\mathcal{X}$ *is both open and closed with respect to the topology* $\mathcal{U}$*.*

Proof. Let A be an arbitrary element of $\mathcal{U}$. Then $-A$ is closed. However, because $\mathsf{cl}(-A) = -A \in \mathcal{U}$, $-A$ is also open. □

The import of Theorem 1 for the New Foundation is that a boolean algebra of events is a highly restrictive and special kind of representation space for cognitive representations. For the issues presented by Support Theory, it is too restrictive.

Instead of a boolean algebra of events, the New Foundation is based on an algebra of events with a more general form of complementation.

Definition 3 Let $\mathcal{U}$ be a topology and $A \in \mathcal{U}$. Then, by definition, the $\cap$-*complement of A (with respect to $\mathcal{U}$), in symbols, $\ominus A$,* is $\mathsf{int}(-A)$.

That $\ominus A$ is a more general form of complementation than the complementation operator in an algebra of events follows from later theorems. (In particular, see Theorem 4 in Subsection 4.1.)

Let $\mathcal{U}$ be a topology and $A \in \mathcal{U}$. Note that by Definition 3, $\ominus A$ is the largest (in terms of the $\subseteq$ relation) element of $\mathcal{U}$ such that $A \cap \ominus A = \varnothing$. This justifies the use of the term "$\cap$-complement."

Definition 4 $\langle \mathcal{X}, \subseteq, \cup, \cap, X, \varnothing \rangle$ is said to be a *topological algebra of events* if and only if for some topology $\mathcal{U}$,

(1) X is the universe of $\mathcal{U}$ and $\ominus$ is the operation of $\cap$-complementation with respect to $\mathcal{U}$;

(2) $\mathcal{X} \subseteq \mathcal{U}$;

(3) X and $\varnothing$ are in $\mathcal{X}$; and

(4) For all E and F in $\mathcal{X}$, $E \cup F$, $E \cap F$, and $\ominus E$ are in $\mathcal{X}$.

When $\alpha|\beta$ is presented to the participant, she forms *cognitive representations* $\mathbf{CR}(\beta)$ corresponding to β and $\mathbf{CR}(\alpha)$ corresponding α that she employs as the bases for making a probabilistic judgment of $\alpha|\beta$. These representations are modeled as open sets within a topology $\mathcal{U}$. Unless otherwise specified, the universal set of this topology is taken to be the cognitive representation B corresponding to the conditioning event β. More generally, all *descriptions* used in the computation of the probability of $\alpha|\beta$ are represented cognitively as open subsets of B. Of course, in making the computation, the participant may need to create additional cognitive representations.

In the New Foundation, the relationship between the semantic and cognitive representations of α is very minimal: It is assumed that

$$\mathbf{CR}(\alpha) = \mathbf{SR}(\alpha) \text{ iff } \mathbf{SR}(\alpha) = \varnothing\,,$$

that is, only the empty set is common to both the semantic and cognitive representations. This assumption is possible because of the relatively minor role the semantics of descriptions plays in the modeling of the kinds of judgments presented in this article. Of course, other kinds of judgments may require additional structure between between semantic and cognitive representations.

For the kinds of judgments presented in this article, **SR** and **CR** are so unrelated that one can have for particular descriptions α and γ,

$$\mathbf{SR}(\alpha) \subset \mathbf{SR}(\gamma) \text{ and } \mathbf{CR}(\gamma) \subset \mathbf{CR}(\alpha)\,.$$

The following is a formulation within the New Foundation of Support Theory's *use* of "non-extensionality:" Descriptions α and δ are said to exhibit *non-extensionality* if and only if

$$\mathbf{SR}(\alpha) = \mathbf{SR}(\delta) \;\text{ iff }\; \mathbf{CR}(\alpha) \neq \mathbf{CR}(\delta)\,.$$

(It should be noted that this version of "non-extensionality" is based on a distinction between semantic and cognitive representations, whereas Support Theory's version of "non-extensionality" is based on a completely different kind of distinction. See Section 3 for a discussion of this issue.)

For some specified kinds of judgments, additional relationships between the semantic and cognitive representations are assumed. The following is an example of such a relationship that is used later in frequency based judgments:

Definition 5 *Cognitively preserve semantical disjointness* is said to hold for a judgment of probability if and only if for all descriptions σ and τ used in the judgment,

$$\text{if } \mathbf{SR}(\sigma) \cap \mathbf{SR}(\tau) = \varnothing, \text{ then } \mathbf{CR}(\sigma) \cap \mathbf{CR}(\tau) = \varnothing\,. \quad \square$$

In the theory of cognitive representations presented here, it is only operations performed on cognitive representations and simple relationships between representations that are designed to capture essential features of the cognitive processing in probability judgments. For this purpose, representing descriptions as open sets is convenient, because the richness of the mathematical theory of topology can be exploited. This includes the rich varieties of topological spaces, which can be exploited to better model certain ideas in support theory and to characterize properties of particular event spaces.

The operation of $\cap$-complementation, $\ominus$ plays a central role in the New Foundation's modeling of judgments of probability. The New Foundation uses it for modeling for all but a few situations in the Support Theory literature. In the few exceptions, the operation of set-theoretic complementation is used instead.

Convention 1 It is assume that in evaluating the probability of $\alpha|\beta$, "the probability of α occurring given β," the participant, through the use of information presented to her and her own knowledge, creates the *cognitive complement* of $A = \mathbf{CR}(\alpha)$ with respect to the universe under consideration, $B = \mathbf{CR}(\beta)$. The cognitive complement of A need not correspond to any description. In particular need not correspond to the description "β and not α." *Unless otherwise stated, the operation of cognitive complementation is assumed to be the operation of* $\cap$*-complementation,* $\ominus$*.*

In some cases, classical complementation, $-A$, will coincide with $\ominus A$. There are a few special situations, generally involving unusual probability estimation tasks, where the operation of cognitive complementation of A is $-A$ in a situation where $-A \neq \ominus A$. In such special circumstances it will be stated explicitly that $-$ is the operation of cognitive complementation. $\Box$

A primary difference between the $-$ and $\ominus$ is that $-A$ is a closed set in the underlying topology, whereas, $\ominus A$ is open. There are topologies where all open sets are closed and visa versa. A boolean algebra of sets with each set in the algebra being open is such a topology, i.e., the event space of classical probability theory can be viewed as a special kind of topological space where $- = \ominus$.

The New Foundation assumes that the participant uses methods of evaluations and heuristics to determine the support for A and the support against A and reports her probability of $\alpha|\beta$ as the number,

$$\frac{\text{support for } A}{\text{support for } A \ + \text{support against } A}\,.$$

The New Foundation assumes that the support against A = the support for the cognitive complement of A; that is,

$$\text{support against } A = \text{support for } \ominus A.$$

Many of these considerations are summarized as parts of the following convention and "Cognitive Model" subsection.

Convention 2 Throughout this article, $S^+(A)$ stands for the support for A, $S^-(A)$ for the support against A, and $\mathbb{P}(\alpha|\beta)$ for the participant's judged probability when presented $\alpha|\beta$, where $\alpha \neq \varnothing$, $\beta \neq \varnothing$, and $\alpha \subseteq \beta$.

The following assumptions are made, unless explicitly stated otherwise:

(1) S^+ and S^- take values in the positive reals.

(2) $S^-(A) = S^+(\ominus A)$.

(3) The participant judges $\mathbb{P}(\alpha|\beta)$ in a manner consistent with the equation,

$$\mathbb{P}(\alpha|\beta) = \frac{S^+(A)}{S^+(A) + S^+(\ominus A)}\,.$$

(4) For all cognitive representations E and F, if $E \subset F$ and $S^+(E)$ and $S^+(F)$ have been determined, then $S^+(E) < S^+(F)$. □

2.3 Cognitive Model

In the New Foundation, there are two kinds of situations, each requiring its own set of modeling ideas. The first is the modeling of a single probabilistic judgment $\mathbb{P}(\alpha|\beta)$. The modeling goal is to explain the single judgment in terms judgmental heuristics and support. The second situation is modeling of multiple probabilistic judgments with a common conditioning event, e.g., the simultaneous modeling of $\mathbb{P}(\alpha|\beta)$ and $\mathbb{P}((\gamma \vee \delta)|\beta)$, where $\mathbb{P}((\gamma \vee \delta))$ is an unpacking of α. The second situation's modeling is useful for describing theoretical relationships among the multiple cognitive representations and their cognitive complements, where each representation and its cognitive complement were obtained from the modeling of a single probabilistic judgment.

Both kinds of situations use the concept of "clear instance of α," but in different ways.

Definition 6 Let α be a description. Then i is said to be a *clear instance of* α if and only if the following is true: if the participant were to make an independent judgment as to whether i was a "clear exemplar" of α, then the participant would judge i to be a clear exemplar of α.

The set of clear instances of α is denoted by $\mathsf{CI}(\alpha)$. □

Note that in Definition 6, "clear instance" is defined counterfactually; that is, it is not asserted that the participant is actually asked to make the judgment of "whether i is a 'clear exemplar' of α." Thus the usage of "clear instance" here is theoretical in nature. In philosophy it is a modal concept, because it does not refer to what actually happens in the experiment (which is a probability judgment), but what would happen if the experiment were replaced by another experiment requiring a different judgment (i.e., a clear exemplar judgment). In the experiments reported in this article, participants were not asked to make "clear instance" judgments. Of course, in principle, the paradigms could be expanded to experimentally approximate the modal predicate CI by having some participants make clear instance judgments.

Notice also that in terms of psychology, the occurrence of a clear instance i during the making of the probability judgment of $\alpha|\beta$ is somewhat like the recalling an item in a memory experiment in that the participant generates i, and the occurrence of i in the paradigm where the participant is asked, "Is i a clear instance of α?" is somewhat like making a Yes-No recognition judgment in that i is presented to the subject for evaluation.

In the New Foundation, the concept of "clear instance" plays an important role. Its modal character makes it a "nonextensional" (or "intensional") concept according to the standard philosophical usage.

The principle difference in the two kinds of modelings is that in a multiple judgment situation including the single probabilistic judgments of $\mathbb{P}(\alpha|\beta)$ and $\mathbb{P}(\tau|\beta)$, the judging of $\mathbb{P}(\tau|\beta)$ may have brought to mind instances a of $A = \mathbf{CR}(\alpha)$ that were not brought to mind when $\mathbb{P}(\alpha|\beta)$ was judged. The New Foundation calls such an element a an *external clear instance* of A. Note that a is not an element of A, because τ and its cognitive representation was not part of the judging of $\mathbb{P}(\alpha|\beta)$. Also note that external clear instances do not occur in the modeling of single probabilistic judgments.

In making a single probabilistic judgment $\mathbb{P}(\alpha|\beta)$, the participant may have a clear instance i of α that is not in $\mathbf{CR}(\alpha)$, for example, an instance i that was not available in the judging of the support for α, but was available as a *bad* exemplar of $\ominus \mathbf{CR}(\alpha)$ in the judging of the support for $\ominus \mathbf{CR}(\alpha)$. Such an element i is called an *unrealized clear instance of* $\mathbf{CR}(\alpha)$ *(in the judgment of* $\mathbb{P}(\alpha|\beta)$*).*

2.3.1 The model for a single probabilistic judgment

When $\alpha|\beta$ is presented to the participant for probabilistic judgment, she forms a cognitive representation B corresponding to β. B is taken as the universal set of a topology. She also forms a cognitive representation A corresponding to α that is an open subset of B. She employs cognitive heuristics to find $S^+(A)$ and $S^-(A)$, and produces a judgment of probability, $\mathbb{P}(\alpha|\beta)$, consistent with the formula,

$$\mathbb{P}(\alpha|\beta) = \frac{S^+(A)}{S^+(A) + S^-(A)}\,.$$

It is assumed that in finding $S^-(A)$, she creates the cognitive event $\ominus A$ and uses,

$$S^-(A) = \text{support for } \ominus A = S^+(\ominus A)\,.$$

Convention 3 Unless otherwise explicitly stated, it is a assumed throughout this subsection that a case of a the single probability judgment $\mathbb{P}(\alpha|\beta)$ is being considered and that $A = \mathbf{CR}(\alpha)$, $B = \mathbf{CR}(\beta)$, and $\ominus A$ is the cognitive complement of A with respect to $B^{\cdot}$ $\square$

Definition 7 Elements of A are called *<u>clear instances</u> of A (generated by $\mathbb{P}(\alpha|\beta)$.)* $\square$

Obviously,

$$A \subseteq \mathsf{Cl}(\alpha)\,;$$

however, as previously discussed, elements in B may contain clear instances i of α that are not in A, e.g., i may not have been available in the generation of A but available in the generation of $\ominus A$ and rejected as being part of $\ominus A$.

