Constraining Meanings With Contextuality

J. Acacio de Barros[1*], Carlos Montemayor[2*], Leonardo P.G. De Assis[1,3], Paul Skokowsi[4,5] and John Perry[5]

[1*]School of Humanities and Liberal Studies, San Francisco State University, San Francisco, 94132, CA, USA.
[2]Philosophy Department, San Francisco State University, San Francisco, 94132, CA, USA.
[3]Stanford Pre-Collegiate University-Level, Stanford University, Redwood City, 94063, CA, USA.
[4]Department of Philosophy, Oxford University, Oxford, UK.
[5]Center for the Explanation of Consciousness, Stanford University, Stanford, 94305, CA, USA.

*Corresponding author(s). E-mail(s): barros@sfsu.edu; cmontema@sfsu.edu;
Contributing authors: lpgassis@stanford.edu; pskokowski@gmail.com; johnperry43@gmail.com;

Abstract

In this paper, we defend two claims. First, we argue that a notion of contextuality that has been formalized in physics and psychology is applicable to linguistic contexts. Second, we propose that this formal apparatus is philosophically significant for the epistemology of language because it imposes homogeneous rational constraints on speakers. We propose a Contextuality Principle that explains and articulates these two claims. This principle states that speakers update contextual information by significantly reducing the space of probabilities and variables in a non-commutative way. Some contexts affect other contexts not merely in terms of the information they contain, but also on the basis of their sequential order. In particular, we argue that the Contextuality by Default (CBD) theory provides a formalism that helps explain the role of contextuality in rational linguistic exchanges.

Keywords: linguistic contextuality, quantum contextuality, contextuality by default, semantics, pragmatics

1 Introduction

Contemporary theories of linguistic meaning or semantic content depend fundamentally on a notion of "information updates," which determine how speakers solve problems concerning what is meant by an utterance in a conversation. These updates require assessments about information under considerable uncertainty. However, there is no agreement about what kind of representational or formal constraints should inform this updating process. According to traditional correspondence or truth theories, semantic content can be determined by recursively defining the truth of a sentence in a metalanguage. In particular, classic truth-functional approaches presuppose that propositions are the main bearers of meaning, and that they have a static or insensitive content across contexts and speakers intentions. According to these theories, sentences have content in virtue of the proposition they express, which determines the sentence's truth conditions. They can then be used in an ongoing conversation in a more flexible and contextual manner. These theories explain contextual information and the speaker's intentions in terms of pragmatic commitments or features of a conversation that are not constitutive of the meaning of expressions. Thus, classic semantic theories require a notion of information update for conversations that allows speakers to i) identify the content of the sentences expressed and ii) decide how aspects of the conversational context contribute to the further elucidation of the meaning of an expression. Both processes presuppose non-trivial informational frameworks at the content and context level.

An alternative approach is to begin with the informational updates of a conversational context and then build semantic content out of these contextual informational exchanges. This dynamical approach represents stages of discourses and their values according to functions from context to context, or context change potentials. The exact nature of these functions is debated. They serve as instructions for updating the context, but it is unclear how much information is contained in the context. Crucially, while there is controversy regarding how much information is contained in a context, everyone agrees that they are representations of information that is constantly updated, including informational states of the speakers, and that they are richer than a Kaplanian index, which identifies a speaker, time, place, and related environmental information (see [1]). However, it is unclear how much information in the context must be shared among speakers. More importantly, a key difficulty is how to characterize this information in terms of mental states, abstract representations, or objective aspects of a conversation. Since contexts are also essential to the static views explanation of, for instance, anaphora cases and pragmatic implicature, it is a pressing issue to characterize contexts more fully.

In this paper, we defend two claims. First, we argue that a notion of contextuality that has been formalized in physics and psychology applies to linguistic contexts. This application has the advantage that a formally developed model for contextuality that has been studied in detail could be investigated in a new setting. This opens the possibility for further insights regarding how meaning is determined in a conversation, independently of the semantic approach one ultimately favors.

In our second claim, we propose that this formal apparatus is philosophically significant for the epistemology of language because it imposes homogeneous rational constraints on speakers. We introduce a Contextuality Principle (CP) that explains and articulates these two claims. This principle states that speakers update contextual information by significantly reducing the space of probabilities and contents of variables in a mutually dependent and even non-commutative way. As contextual information is updated diachronically, contexts affect other contexts not merely in terms of the information they contain, but also on the basis of their order.

A motivation for this approach is based on Leibniz's Principle of Sufficient Reason (PSR) [2]. We see our CP as an extension of PSR because underlying it is the assumption that utterances are meaningful to the speaker in a non-accidental way. Two speakers communicating with each other assume the other is saying something that makes sense to them. Like the PSR, we think of the CP as not normative, but as a constraint for explanations in language. This view differs from other approaches such as the Gricean maxims [3], which are directives speakers should follow somehow. The reader is left with the exercise of reflecting on how much of the Gricean maxims derive from CP.

In particular, since information updates under uncertainty and increasing possible sources of data are essential and unavoidable for any theory of meaning, a preliminary step is to provide a convenient way of formalizing such complexity. Speakers have limited cognitive and informational resources, and the exploding possibilities that informational exchanges open up for speakers need to be handled in a psychologically realistic way. This situation is analogous to many kinds of decision-making for which Bayesian approaches may become problematic [4]. For our purposes, independently of which semantic framework one favors, there is always an informational process in which the content of a sentence is true at a world, or in which a sentence specifies an update at a context of discourse. Our main focus is the general structure of the informational content among speakers, and how they manage to realistically converge on information updates in spite of the increasing possibilities and variables produced by such updates, which determines the interaction between foreground and background.

A formal notion of context is key here. There are two relevant options here. One of them assigns a totality of possibilities and probabilities to a set; the other makes updates variable-dependent according to the circumstances and individuals involved. The second approach is the one we propose, which is also the standard framework for understanding quantum probabilities and non-commutative effects in psychology. For any discussion on pragmatics or semantics, static views with a rich notion of context or a dynamic and pragmatically driven informational updating system, context plays the same crucial role. This is true for all semantic theories. The relevant possibilities can be understood as possible worlds, situations, or representations but updates are a central commitment on all accounts. Some relation must exist between utterances, mental states, and informational changes to content, however construed by the different semantic theories. Our proposal is that these are formally equivalent to informational changes, as specified in other contextual approaches. Truth-making or information-updating both require an account of context. Thus, regardless of the choice of semantics, and the intricate issue of the relation between sentences and propositions, or between semantics and pragmatics, our main point is that the general formal framework for contextual information updating in semantics will be the same.

Contextuality is a fundamental concept in physics, where it is key to understanding the strangeness of quantum mechanics [5–7]. Contextuality is related to the complementarity principle, which states that there are some quantities that cannot be simultaneously observed, which allows for a given property to be observed in a given experimental contexts but not in another. For example, one of the most conceptually important results in quantum mechanics is Kochen-Specker's theorem (KST) [8], where it is shown that for certain quantum systems it is possible to find a set of mutually exclusive experimental conditions, a.k.a. contexts, where a property that is observed more than twice cannot have context-independent values assigned to them. KST is at the core of John Bell's famous inequalities, where the local truth-values of properties are affected by contexts determined by far away experiments. Furthermore, contextuality may be at the heart of what makes quantum mechanics different from classical physics [7], and seems to be the reason for quantum algorithms' superiority [5].

Contextuality, unsurprisingly, also appears in psychology, but it was not until recently that researchers realized that there was a connection between the concept in psychology and in physics. This connection was, as far as we know, made explicit in research concerning selective influences on the basis of classic probability theory [9], which was then applied and further developed in the setting of quantum cognition (for a review of quantum cognition, see [10–12]). The idea of quantum cognition is to use the mathematical apparatus of quantum mechanics to describe experimental situations. In those cases, it can be argued that the central feature used is contextuality [13]. Afterwards, a mathematical theory of contextuality has been perfected by Dzhafarov and Kujala [14], and it was shown that the idea of contextuality in psychology could also be used in physics. Their theory was applied not only to quantum systems [15] but also to psychology experiments [16, 17].

