Article

Contextuality with Disturbance and without: Neither Can Violate Substantive Requirements the Other Satisfies

Ehtibar N. Dzhafarov [1,*] and Janne V. Kujala [2]

[1] Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA
[2] Department of Mathematics and Statistics, University of Turku, FI-20014 Turun yliopisto, Finland
[*] Correspondence: ehtibar@purdue.edu

Abstract: Contextuality was originally defined only for consistently connected systems of random variables (those without disturbance/signaling). Contextuality-by-Default theory (CbD) offers an extension of the notion of contextuality to inconsistently connected systems (those with disturbance) by defining it in terms of the systems' couplings subject to certain constraints. Such extensions are sometimes met with skepticism. We pose the question of whether it is possible to develop a set of substantive requirements (i.e., those addressing a notion itself rather than its presentation form) such that (1) for any consistently connected system, these requirements are satisfied, but (2) they are violated for some inconsistently connected systems. We show that no such set of requirements is possible, not only for CbD but for all possible CbD-like extensions of contextuality. This follows from the fact that any extended contextuality theory T is contextually equivalent to a theory T′ in which all systems are consistently connected. The contextual equivalence means the following: there is a bijective correspondence between the systems in T and T′ such that the corresponding systems in T and T′ are, in a well-defined sense, mere reformulations of each other, and they are contextual or noncontextual together.

Keywords: contextual equivalence; contextuality; consistent connectedness; consistification; connections; disturbance; signaling

1. Introduction

A formal theory T of contextuality is defined by a class $\Re$ of possible systems of random variables and a rule by which these systems are divided into noncontextual and contextual ones. In the original theory of contextuality (a term in which we include both the Kochen–Specker contextuality and the contextuality in distributed systems, referred to as nonlocality [1–8]), the class $\Re$ is confined to consistently connected systems, or a subclass thereof. These are the systems with no "disturbance" or "signaling," which means that the variables representing the same property (answering the same question) in different contexts are identically distributed. The Contextuality-by-Default theory (CbD) extends the notion of contextuality to all systems of random variables, including those with disturbance [9,10], and it has been applied to several experimental and theoretical situations [11–18]. A recent workshop on contextuality [19] exhibited a renewed interest to studying contextuality in inconsistently connected systems, including approaches that are distinctly non-CbD-like [20–22], and some work directly critical of CbD ([23], responded to in Ref. [24]).

The present paper is not about CbD specifically. Rather, it is about a broad class of all possible *CbD-like theories*, as defined below. The plan and the main message of the paper are as follows. In Section 2, we present the terminology and notation to be used and define the notion of a system of random variables modeling (representing or describing) another system. In Section 3, we define the traditional notion of contextuality in the language of probabilistic couplings [25], and we introduce the notion of C-contextuality as a very broad

generalization of both traditional contextuality and CbD-contextuality. In Section 4, we introduce the notion of consistification of a system and show that any theory T, irrespective of its class $\mathfrak{R}$ of systems and a specific version of the C-contextuality it uses, can be redefined as a theory T′, whose systems are consistently connected, and that uses the traditional notion of contextuality. Because of this, we conclude that there can be no set of substantive requirements X for the notion of contextuality that are satisfied by all consistently connected systems but contravened by some inconsistently connected ones. Indeed, if such a set of requirements existed, one could form a theory T whose class $\mathfrak{R}$ includes some systems contravening X. However, X would then be satisfied by the theory T′ that is contextually equivalent to T and a mere reformulation thereof. Consequently, requirements X cannot be substantive: they address a form rather than the substance of the notion of contextuality. In Section 5, we discuss some issues related to the consistified systems (the term used for the consistently connected systems in T′), including the representability thereof by hidden variable models. We also briefly discuss there a still more general (in fact, maximally general) notion of C-contextuality, one that does not have the existence-and-uniqueness property postulated for C-contextuality. In the final analysis, this does not alter the main conclusion of the paper.