Definition 8 Clear instances of α that are elements of $B - A$ are called *unrealized clear instances of A.* Elements in $\mathsf{bd}(A)$ are called *poor instances* of A generated by the judging of $\mathbb{P}(\alpha|\beta)$; and elements of $\mathsf{bd}(A) \cap \mathsf{bd}(\ominus A)$ are called *weakly ambiguous instances* of A generate by the judging of $\mathbb{P}(\alpha|\beta)$. $\square$

The following is the reason why I thought "weakly ambiguous" is an appropriate use of the term here: Suppose a is a weakly ambiguous element of A, U is an arbitrary open set, and $a \in U$. Then,

$$U \cap A \neq \varnothing \ \text{ and } \ U \cap \ominus A \neq \varnothing\,.$$

Suppose in this probabilistic judgment or in a different probabilistic judgment U is the cognitive representation of some description θ. Then $\mathsf{Cl}(\theta)$ has instances of A and $\ominus A$. Because U is arbitrary, it then follows that for any description σ in any probabilistic judgment for which $a \in \mathbf{CR}(\sigma)$, there exist a clear instance of $\mathbf{CR}(\sigma)$ that is in A and another clear instance $\mathbf{CR}(\sigma)$ that is in $\ominus A$.

The cognitive model assumes that *unrealized clear instances of A are not weakly ambiguous instances of A.* It follows from Definition 8 and this assumption that the unrealized clear instances of A are in $\mathsf{bd}(A) - (\mathsf{bd}(A) \cap \mathsf{bd}(\ominus A))$. As an illustration, let B be the cartesian plane and A the open unit disk minus the origin. Then

- $\ominus A = (B -$ the closed unit disk$) \cup \{$the origin$\}$,
- $\mathsf{bd}(A) =$ the unit circle $\cup\, \{$the origin$\}$,
- the set of weakly ambiguous elements of $A =$ the unit circle, and
- the set of unrealized clear instances of $A = \{$the origin$\}$.

The set of poor instances include the unrealized clear and weakly ambiguous instances and may include other kinds of instances.

In classical probability theory, which is based on measure theory, the boundaries of events have measure 0, that is, they have probability 0. In this sense, they can be "ignored" in calculations involving positive probabilities. In the Cognitive Model, the boundaries may contain subsets that have positive support. However, the Cognitive Model assumes that poor instances are ignored in the computations of $S^+(A)$ and $S^+(\ominus A)$, and therefore they have no impact on the number $\mathbb{P}(\alpha|\beta)$ which is computed by the equation

$$\mathbb{P}(\alpha|\beta) = \frac{S^+(A)}{S^+(A) + S^+(\ominus A)} \, .$$

Ignoring poor instances is a plausible cognitive strategy, because the clear instances have more impact than poor ones, and the ignoring of the poor ones greatly reduces the complexity of the calculation of $S^+(A)$ and $S^+(\ominus A)$.

2.3.2 The model for multiple probabilistic judgments

The discussion of multiple probabilistic judgments of this subsection is restricted to describing unpacking experiments like those occurring in Tversky and Koehler (1994). The general case for multiple probabilistic judgments is not presented in this article. I believe the strategy for analyzing the restricted form of unpacking of this subsection extends to the general multiple probabilistic case.

Many experiments in support theory compare judged probabilities of descriptions of the forms $\alpha|\beta$ and $(\gamma \vee \delta)|\beta$, where $(\gamma \vee \delta)$ is an unpacking of α. In some paradigms the probability estimates $\mathbb{P}(\alpha|\beta)$ and $(\mathbb{P}(\gamma \vee \delta)|\beta)$ come from the same participant; in other paradigms, they come from different participants, or from averaging across participants. It is useful to be able to describe theoretical relationships among the cognitive representations α, β, γ, δ, $\gamma \vee \delta$, and their cognitive complements in order to provide theoretical explanations for the observed relationships between $\mathbb{P}(\alpha|\beta)$, $\mathbb{P}(\gamma|\beta)$, $\mathbb{P}(\delta|\beta)$, and $(\mathbb{P}(\gamma \vee \delta)|\beta)$. For such considerations, the topologically based event space of the Cognitive Model provides a richer language for describing such relationships than the boolean algebra event space. As described in Theorem 1, a boolean event space is a special case of a Cognitive Model event space in which each open set is also a closed set, or equivalently, each open set contains it own boundary. This effectively eliminates the use of the boundary of a cognitive representation as a modeling tool, and restricts the kinds of interpretations that can be given to empirical findings.

Convention 4 Throughout this subsection, Let $\gamma \vee \delta$ be an unpacking of α in the multiple probabilistic situation based on the two single probabilistic judgments $\mathbb{P}(\alpha|\beta)$ and $\mathbb{P}(\gamma \vee \delta|\beta)$. Let

$$A = \mathbf{CR}(\alpha)$$

in the judging of $\mathbb{P}(\alpha|\beta)$, and

$$E = \mathbf{CR}(\gamma \vee \delta)$$

in the judging of $\mathbb{P}(\gamma \vee \delta|\beta)$ $\square$

The Cognitive Model allows for different cognitive representations of β in the modelings of the single probabilistic judgments $\mathbb{P}(\alpha|\beta)$ and $\mathbb{P}(\gamma \vee \delta|\beta)$. In the multiple probabilistic situation, the universal set is assumed to be,

$$B = B_\alpha \cup B_{\gamma\vee\delta} \,,$$

where B_α is the cognitive representation of β used in the cognitive modeling of $\mathbb{P}(\alpha|\beta)$ (as a single probabilistic judgment), and where $B_{\gamma\vee\delta}$ is the cognitive representation of β used in the cognitive modeling of $\mathbb{P}(\gamma \vee \delta|\beta)$ (as a single probabilistic judgment).

To apply the Cognitive Model to capture the idea of "subadditivity," some assumptions have to be made about the relationships among B_α, $B_{\gamma\vee\delta}$, A, and E. The following assumptions appear to me to capture, within the cognitive model, the the idea of "subadditivity" inherent in Tversky and Koehler (1994):

1. $B_\alpha \subseteq B_{\gamma\vee\delta} = B$.
2. $A \subseteq E$.
3. $\ominus A = \ominus E$.

The unpacking of α into $\gamma \vee \delta$ often bring about clear instances of α in E that are not in A. Such instances are called *external clear instances of* A. Note that an "external clear instances of A" may or may not be an "unrealized clear instances of A" that arose in the single probabilistic judgment of $\mathbb{P}(\alpha|\beta)$. While a later example will provide an instance of neither $A \subseteq E$ nor $E \subseteq A$, Statement 2 above is adopted in this subsection, in part because I interpret Support Theory assuming that the repacking of $\gamma \vee \delta$ into α produces no external clear instances of E.

The following example captures some of the intuition for "external clear instance" in the multiple probabilistic judgment situation. It is taken from a study of Sloman, Rottenstreich, Wisniewski, Hadjichristidis, and Fox discussed later.

> θ: Consider all the people that will die in the U.S. next year. Suppose we select one of these people at random.
>
> Please estimate the *probability* that
>
> > ζ: this person's death will be attributed to disease
> >
> > τ: this person's death will be attributed to cirrhosis

Let Z and T be respectively the cognitive representations of ζ and τ in the single probabilistic judgments of $\zeta|\theta$ and $\tau|\theta$. Suppose that i is an instance of T that came to mind when the participant(s) judged $\mathbb{P}(\tau|\theta)$ but did not come to mind when participant(s) judged $\mathbb{P}(\zeta|\theta)$. Then i is an "external clear instance" of Z, that is, if i were available when judging $\mathbb{P}(\zeta|\theta)$, the participant would consider i to be a clear instance of Z.

In terms of the formalism presented here and the above special assumptions about A, E, and their cognitive complements, the type implicit subadditivity assumed by Tversky and Koehler (1994) to result from the unpacking α into $\gamma \vee \delta$ is expressed by the following equation:

$$A \subseteq E \quad \text{and} \quad \ominus\, A = \ominus\, E\,. \tag{2}$$

Applying S^+ and $\mathbb{P}$ to Equation 2 yields,

$$S^+(A) \leq S^+(E) \quad \text{and} \quad S^+(\ominus\, A) = S^+(\ominus\, E),$$

and

$$\mathbb{P}(\alpha|\beta) = \frac{S^+(A)}{S^+(A) + S^+(\ominus\, A)} \leq \frac{S^+(E)}{S^+(E) + S^+(\ominus\, E)} = \mathbb{P}(\gamma \vee \delta\,|\,\beta)\,.$$

2.4 Frequency Based Judgments

This subsection presents two studies involving unpacking. The first shows the unpacking of an implicit description into two explicit disjunctions. Both disjunctions demonstrate explicit strict subadditivity. They, however, differ on implicit subadditivity—one showing strict subadditivity, and the other only additivity. The second study also unpacks an implicit description into two explicit disjunctions. Here, one disjunction displays additivity, while the other demonstrates explicit superadditivity. Both studies involve frequency judgments, and their results are explained in terms of the Cognitive Model.

Rottenstreich and Tversky (1997) presents the following study involving 165 Standford undergraduate economic students. They were presented Case 1 and Case 2 for evaluation, with Case 2 taking place a few weeks after Case 1. Both cases consisted of a questionnaire in which the participants were informed the following:

> Each year in the United States, approximately 2 million people (or 1% of the population) die from a variety of causes. In this questionnaire you will be asked to estimate the probability that a randomly selected death is due to one cause rather than another. Obviously, you are not expected to know the exact figures, but everyone has some idea about the prevalence of various causes of death. To give you a feel for the numbers involved, note that 1.5% of deaths each year are attributable to suicide.

In terms of our notation the following were presented for probabilistic judgment:

$$\alpha|\beta,\ \alpha_s|\beta,\ \alpha_a|\beta,\ (\alpha_s \vee \alpha_a)|\beta,\ \alpha_d|\beta,\ \alpha_n|\beta,\ (\alpha_d \vee \alpha_n)|\beta,$$

where

β is death,

α is homicide,

α_s is homicide by a stranger,

α_a is homicide by an acquaintance,

$\alpha_s \vee \alpha_a$ is homicide by a stranger or homicide by an acquaintance,

α_d is daytime homicide,

α_n is nighttime homicide,

$\alpha_d \vee \alpha_n$ is homicide during the daytime or homicide during the nighttime.

In both Cases 1 and 2 the participants were randomly divided into three groups.

In Case 1, approximately one third of the participants judged $\alpha|\beta$, approximately one third judged $(\alpha_s \vee \alpha_a)|\beta$, and approximately one third judged both and $\alpha_s|\beta$ and $\alpha_a|\beta$.

In Case 2, approximately one third of the participants again judged $\alpha|\beta$, approximately one third judged $(\alpha_d \vee \alpha_n)|\beta$, and approximately one third judged both and $\alpha_d|\beta$ and $\alpha_n|\beta$.