Finally, as seen above, contextuality is also a central concept in semantics [18] and pragmatics [19], where many examples of context dependence exist. Presently it is not clear how contextuality in language relates to contextuality in physics and the social sciences, and the goal of this paper is to bridge this gap. We plan on doing this by first showing that Dzhafarov and Kujala's formalization of contextuality in physics and psychology, known as Contextuality by Default [14], applies to the concept of contextuality in linguistic exchanges.

Our second goal is argue that contextuality in language, defined in such a precise mathematical way, is not only an interesting result, but that it is also relevant to discussions related to rational constraints on linguistic exchanges.

Our reasoning for the relevance of a formal definition in semantics is based on three points. First, it can be argued that speakers and listeners engaged in a conversation have rationally constrained expectations about the probable meaning of utterances in a context. These constraints on possible content variables are necessary for the effective and consistent exchange of information between the participants of a conversation. For example, the lack of consistency in utterances of a speaker is often perceived by the listener as a misinterpretation of the intended meanings, and if the listener becomes convinced that the speaker is herself inconsistent, then the conversation often ceases (or one becomes either amused or enraged by the absurdity of the situation, as it happens sometimes in Thanksgiving dinners). These rational expectations constrain what speakers intend to say (their implicit and explicit communicative motivations) and how they can say what they intend to say (how they utter expressions at a context and a time, as well as how they update the contextual information). Thus, these constraints shape how speakers (and listeners) construct sentence meaning in a context.

Second, successful updates to the context and accuracy conditions for propositions or expressions are also restricted by these rational expectations. For example, if someone says “my car might have been stolen” in a context in which they must urgently drive somewhere, what their expression is about cannot be merely captured by the conditions that would make it true (for instance, the extent of their ignorance about the location of the car, or about their ignorance in general). Rather, speakers must expect that the conditions of relevance saliently include the immediate need to drive somewhere in this specific context.

Third, contextuality is needed to not only provide such conditions for communicative success and accuracy, but also to specify transitions concerning relevance between foreground and background. This is also critical for identifying the speech acts of other speakers in a cooperative way. For instance, the sentence “They are always so charming,” depending on context, may be either a compliment or a criticism. A competent speaker must pay attention to shifts in the contextual cues that determine speaker-intentions and this process fundamentally depends on content variable-assignments.

As mentioned, our goal is to show that it is possible to represent the role of contextuality in language by appealing to a formal concept of contextuality in physics and psychology formalized by the theory of Contextuality by Default (CbD) [14]. Thus, linguistic contextuality can be studied as a mathematically well-defined phenomenon. We should emphasize that we are not claiming that such formal rational constraints are explicitly worked out in the head of speakers (or listeners). Rather, we hypothesize that information-updating depends upon the abilities of speakers to identify relevant information in a contextual way.

For the sake of clarity, we are not proposing CbD as a formalism that could serve as a unifying theory for quantum contextualism, linguistic contexts, and rational contextualization. The scope of our argument is much more narrow, and it concerns a lacuna in the extant literature concerning the lack of formal accounts in linguistics that appeal to CbD in particular. We argue that this is an important gap in our understanding of contextuality in linguistics. There might be further implications of our proposal for a unifying theory of contextuality, but these implications are beyond the scope of this paper.

This paper is organized the following way. In Section 2 we start with some examples of contextuality in language, and provide a basic mathematical definition of contextuality that will be expanded in the following section. In Section 3 we discuss the concept of contextuality in physics and psychology, and introduce a simplified version of Dzhafarov and Kujala's Contextuality by Default (CbD) theory, highlighting only the necessary mathematical features. Finally, in Section 4 we provide a more refined model that formalizes the standard concepts of contextuality in language in terms of a mathematical structure that can be translated into a probabilistic theory. We show that linguistic exchanges are contextual in a formally precise sense, specified by CbD.

2 Contextuality in Language

In this section we introduce a general concept of contextuality that fits some of the usual discussions of contextuality in linguistics. This will allow us to put forth the ideas within the realm of a more general theory of contextuality, to be developed in Section 4. As previously mentioned, we propose a Contextuality Principle, according to which speakers update contextual information by significantly reducing the space of probabilities and variables in a non-commutative way. When discussing contextuality, we will presume that there is a theory of meaning, i.e. a way to attribute truth values to an utterance or update information from context to context according to some dynamical process. We will also impose on this theory that it abides to the standard rules of reasoning, e.g. classical logic[1]. The key idea is that speakers must reduce a vast array of possibilities regarding the contents of linguistic variables in a way that is consistent with the basic rules of rational coherence.

We do not mean, of course, that speech is always logically consistent, or that humans are infallible. For example, it seems to be the case that human decision makers do not follow a classical logic [11, 20, 22]. So, it is possible that actual semantics follow some other type of logic, with some researchers even proposing the use of quantum lattices [23]. However, when this is the case, we argue that such logics are also concerned with contextual situations, in the sense that truth value assignments or conversational updates are not the same in two different contexts (see our discussion below in Section 2). Basic compatibility with classical logic simply means that in the process of meaning

[1] It may be possible to extend our discussions in this paper to intuitionist logic, if we follow [20], or even a more general lattice structure, such as orthomodular lattices [21], instead of a classical Boolean structure, but this would go beyond the scope of this paper.

construction, two participants in a conversation have as common ground the assumption of consistency regarding what the other speaker is saying. If, in the process of a conversational exchange, a contradiction is found, it needs to be resolved. We take this to be a basic constraint on successful communication. There are many ways in which this can happen; for example, a statement may be recanted, or the meaning of a content variable may be modified or updated. It is not until the whole set of utterances is put in a consistent logical form, through the type of processes mentioned, that the participants in a conversation will achieve coordination through successfully constrained updates.

With that requirement of consistency in mind, we start with the question of when is an utterance contextual? To a first approximation, a condition that specifies contextual utterances can help elucidate the nature of contextuality in linguistic exchanges more generally. For example, according to Cappelen & Lepore [24][page 17], "Suppose someone suspects that an expression e [...] is context-sensitive. How could he go about establishing this? One way that philosophers of language do so is to think about (or imagine) various utterances of sentences containing e. If they have intuitions that a *semantically relevant* feature of those utterances varies from context to context, then that, it is assumed, is evidence [that] e is context-sensitive." If we restrict ourselves to formal semantics, then how do we think about "intuitions" about variation with context? Given an utterance, what are the conditions under which we *must* call such utterance contextual? How could we determine that e is contextual in a principled manner?

Before addressing this issue, a disclaimer is necessary. Our proposal is not an attempt at providing a comprehensive theory of contextuality that includes all the possible forms and nuances of contextual expressions in linguistics or philosophy of language. Instead, it is our goal to show that the most common uses of contextuality are suitably described within the framework developed in physics and psychology (though we will not need to use the full blown theory developed in [14]). In order to achieve this more formal goal, we will use a few well-defined and standard views of contextual expressions in semantics and pragmatics to motivate our definitions.

Let us examine some simple examples of contextuality in language using the truth-conditional approach to semantics as one salient possible approach to meaning assignments (see the introduction for a justification of how our approach applies to dynamical approaches as well). A paradigmatic example of contextuality in language is that of indexicals [18]. Take the following sentence:

P. I went to the beach yesterday.

According to a truth-conditional approach to semantics, to attribute meaning to P means specifying its truth conditions in order to specify a truth value. But in order to assign a truth value, simply hearing the sentence above is not sufficient. We need to know, for example, who uttered it, when was it uttered, and whether the person who uttered it actually went to the beach that day

[18]. Contexts require at the very least this amount of information, and clearly some additional information. But this is enough to appreciate the nature of contextuality in language. Take for instance the two cases below.