The idea that consistification precludes the possibility of rejecting extended contextuality while accepting the traditional one was previously mentioned in Ref. [24]. However, it was confined to CbD only and mentioned without elaborating. The consistification procedure was first described in Ref. [13] for an older version of the CbD approach, and it was elaborated and adapted to the current version of CbD in Ref. [26]. Finally, the C-contextuality in our paper generalizes a more limited version of C-contextuality that was used in Ref. [27] as a generalization of the CbD approach.

2. Basic Notions

A *system* of random variables is a set of double-indexed random variables

$$\mathcal{R} = \left\{ R_q^c : c \in C, q \in Q, q \prec c \right\}, \tag{1}$$

where $q \in Q$ identifies what the random variable R_q^c represents (measures, responds to, or describes); $c \in C$ identifies circumstances under which R_q^c is recorded (including what other random variables are recorded together with R_q^c); q and c are referred to as, respectively, the *content* and the *context* of the random variable R_q^c; and the relation $q \prec c$ indicates that a variable with content q is recorded in context c. As an example, this is a system with $Q = \{1, 2, 3\}$ and $C = \{1, 2, 3, 4\}$:

$$
\begin{array}{|c|c|c||c|}
\hline
R_1^1 & R_2^1 & & c = 1 \\
\hline
 & R_2^2 & R_3^2 & c = 2 \\
\hline
R_1^3 & & R_3^3 & c = 3 \\
\hline
R_1^4 & R_2^4 & R_3^4 & c = 4 \\
\hline\hline
q = 1 & q = 2 & q = 3 & \mathcal{R} \\
\hline
\end{array}
. \tag{2}
$$

The subset $R^c = \left\{ R_q^c : q \in Q, q \prec c \right\}$ of random variables recorded in the same context c (a row in the matrix above) is termed a *bunch*, and the subset $\mathcal{R}_q = \left\{ R_q^c : c \in C, q \prec c \right\}$ of random variables sharing a content q (a column in the matrix above) is termed a *connection*. The difference in font (R^c vs. $\mathcal{R}_q$) reflects the fact that R^c is a random variable in its own right (i.e., all its components are jointly distributed), whereas the components of $\mathcal{R}_q$ are not jointly distributed. In fact, no two random variables R_q^c and $R_{q'}^{c'}$ are jointly distributed unless they are in the same bunch, $c = c'$. The measurable space on which R_q^c is distributed is assumed to be the same for all elements of a connection and can be denoted $\left(A_q, \Sigma_q \right)$.

The triple $(Q, C, \prec)$ is called the *format* of the system. It is essentially the mathematical depiction of "what the system is about," what kind of empirical or theoretical situation it represents. Thus, the format of the system in (2) can be presented as

$\star$	$\star$		$c = 1$
	$\star$	$\star$	$c = 2$
$\star$		$\star$	$c = 3$
$\star$	$\star$	$\star$	$c = 4$
$q = 1$	$q = 2$	$q = 3$	$\mathcal{R}$'s format

$$\text{(3)}$$

where $\star$ indicates the elements of the relation $\prec$. To define a system of a given format, one has to specify the distributions of its bunches.

As should be clear from Abstract and Introduction, in this paper, we use the notion of one system of random variables, $\mathcal{B}$, being a "mere reformulation" of another, $\mathcal{A}$. Intuitively, this means that regardless of what empirical or theoretical situation is modeled (described, represented) by $\mathcal{A}$, it is also modeled by $\mathcal{B}$. The relation between a system and a situation it depicts is difficult to formalize directly, as one would have then to impose some formal structure on the situation being represented before it is represented (as in the representational theory of measurement, [28,29]). However, it is sufficient for our purposes to formalize a simpler relationship: between a system $\mathcal{A}$ and another system that models (describes, represents) the system $\mathcal{A}$. Moreover, rather than presenting this relationship in a most general possible way, it will suffice to describe one special, universally applicable construction of the modeling systems $\mathcal{B}$. We will refer to this construction as *canonical modeling*.