Rottenstreich and Tversky predicted that $\alpha_s \vee \alpha_a$ was "more likely to bring to mind additional possibilities than $\alpha_d \vee \alpha_n$." They reasoned,

> Homicide by an acquaintance suggests domestic violence or a partner's quarrel, whereas homicide by a stranger suggests armed robbery or drive-by shooting. In contrast, daytime homicide and nighttime homicide are less likely to bring to mind disparate acts and hence are more readily repacked as ["homicide"]. Consequently, we expect more implicit subadditivity in Case 1,
>
> $$[\text{i.e., } \mathbb{P}(\alpha_s \vee \beta_a) - \mathbb{P}(\alpha|\beta) > \mathbb{P}(\alpha_d \vee \alpha_n) - \mathbb{P}(\alpha|\beta)\,,]$$
>
> due to enhanced availability, and more explicit subadditivity in Case 2,
>
> $$[\text{i.e., } \mathbb{P}(\alpha_d|\beta) + \mathbb{P}(\alpha_n|\beta) - \mathbb{P}(\alpha_d \vee \alpha_n) > \mathbb{P}(\alpha_s|\beta) + \mathbb{P}(\alpha_a|\beta) - \mathbb{P}(\alpha_s \vee \alpha_a)\,,]$$
>
> due to repacking of the explicit disjunction.

Rottenstreich and Tversky found that their predictions held: Letting $\mathbb{P}$ stand for the median probability judgment, they found

Case 1 $\mathbb{P}(\alpha|\beta) = .20,\;\; \mathbb{P}(\alpha_s \vee \alpha_a) = .25,\;\; \mathbb{P}(\alpha_s) = .15,\;\; \mathbb{P}(\alpha_a) = .15\;;$

Case 2 $\mathbb{P}(\alpha|\beta) = .20,\;\; \mathbb{P}(\alpha_d \vee \alpha_n) = .20,\;\; \mathbb{P}(\alpha_d) = .10,\;\; \mathbb{P}(\alpha_a) = .21\,.$

Note that Rottenstreich and Tversky's analysis corresponds in the New Foundation to the multiple probabilistic situation like the one discussed in the previous subsection. Their prediction $\alpha_s \vee \alpha_a$ was "more likely to bring to mind additional possibilities than $\alpha_d \vee \alpha_n$," and their subsequent reasoning for the prediction, corresponds to the assumptions about unpacking and external clear instances in the previous subsection. However, for their argument to work they must make some assumption relating supports against α and $\alpha_s \vee \alpha_a$. In the previous subsection it was assumed that

$$\text{the support against } \alpha = \text{the support against } \alpha_s \vee \alpha_a. \tag{3}$$

We were able to make this assumption while keeping the cognitive representation of β constant. In Support Theory, which has a boolean rather than a topological approach to events, it appears to me that Equation 3 is an unreasonable approach because it would require

$$\text{the support against } \alpha < \text{the support against } \alpha_s \vee \alpha_a,$$

which tends to cancel the effect of

$$\text{the support for } \alpha < \text{the support for } \alpha_s \vee \alpha_a\,,$$

in probability judgments derived from support, i.e., Equation 3 is unreasonable for Support Theory because the intuitive justification that the higher value of $\mathbb{P}(\alpha_s \vee \alpha_a \,|\, \beta)$ in the above example is due to implicit subadditivity is greatly weakened. The most reasonable path within Support Theory appear to me is to assume

$$\text{the support against } \alpha > \text{the support against } \alpha_s \vee \alpha_a.$$

This path is unreasonable if

> the (non-extensional) event corresponding to β in "$\alpha|\beta$" = the (the non-extensional) event corresponding to β in "$\alpha_s \vee \alpha_a \,|\, \beta$",

because (the non-extensional) event corresponding to $\alpha_s \vee \alpha_a \supseteq$ (the non-extensional) event corresponding to α. But it becomes reasonable if

> the (the non-extensional) event corresponding to β in "$\alpha|\beta$" $\subset$ the (the non-extensional) event corresponding to β in "$\alpha_s \vee \alpha_a \,|\, \beta$".

However, it appears to me that such a "reasonable" approach *for explaining the implicit subadditivity observed in the above example* is likely to open a can of worms in Rottenstreich and Tversky's version of Support Theory.

Recall that cognitively preserved semantical disjointness (Definition 5) holds for a judgment if and only if for all descriptions σ and τ used in the judgment, if σ and τ are semantically disjoint, then they are cognitively disjoint. The Cognitive Model makes the following assumption:

Additional Cognitive Model Assumption *For frequency judgments based on availability, cognitively preserved semantical disjointness holds.* □

Note that the above Cognitive Model assumption is only for frequency judgments. Cognitively preserved semantical disjointness may fail for other kinds of judgments. In fact, it often fails for situations when the probability judgment is based on similarity. The reason for this failure is discussed in the next subsection.

Consider the following example from Sloman, Rottenstreich, Wisniewski, Hadjichristidis, and Fox (2003). University of Chicago undergraduates were asked to complete the following item:

> Consider all the people that will die in the U.S. next year. Suppose we select one of these people at random. Please estimate the *probability* that this person's death will be attributed to the following causes.

The students were divided into three groups, *packed, typical,* and *weak-atypical.* These groups judged the following descriptions:

packed: disease
typical: heart disease, cancer, stroke, or any other disease
weak-atypical: pneumonia, diabetes, cirrhosis, or any other disease

The typical group judged the probability of a description mentioning the three most common causes of death; the weak-atypical judged a description mentioning less common causes. Sloman et al. theorized that in judging the packed condition, participants would naturally unpack it into typical exemplars, and thus predicted that the packed condition and the (unpacked) typical condition should yield approximately the same judged probabilities, producing additivity for the unpacking. They also theorized that unpacking the packed condition into atypical exemplars with weak support (i.e., into the weak-typical condition) would capture attention away from the more typical exemplars thus yielding lower support and therefore a lower judged probability. Thus for the weak-packed condition, they predicted *superadditivity,* i.e., that the sum of the probabilities for the partition will be less than 1, contradicting a basic principle of the support theories of Tversky and Koehler (1994) and Rottenstreich and Tversky (1997). The data from the above study supported these predictions: they found additivity for the typical condition and superadditivity for the weak-atypical condition.

Note that in terms of the Cognitive Model, the unpacking of the packed case into the weak-atypical case provides an example of cognitive representations D (for the packed case) and E (for the unpacked case) such that neither $D \subseteq E$ nor $E \subseteq D$. That is, in the multiple probabilistic judgment situation, D has elements that are external clear instances of the unpacked case, and E has external clear instances of the packed case.

Sloman et al. also theorized that in the rare cases where a condition is unpacked into atypical exemplars that have stronger support than typical exemplars, one may observed subadditivity.

The above studies of Sloman et al. and Tversky and Rottenstreich involving frequency judgments show that the method of unpacking matters with respect to additivity.

2.5 Judgments Based on Representativeness

Kahneman and Tversky (1982) gave participants the following description β:

> β: Linda is 31 years old, single outspoken and very bright. She majored in philosophy. As a student she was deeply concerned with the issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

participants were asked to rank order the following statements by their probability, using 1 for the most probable and 8 for the least probable. The descriptions denoted by γ and α below are the ones that play the important roles in the discussion presented here.

> Linda is a teacher in elementary school
> Linda works in a bookstore and takes Yoga classes
> Linda is active in the feminist movement
> Linda is a psychiatric social worker
> Linda is a member of the League of Women voters
> γ: Linda is a bank teller
> Linda is an insurance salesperson
> α: Linda is a bank teller and is active in the feminist movement

Kahneman and Tversky found that over 85% of participants believed it was more likely that Linda was both a bank teller and a feminist (α) than just a bank teller (γ). This is an example of what has become known as the *conjunction fallacy*. According to Kahneman and Tversky, it is due to representativeness: "bank teller and is active in the feminist movement" is more a "representative" description of Linda than just "bank teller."

The New Foundation allows cognitive representations to take many forms. All forms are sets, but the sets can have different kinds of elements. For representativeness, the elements of the sets are taken to be exemplifying properties. This choice allows for a better modeling of the similarity concept.

For the purposes of this article, it is assumed that the representativeness heuristic is employed for making probability judgments about Linda in the above example. Accordingly, the description of Linda, β, makes available to the participant a set of properties, L, that exemplifies people fulfilling that description. Similarly, the predicate "is a bank teller" makes available a set of properties, T, exemplifying bank tellers, and the predicate "is a bank teller and is active in the feminist movement" makes available a set of properties, TF, exemplifying people who are bank tellers and are active in the feminist movement. β, α, and γ are assumed to have the following cognitive representations:

- $\mathbf{CR}(\beta) = L$.

- $\mathbf{CR}(\alpha) = L \cap TF$.
- $\mathbf{CR}(\gamma) = L \cap T$.

The "conjunction fallacy" arises because $L \cap TF$ has greater support than $L \cap T$, that is the properties in $L \cap TF$ are more available to the participant than those in $L \cap T$. Notice that because T and TF depend on the availability, it is most likely that for most participants, $TF - T \neq \varnothing$ and $T - TF \neq \varnothing$.

Notice that β does not completely characterize a person (real or fictitious). It gives some characteristics that a person may satisfy. For the purposes of making frequency judgments, the cognitive representation of β may be thought of as the set D of exemplars d that (cognitively) satisfies β when d is appropriately substituted for "Linda". For similarity judgments, the cognitive representation of β is viewed as the set of properties, L. L may be generated in different ways, e.g., as the set of properties common to the elements of D, or as the set of properties cognitively derivable from the description β.

For theorizing about *similarity,* the similarity interpretation of proper noun "Linda" is best viewed as the set properties L. This puts "Linda" at the same level as the similarity interpretation of the noun phrase "bank teller," which has the set T of properties as its cognitive representation. The similarity interpretation of "Linda is a bank teller," γ, is then $L \cap T$. Note how this differs from the semantical representation of γ: In the Semantic Interpretation, (i) "Linda" is interpreted as an individual, l, not as a set; (ii) the predicate, "is a bank teller," is interpreted as a set t (i.e., the set of bank tellers); and (iii) the statement, "Linda is a bank teller," is interpreted as the statement $l \in t$. Thus for similarity judgments involving a propositional description, the cognitive representation relates subject and predicate through set theoretic intersection, whereas the Semantic Interpretation relates subject and predicate through set theoretic membership.

The following Cognitive Model assumption summarizes the primary theoretical difference between frequency and similarity judgments.

Additional Cognitive Model Assumption *Probability judgments involving frequency are based on the* availability of exemplars *that come to mind; whereas, probability judgments involving similarity are based on the* availability of exemplifying properties *that come to mind.* □

Note that, in general, cognitively preserved semantical disjointness (Definition 5) is going to fail for many pairs of descriptions ζ and τ when similarity is used, because, for example, the semantical disjointness of $\mathbf{SR}(\zeta)$ and $\mathbf{SR}(\tau)$ does not preclude cognitive exemplars of ζ and τ from having available properties in common. In the multiple probabilistic representation condition, this leads to the concept of "strong ambiguity:"

Definition 9 Suppose ζ and τ are semantical disjoint and representativeness is used in the single probabilistic judgments of $\mathbb{P}(\zeta|\beta)$ and $\mathbb{P}(\tau|\beta)$. In the judgments $\mathbb{P}(\zeta|\beta)$ and $\mathbb{P}(\tau|\beta)$, let

$$Z = \mathbf{SR}(\zeta) \ \text{ and } \ T = \mathbf{SR}(\tau)\,.$$

Consider a multiple probabilistic judgment situation containing the judgments $\mathbb{P}(\zeta|\beta)$ and $\mathbb{P}(\tau|\beta)$. Then i is said to be *strongly ambiguous with respect to Z and T* if and only if $i \in (Z \cap T)$, i.e., if and only if i is a clear instance of Z and a clear instance of T. $\square$

The following example illustrates the impact of strongly ambiguous elements on probability judgments.

Example 1 Suppose that the participant is instructed to "judge the probability of ζ occurring rather than τ," and at later time is instructed to "judge the probability of τ occurring rather than ζ;" that is, judge the probabilities of $\zeta|\zeta \vee \tau$ and $\tau|\tau \vee \zeta$. We assume ζ and τ are semantically disjoint and that the heuristic of representativeness is being used to obtain the probability judgments. We also assume a multiple probabilistic situation where

$$\zeta \vee \tau \;=\; \tau \vee \zeta \;=\; \beta\,,$$

and its only cognitive representations are

(*i*) $Z = \mathbf{CR}(\zeta)$ and $\ominus Z$ from the judging of $\mathbb{P}(\zeta|\beta)$,

(*ii*) $T = \mathbf{CR}(\tau)$ and $\ominus T$ form the judging $\mathbb{P}(\tau|\beta)$, and

(*iii*) $B = \mathbf{CR}(\beta)$.