Bob. Bob is a graduate student at UC Berkeley, and he likes to surf and go to the beach. Because it is summer, Bob went to the beach on July 27th, 2020.
Alice. Alice is also a graduate student, but she goes to Stanford. Alice is Bob's friend. When he invited her to go to the beach on July 27th, 2020, she declined, stating that she had plans to go to the movies with Carlos instead. Bob was devastated.

Now, consider the following two contexts.

Context 1. On the evening of July 28th, 2020, P is uttered by Alice.
Context 2. On the morning of July 28th, 2020 P is uttered by Bob.

It is clear that P in Context 1 is false, whereas P in context 2 is true. The truth-value of the statement P depends on the context because P itself depends on the context: P in one context is a proposition about Alice, whereas in the other it is about Bob. If we were, naively, to insist that a proposition P is the same in all contexts, we would reach a contradiction (P is true and P is false). This contradiction, of course, is eliminated when we attach to P some ancillaries, or labels: the speaker, a time, and a location, thus leading to two distinct propositions:

P1. "I went to the beach yesterday," uttered by Alice at 7:32pm of July 28th, 2020 sitting at a table in Stanford's Green Library.
P2. "I went to the beach yesterday," uttered by by Bob, while he sobbed, at 2:11am of July 28th, 2020 drinking a beer at Antonio's Nut House.

P1 and P2 are now, clearly, different propositions.

Let us examine a similarly simple example. Imagine the following conversation:

Bob: I read about Socrates yesterday.

Alice: What was it?

Bob: The guy had some interesting ideas. I mean, I knew he wrote about philosophy, but I didn't know he also wrote about economics. Did you know this?

Alice: Not really, but it doesn't surprise me. What did he write? I thought very little was left of what he wrote.

B: No, he used to have a column in some Brazilian newspapers and magazines. You can find them online easily.

A: Ah, now it makes sense. You must be talking about the Brazilian footballer, Socrates, and not really about *Socrates*! I though you were talking about *Socrates*.

Again, in this example a *confusion* happens when one of the participants in the conversation attributes to the term "Socrates" a different referent than the other participant, generating ambiguity in their conversation (many well-known people were named Socrates, but the best known was the Greek philosopher who lived more than two thousand years ago). In this case, the context of the conversation starts to dispel the ambiguity by narrowing the possible referents of "Socrates". Notice, however, that the contradictions show up in a different way than before: it is not until the end that Alice notices that there was an incoherence in their conversation by observing an apparent conflict between her views about Socrates and Bob's, which consequently required her to update the meaning-assignments for Socrates. But, more importantly, it is not until Alice notices an apparent contradiction (vis a vis her view of Socrates, the philosopher, and what she knew about his life) that she changes her thoughts and rational expectations regarding this conversation about Socrates (e.g. "Socrates is a philosopher", "Socrates was a player in the Brazilian National team", etc.). This toy example shows how the contextual specification of the content of linguistic terms or variables helps speakers coordinate and specify their referents in a rationally coherent way.

Thus, we want to emphasize that, as we read the dialogue, it should become clear to us (if we ever heard of Socrates the footballer) from the text itself that Alice's meaning-updates become inevitable from Alice's point of view. Here we see again the same type of contextuality as before: the attribution of the participant in the conversation to "Socrates is the Greek philosopher who lived circa 469 BCE–399 BCE," together with some ancillary information about Socrates, were in conflict with thinking this as T and the other as F. Once again, to prevent a contradiction and to give meaning to the dialogue, participants have to attach to the word "Socrates" some specifying labels. Variables and quantification alone cannot account for how these updates occur at a context in order to guarantee flexibility in the updating processes and stability for the rational and coherence constraints that must shape conversations among speakers. For a conversation to be intelligible for all participants, referent and value assignments must be coordinated consistently. Thus, we have a justification to put forth the following (preliminary) definition of contextuality (using truth values, which could be extended to other semantic approaches concerning contextual information updates).

Definition 1 Let $P_1, \ldots, P_N$ be a collection of propositions (which can also be statements in a language), and let $C_1, \ldots C_M$ be a collection of contexts in which some of the propositions P_i, $i = 1 \ldots N$, are assigned truth values. Let $\mathfrak{B}_j, j = 1 \ldots M$, be a collection of Boolean algebras associated to the set of propositions and truth values in context C_j. The propositions are *contextual type I* if and only if there does not exist any Boolean algebra $\mathfrak{B}$ over all $P_1, \ldots, P_N$ such that the $\mathfrak{B}_j$ are sub-algebras of $\mathfrak{B}$.

The idea of Definition 1 is that, somehow, all the propositions can be put together in a consistent way, without any contradictions. This is clearly not the case for the examples discussed above. We see, for instance, that according to our definition, we have P uttered by Alice and Bob in two different contexts, and they are incompatible because of their different referents. It is the recognition that each P uttered is different that leads to the resolution of this apparent conflict. So, the example with Alice and Bob clearly satisfies the definition above. Since there is a contradiction between the two truth-value assignments of the propositions in Context 1 and Context 2, there is no global Boolean algebra that comprehends both Contexts. What is the alternative? Formally speaking, for this problem to be solved it suffices that speakers label the propositions in terms of contexts, i.e., they should recognize that P in context 1 is different from P in context 2. The same is true for the sentence "Socrates wrote about economics" which is the focus of the dialogue between Alice and Bob. It is not until Alice realizes that there are two different referents of "Socrates" the conflict in meaning is resolved.

Other examples abound in language, some quite different from and more complex than the case of indexicals. For instance, the sentence "Do your homework" takes completely different meanings when it is uttered by someone to their kids in an evening at home or when uttered by a colleague in the context of preparing for a department meeting. However, these examples are similar to what we have above: if we were to attach full meanings and their implicatures to the sentences as well as ancillary propositions associated to them we would reach contradictions. Take, instead of a Boolean algebra, a set of semantic meanings to words and sentences. The purpose of semantic coordination is to find a function from utterances to references in a contextually flexible and rational manner.

However, a simple function from utterance to a reference cannot guarantee speaker coordination when there is contextuality or ambiguity, since a given utterance can be assigned many different references or values. If we focus on contextuality, this non-univocity of a function exists because in addition to the utterance, we need to add context[2]. When we maximally specify not only the utterance, but also the context, we should be able to create such a function. Contextuality is, therefore, a type of epistemic impossibility concerning univocal meaning determination, given just the utterance and a restricted set of contextual (and observable) cues. In other words, even if we have more than the utterance, such as the speaker or time, sometimes it is not possible to attach a well-defined meaning to the utterance. When this is the case, we think of the "utterance + partial context" as contextual. In the end, the formalization of these ideas can be accomplished, as we shall see below, with the introduction of a probability space and random variables, since we need something more than truth value assignments in a Boolean algebra.

In this section we made the assumption that speakers, when constructing and coordinating meaning from utterances, require consistency in spite of the constant changes in meaning assignments from context to context. The informational process of meaning construction and value assignment by the agent must be consistent: once the agent realizes there is an inconsistency, they will

[2]Here we are viewing context as an umbrella term meaning whatever we need to add to the utterance in order to allow for a coordinated connection to a reference.

proceed to revise their thinking or understanding concerning the utterance and its contents by updating how it relates to context updates.

3 Contextuality in Physics and Psychology

Up to now we have been talking about contextuality in language, but contextuality does show up in very salient ways in other areas of inquiry, such as psychology, decision making, and physics. A famous example is that of order effects [25]. Simply put, if I were to ask two questions in different orders to a group of people, the statistical distribution of responses would change with the order. This is a well-known effect in social sciences, and researchers have to go through great lengths to prevent it from happening in, for example, surveys or opinion polls. If the present approach is correct, similar effects should be examined in semantics and pragmatics.

There are obvious cases where we can expect order effects in psychology and decision making. Say someone asks the two following questions.

Q1. Are you in favor of imposing restrictions for suspected terrorists to buy guns?
Q2. Are you in favor of more strict gun-control laws?