Consider two classes of systems, $\mathfrak{R}$ and $\mathfrak{R}^\dagger$, in a bijective correspondence to each other, about which we say that any system in $\mathfrak{R}$ is canonically modeled by the corresponding system in $\mathfrak{R}^\dagger$. The following definition gives a precise meaning to this relation.

Definition 1. *We say that a system $\mathcal{R} \in \mathfrak{R}$ with format $(Q, C, \prec)$ is canonically modeled by a system $\mathcal{R}^\dagger \in \mathfrak{R}^\dagger$ with format $\left(Q^\dagger, C^\dagger, \prec^\dagger\right)$ if*

(canonical contents) $Q^\dagger = \{(q, c) : q \prec c\}$,

(canonical contexts) $C^\dagger = \{(\cdot, c) : c \in C\} \sqcup \{(q, \cdot) : q \in Q\}$,

(canonical relation) $(q, c) \prec^\dagger (\cdot, c) \iff (q, c) \in Q^\dagger$, *and* $(q, c) \prec^\dagger (q, \cdot) \iff (q, c) \in Q^\dagger$,

(main bunches) $R^{(\cdot, c)} = \left\{ R^{(\cdot, c)}_{(q, c)} : (q, c) \prec^\dagger (\cdot, c) \right\} \overset{d}{=} \left\{ R^c_q : q \prec c \right\} = R^c$

(auxiliary bunches) $R^{(q, \cdot)} = \left\{ R^{(q, \cdot)}_{(q, c)} : (q, c) \prec^\dagger (q, \cdot) \right\}$ *is uniquely determined by the distributions of the corresponding variables in* $\mathcal{R}_{(q, \cdot)} = \left\{ R^{(\cdot, c)}_{(q, c)} : (q, c) \prec^\dagger (q, \cdot) \right\}$.

Here, the symbol $\overset{d}{=}$ stands for "has the same distribution as". The dot symbol in $(\cdot, c)$ and $(q, \cdot)$ should be taken as part of the names of these contexts. We choose this notation to emphasize that every random variable R^c_q of the system $\mathcal{R}$ is placed in $\mathcal{R}^\dagger$ within two contexts, $(\cdot, c)$ and $(q, \cdot)$, whose names are derived from the indices of the variable. Note that the variables in the set $\mathcal{R}_{(q, \cdot)}$ defined here have the same distributions as the corresponding variables in $\mathcal{R}_q = \left\{ R^c_q : q \prec c \right\}$. We use the former set, however, to emphasize that the auxiliary bunches are uniquely determined by the corresponding variables in the main bunches. Note that the variables in $\mathcal{R}_{(q, \cdot)}$ are not jointly distributed, so $R^{(q, \cdot)}$ depends on their individual distributions only.

To give an example, consider the systems

R_1^1	R_2^1	$c=1$
R_1^2	R_2^2	$c=2$
$q=1$	$q=2$	$\mathcal{A}$

$R_{(1,1)}^{(\cdot,1)}$	$R_{(2,1)}^{(\cdot,1)}$			$c=(\cdot,1)$
		$R_{(1,2)}^{(\cdot,2)}$	$R_{(2,2)}^{(\cdot,2)}$	$c=(\cdot,2)$
$R_{(1,1)}^{(1,\cdot)}$		$R_{(1,2)}^{(1,\cdot)}$		$c=(1,\cdot)$
	$R_{(2,1)}^{(2,\cdot)}$		$R_{(2,2)}^{(2,\cdot)}$	$c=(2,\cdot)$
$q=(1,1)$	$q=(2,1)$	$q=(1,2)$	$q=(2,2)$	$\mathcal{B}$

$$(4)$$

Observe that in system $\mathcal{B}$ the contents, contexts, and the relation between them are constructed in accordance with Definition 1. System $\mathcal{B}$ canonically models system $\mathcal{A}$ if

$$\left\{ R_{(1,1)}^{(\cdot,1)}, R_{(2,1)}^{(\cdot,1)} \right\} \overset{d}{=} \left\{ R_1^1, R_2^1 \right\}, \qquad \left\{ R_{(1,2)}^{(\cdot,2)}, R_{(2,2)}^{(\cdot,2)} \right\} \overset{d}{=} \left\{ R_1^2, R_2^2 \right\}, \tag{5}$$