Suppose

$$\Sigma = \{i | i \text{ is strongly ambiguous with respect to } Z \text{ and } T\}\,,$$

and $\Sigma \neq \varnothing$. Let i be an arbitrary element of Σ. Then $i \notin \ominus Z$ (because it is a clear instance of Z) and similarly $i \notin \ominus T$. Thus, when the support for Z is being evaluate in the judging of $\mathbb{P}(\zeta|\zeta \vee \tau)$, all elements of Σ appear as clear instances of Z and as such add to the support for Z; however, because $\Sigma \not\subseteq \ominus Z$, elements of Σ do not add to the support against Z. Similarly, in the judging of $\mathbb{P}(\tau|\zeta \vee \tau)$, elements of Σ add to the support for T, but do not add to the support against T. Therefore in computing the sum

$$\mathbb{P}(\sigma|\sigma \vee \tau) + \mathbb{P}(\tau|\sigma \vee \tau)$$

support for instances in Σ are evaluated twice, while support for other clear instances of Z or of T are evaluated only once, i.e., the support for strongly ambiguous elements are "double counted." □

2.6 Probability Judgments for Binary Partitions

In the recent literature the are accounts of empirical demonstrations of violations of binary complementarity. This subsection presents a brief discussion of some these violations and their relationship to the Cognitive Model.

Throughout the subsection it is assumed that participants (in either within-subject or between-subjects designs) are presented $\alpha|\beta$ and $\gamma|\beta$ for probabilistic judgment. It is also assumed that in the semantics, the extensions of α and γ are disjoint, i.e., $\mathbf{SR}(\alpha) \cap \mathbf{SR}(\gamma) = \varnothing$, and in the semantics, $\alpha \vee \gamma$ exhausts β, i.e., $\mathbf{SR}(\alpha) \cup \mathbf{SR}(\gamma) = \mathbf{SR}(\beta)$. As before A, B, and C are respectively the cognitive representations of α, β, and γ.

The following theorem is a simple algebraic consequence of the equations,

$$\mathbb{P}(\alpha|\beta) = \frac{S^+(A)}{S^+(A) + S^+(\ominus A)} \text{ and } \mathbb{P}(\gamma|\beta) = \frac{S^+(C)}{S^+(C) + S^+(\ominus C)}\,.$$

Theorem 2 *The following three statements are true:*

$$\mathbb{P}(\alpha|\beta) + \mathbb{P}(\gamma|\beta) \text{ is } \begin{cases} > 1 \text{ iff } S^+(A)S^+(C) > S^+(\ominus A)S^+(\ominus C)\,, \\ = 1 \text{ iff } S^+(A)S^+(C) = S^+(\ominus A)S^+(\ominus C)\,, \\ < 1 \text{ iff } S^+(A)S^+(C) < S^+(\ominus A)S^+(\ominus C)\,. \quad \Box \end{cases} \quad \Box \qquad (4)$$

Part of the foundation of Support Theory was the numerous examples of binary complementarity (additive binary partitions) holding. In these examples, it was easy to interpret C as the cognitive complement of A and visa versa, i.e.,

$$C = \ominus A \text{ and } A = \ominus C\,. \qquad (5)$$

Therefore, by Theorem 2 one would expect $\mathbb{P}(\alpha|\beta) + \mathbb{P}(\gamma|\beta) = 1$, which is what was observed empirically.

Note that it follows from Equation 5 that

$$A = \ominus \ominus A \text{ and } C = \ominus \ominus C\,. \qquad (6)$$

This is analogous $A = --A$ in a boolean algebra of sets. However, Equation 6 does not require $A = --A$. For example, let the universe, B, be the Cartesian plane, A the open unit disk, $C = (B -$ the closed unit disk). Equation 6 holds but $\ominus A \neq -A$. In the New Foundation, Equations 5 and 6 represent a special kind of situation. For instance, Statement 2 of Theorem 10 in Subsection 4.3 gives an example of an event E such that $E \neq \ominus \ominus E$.

evaluation of Binary Complementarity were observed by Brenner and Rottenstreich (1999), Macchi, Osherson, and Krantz (1999), and Idson, Krantz, Osherson, and Bonini (2001).

Brenner and Rottenstreich (1999) observed, "Several researchers have found compelling evidence for the descriptive validity of binary complementarity ... Interestingly, however, these researchers examined judgments involving only singleton hypotheses, ..." Brenner and Rottenstreich decided to investigate binary partitions consisting of a singleton and an explicit disjunction (of several items). "We suggest there may be a preference toward singletons in the focal position. For example, judging the probability that *Titanic* will win Best Picture rather than *either As Good As it Gets, Good Will Hunting, L. A. Confidential, or The Full Monty* seems quite natural. However, judging the probability that *either As Good As it Gets, Good Will Hunting, L. A. Confidential, or The Full Monty* will win Best Picture rather than *Titanic* seems more unwieldy. Put differently, it seems natural to compare the likelihood of a single possibility to that of a set of alternatives. On the other hand, it seems awkward to compare the likelihood of a set of possibilities to that of a single alternative. As a result, there may be a greater tendency to repack *either As Good As it Gets, Good Will Hunting, L. A. Confidential, or The Full Monty* or more generally any disjunction, when it is in the focal than in the alternative position." For partitions, (S, D) consisting of a singleton, S, and a disjunction, D, they expected to find departures from binary complementarity of the form $P(S, D) + P(D, S) < 1$.

Brenner and Rottenstreich ran several experiments testing the dual predictions of binary complementarity for singleton-singleton judgments and violations of binary complementarity for singleton-disjunction pairs. They found the following consistent pattern:

> Sums of judgments for complementary hypotheses are close to 1 when the hypotheses are singletons, and are less than 1 when one of the hypotheses is a disjunction. We observed this pattern in judgments of probability and frequency, and for judgments involving both externally and self-generated hypotheses. *(Pg. 146)*

Let α be "either As Good As it Gets, Good Will Hunting, L. A. Confidential, or

The Full Monty will win Best Picture," and let γ be "Titanic will win Best Picture." In the judging of $P(\gamma, \alpha)$ and $P(\alpha, \gamma)$, assume that the similarity heuristic is being employed and that superadditivity results. For a propositional description δ, let $s_f(\delta)$ be the support for δ in the focal position and $s_a(\delta)$ be for δ in the alternate position. I interpret Rottenstreich's and Brenner's reasoning for this case to be that superadditivity is observed because

$$s_a(\alpha) > s_f(\alpha) \ \text{ and } \ s_a(\gamma) = s_f(\gamma) \tag{7}$$

and

$$P(\gamma, \alpha) = \frac{s_f(\gamma)}{s_f(\gamma) + s_a(\alpha)} \ \text{ and } \ P(\alpha, \gamma) = \frac{s_f(\alpha)}{s_f(\alpha) + s_a(\gamma)} \, .$$

I interpret Rottenstreich and Brenner as suggesting that this occurs because α is interpreted differently in the focal position than in the alternative position. If this is the case, then the conditional hypotheses $\alpha \vee \gamma$ (for $P(\gamma, \alpha)$) and $\gamma \vee \alpha$ ($P(\gamma, \alpha)$) for would also be interpreted differently. This conclusion is also permitted by the Cognitive Model. However, because the Cognitive Model uses a non-classical form complementation, the conclusion about subadditivity can reached *without* having the conditional hypotheses change as γ and α change focal and alternate positions:

Let $\beta = \alpha \vee \gamma = \gamma \vee \alpha$, $A = \mathbf{CR}(\alpha)$, and $C = \mathbf{CR}(\gamma)$. Interpret s_a as S^- and s_f as S^+. Then

$$s_a(\alpha) = S^+(\ominus A), \ s_f(\alpha) = S^+(A), \ s_a(\gamma) = S^+(\ominus C), \ \text{ and } \ s_f(\gamma) = S^+(C) \, .$$

Rottenstreich and Brenner's reasoning about support, Equation 7, then becomes,

$$S^+(\ominus A) > S^+(A) \ \text{ and } \ S^+(\ominus C) = S^+(C) \, ,$$

which by Equation 4 yields

$$\mathbb{P}(\alpha|\beta) + \mathbb{P}(\gamma|\beta) < 1 \, .$$

2.7 Probability Judgments for Events with Minuscule Support

A few studies have been conducted that show superadditivity in situations where it is very difficult for participants to find significant support for the cognitive representations of descriptions. For such special situations the Cognitive Model computes "support against" in a different manner than previously:

Cognitive Model Assumption for Cases with Very Small Support for Cognitive Representations When there is very small support for a cognitive representation R, the support against R, $S^-(R)$ is either $S^+(-A)$ or $S^+(\ominus A)$ depending on the size of the boundary. $\square$

The intuition for $S^-(R) = S^+(-R)$ is that in the search for clear instances of R, the participant finds many instances that are not clear instances of R or or come up with "blanks" for searches of clear instances. Such non-clear instances are called collectively *poor instances of* R. The Cognitive model interpret the poor instances of R to be part of the boundary of R. When the boundary of R is huge with respect to R, it is not ignored, but instead is added in into the computation of the support against R, $S^-(R)$, which in this case the Cognitive Model interprets as $S^+(-R)$. A minority of participants may choose to ignore huge boundaries and have $S^-(R) = S^+(\ominus R)$. Also there may be situations in which the support for A is very small but the boundary is also very small. In such a case the Cognitive Model assumes $S^-(R) = S^+(\ominus R)$.

Macchi, Osherson, and Krantz (1999) conducted studies involving ultra-difficult general information questions. For example, participants were presented one of the following:

> The freezing point of gasoline is not equal to that of alcohol. What is the probability that the freezing point of gasoline is greater than that of alcohol? (Alternate version: What is the probability that the freezing point of alcohol is greater than that of gasoline?)

For such items, they found superadditivity. They reasoned that given the ultra-difficulty of the questions there is relatively little evidence in favor of the focal description. Thus, if the participants attended relatively more to the focal rather than the alternative description, then they might not appreciate the fact that the alternative description also has little support, producing superadditivity.

Let α be "The freezing point of gasoline is greater than that of alcohol," A be $\mathbf{CR}(\alpha)$, γ be "The freezing point of gasoline is not equal to that of alcohol," and C be $\mathbf{CR}(\gamma)$.

The Cognitive Model provides a different interpretation: participants realize that there is very little support for A and C. In fact A and C have many poor poor instances (including coming with "blanks"). This produces very large boundaries for A and C. The model predicts that in such a case the judged probability of α is likely to be extra small, because the support against A is $S^+(-A)$, which is extra large because it contains $\mathsf{bd}(A)$, and similarly the the judged probability of γ is likely to be extra small.

In a follow-up study, Macchi et al. attempted to equalize the amounts of attention paid to focal and alternative descriptions by explicit mention of them. participants were presented with,

> The freezing point of gasoline is not equal to that of alcohol. Thus, either the freezing point of gasoline is greater than that of alcohol, or the freezing point of alcohol is greater than that of gasoline. What is the probability that the freezing point of gasoline is greater than that of alcohol?,

or with the above with "gasoline" and "alcohol" permuted.

In this latter study, the typical sum of judged probabilities was about 1.01. Macchi et al. concluded that additivity was observed because the participants attended symmetrically to the focal and alternative descriptions. The Cognitive Model suggests a different reason:

The conditioning hypothesis, β, in the original was changed to a different one, $\beta^\star$ in the follow-up. Although β and $\beta^\star$ have the same interpretation in the semantic model, they have different cognitive representations, respectively, B and $B^\star$. The the likelihood of generating poor instances of A relative to $B^\star$ is much less than that of A relative to B; i.e., in the followup study, a reasonable amount of the ambiguity present in A has been squeezed out through the incorporation of clarifying text, thus reducing the number of poor instances generate in determining the support for A, i.e., reducing the size of the boundary of A relative to A. Therefore, the support against A in the followup is $S^+(\ominus A)$. Similarly, $\ominus C$ in the followup is $\ominus C$. Additivity is observed because $\ominus A\ = C$ and $\ominus C\ = A$. This suggests a symmetric relationship between focal and alternative descriptions, but not necessarily of the kind theorized by Macchi et al.