Because Q1 reflects an often held view of conditions under which gun control are considered reasonable by most people, we can imagine how the order of questions would affect the rates for Q2. Notice that the *meaning* of those sentences did not change, but the *context* did.

A famous example of order effects discussed in psychology is the following less obvious case [25].

Q3. Do you think Bill Clinton is honest and trustworthy?
Q4. Do you think Al Gore is honest and trustworthy?

In order Q3→Q4, 50% of respondents say "yes" to Q3 and 60% to Q4; in the reverse order Q4→Q3 57% answered "yes" to Q3 and 68% to Q4. We now have a different case, though: instead of a single instance where the truth value changes from one context to another, we have probability measures over this truth values. However, we can easily see that all these value assignments are somewhat related: the beliefs expressed by the responses in opinion pools, considering that samples are not biased, change when we change the order of the questions. We will see how this is so when we formalize this idea in terms of random variables, but for the moment let us consider a simple case where ω_3 and ω_4 are partitions in a probability space corresponding to the propositions Q3 and Q4 being true, respectively, and we attribute to them a probability measure. It then follows that $p(\omega_3) = p(\omega_3 \cap \omega_4) + p(\omega_3 \cap \omega_4^*)$, where ω^* is the set-complement of ω. It is clear from the structure of probabilities that $p(\omega_3)$ is independent of the order in which Q3 was asked. Contextuality here shows up as the need, once again, to say that Q3 by itself is different from Q3 in the context of Q4.

Because order effects provide a prototype for many of the applications of CbD and non-standard probabilities, it is helpful for us to discuss it in some detail. Here, we will focus on some mathematical approaches for them. Since a single probability space cannot model order effects, it is natural to ask how one can describe it mathematically. We can organize mathematical modeling of order effect in two categories: standard probability models and non-standard probability ones. Let us examine each separately.

Non-standard probability models are based on probability theory modifications that allow for standard results violations. For instance, the usual result that $P(A \cup B) = P(B \cup A)$ is an immediate consequence of representing probabilities in the language of set theory. Suppose one did not define probabilities over sets, but instead, one used other mathematical objects. In that case, the result of conjunctions of disjunctions could be different and even depend on the order. This move is at the core of using the quantum mechanical formalism to describe order effects (such as in the works of [26–32]). For example, Busemeyer and Wang obtained probabilistic order effect models that fit experimental data using a vector space formalism.

The idea behind most of such models is straightforward. They assume that subspaces of vector spaces mathematically represent mental states. An answer to a question is the equivalent of a quantum measurement, which leads to the projection of the original mental state into the subspace corresponding to the question, akin to what happens to quantum measurements in physics. Following Borns rule, probabilities are computed by the square of the vectors components absolute values before a collapse. Order effects come from the fact that projections of a vector v into a space A and then B may lead to different outcomes from projections of v into B and then A. Many researchers used a variation of the above ideas to account for order effects and other violations of probabilistic models in social sciences (the reader is referred to references [11, 12, 33] as good surveys of such applications). Additionally, quantum-like models have been used in linguistics to account for context-dependency, such as the following works: [23, 34–37].

One issue with the projection models is that it predicts that a question A, when presented before and after another question B, will have no repeatability [38]. This issue occurs because, at each answer, there is a projection. Since A and B do not correspond to the same vector space, A followed by B does not necessarily lead to the same A as before. This effect on previous observations is similar to quantum physics, where measurements of two incompatible observables are not consecutively reproducible. This lack of repeatability is known in the literature as the ABA problem [39].

To address the ABA problem, Ozawa and Khrennikov recently modified projection models to account for the repeatability of answers [40, 41]. In their model, they include as part of the measurement a back action from the measurement apparatus, itself considered in a superposition of quantum states. The idea is that the participants mental state comprised not only of the vector associated with the question but also of other mental correlates is as much a part of the measurement process as everything else. Since quantum-like formalisms model mental states, a measurement is not a simple projector but a more sophisticated quantum process involving a mental environment. With this model, the authors were able to overcome the ABA problem.

Another alternative approach is using extended probability theories, such as negative probabilities [42–46]. The idea of this approach is to simply relax the axioms of probability to allow for non-observable events to take negative values. Negative probabilities, defined this way, may be non-monotonic, as is required of order effects or contextual measurements. Interested readers can refer to the references mentioned above.

For standard probability approaches, we single out two main ones. One is CbD, or Contextuality by Default, which we discuss elsewhere in this paper. Another is a Bayesian approach proposed by [47]. We will discuss the Bayesian approach briefly to contrast with the quantum, primarily as it refers to the ABA models discussed by Ozawa and Khrennikov [40, 41].

In the Bayesian approach, a question is seen as self-reflection. Self-reflections, like experiments, lead to updated prior probabilities. Since internal biases or correlations will change the posterior assessment of a response, the order of questions will affect the expected beliefs. Order effects come in this model from priors that are correlated. For example, if one asks a participant what they think about whether Clinton is honest and trustworthy, they will likely not have a pre-answered response. They will probably use several heuristics to reach a yes or no answer. Once they get a response, this is new information, and their belief about Clintons honesty and trustworthiness are updated. Since it is expected that a participants belief about Clinton is correlated to a belief about Gore (who was his vice-president), it follows that Gores probability is also updated (the reader is referred to [47] for detail). The updated probability shows an order effect: asking A before B changes B as a simple Bayesian update of beliefs about B, as B is correlated to A.

CbD and Bayesian order effect have an essential similarity. Both resolve the problem of order effect by changing the probability space. In CbD, the probability space is expanded a priori by assuming that each question is different for each order, thus leading to larger sample spaces. The Bayesian approach also has larger spaces: the prior and posterior ones. However, a significant difference is that CbD only describes order effects and has numerous parameters. In contrast, the Bayesian approach has fewer parameters and makes definite predictions about certain experimental conditions (see [47]).

Comparing CbD and the Bayesian model to Ozawa and Khrennikovs (OK) work is worthwhile. All three approaches can model ABA, with CbD being the most flexible simply because of its large number of parameters and its philosophical approach that each case is a different random variable. So, just for ABA, in CbD, we would have A_1, B_2, and A_3 as random variables instead of A and B. This large number of parameters allows for almost any fitting related to order effects. On the other hand, OK has constraints that must be satisfied by the model and have fewer parameters to work with than CbD.

Finally, the Bayesian model has the fewest parameters and makes definite predictions about ABA. Surprisingly, it does not predict that A will have the same value after AB since the prior has been updated. Still, this lack of exact repeatability is not inconsistent with empirical evidence (see [39]). It would be interesting to see a direct comparison of those models, but this task goes well beyond the scope of this paper.

Another example of contextuality comes from the violation of Savage's Sure Thing Principle (STP). Savage formulates the STP in the following way [48].

> "A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he should buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say. It is all too seldom that a decision can be arrived at on the basis of the principle used by this businessman, but, except possibly for the assumption of simple ordering, I know of no other extralogical principle governing decisions that finds such ready acceptance."

Though, as Savage states, such principle finds "ready acceptance," empirical data suggests that humans violate the STP in decision making [49, 50]. The essence of this violation is that, under uncertainty, one does not see the propositions as the same. In other words, the uncertainty of the context causes decision-makers to think differently about the questions. As Daniel Kahneman said, "when the question is difficult and a skilled solution is not available, intuition still has a shot: an answer may come to mind quickly – but it is not an answer to the original question."