Idson, Krantz, Osherson, and Bonni (2001) conducted experiments that in addition to probability judgments had within-respondent's judgments of *support for.* They classified items as "high knowledge" if both elements of binary partition had high judged support, "and "low knowledge" if both elements had low support. They found that for binary partitions consisting of high knowledge items, subadditivity is regularly seen, and that for those consisting of low knowledge items, superadditivity is seen, but less regularly.

Idson, et al. presented an equation that related judgments of probability to judgments of support. The form of their equation in terms of judged support is very different from that used in Support Theory and the Cognitive Model. I do not see any reason to assume that the cognitive operations in forming support for a cognitive event as part of the making of a probability judgment should be the same as those for making a support judgment. Thus the support functions that are theoretically employed in making probability judgments should not be automatically identified with those that result from direct judgments of support. More research is needed to establish the relationship between the two kinds of support functions.

3 Discussion

Although Support Theory has made a number of advances experimentally and conceptually from the founding articles of Tversky and Koehler (1994) and Rottenstreich and Tversky (1997), the original concepts of "extensionality" and "nonextensionality" used by Tversky and Koehler have remained unchanged. I find a number of problems with these two foundational concepts.

Tversky and Koehler (1994) formulated as follows their approach to probability assignments, noting its differences from the traditional approach from probability theory:

> Let $\mathbf{T}$ be a finite set including at least two elements, interpreted as states of the world. We assume exactly one state obtains but it is generally not known to the judge. Subsets of $\mathbf{T}$ are called *events.* We distinguish between events and descriptions of events, called *Hypotheses.* Let $\mathbf{H}$ be a set of hypotheses that describe the events in $\mathbf{T}$. Thus, we assume that each hypothesis $A \in \mathbf{H}$ corresponds to a unique event $A' \subseteq \mathbf{T}$. This is a many-to-one mapping because different hypotheses, say A and B, may have the same extension (i.e., $A' = B'$). For example, suppose one rolls a pair of dice. The hypotheses "The sum is 3" and "The product is 2" are different descriptions of the same event; namely, one die shows 1 and the other shows 2. We assume that $\mathbf{H}$ is finite and that it includes at least one hypothesis for each event. ...
>
> Thus, $P(A, B)$ is the judged probability that A rather than B holds assuming that one and only one of them is valid. Obviously, A and B may each represent and explicit or an implicit disjunction. The extensional counterpart of $P(A, B)$ in the standard theory is the conditional probability $P(A'|A' \cup B')$. The present treatment is nonextensional because it assumes that the probability judgment depends on the descriptions A and B, not just the events A' and B'. We wish to emphasize that the present theory applies to the hypotheses entertained by the judge, which do not always coincide with the given verbal descriptions. A judge presented with an implicit disjunction may, nevertheless, think about it as

an explicit disjunction and vise versa. *pg. 584*

Notice first that Tversky and Koehler are interpreting **T** in the same manner of classical probability theory as "states of the world," where exactly one state obtains. This is reasonable for scientific applications of probability, but I and others believe it is inappropriate for modeling subjective probabilities of *belief.* (For example, subjective probability estimates of counterfactual propositions cannot be modeled as the kind of "states of the world" Tversky and Koehler had in mind. In general "beliefs" require a different kind of modeling than "facts.") Second, notice that near the end of the above quotation they say, "The present treatment is nonextensional because it assumes that the probability judgment depends on the descriptions A and B, not just the events A' and B'. We wish to emphasize that the present theory applies to the hypotheses entertained by the judge, which do not always coincide with the given verbal descriptions." I interpreted the last quote as saying that the probability judgment depends on interpretations that the judge gives to the descriptions. So why should Tversky and Koehler concerned with "extensions" at all, since it doesn't enter in the judging of probabilities? I believe part of the reason is practical: Most of the Support Theory's experimental research is based on giving different descriptions of the same event for probabilistic judgment and finding that different probabilities are given to descriptions and there is some consistency to how the probabilities vary with descriptions, e.g., the more fine a description of a partition of an event is the higher is the sum of probabilities of descriptions of the elements of the partition. To carry out this strategy, "descriptions of the same event" have to be defined, and this is done in terms of "extension." The New Foundation provides a more psychologically based foundation for the above kind of experimental paradigm, avoiding the metaphysics inherent in Tversky and Koehler's use of "extension." This is done by including a natural language semantics of descriptions in place of their realistic interpretation of the descriptions. Another part of the reason for including "extension" in their formulation is that it provides them a means for describing various experimental results as normative failures, that is, as "failures of extensionality," and a means for giving reasons for such "failures." The New Foundation is not founded on normative concepts, and in it such experimental results cannot be looked at as "failures" of anything.

A primary nonextensional aspect of the New Foundation is its use of the modal concept of "clear instance" (Definition 6). In terms of this modal concept, "unrealized clear instance," "external clear instance," and "strongly ambiguous clear instance" are defined and used to describe features of probability judgments that deviate from the standard probability theory.

One of the key features of the New Foundation is the sharp distinction between the semantic processing employed in the use of language and cognitive processing employed specifically for probability judgments. The lack of such a distinction has, in my view, generated some misunderstanding and controversies in the literature.

It is essentially a tautology to say that when the participant assigns significantly different probabilities α and γ, she "understands" α differently from γ. However, it is important to characterize the nature of the difference. For the purposes of the present discussion, the difference may arise in two ways: (1) $\mathbf{SR}(\alpha) \neq \mathbf{SR}(\gamma)$, and (2) $\mathbf{SR}(\alpha) = \mathbf{SR}(\gamma)$ but $\mathbf{CR}(\alpha) \neq \mathbf{CR}(\gamma)$. (1) is basically a problem of psycholinguistics and is uninteresting from the point of view of probability estimation. (2), of course, is the basis discussed above for much of the experimental research of Support Theory. The problem is that both (1) and (2) are mental, and therefore can only be assessed by indirect means. In particular, (1) cannot be distinguished from (2) by only looking at the estimated probabilities of α and γ. The Cognitive Model suggests that in situations where the validity of $\mathbf{SR}(\alpha) = \mathbf{SR}(\gamma)$ is in question, the validity should be evaluated directly through a psycholinguistic paradigm, rather than as part of a probabilistic judgment paradigm.

In the simplified form of the semantics presented here, the logical connectives "and", "or", and "not," which act on propositional descriptions, are semantically interpreted as, respectively, $\cap$, $\cup$ and $-$, which act on sets. The Cognitive Model does not currently interpret these logical connectives. Obviously, it would be greatly enhanced with such interpretations. But first a great deal of empirical research is needed to establish basic facts about them and their relationship to the cognitive complementation operator $\ominus$.

The New Foundation was designed for the kinds of studies generally conducted by support theorists. Some probabilistic estimation tasks do not fall into this paradigm. For example, presenting a partition and asking the participant "to assign probabilities to each of the alternatives so that the probabilities add to 1." The central feature of the Cognitive Model is that the participant creates a complement $\ominus A$ of a cognitive event A and assigns probabilities through some comparison between A and $\ominus A$. What distinguishes the Cognitive Model from other models in the literature is that the created complement is $\ominus A$ instead of $-A$ (except possibly in situations with events that have minuscule support.)

Narens (2003) provides a qualitative generalization of probability theory designed to allow for the kind of ambiguity inherent in the Ellsberg paradox (Ellsberg, 1961). Narens shows that his generalization also accounts several phenomena of Support Theory.

The generalization is arrived at by first providing a qualitative axiomatization of a version of conditional probability. This is done in such a way that one of the axioms captures a consequence of conditional probability that some theorists have suggested to be invalid for a general, rational theory of belief. This axiom is eliminated, and the most general quantitative model corresponding to the remaining axioms is found. The result is a generalized probability function $\mathbb{B}(a|b)$ having the form,

$$\mathbb{B}(a|b) = P(a|b)v(a)\,,$$

where P is a uniquely determined conditional probability function and v is a function into $\mathbb{R}^+$ that is unique up to multiplication by a positive real. Narens uses this result to provide alternative accounts of various findings about psychological judgments of probability.

The axiom of qualitative conditional probability that is eliminated is called *Binary Symmetry.* It is formulated as follows, where "$\sim$" stands for "is equally likely as" and "$(e|f)$" for the conditional event of "e occurring given f has occurred:" Suppose a, b, c, and d are mutually disjoint, nonempty events and

$$(a|a \cup b) \sim (b|a \cup b) \text{ and } (c|c \cup d) \sim (d|c \cup d)\,. \tag{8}$$

Then

$$(a|a \cup b) \sim (c|c \cup d) \tag{9}$$

and

$$(a|a \cup c) \sim (b|b \cup d)\,. \tag{10}$$

Note that if uncertainty is measured by the probability function P and Equation 8 holds, then $P(a) = P(b)$ and $P(c) = P(d)$, from which Equations 9 and 10 follow.

Also note that if a, b, c, and d in Equation 8 are chosen in an Idson et al. paradigm so that a and b have strong support and c and d have weak support, then the results of Idson et al. would suggest that the participant would consider $(a|a \cup b)$ to be more likely than $(c|c \cup d)$, thus invalidating Equation 9.

One interpretation Narens gives to his generalized probability theory is that uncertain events contains two dimensions, a probabilistic dimension, and another dimension that he calls *definiteness.* Definiteness has a number of substantive interpretations, one of which is the inverse of ambiguity (i.e., greater ambiguity = less definiteness, and visa versa).

Narens (2004) applies the generalized theory of probability to a normative theory empirical events. It is argued that empiricalness founded on verifiability and refutability require empirical events to have the logical structure of the same type of the event space used in the New Foundation. In this situation, "$\ominus\, a$" is interpreted as the event, "The event a is empirically refutable." As a consequence, the sum of the probabilities of an empirical event and its *empirical negation* is ≤ 1, i.e., superadditivity obtains.

The $\ominus$ operator allows greater flexibility for modeling than the $-$ operator. This permits greater flexibility for interpreting results and allows for a wider range of phenomena. For example, it is permitted to have a situation where $\mathbf{SR}(\alpha) = \mathbf{SR}(\alpha_1)$, $\mathbf{CR}(\alpha) = \mathbf{CR}(\alpha_1)$, $\mathbf{SR}(\beta) = \mathbf{SR}(\beta_1)$, $\mathbf{CR}(\beta) = \mathbf{CR}(\beta_1)$, and $\mathbb{P}(\alpha|\beta) \neq \mathbb{P}(\alpha_1|\beta_1)$. This may occur when $\alpha|\beta$ gives rise to a different topology than $\alpha_1|\beta_1$, producing different complementation operations used in judging the probabilities of $\alpha|\beta$ and $\alpha_1|\beta_1$.

The $\ominus$ operator assumes an underlying topology. Topology is a very general subject with spaces that vary enormously in terms of their features. The Cognitive Model makes no special assumptions about the underlying topology. This allows for the Model to be improved through the incorporation of additional topological assumptions. These can be either general assumptions applying to all cognitive representations, or special assumptions that apply to restricted class of representations, for example, representations based on representativeness.

As discussed later in Subsection 4.2, the topological event space of the New Foundation shares many features with intuitionistic logic. The principle difference is that intuitionistic logic is based primarily on an implication connective that is not part of the topological event space of the New Foundation. Logical implications do not play a role in theory of probabilistic judgments presented here, because probabilities are computed directly in terms of the supports for a cognitive event and its cognitive complement.

The algebraic structure of topological event spaces is discussed in detail in Section 4. An investigation of the kinds of event spaces that generalize boolean lattices of events is also undertaken in Section 4, and it is concluded that for Support Theory Phenomena that event spaces like those employed in the New Foundation are most likely the best event spaces for this purpose.

Throughout this article the New Foundation has been used to provide a basis for understanding empirical results about human probability judgments, sometimes providing new, alternative explanations. It is based on concepts from a non-classical logic and from topology. Together these two subject matters provide a new and rich language for organizing and describing theoretical and empirical research about probability judgments.