It is straightforward to see that a violation of STP implies contextuality. Let A be the proposition "Buy the property in question," and B be the proposition "The Democrat candidate will win the election." In those terms, STP can be formulated as the statement "If $p(A \mid B) > p(\neg A \mid B)$ and $p(A \mid \neg B) > p(\neg A \mid \neg B)$, then $p(A) > p(\neg A)$," where $p(A \mid B)$ is the conditional probability that A is true if B is true, defined as

$$p(A \mid B) = \frac{p(A\&B)}{p(B)}, \tag{1}$$

if $p(B) > 0$, and similarly for $p(A \mid \neg B)$, and where "$\neg X$" is the negation of X. From $p(A \mid B) > p(\neg A \mid B)$ and $p(A \mid \neg B) > p(\neg A \mid \neg B)$, if follows that

$$p(A \mid B)p(B) + p(A \mid \neg B)p(\neg B) > \\ p(\neg A \mid B)p(B) + p(\neg A \mid \neg B)p(\neg B), \tag{2}$$

and that

$$p(A) > p(\neg A). \tag{3}$$

In other words, the assumption that there exists a probability space that represents the propositions "A," "B," and "A and B" imply Savage's STP. Therefore, a violation of STP implies that no such probability space exists for those random variables, and the system is considered contextual. This notion of contextuality suggests, therefore, that conversational updates and transitions from foreground to background may operate under similar contextual order effects thereby helping speakers eliminate possible assignments in a heuristic manner.

All examples presented above are of direct influences or what we call explicit contextuality. As a last example, let us consider a more subtle case of proper contextuality given in [17]. In a crowdsourced experiment, Cervantes and Dzhafarov asked participants questions in the context of Hans Christian Andersen's Snow Queen. Participants in the experiment were informed that in the story the Snow Queen is beautiful and evil, that Gerda is beautiful and kind, that the Troll is unattractive and evil, and that the old Finn woman is unattractive and kind. In their experiment, each context corresponded to two choices of characters (either Gerda/Troll or Snow Queen/Old Finn Woman) and two choices of characteristics (Beautiful/Unattractive or Kind/Evil), for a total of four contexts. We will not go into the details here, but it suffices to say that this experiment exhibits both types of direct and proper influences concerning random variables (see [17] for detail).

4 Truth and Beliefs

Since a speaker's linguistic capacities are not necessarily based on certainty, but may be grounded on degrees of belief, it is important to extend the above arguments to propositions that are thought to be uncertain by the agent. Let us elaborate on what we mean here. Let two agents, Alice (A) and Bob (B), engage in a conversation that includes a series of utterances. The purpose of such conversation is to share information about their own beliefs. As such, each agent starts with a series of beliefs about the veracity of the propositions expressed, and given that one of the purposes of a conversation is to understand each other (although this may be not the case in some instances), as the conversation progresses, such beliefs may change. This is the case in the examples above, where utterances have an assumed value assignment that was revised after the agents realized that their beliefs were inconsistent. In the process of belief (or truth) construction, there is an underlying assumption by each agent that the other person involved in the conversation is consistent in their beliefs, since if this were not the case there would be no point in engaging in collaborative meaning construction.

So, meaning coordination or construction depends, at least partly, on rational attributions of truth values to beliefs. For example, if Alice holds a belief that the utterance U is true, then she must also hold the belief that utterance "not U" is false. Similarly, if Alice believes that utterances $U_1, U_2, \ldots, U_n$ are true, the logical consequences of this belief should also be true, e.g. that "U_1 and U_2" and "not U_1 & not U_2" are also true. If Alice is uncertain about an utterance U', she may consider the two "universes" (or consequences) where in one U' is true and another where U' is false, together with the logical consequences of such truth values. The underlying assumption here is that truth values are assigned in a consistent way when trying to make sense of utterances.

Let us examine in more detail the situation where Alice is not certain about the truth value of an utterance. In this case, if Alice wants to be consistent, as in the case of certainty, the attribution of a high degree of belief to an utterance implies that its negation must have a low degree of belief. For example, Alice cannot believe with high confidence that it will rain all day tomorrow and also that it will be sunny all day. Furthermore, to be consistent, Alice cannot hold strong beliefs that contradict other standing beliefs that must be assumed as background. For example, Alice cannot believe that Bob is an "excellent tennis player" and a "first rate philosopher," while at the same time having a strong belief that Bob "cannot do anything right." Alternatively, Alice cannot simultaneously believe that Bob is a millionaire and also believe that Bob cant afford to buy a Frappuccino at Starbucks. Another way to think of this is the following. For each set of truth values Alice attributes to a series of propositions (or utterances), she also needs to attribute truth values to their logical consequences (although we are not endorsing strong versions of the closure principle, which is orthogonal to the present discussion, which focuses on the information that normal, cognitively limited human beings, have at any point in time). Then, for each set of truth tables constructed in such a way, Alice needs to assign to them a measure of belief[3].

Probabilities are the standard way to formally represent rational beliefs, which is consistent with the idea that agents create meaning and assign values to propositions or sentences in a consistent and coherent way. Beliefs are not always either T or F, but may have a value that falls in between. From Cox's theorem, it follows that rationality, in the form of logical coherency and consistency, implies adopting the theory of probabilities as a measure for rational belief [51, 52]. In other words, if Alice is trying to be rational, or to at least try to make sense of Bob's utterances in a rational way by assuming that Bob is rational, then Alice would construct and coordinate meaning assignments with Bob by using some approach consistent with classical probability theory.

This epistemic commitment leads to a generalization of our definition of contextuality above. We can now accommodate cases in which Alice and Bob are not certain of the meaning of the sentences they are uttering, or in which their beliefs may not be coordinated with respect to value assignments.

[3]We do not mean that explicit rule-following, i.e. coherentism, or explicit conscious inference might be important here. All we assume is that speakers have the capacities to identify meaning in a rational way, however those capacities are construed.

Definition 2 Let $\{C_j\}$, $j = 1, \ldots, M$, be a collection of contexts, and for each context C_j let $P = \{P_{i,j}\}$, $i = 1 \ldots N$, be a collection of propositions. Let $\mathfrak{P} = (\Omega, \mathcal{F}, p)$ be a probability space, and $\mathbf{P}_{i,j}$, $j = 1, \ldots, M$, $i = 1 \ldots N$ a collection of ± 1-valued random variables on $\mathfrak{P}$ such that $\mathbf{P}_{i,j}$ represents the truth-values of $P_{i,j}$'s (1 corresponding to T and -1 to F). Then P is non-contextual if and only if there exists a $\mathfrak{P}$ such that $p(\mathbf{P}_{i,j} = \mathbf{P}_{i,j'}) = 1$, $\forall j, j'$, and P is contextual if no such $\mathfrak{P}$ exists.

Additionally, we can introduce some further terminology that comes from Contextuality-by-Default [14].

Definition 3 Let $\{C_j\}$, $j = 1, \ldots, M$, be a collection of contexts, and for each context C_j let $P = \{P_{i,j}\}$, $i = 1 \ldots N$, be a collection of contextual propositions. Let $\mathfrak{P} = (\Omega, \mathcal{F}, p)$ be a probability space, and $\mathbf{P}_{i,j}$, $j = 1, \ldots, M$, $i = 1 \ldots N$ a collection of ± 1-valued random variables on $\mathfrak{P}$ such that $\mathbf{P}_{i,j}$ represents the truth-values of $P_{i,j}$'s (1 corresponding to T and -1 to F). If $p(\mathbf{P}_{i,j} = \mathbf{P}_{i,j'}) = c_{i,j,j'} \leq 1$, then P is said to be *influenced directly by the context* or that it has *contextuality by direct influences*, where $c_{i,j,j'}$ is the maximal probability that $\mathbf{P}_{i,j}$ can equal $\mathbf{P}_{i,j'}$ if the two variables were jointly distributed.

The idea of the above definition is that the context of a proposition directly influences its truth-value. However, we can refine contextuality even further, which is the core of the following definition.