4 Lattice Formulation

In the preface to his first edition of *The Logic of Chance* (1866), John Venn observed the following about the development of probability theory:

> Probability has been very much abandoned to mathematicians, who as mathematicians have generally been unwilling to treat it thoroughly. They have worked out its results, it is true, with wonderful acuteness, and the greatest ingenuity has been shown in solving various problems that arose, and deducing subordinate rules. And that was all that they could in fairness be expected to do. Any subject which has been discussed by such men as Laplace and Poisson, and on which they have exhausted their powers of analysis, could not fail to be profoundly treated, so far as it fell with in their province. But from this province the real principles of the science have generally been excluded, or so meagrely discussed that they had better have been omitted altogether. Treating the subject as mathematicians such writers have naturally taken it up at the point where their mathematics would best come into play, and that of course has not been at the foundations. In the work of most writers on the subject we should search in vain for anything like a critical discussion of the fundamental principles upon which its rules rest, the class of enquiries to which it is most properly applicable, or the relation it bears to Logic and the general rules of inductive evidence.

By substituting in the above quotation "probability judgments" for "probability," "experimentalists" for "mathematicians," "experiments" for "mathematics," and "Tversky and colleagues" for "Laplace and Poisson," the resulting description provides, in my opinion, an accurate account today's probability judgment literature.

In mid 19th Century, George Boole provided a partial foundation for probability theory by identifying the propositional calculus of logic with a boolean algebra of sets. This paved the way for the integration of measure theory into probability theory and the current standard approach to probability as formulated by Kolmogorov (1933). I believe analogous developments are needed for the scientific understanding of human probability judgments, and I see this article's use of topological algebras of events as event spaces for probability judgments as a step in this direction.

One way of investigating the possibilities for modeling the logical structure inherent probability judgments is through the use of lattice theory. In this section it is shown that a topological algebra of events has the same kind of structure as the event space generated by the intuitionistic propositional calculus. This section shows how the topological algebra of events fit into a lattice theoretic analysis of event spaces that allow for the existence of probability functions. Also in this section, lattice theoretic formulations are given the event spaces of classical probability theory and quantum mechanics. These lattice formulations allow for simple and direct comparisons of the event spaces of standard probability theory, of the New Foundation, and of quantum mechanics. Through the use of these comparisons and other lattice theory results, it is argued that topological event space appear to be best kind of event space for modeling the kind of probabilistic phenomena considered by support theorists.

4.1 Distributive Lattices

Definition 10 $\preceq$ is said to be *partial ordering on* A if and only if A is a nonempty set and the following three conditions hold for each a, b, and c in A:

(*i*) $a \preceq a$;

(*ii*) if $a \preceq b$ and $b \preceq a$, then $a = b$; and

(*iii*) if $a \preceq b$ and $b \preceq c$, then $a \preceq c$. □

Definition 11 $\langle A, \preceq, u, z \rangle$ is said to be a *lattice (with unit element* u *and zero element* z*)* if and only if

(*i*) $\preceq$ is a partial ordering on a;

(*ii*) for each a and b in A, there exists a unique c in A, called the *join* of a and b and denoted by $a \sqcup b$, such that $a \preceq a \sqcup b$, $b \preceq a \sqcup b$, and for all d in A,

$$\text{if } a \preceq d \text{ and } b \preceq d\,, \text{ then } a \sqcup b \preceq d\,;$$

(*iii*) for each a and b in A, there exists a unique element c in A, called the *meet* of a and b and denoted by $a \sqcap b$, such that $a \sqcap b \preceq a$, $a \sqcap b \preceq b$, and for all d in A,

$$\text{if } d \preceq a \text{ and } d \preceq b\,, \text{ then } d \preceq a \sqcap b\,;$$

and

(*iv*) for all a in A, $z \preceq a$ and $a \preceq u$. □

Let $\langle A, \preceq, u, z \rangle$ be a lattice. Then it easily follows that $\sqcup$ and $\sqcap$ are commutative and associative operations on A.

Convention 5 A lattice $\langle A, \preceq, u, z\rangle$ is often written as

$$\langle A, \preceq, \sqcup, \sqcap, u, z\rangle . \quad \square$$

Definition 12 Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle$ be a lattice. Then $\mathfrak{A}$ is said to be *distributive* if and only if for all a, b, and c in A,

$$a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c) . \quad \square$$

Example 2 (Distributive Lattice of Propositions) The propositional calculus from logic is an important example of a distributive lattice. Let $\rightarrow$, $\leftrightarrow$, $\vee$, and $\wedge$ stand for, respectively, the logical connectives of implication ("if ... then"), logical equivalence, ("if and only if"), disjunction ("or"), and conjunction, ("and"). Consider two propositions α and β to be equivalent if and only if $\alpha \leftrightarrow \beta$ is a tautology. Equivalent propositions partition the the set of propositions into equivalence classes. Let A be the set of equivalence classes. The equivalence class containing a tautology is denoted by u, and the equivalence class containing a contradiction is denoted by z. By definition, equivalence class $a \leq$ equivalence class b if and only if for some elements α in a and β in b, $\alpha \rightarrow \beta$ is a tautology. It is easy to check that $\leq$ is a partial ordering on A.

Let a and b be arbitrary elements of A, α and β be respectively arbitrary elements of a and b, j be the equivalence class of $\alpha \vee \beta$, and m be the equivalence class of $\alpha \wedge \beta$. Then it is easy to check that j is the join of a and b, m is the meet of a and b, and

$$\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle ,$$

is a distributive lattice. $\square$

Example 3 (Distributive Lattice of All Subsets of a Set) Let X and a nonempty set and $\mathcal{X}$ be the set of all subsets of X. Then $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, X, \varnothing\rangle$ is a distributive lattice.

Definition 13 Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle$ be a lattice and a be an arbitrary element of A.

b in A is said to be the $\sqcap$-*complement of* a if and only if

(i) $b \sqcap a = z$, and

(ii) for all c in A, if $c \sqcap a = z$ then $c \leq b$.

It easily follows that if the $\sqcap$-complement of a exists, then it is unique.

$\mathfrak{A}$ is said to be $\sqcap$-*complemented* if and only if each of its elements has a $\sqcap$-complement. $\sqcap$-complementation is an important concept in lattice theory and logic, and it plays a central role in this article.

The $\sqcap$-complement of a, when it exists, is denoted by $\ominus A$. $\square$

The following related concept of "complement" plays a minor role in this article.

Definition 14 The $\sqcup$-*complement,* c, of a is defined by

(i') $c \sqcup a = u$, and

(ii'') for all d in A, if $d \sqcup a = u$ then $d \geq c$,

and $\mathfrak{A}$ is said to be $\sqcup$-*complemented* if and only if each of its elements has a $\sqcup$-complement. $\square$

Theorem 3 *Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z \rangle$ be a distributive lattice, a, an arbitrary element of A, b the $\sqcap$-complement of a, and c the $\sqcup$-complement of a. Then $b \leq c$.*

Proof:

$$b = b \sqcap u = b \sqcap (a \sqcup c) = (b \sqcap a) \sqcup (b \sqcap c) = z \sqcup (b \sqcap c) = b \sqcap c \leq c\,. \ \square$$

Definition 15 Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z \rangle$ be a lattice and a be an element of A. Then an element b in A is said to be a *complement* of a if and only if

$$a \sqcap b = z \ \text{ and } \ a \sqcup b = u\,.$$

$\mathfrak{A}$ is said to be *complemented* if and only if a complement exists for each element in A. $\mathfrak{A}$ is said to be *uniquely complemented* if and only if each element in a has a unique complement. $\square$

Definition 16 A complemented distributive lattice is called a *boolean lattice.* $\square$

The following theorem is easy to show.

Theorem 4 *Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z \rangle$ be a boolean lattice and a be an arbitrary element of A. Then the following two statements are true:*

1. *a has a unique complement.*

2. *The complement of a = the $\sqcap$-complement of a = the $\sqcup$-complement of a.* $\square$

Convention 6 Usually boolean lattices $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle$ are described by the notation,

$$\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, -, u, z\rangle ,$$

explicitly listing the boolean complementation operation $-$. $\square$

Example 4 (Boolean Lattices of Propositions and Subsets) Example 2 reveals the distributive lattice structure of the propositional calculus. Using the notation and concepts of Example 2, it is easy to show that the distributive lattice of propositions $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, -, u, z\rangle$ is boolean with $-a =$ the equivalence containing $\neg\,\alpha$, where α is in a and $\neg$ is logical negation. Example 3 of a distributive lattice of all subsets of a nonempty set X,

$$\langle \mathcal{X}, \subseteq, \cup, \cap, -, X, \varnothing\rangle ,$$

is boolean with $-$ being set-theoretic complementation. $\square$

4.1.1 Representation Theorems of Stone

Definition 17 $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle$ is said to be a *lattice of sets* if and only if $\mathfrak{A}$ is a lattice, u is a nonempty set, each element of A is a subset of u, $z = \varnothing$, $\preceq\, = \,\subseteq$, $\sqcup = \cup$, and $\sqcap = \cap$. $\square$

Note that it follows from properties of $\cap$ and $\cup$ that a lattice of sets is automatically distributive. The following two theorems are due to Stone (1936, 1937).

Theorem 5 (Representation Theorem for Distributive Lattices) *Suppose* $\mathfrak{A}$ *is a lattice. Then the following two statements are equivalent:*

1. $\mathfrak{A}$ *is distributive.*
2. $\mathfrak{A}$ *is isomorphic to a lattice of sets.* $\square$

Theorem 6 (Stone's Representation Theorem) *Suppose* $\mathfrak{A}$ *is a lattice. Then the following two statements are equivalent:*

1. $\mathfrak{A}$ *is boolean.*
2. $\mathfrak{A}$ *is isomorphic to a lattice of sets, and under the isomorphism the complementation operation of* $\mathfrak{A}$ *is mapped onto the set-theoretic complementation operation.* $\square$

Probability functions often have as their domain a set of propositions (formulated in classical logic). Mathematicians and others often interpret such situations as probability functions on an event space, where the events form a boolean lattice of sets. For most probabilistic situations, the Stone Representation Theorem justifies this practice.

4.1.2 $\cap$-Complemented Open Set Lattices

$\sqcap$-complemented distributive lattices may be looked at as describing a logic that applies to more situations than classical propositional logic. This allows for a coherent interpretation of "probability" that generalizes classical probability theory, and, as it is argued in this article provides for a better accounting of human probability judgments. $\sqcap$-complemented distributive lattices also have representations as lattices of events. But unlike the boolean case, these events require additional structure.

Definition 18 $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing \rangle$ is said to be a $\cap$-*complemented open set lattice of* $\mathcal{U}$ if and only $\langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing \rangle$ is a lattice of sets, $\mathcal{U}$ is a topology, $\mathcal{X} \subseteq \mathcal{U}$, and *with respect to* $\mathcal{U}$,

$$\ominus A = \mathsf{int}(\mathsf{cl}(-A)),$$

for all A in $\mathcal{X}$.