Definition 4 Let $\{C_j\}$, $j = 1, \ldots, M$, be a collection of contexts, and for each context C_j let $P = \{P_{i,j}\}$, $i = 1 \ldots N$, be a collection of propositions influenced directly or not by the context. Let $\mathfrak{P} = (\Omega, \mathcal{F}, p)$ be a probability space, and $\mathbf{P}_{i,j}$, $j = 1, \ldots, M$, $i = 1 \ldots N$ a collection of ± 1-valued random variables on $\mathfrak{P}$ such that $\mathbf{P}_{i,j}$ represents the truth-values of $P_{i,j}$'s (1 corresponding to T and -1 to F). If there is no $\mathfrak{P}$ such that $p(\mathbf{P}_{i,j} = \mathbf{P}_{i,j'}) = c_{i,j,j'} \leq 1$, where $c_{i,j,j'}$ is the maximal probability that $\mathbf{P}_{i,j}$ can equal $\mathbf{P}_{i,j'}$ if the two variables were jointly distributed, then P exhibits *proper contextuality.*

Proper contextuality is subtler than explicit or direct influences, as the name suggests. Suppose Alice, Bob and Carol are good friends, and suppose further that they enjoy disagreeing with each other in class, trying to make one another look wrong. Because they are interested only in disagreeing, it is impossible to predict what side they will choose to defend. So, for example, if a Professor in a class puts forth a point, and Alice is in this class, half of the time Alice agrees with the Professor. The same behavior is observed for Bob and Carol, with them also disagreeing with the Professor 50% of the time. Now, Alice, Bob, and Carol are very good friends, and one of their favorite way of having fun is to disagree with each other in class. It so happens that in a particular semester they have plenty of chances of doing so: Alice is enrolled in classes X and Y, Bob is enrolled in classes X and Z, and Carol is enrolled in classes Y and Z. In class X, whenever the Professor says something, Alice disagrees with her 50% of the time, and so does Bob, but Alice and Bob disagree among themselves all the time. This is also true for Alice and Carol in class Y, and for Bob and Carol in class Z.

To see how this example can provide contextuality, imagine the following three propositions: A="Alice agrees with the professor," B="Bob agrees with the Professor," and C="Carol agree with the Professor." We can code such propositions in terms of ± 1-valued random variables $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ for Alice, Bob, and Carol respectively, where the variables take value $+1$ if the answer to the proposition is true, and -1 if the proposition is false. From the fact that they randomly agree with the Professor, we can at once infer that $\langle \mathbf{A} \rangle = \langle \mathbf{B} \rangle = \langle \mathbf{C} \rangle = 0$. But from the fact that they all pairwise disagree when in their classes, we have that $\langle \mathbf{AB} \rangle = \langle \mathbf{AC} \rangle = \langle \mathbf{BC} \rangle = -1$. This, of course, leads to a mathematical contradiction: $\mathbf{A} = 1$ implies $\mathbf{B} = -1$ and $\mathbf{C} = -1$ from the two first correlations, in contradiction with the last correlation, which states that the product of $\mathbf{B}$ and $\mathbf{C}$ is always -1. This is, of course, a logical contradiction.

Now, let us see how this example fits within Definition 4. First, we need to label the random variables with respect to the context, X, Y, or Z.

Context X: Alice and Bob are in this class, but not Carol. The random variables here are $\mathbf{A}_X$ and $\mathbf{B}_X$.
Context Y: Alice and Carol are in this class, but not Bob. The random variables here are $\mathbf{A}_Y$ and $\mathbf{C}_Y$.
Context Z: Bob and Carol are in this class, but not Alice. The random variables here are $\mathbf{B}_Z$ and $\mathbf{C}_Z$.

We should note that there are no direct influences, since $\langle \mathbf{A}_X \rangle = \langle \mathbf{A}_Y \rangle = 0$, $\langle \mathbf{B}_X \rangle = \langle \mathbf{B}_Z \rangle = 0$, and $\langle \mathbf{C}_Y \rangle = \langle \mathbf{C}_Z \rangle = 0$, and therefore $p(\mathbf{A}_X = \mathbf{A}_Y)$ can have maximal value 1, and so can $p(\mathbf{B}_X = \mathbf{B}_Z)$ and $p(\mathbf{C}_Y = \mathbf{C}_Z)$. It then follows that we can simply set $\mathbf{A}_X = \mathbf{A}_Y \equiv \mathbf{A}$, $\mathbf{A}_X = \mathbf{A}_Y \equiv \mathbf{A}$, and $\mathbf{A}_X = \mathbf{A}_Y \equiv \mathbf{A}$. As we saw before, this, together with $\langle \mathbf{A}_X \mathbf{B}_X \rangle = \langle \mathbf{A}_Y \mathbf{C}_Y \rangle = \langle \mathbf{B}_Z \mathbf{C}_Z \rangle = -1$, implies a mathematical contradiction, and as such it is not possible to have a probability space (and a corresponding Boolean algebra) that yields those correlations. Therefore no joint probability exists that is consistent with those assumptions without a contextual labeling of variables.

However, it is interesting to note that contextuality does not require perfect correlations, which rarely appear in real situations (e.g., it may be possible that eventually Alice finds a position so clearly ridiculous that she might feel forced to agree with Carol). This is when probabilities come into play. For example, for a three random-variable system like the one above, Suppes and Zanotti [53] showed that a necessary and sufficient condition for the existence of a joint probability distribution is the satisfaction of the following inequalities

$$\begin{aligned} -1 \leq \langle \mathbf{AB} \rangle + \langle \mathbf{AC} \rangle + \langle \mathbf{BC} \rangle \\ \leq 1 + 2 \min\{\langle \mathbf{AB} \rangle, \langle \mathbf{AC} \rangle, \langle \mathbf{BC} \rangle\}. \quad (4) \end{aligned}$$

The perfect negative correlations clearly violate those inequalities, but to understand where the inequalities come from, let us look at the left hand side of the inequality. Consider the following mathematical implications: if the outcomes are $\mathbf{AB} = 1$ and $\mathbf{AC} = 1$, then $\mathbf{BC} = 1$; if the outcomes are $\mathbf{AB} = -1$ and $\mathbf{AC} = 1$, then $\mathbf{BC} = -1$; if $\mathbf{AB} = 1$ and $\mathbf{AC} = -1$, then $\mathbf{BC} = -1$; and if $\langle \mathbf{AB} \rangle = -1$ and $\langle \mathbf{AC} \rangle = -1$, then $\langle \mathbf{BC} \rangle = 1$. For each of these mathematical possibilities listed, we can compute the expression $\mathbf{AB} + \mathbf{AC} + \mathbf{BC}$, and find that it is always greater or equal than -1. Therefore, if we now average all possible consistent outcomes, we get that $\langle \mathbf{AB} + \mathbf{AC} + \mathbf{BC} \rangle = \langle \mathbf{AB} \rangle + \langle \mathbf{AC} \rangle + \langle \mathbf{BC} \rangle \geq -1$, which corresponds to the left hand of the inequality. For the right hand of the inequality, the argument is a little more complicated, and we refer readers to the proof in [53] .

To see how the Suppes and Zannoti inequality works, imagine that Alice and Bob agree only 10% of the time, Alice and Carol 15%, and Bob and Carol 30%. It is easy to compute the joint expectations as $\langle \mathbf{AB} \rangle = -0.80$, $\langle \mathbf{AC} \rangle = -0.7$, and $\langle \mathbf{AB} \rangle = -0.4$. With such values, $\langle \mathbf{AB} \rangle + \langle \mathbf{AC} \rangle + \langle \mathbf{BC} \rangle = -1.9 < -1$, which clearly violates the inequality. In this case, even though it is not immediate that the variables $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are contextual, the assumption that they are not is not consistent with an underlying Boolean algebra required by the existence of a probability space unless the variables are contextually labeled, as we saw from the sketch of proof for the inequalities.

It is interesting to note that this type of proper contextuality is not really part of most everyday language, but in the current context of political and social polarization it presents intriguing implications for more complex social dynamics in linguistic exchanges. Unless the speakers are part of a very careful and detailed conversation, most situations in which proper contextuality appears in language would pass unnoticed. However, this type of contextuality is exactly what is at play, together with direct contextual influences, in meaning formation and linguistic cooperation at a larger social level. For example, in hermeneutics, it is the recognition that the existing concepts, symbols, or propositions lead to contradictions that demands a re-examination of such concepts, and often the splitting of one concept into others with different assigned meanings, equivalent to the relabeling of variables in different contexts. In this sense, though less part of the usual discussions about contextuality in language, proper contextuality is essential for communication and belief formation.

5 Conclusions

In this paper we defended the claim that contextuality can be formalized in linguistics by borrowing models used in other disciplines. For instance, we described many examples of contextuality in linguistics that can be represented mathematically by a theory that has been also used in physics and psychology: Contextuality by Default (CbD). In CbD, a system of propositions or experimental outcomes is contextual if their probabilistic representation does not allow for a description by a single probability space or a joint probability distribution unless the random variables are labeled contextually. As we argued, this definition fits well with linguistics.

Contextuality in language is widely discussed in the literature, but as far as the authors know, the only attempt to provide a detailed mathematical theory that describes it is Abramsky and Sadrzadeh's paper on sheaf theory and semantics [54]. However, these authors focus on the formalization of semantic relations and contextuality without addressing its connection to the rational beliefs of speakers. To include belief, we used instead the mathematical framework of Contextuality by Default. There are similarities in the definition of contextuality between CbD and sheaf theory [9], but CbD provides a more immediate framework where meaning and belief can be readily connected. In other words, here we show how the assignment of beliefs by the speakers and listeners at a context may follow rational consistent norms. This has the advantage of directly connecting contextuality in language with decision-making under uncertainty, and more generally, with the dynamics of how speakers update linguistic information from context to context.

Under the assumption that rationality constrains such language exchanges, we argued that contextuality in language is not different from some situations of contextuality in physics or psychology, and context dependence described by CbD can be studied and expanded to other domains in which communication is involved. A potential application is of this framework is for example, in philosophy of mind language, where contextuality plays a fundamental role in many discussions concerning centered worlds and the two dimensions of meaning. In our approach, we achieve a characterization of semantic context dependence that is similar to the two dimensional semantic theory. In fact, CbD's use of an index for random variables can be conceived of as the additional dimension introduced by Stalnaker [19]. We emphasize, however, that this does not come, as proposed by some, with a corresponding dualistic perspective. A further empirical implication of the present approach is the identification of order effects in communication.

Another issue, raised by one of the reviewers, is whether our Contextuality Principle provides an optimal solution to the issue of linguistic contextuality. For example, could non-commutative algebras provide a better model than CbD? We do not know. Our paper proposes a specific approach that fills a lacuna in the literature, where no other detailed model exists. As such, we cannot infer whether our approach is optimal without a precise formulation of alternatives. It is not until then that one could approach the issue of optimality. We hope our contribution allows for such future research to happen.

Our main goal, however, is to highlight the relation between rational principles and contextuality in conversational exchanges. Consider the Gricean Maxims concerning rational cooperation in conversations (e.g., quantity, quality, relation, and manner). Our Contextuality Principle may place enough constraints to satisfy most, in not all of these maxims. If so, a more concise and formally rigorous approach to rational cooperation in communication would be on offer, with potential consequences for debates concerning the psychology of communicative cooperation. We mention this as a concluding thought on the implications of the Contextuality Principle. Developing this idea in detail is material for another paper.

Acknowledgments. The authors thank the two anonymous referees for their important questions and suggestions. We feel our manuscript improved substantially because of their reviews. JAB thanks John Perry and Paul Skokowski for their hospitality at the Center for the Explanation of Consiousness at Stanford. JAB also thanks Otavio Bueno and Adonai Sant'Anna for discussions. LPGA acknowledge support from the Patrick Suppes Gift Fund, and the Byrne Gift Funds at Stanford University.

Conflict of Interests. On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

[1] Lewis, K.S.: Dynamic Semantics. In: Oxford Handbooks Online. Oxford University Press, Oxford, UK (2018)

[2] Melamed, Y.Y., Lin, M.: Principle of Sufficient Reason. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy, Summer 2021 edn. Metaphysics Research Lab, Stanford University, Stanford, CA (2021). https://plato.stanford.edu/archives/sum2021/entries/sufficient-reason/

[3] Grice, P.: Studies in the Way of Words. Harvard University Press, Cambridge, Massachusetts (1989)

[4] Atmanspacher, H., Basieva, I., Busemeyer, J.R., Khrennikov, A.Y., Pothos, E.M., Shiffrin, R.M., Wang, Z.: What are the appropriate axioms of rationality for reasoning under uncertainty with resource-constrained systems? Behavioral and Brain Sciences **43** (2020). https://doi.org/10.1017/S0140525X19001535

[5] Howard, M., Wallman, J., Veitch, V., Emerson, J.: Contextuality supplies the 'magic' for quantum computation. Nature **510**(7505), 351–355 (2014). https://doi.org/10.1038/nature13460

[6] Cabello, A.: Exclusivity principle and the quantum bound of the Bell inequality. Physical Review A **90**(6), 062125 (2014). https://doi.org/10.1103/PhysRevA.90.062125

[7] Oas, G., de Barros, J.A.: A Survey of Physical Principles Attempting to Define Quantum Mechanics. In: Dzhafarov, E., Zhang, R., Jordan, S.M. (eds.) Contextuality From Quantum Physics to Psychology. World Scientific, Singapore (2015)

[8] Kochen, S., Specker, E.P.: The Problem of Hidden Variables in Quantum Mechanics. Journal of Mathematics and Mechanics **17**, 59–87 (1967)

[9] Dzhafarov, E.N.: The Contextuality-by-Default View of the Sheaf-Theoretic Approach to Contextuality. arXiv:1906.02718 [quant-ph] (2019). arXiv: 1906.02718. Accessed 2019-09-17

[10] Khrennikov, A.: Ubiquitous Quantum Structure. Springer, Heidelberg (2010)

[11] Busemeyer, J.R., Bruza, P.D.: Quantum Models of Cognition and Decision. Cambridge Univ. Press, Cambridge (2012)

[12] Haven, E., Khrennikov, A.: Quantum Social Science. Cambridge Univ. Press, Cambridge (2013)

[13] de Barros, J.A., Suppes, P.: Quantum mechanics, interference, and the brain. Journal of Mathematical Psychology **53**(5), 306–313 (2009). https://doi.org/10.1016/j.jmp.2009.03.005. Accessed 2015-01-08

[14] Dzhafarov, E.N., Kujala, J.V.: Contextuality-by-Default 2.0: Systems with Binary Random Variables. In: de Barros, J.A., Coecke, B., Pothos, E. (eds.) Quantum Interaction: 10th International Conference, QI 2016. Lecture Notes in Computer Science, vol. 10106. Springer, Cham, Switzerland (2017)

[15] Kujala, J.V., Dzhafarov, E.N., Larsson, J.-A.: Necessary and Sufficient Conditions for an Extended Noncontextuality in a Broad Class of Quantum Mechanical Systems. Physical Review Letters **115**(15), 150401 (2015). https://doi.org/10.1103/PhysRevLett.115.150401. Accessed 2015-10-12

[16] Cervantes, V.H., Dzhafarov, E.N.: Advanced analysis of quantum contextuality in a psychophysical double-detection experiment. Journal of Mathematical Psychology **79**, 77–84 (2017). https://doi.org/10.1016/j.jmp.2017.03.003

[17] Cervantes, V.H., Dzhafarov, E.N.: Snow queen is evil and beautiful: Experimental evidence for probabilistic contextuality in human choices. Decision **5**(3), 193–204 (2018). https://doi.org/10.1037/dec0000095

[18] Perry, J.: The Semantics and Pragmatics of Indexicals. In: Hale, B., Wright, C., Miller, A. (eds.) The Blackwell Guide to the Philosophy of Language, 2nd edn. Wiley Blackwell, Oxford (2017)

[19] Stalnaker, R.: Context and Content: Essays on Intentionality in Speech And Thought. Clarendon Press, Oxford ; New York (1999)

[20] Narens, L.: Probabilistic Lattices: With Applications to Psychology. World Scientific, Singapore (2014). Google-Books-ID: Bh23CgAAQBAJ