$\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing \rangle$ is said to be a $\cap$-*complemented open set lattice* if and only if for some $\mathcal{U}$, $\mathfrak{X}$ is a $\cap$-complemented open set lattice of $\mathcal{U}$. $\square$

Theorem 7 *Suppose* $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing \rangle$ *is a* $\cap$-*complemented open set lattice in the sense of Definition 18. Then* $\ominus$ *is the* $\cap$-*complementation operation of* $\mathfrak{X}$*, and thus* $\mathfrak{X}$ *is a* $\cap$-*complemented set lattice in the sense of Definition 13.*

Proof. Immediate from Definitions 18 and 13. $\square$

Example 5 (Topological Lattices on Open Sets) Let $\mathcal{U}$ be a topology with universe U. Then

$$\langle \mathcal{U}, \subseteq, \cup, \cap, \ominus, U, \varnothing \rangle$$

is a $\cap$-complemented open set lattice. More generally,

$$\langle \mathcal{V}, \subseteq, \cup, \cap, \ominus, U, \varnothing \rangle$$

is a $\cap$-complemented open set lattice, where $\mathcal{V}$ is a subset of of $\mathcal{U}$ such that U and $\varnothing$ are in $\mathcal{V}$, and whenever E and F are in $\mathcal{U}$, $E \cup F$, $E \cap F$, and $\ominus E$ are in $\mathcal{V}$. $\square$

Theorem 8 *Each $\sqcap$-complemented distributive lattice is isomorphic to a $\cap$-complemented open set lattice (Definition 18).*

Proof. Let $\mathfrak{L} = \langle \mathcal{L}, \subseteq, \sqcup, \sqcap, \ominus_{\mathfrak{L}}, u, z\rangle$ be an arbitrary $\sqcap$-complemented distributive lattice. By Theorem 5, let ϕ be an isomorphism of $\langle \mathcal{L}, \subseteq, \sqcup, \sqcap, u, z\rangle$ onto a lattice of sets, $\langle \mathcal{X}, \subseteq, \cup, \cap, X, \varnothing\rangle$. Define the following operation $\ominus$ on $\mathcal{X}$: For all A, B and C in $\mathfrak{X}$,

$$A \ominus B \;=\; C \;\text{ iff }\; \phi^{-1}(A) \ominus_{\mathfrak{L}} \phi^{-1}(B) = \phi^{-1}(C)\,.$$

It then easily follows that

$$\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing\rangle$$

is isomorphic to $\mathfrak{L}$, and is a $\mathfrak{X}$ is $\cap$-complemented lattice. Thus to show the theorem, we need to only find a topology $\mathcal{U}$ such that $\mathfrak{X}$ is a $\cap$-complemented open set lattice of $\mathcal{U}$; that is, find a topology $\mathcal{U}$ such that $\ominus_1$ defined on $\mathcal{X}$ by,

$$\text{for all } A \text{ in } \mathcal{X},\ \ominus_1 A = \mathsf{int}(\mathsf{cl}(-A)),$$

is the operation $\ominus$ of $\cap$-complementation for $\mathfrak{X}$.

Let

$$\mathcal{U} = \{\bigcup \mathcal{F} \,|\, \mathcal{F} \neq \varnothing \text{ and } \mathcal{F} \subseteq \mathcal{X}\}\,.$$

Then, because for nonempty subsets $\mathcal{F}$ and $\mathcal{G}$ of $\mathcal{X}$,

$$\bigcup \mathcal{F} \,\cap\, \bigcup \mathcal{G} \;=\; \bigcup \{F \cap G \,|\, F \in \mathcal{F} \text{ and } G \in \mathcal{G}\}\,,$$

it easily follows that $\mathcal{U}$ is a topology and $\mathfrak{X}$ is an open set lattice of $\mathcal{U}$.

Let A be an arbitrary element of $\mathcal{X}$ and, with respect to $\mathcal{U}$, let

$$\ominus_1 A = \mathsf{int}(\mathsf{cl}(-A))\,.$$

Because $\ominus A$ is in $\mathcal{X}$, it is in $\mathcal{U}$, i.e., $\ominus A$ is open. Therefore $-A$ is closed, and it follows that

$$\mathsf{cl}(-A)) = -A\,.$$

Therefore, $A \cap \ominus_1 A = \varnothing$. Because $\ominus A \subseteq -A$ and $\ominus_1$ is the interior of $\mathsf{cl}(-A))$, it follows that $\ominus A \subseteq \ominus_1 A$. Thus to show the theorem, it needs to only be shown that $\ominus_1 A \subseteq \ominus A$. This will be done by contradiction. Suppose x is in $\ominus_1 A$ but x is not in $\ominus A$. Because $\ominus_1$ is open, let $\mathcal{F} \subseteq \mathcal{X}$ be such that

$$\ominus_1 A = \bigcup \mathcal{F}\,.$$

Then because x is in $\ominus_1 A$, let F in $\mathcal{F}$ be such $x \in F$. Then $F \in \mathcal{X}$. Because $\ominus_1 A \subseteq -A$, $F \cap A = \varnothing$. Therefore, because $\ominus$ is the operation $\cap$-complementation for $\mathfrak{X}$, $F \subseteq \ominus A$, and therefore x is in $\ominus A$, a contradiction. $\square$

4.2 Relationship to Intuitionistic Logic

The Dutch mathematician L. L. J. Brouwer introduced intuitionism as an alternative form of mathematics. It followed from his philosophy of mathematics that the methods of derivation used in intuitionistic mathematical theorems were in theory non-formalizable. However, the intuitionistic theorems he produced displayed sufficient regularity in their proofs that an axiomatic approach to his methods of proof appeared feasible. Heyting (1930) produced such an axiomatic approach. Today logics that satisfy equivalents of Heyting's axiomatization are generically called *intuitionistic logic.* Later researchers showed intuitionistic logic to have other interpretations. For example, Kolmogorov (1932) showed that it had the correct formal properties for a theory for mathematical constructions, and Gödel (1933) showed that it provided a logical foundation for proof theory of mathematical logic. Kolmogorov and Gödel achieved their results by giving interpretations to the logical primitives that were different from Heyting's. Narens (2004) provides an interpretation for the intuitionistic connectives of conjunction, disjunction, and negation for a probabilistic theory of empirical events, and the New Foundation provides an interpretation of those primitives as part of an event space for mental representations of propositions for probabilistic judgment.

In a manner similar to Example 2 (distributive lattice of prpositions), the intuitionistic version of "if and only if" produces an equivalence relation on intuitionistic propositions, and the induced algebra on its equivalence classes yields a lattice, called a *pseudo boolean algebra.*

Definition 19 $\mathfrak{P} = \langle \mathcal{P}, \subseteq, \cup, \cap, \Rightarrow, X, \varnothing \rangle$, where $\Rightarrow$ is a binary operation on $\mathcal{P}$, is said to be a *pseudo boolean algebra of subsets* if and only if the following three conditions hold for all A and B in $\mathcal{P}$:

(1) $\mathfrak{P} = \langle \mathcal{P}, \subseteq, \cup, \cap, X, \varnothing \rangle$ is a lattice of sets.

(2) $A \cap (A \Rightarrow B) \subseteq B$.

(3) For all C in $\mathcal{P}$, if $A \cap C \subseteq B$ then $C \subseteq (A \Rightarrow B)$.

$\Rightarrow$ is called the operation of *relative pseudo complementation.* □

Let $\mathfrak{B} = \langle \mathcal{B}, \subseteq, \cup, \cap, -, X, \varnothing \rangle$ be a boolean algebra of subsets. Then it is easily follows that

$$\langle \mathcal{B}, \subseteq, \cup, \cap, \Rightarrow, X, \varnothing \rangle$$

is pseudo boolean algebra of subsets, where for all A and B in $\mathcal{B}$,

$$A \Rightarrow B \ = \ -A \cup B.$$

Thus each boolean algebra of subsets is a pseudo boolean algebra of subsets.

In a boolean algebra of subsets, $(-\,A) \cup B$ corresponds to the implication operation, $\rightarrow$, of classical logic. Similarly, $\Rightarrow$ is the pseudo boolean algebra of subsets version of the implication operation of intuitionistic logic. In a pseudo boolean algebra of subsets, the negation operator $\neg$ has the following definition:

$$\neg\, A \ = \ A \Rightarrow \varnothing.$$

It is not difficult to verify that $\mathbb{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \neg, X, \varnothing \rangle$ is a $\cap$-complemented open set lattice with $\neg\ =\ \ominus$. In terms of the corresponding intuitionistic logic, $\neg\ (=\ominus)$ corresponds to the negation operator of the logic.

$\cap$-complemented open set lattices are lattices of subsets with a weakened form of complementation that has some of the fundamental properties of intuitionistic negation. They differ from pseudo boolean algebras of subsets in that the operation of relative pseudo complementation, $\Rightarrow$ is not a primitive.

4.3 Properties of $\cap$-complemented open set lattices

Theorem 9 *Suppose $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing \rangle$ is a $\cap$-complemented open set lattice. Then the following eight statements are true for all A and B in $\mathcal{X}$:*

1. $\ominus X = \varnothing$ *and* $\ominus \varnothing = X$.
2. *If* $A \cap B = \varnothing$, *then* $B \subseteq \ominus A$.
3. $A \cap \ominus A = \varnothing$.
4. *If* $B \subseteq A$, *then* $\ominus A \subseteq \ominus B$.
5. $A \subseteq \ominus \ominus\, A$.
6. $\ominus A = \ominus \ominus \ominus A$.
7. $\ominus(A \cup B) = \ominus A\ \cap\ \ominus B$.
8. $\ominus A \cup \ominus B \subseteq \ominus(A \cap B)$.

Proof. Statements 1 to 3 are immediate from Definition 13.

4. Suppose $B \subseteq A$. By Statement 3, $A \cap \ominus A = \varnothing$. Thus $B \cap \ominus A = \varnothing$. Therefore, by Statement 2, $\ominus A \subseteq \ominus B$.

5. By Statement 3, $\ominus A \cap A = \varnothing$. Thus by Statement 2, $A \subseteq \ominus \ominus A$, showing Statement 5.

6. By Statement 5, $A \subseteq \ominus \ominus A$. Thus by Statement 4,

$$\ominus \ominus \ominus A \subseteq \ominus A.$$

However, by Statement 5,

$$\ominus A \subseteq \ominus \ominus (\ominus A) = \ominus \ominus \ominus A.$$

Therefore, $\ominus A = \ominus \ominus \ominus A$.

7. By Statement 3, (i) $\ominus A \subseteq -A$ and (ii) $\ominus B \subseteq -B$. Thus

$$\ominus A \cap \ominus B \subseteq -A \cap -B = -(A \cup B).$$

Therefore $(A \cup B) \cap (\ominus A \cap \ominus B) = \varnothing$, and thus by Statement 2,

$$\ominus A \cap \ominus B \subseteq \ominus(A \cup B). \tag{11}$$

Because $A \subseteq A \cup B$ and $B \subseteq A \cup B$, it follows from Statement 4 that

$$\ominus(A \cup B) \subseteq \ominus A \text{ and } \ominus(A \cup B) \subseteq \ominus B.$$

Therefore

$$\ominus(A \cup B) \subseteq \ominus A \cap \ominus B. \tag{12}$$

Equations 11 and 12 show that

$$\ominus(A \cup B) = \ominus A \cap \ominus B.$$

8. From

$$A \cap B \subseteq A \text{ and } A \cap B \subseteq B,$$

it follows from Statement 4 that

$$\ominus A \subseteq \ominus(A \cap B) \text{ and } \ominus B \subseteq \ominus(A \cap B),$$

and thus

$$\ominus A \cup \ominus B \subseteq \ominus(A \cap B). \quad \square$$

The following theorem gives some fundamental properties of boolean lattices of sets that fail for some $\cap$-complemented open set lattices.

Theorem 10 *There exists a $\cap$-complemented open set lattice $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \cup, \cap, \ominus, X, \varnothing \rangle$ such that the following three statements are true about $\mathfrak{X}$.*

1. *For some A in $\mathcal{X}$, $A \cup \ominus A \neq X$.*

2. *For some A in $\mathcal{X}$, $\ominus \ominus A \neq A$.*

3. *For some A and B in $\mathcal{X}$, $\ominus(A \cap B) \neq \ominus A \cup \ominus B$.*

Proof. Let X be the set of real numbers, $\mathcal{X}$ be the be usual topology on X determined by the usual ordering on X, C be the infinite open interval $(0, \infty)$, and D be the infinite open interval $(-\infty, 0)$. Then the reader can verify that Statement 1 follows by letting $A = C$, Statement 2 by letting $A = C \cup D$, and Statement 3 by letting $A = C$ and $B = D$. $\square$

4.4 Quantum Logic

Boolean lattices and $\cap$-complemented distributive lattices are two examples of event spaces of rich propositional logics that are useful for formulating ideas about and constructing models of Support Theory phenomena. Interestingly, they both allow the existence of (generalized) probability functions.

Definition 20 Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z \rangle$ be a lattice. A probability function $\mathbb{P}$ on $\mathfrak{A}$ has the formal properties of a finitely additive measure, that is,

(1) $\mathbb{P}(u) = 1$ and $\mathbb{P}(z) = 0$;

(3) for all a and b in A, if $a \sqcap b = z$, then $\mathbb{P}(a \sqcup b) = \mathbb{P}(a) + \mathbb{P}(b)$.

For boolean lattices, Condition 3 is equivalent to

$$\text{For all } a \text{ and } b \text{ in } A, \mathbb{P}(\text{a}) + \mathbb{P}(\text{b}) = \mathbb{P}(a \sqcup b) + \mathbb{P}(a \sqcap b). \qquad (13)$$

Obviously Equation 13 implies Condition 3. The equivalence of Condition 3 and Equation 13 uses the that $\mathfrak{A}$ is complemented and distributive. Both Condition 3 and Equation 13 can be used for formulating generalizations of probability functions that apply to lattices more general than the boolean ones. Because Equation 13 retains more structure of probability functions, it is used in the generalizations discussed in this section, and is referred to as simply a "probability function" when applied to lattices.

There is an extensive literature of a third kind of propositional logic and event space that grew out von Neumann's mathematical foundations of quantum mechanics known as "quantum logic." Its event space also allows for the existence of probability functions. Outside of these three, there are, to my knowledge, no other rich interesting propositional calculi with event spaces that support probability functions. Theorems below provide a reasonably good explanation for this.

We first note that it is exceedingly difficult to show instances of uniquely complemented lattices that are not boolean. That is, unique complementation with a little other structure often implies distributivity. The Birkhoff-Ward Theorem below is an example of this.

Definition 21 Let $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z \rangle$ be a lattice. a is said to be an *atom* of $\mathfrak{A}$ if and only if a is in A and for all b in A, if $b \leq a$ and $b \neq z$, then $b = a$. $\mathfrak{A}$ is said to be *atomic* if and only if for each element e in A there exist an atom f of A such that $f \leq e$.

Theorem 11 (Birkhoff-Ward Theorem) *An atomic uniquely complemented lattice is distributive.*

Proof. See Saliĭ (1988) pg. 41.

This is only one of many theorems characterizing boolean lattices in terms of unique complementation. Saliĭ (1988) who presents many examples of this writes,

> The results of the preceding section cast great doubt on the existence of nondistributive uniquely complemented lattices. At the end of the 1930's the doubt (with a much smaller set of corroborating facts) gained widespread conviction. Thus the appearance in 1945 of [Dilworth's Theorem] was completely unexpected. *(pg. 51)*

Theorem 12 (Dilworth's Theorem) *Any lattice can be embedded in some uniquely complemented lattice.*

Proof. See Saliĭ (1988) pg. 51.

Dilworth's proof and subsequent proofs of his theorem did not produce an explicit example of a non-distributive complemented lattice. Saliĭ (1988) comments,

> *We know that nondistributive uniquely complemented lattices exist, but a present we do not have a single explicit example.* Such lattices have not yet been encountered in mathematical practice.

From Saliĭ's observation and theorems given below, I conclude that nondistributive uniquely complemented lattices are very likely to be of little use for providing a foundation for support theory.

The following kind of complement has many characteristics of the boolean complement operation, and because of this may be useful in modeling probabilistic phenomena.

Definition 22 $'$ is said to be the operation of *orthocomplementation* on the lattice $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle$ if and only if for each a and b in A,

(i) a' is the complement of a,

(ii) $a'' = a$, and

(iii) If $a \leq b$ then $b' \leq a'$

The equivalents to Conditions (i) and (ii) above for $\cap$-complementation, i.e., $a \cup \ominus a = u$ and $a = \ominus \ominus a$ fail for some $\cap$-complemented distributive lattices (Theorem 10). The equivalent of Condition (iii), i.e., if $a \leq b$ then $\ominus b \leq \ominus a$, holds for all $\cap$-complemented lattices (Statement 4 of Theorem 9).

Conditions (i), (ii), and (iii), above of an othrocomplement implies De Morgan's laws, i.e.,

$$(a \sqcap b)' = a' \sqcup b' \text{ and } (a \sqcup b)' = a' \sqcap b' .$$

Thus orthocomplementation has the characteristic and essential algebraic features of boolean complementation. However, the following theorem shows that it adds little new when it is a unique complementation operation.

Theorem 13 *Let* $\mathfrak{A} = \langle A, \preceq, \sqcup, \sqcap, u, z\rangle$ *be a lattice with a unique complementation operation* $'$ *such that for all* a *and* b *in* A,

$$\text{if } a \leq b \text{ then } b' \leq a' .$$

Then $\mathfrak{A}$ *is a boolean lattice.*

Proof. See Corollary 1 on pg. 48 of Saliĭ (1988).

I view Support Theory as a generalization of probability theory, in the sense that it is possible for probability judgments in some cases to have the form of classical probability functions. Because of this, it seems reasonable to restrict the modeling of Support Theory phenomena to classes of lattices that have boolean lattices as a special (or degenerate) case.

It is well-known that lattices that have probability functions that satisfy Equation 13 must satisfy the following condition:

Modularity: for all a, b, and c in A, if $a \leq b$, then $a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c)$.

Modularity is a generalization of distributivity. Birkhoff and von Neumann (1936) use non-distributive, complemented, modular lattices to describe the propositional logic of quantum mechanics. Birkhoff and von Neumann wrote,

> Whereas logicians have usually assumed that properties of negation were the ones least able to withstand a critical analysis, the study of mechanics points to the distributive identities as the weakest link in the algebra of logic. *(pg. 839)*

The following theorem shows that non-distributive, complemented, modular lattices must have some elements that do not have unique complements:

Theorem 14 (Birkhoff-von Neumann Theorem) *A modular uniquely complemented lattice is distributive.*

Proof. See Saliĭ (1988) pg. 40.

In their development of quantum logic, Birkhoff and von Neumann used non-distributive orthocomplemented modular lattices as propositional logics for their event spaces. Such event spaces permit probability functions $\mathbb{P}$ with $\mathbb{P}(A') = 1 - \mathbb{P}(A)$, where A' is the othrocomplement of A. In von Neumann's foundational development of quantum mechanics, closed subspaces of an infinite dimensional hilbert space was central to his construction. Subspaces of a finite dimensional hilbert space also form an orthocomplemented modular lattice. For the purposes of this discussion, we use finite dimensional case to illustrate our ideas. Similar results and conclusions hold for the infinite dimensional case.

Just as lattices of open sets allows for more structure than is realized in boolean lattices of sets, the lattice of subspaces of a finite dimensional vector space also allows more structure than is realized in a boolean lattice of sets. It's structure is that of a projective geometry, with the lattice elements corresponding to subspaces of the geometry. Its meet operator is ordinary set intersection, $\cap$. It is its join operator $\sqcup$ that distinguishes it from a boolean algebra of events: Suppose A and B are disjoint nonempty subspaces with basis elements respectively $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$. Then

$$A \cup B \subset A \sqcup B \,, \tag{14}$$

because $A \sqcup B$ not only contain A (the vectors generated by $a_1, \ldots, a_n$) and B (the vectors generated by $b_1, \ldots, b_m$) but also vectors generated by mixes of $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$. Although Equation 14 is a desirable lattice property to have, for example, for modeling subadditivity, achieving it through generation by mixes of elements of A with elements of B does not reflect in a natural way most uses of the availability heuristic. From this I conclude that such lattices are likely to be of little use in the modeling of support theory phenomena, and more generally through representation theorems concerning complemented modular lattices (including orthocomplemented ones), that non-distributive complemented modular lattices are likely to be of little use for similar reasons.

Thus if we are looking for classes lattices that generalize probability theory in a way such that (*i*) it allows boolean lattices as a special case, (*ii*) it is complemented or has a reasonable generalization of complementation, and (*iii*) it supports finitely additive probability functions, then there are essentially three ways to go: (1) a boolean algebra; (2) a distributive lattice with a weakened form of complementation, and (3) a modular lattice with a form of complementation. For (2), I believe that for modeling Support Theory phenomena, $\cap$-complementation is the obvious and correct choice. I believe (3) is unlikely to be of much use for modeling such phenomena, because complemented modular lattices produce spaces with the wrong kind of structures for modeling availability, and I believe that trying a weaker form complementation for modular lattices will not improve things. (1) is still a possibility; however the method of moving from support to probability, e.g., the use formula

$$\mathbb{P}(\alpha|\beta) = \frac{S^+(\mathbf{CR}(\alpha))}{S^+(\mathbf{CR}(\alpha)) + S^-(\mathbf{CR}(\alpha))}$$

for determining subjective probability would have to be changed.

I believe that the New Foundation—(2) above with $\ominus$ as the weakened form of complementation—is the most likely candidate to produce useful event spaces for Support Theory phenomena, because it allows for a rich set of concepts that can interlink with uses of the heuristics of availability and representativeness.

REFERENCES

Brenner, L., and Rottenstreich, Y. (1999). Focus, Repacking, and the Judgment of Grouped Hypotheses. *Journal of Behavioral Decision Making, 12,* 141–148.

Birkhoff, G. and von Neumann, J. (1936). The logic of quantum mechanics. *Annals of Mathematics, 37(4),* 823–843.

Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics, 75,* 643–649.

Fox, C., Rogers, B., and Tversky, A. (1996). Options traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty, 13,* 5–19.

Heyting, A. (1930). Die Formalen Regeln der intuitionistischen Logik. *Sitzungsberichte der Preussichen Akademie der Wissenschaften,* 42–56. English translation in *From Brouwer to Hilbert : the debate on the foundations of mathematics in the 1920s,* Oxford University Press, 1998.

L. C. Idson, D. H. Krantz, D. Osherson, and N. Bonini. (2001). The Relation between Probability and Evidence Judgment: An Extension of Support Theory. *Journal of Risk and Uncertainty, 22,* 227–249.

Kahneman, D., Slovic, P., and Tversky, A. (Eds.) (1982).*Judgment under Uncertainty: Heuristics and Biases.* New York: Cambridge University Press.

Kahneman, D., and Tversky, A. (1982). Judgment of and by representativeness. In Kahneman, D., Slovic, P., and Tversky, A. (Eds.),*Judgment under Uncertainty: Heuristics and Biases.* New York: Cambridge University Press.

Kolmogorov A. (1932) Zur Deutung der intuitionistischen Logik. *Mathematische Zeitschrift, 35,* 58–65. English translation in P. Mancosu, Ed., *From Brouwer to Hilbert : the debate on the foundations of mathematics in the 1920s,* Oxford University Press, 1998.

Kolmogorov A. (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Republished New York, Chelsea, 1946.

Macchi, L., Osherson, D., and Krantz, D. H. (1999). A note on superadditive probability judgments. *Psychological Review, 106,* 210–214.

Narens, L. (2003). A theory of belief. *Journal of Mathematical Psychology, 47,* 1–31.

Narens, L. (2004). A theory of belief for empirical refutations. To be published in *Synthese.*

Redelmeier, D., Koehler, D., Liberman, V., and Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified alternatives. *Medical Decision Making, 15,* 227–230.

Rasiowa, H. and Sikorski, R. (1968). *The Mathematics of Metamathematics.* Warsaw: Panstwowe Wydawn. Naukowe.

Rottenstreich, Y., and Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in Support Theory. *Psychological Review, 104,* 203–231.

V. N. Saliĭ. (1988). *Lattices with Unique Complements.* Providence, RI: American Mathematical Society.

Sloman, S., Rottenstreich, Y., Wisniewski, C., Hadjichristidis, C., and Fox, C. R. (2003). Unpacking Implicit Probability Judgment. Manuscript.

Stone, M. H. (1936). The theory of representations for boolean algebras. *Trans. of the Amer. Math. Soc., 40,* 37–111.

Stone, M. H. (1937). Topological representations of distributive lattices and Brouwerian logics. *Čat. Mat. Fys. 67,* 1–25.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., and Koehler, D. (1994). Support Theory: A nonextensional representation of subjective probability. *Psychological Review, 101,* 547–567.

Wallsten, T., Budescu, D., and Zwick, R. (1992). Comparing the calibration and coherence of numerical and verbal probability judgments. *Management Science, 39,* 176–190.