[21] Holik, F., Saenz, M., Plastino, A.: A discussion on the origin of quantum probabilities. Annals of Physics **340**(1), 293–310 (2014). https://doi.org/10.1016/j.aop.2013.11.005. Accessed 2014-02-22

[22] Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica: Journal of the Econometric Society, 263–291 (1979)

[23] Bruza, P., Kitto, K., Ramm, B.J., Sitbon, L.: A probabilistic framework for analysing the compositionality of conceptual combinations. Journal of Mathematical Psychology **67**, 26–38 (2015)

[24] Cappelen, H., Lepore, E.: Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism. Blackwell Publishing Ltd, Malden, MA (2005)

[25] Moore, D.W.: Measuring New Types of Question-Order Effects: Additive and Subtractive. The Public Opinion Quarterly **66**(1), 80–91 (2002). https://doi.org/10.2307/3078697. Accessed 2013-06-16

[26] Atmanspacher, H., Rmer, H.: Order effects in sequential measurements of non-commuting psychological observables. Journal of Mathematical Psychology **56**(4), 274–280 (2012). https://doi.org/10.1016/j.jmp.2012.06.003. Accessed 2015-03-11

[27] Pothos, E.M., Busemeyer, J.R.: Can quantum probability provide a new direction for cognitive modeling? Behavioral and Brain Sciences **36**(03), 255–274 (2013). https://doi.org/10.1017/S0140525X12001525

[28] Wang, Z., Busemeyer, J.R.: A Quantum Question Order Model Supported by Empirical Tests of an A Priori and Precise Prediction. Topics in Cognitive Science **5**(4), 689–710 (2013). https://doi.org/10.1111/tops.12040. Accessed 2014-10-31

[29] Wang, Z., Busemeyer, J.: Reintroducing the Concept of Complementarity into Psychology. Cognition **6**, 1822 (2015). https://doi.org/10.3389/fpsyg.2015.01822. Accessed 2015-11-19

[30] Wang, Z., Solloway, T., Shiffrin, R.M., Busemeyer, J.R.: Context effects produced by question orders reveal quantum nature of human judgments. Proceedings of the National Academy of Sciences **111**(26), 9431–9436 (2014). https://doi.org/10.1073/pnas.1407756111. Accessed 2015-06-04

[31] Yearsley, J.M., Busemeyer, J.R.: Quantum cognition and decision theories: A tutorial. Journal of Mathematical Psychology **74**, 99–116 (2016). https://doi.org/10.1016/j.jmp.2015.11.005. Accessed 2021-12-09

[32] Khrennikov, A.Y.: Contextual Approach to Quantum Formalism. Springer, Cham, Switzerland (2009)

[33] Haven, E., Khrennikov, A.: The Palgrave Handbook of Quantum Models in Social Science. Palgrave MacMillan, London (2017)

[34] Aerts, D., Broekaert, J., Gabora, L., Veloz, T.: The Guppy Effect as Interference. In: Busemeyer, J.R., Dubois, F., Lambert-Mogiliansky, A., Melucci, M. (eds.) Quantum Interaction. Lecture Notes in Computer Science, pp. 36–47. Springer, Berlin, Heidelberg (2012). http://link.springer.com/chapter/10.1007/978-3-642-35659-9_4 Accessed 2015-03-11

[35] Aerts, D., Sozzo, S.: Entanglement of Conceptual Entities in Quantum Model Theory (QMod). In: Busemeyer, J.R., Dubois, F., Lambert-Mogiliansky, A., Melucci, M. (eds.) Quantum Interaction. Lecture Notes in Computer Science, pp. 114–125. Springer, Berlin, Heidelberg (2012). http://link.springer.com/chapter/10.1007/978-3-642-35659-9_11 Accessed 2015-09-16

[36] Aerts, D., Sozzo, S.: Quantum Entanglement in Concept Combinations. International Journal of Theoretical Physics **53**(10), 3587–3603 (2013). https://doi.org/10.1007/s10773-013-1946-z. Accessed 2015-09-16

[37] Bruza, P., Kitto, K., Nelson, D., McEvoy, C.: Is there something quantum-like about the human mental lexicon? Journal of Mathematical Psychology **53**(5), 362–377 (2009)

[38] Khrennikov, A., Basieva, I., Dzhafarov, E.N., Busemeyer, J.R.: Quantum Models for Psychological Measurements: An Unsolved Problem. PLoS ONE **9**(10), 110909 (2014). https://doi.org/10.1371/journal.pone.0110909. Accessed 2015-03-11

[39] Busemeyer, J., Wang, Z.: Is there a problem with quantum models of psychological measurements? PLOS ONE **12**(11), 0187733 (2017). https://doi.org/10.1371/journal.pone.0187733. Publisher: Public Library of Science. Accessed 2021-02-02

[40] Ozawa, M., Khrennikov, A.: Application of theory of quantum instruments to psychology: Combination of question order effect with response replicability effect. Entropy **22**(11), 37 (2020). https://doi.org/10.3390/e22010037

[41] Ozawa, M., Khrennikov, A.: Modeling combination of question order

effect, response replicability effect, and qq-equality with quantum instruments. Journal of Mathematical Psychology **100**, 102491 (2021). https://doi.org/10.1016/j.jmp.2020.102491

[42] de Barros, J.A., Oas, G.: Quantum Cognition, Neural Oscillators, and Negative Probabilities. In: Haven, E., Khrennikov, A. (eds.) The Palgrave Handbook of Quantum Models in Social Science: Applications and Grand Challenges. Palgrave MacMillan, London (2015)

[43] de Barros, J.A.: Decision Making for Inconsistent Expert Judgments Using Negative Probabilities. Lecture Notes in Computer Science, pp. 257–269. Springer, Berlin/Heidelberg (2014)

[44] de Barros, J.A., Oas, G.: Negative probabilities and counter-factual reasoning in quantum cognition. Physica Scripta **T163**, 014008 (2014). Accessed 2014-04-16

[45] de Barros, J.A., Oas, G., Suppes, P.: Negative probabilities and Counterfactual Reasoning on the double-slit Experiment. In: Beziau, J.-Y., Krause, D., Arenhart, J.B. (eds.) Conceptual Clarification: Tributes to Patrick Suppes (1992-2014). College Publications, London (2015)

[46] de Barros, J.A., Kujala, J.V., Oas, G.: Negative probabilities and contextuality. Journal of Mathematical Psychology **74**, 34–45 (2016). https://doi.org/10.1016/j.jmp.2016.04.014. Accessed 2016-11-27

[47] Moreira, C., de Barros, J.A.: Order Effects in Bayesian Updates. Proceedings of the Annual Meeting of the Cognitive Science Society **43**(43) (2021). Accessed 2021-08-03

[48] Savage, L.J.: The Foundations of Statistics, 2nd edn. Dover Publications Inc., Mineola, New York (1972)

[49] Tversky, A., Shafir, E.: The Disjunction Effect in Choice Under Uncertainty. Psychological Science **3**(5), 305–309 (1992)

[50] Shafir, E., Tversky, A.: Thinking through uncertainty: Nonconsequential reasoning and choice. Cognitive Psychology **24**(4), 449–474 (1992)

[51] Cox, R.T.: The Algebra of Probable Inference. The John Hopkins Press, Baltimore (1961)

[52] Jaynes, E.T.: Probability Theory: the Logic of Science. Cambridge Univ Press, Cambridge, Great Britain (2003)

[53] Suppes, P., Zanotti, M.: When are probabilistic explanations possible? Synthese **48**(2), 191–199 (1981)

[54] Abramsky, S., Sadrzadeh, M.: Semantic Unification. In: Casadio, C., Coecke, B., Moortgat, M., Scott, P. (eds.) Categories and Types in Logic, Language, and Physics: Essays Dedicated to Jim Lambek on the Occasion Of His 90th Birthday. Lecture Notes in Computer Science, pp. 1–13. Springer, Berlin, Heidelberg (2014). https://doi.org/10.1007/978-3-642-54789-8_1