and if there is a rule by which the distribution of

$$R^{(q,\cdot)} = \left\{ R_{(q,1)}^{(q,\cdot)}, R_{(q,2)}^{(q,\cdot)} \right\}, \qquad q=1,2, \tag{6}$$

is uniquely determined by the distributions of the corresponding variables in

$$\mathcal{R}_{(q,\cdot)} = \left\{ R_{(q,1)}^{(\cdot,1)}, R_{(q,2)}^{(\cdot,2)} \right\}, \qquad q=1,2. \tag{7}$$

Observe the following properties of canonical modeling.

1. The formats of $\mathcal{R}$ and $\mathcal{R}^\dagger$ are reconstructible from each other, and so are the bunches of the two systems. Moreover, $\mathcal{R}^\dagger$ faithfully replicates the bunches of $\mathcal{R}$. This allows one to say that $\mathcal{R}$ and $\mathcal{R}^\dagger$ describe the same empirical or theoretical situation.

2. One might wonder why we need the auxiliary contexts at all, and they are indeed unnecessary if all one wants is a system modeling another system, e.g.,

$R_{(1,1)}^{(\cdot,1)}$	$R_{(2,1)}^{(\cdot,1)}$			$c=(\cdot,1)$
		$R_{(1,2)}^{(\cdot,2)}$	$R_{(2,2)}^{(\cdot,2)}$	$c=(\cdot,2)$
$q=(1,1)$	$q=(2,1)$	$q=(1,2)$	$q=(2,2)$	$\mathcal{B}'$

However, we will see the utility of the auxiliary contexts when we introduce consistifications and contextual equivalence, in Section 4.

3. The contents in the modeling system are "contextualized". For instance, system $\mathcal{A}$ in (4) may be describing an experiment in which two questions, $q=1$ and $q=2$, are asked in two orders, $c=1$ indicating "1 then 2" and $c=2$ indicating "2 then 1" [30,31]. In this case, in the modeling system, the content $q=(1,2)$ should be interpreted as "question 1 asked second", and $q=(1,1)$ should be interpreted as "question 1 asked first". We return to the issue of interpretation in Section 5.1.

4. The indexation of the variables in a canonical model is clearly redundant, and it can be simplified. It is more important, however, to maintain the general logic of indexing the variables by their contents and contexts.

3. Traditional and Extended Contextuality

A system $\mathcal{R}$ is *consistently connected* if in every connection $\mathcal{R}_q$ all its constituent variables have one and the same distribution. Otherwise, the system is *inconsistently connected*. (The latter term is also used to designate arbitrary systems, i.e., in the meaning of "not necessarily consistently connected".)

An *overall coupling* of a system $\mathcal{R}$ in (1) is an identically labeled system

$$S = \left\{ S_q^c : c \in \mathsf{C}, q \in \mathsf{Q}, q \prec c \right\} \tag{8}$$

of *jointly distributed* random variables such that its bunches S^c are distributed as the corresponding bunches R^c,

$$S^c \stackrel{d}{=} R^c. \tag{9}$$

Clearly, S has the same format as $\mathcal{R}$. A *coupling* S_q of a connection $\mathcal{R}_q$ is a set

$$S_q = \left\{ S_q^c : c \in C, q \prec c \right\} \tag{10}$$

of jointly distributed random variables such that $S_q^c \stackrel{d}{=} R_q^c$ for all its elements. A connection coupling S_q is said to be an *identity coupling* if $S_q^c = S_q^{c'}$ for any two of its elements. Obviously, such a coupling exists if and only if all of its elements (equivalently, all elements of the connection $\mathcal{R}_q$) have one and the same distribution. Moreover, the identity coupling is unique if it exists. (The uniqueness of a coupling should always be understood as the uniqueness of its distribution. In other words, it is irrelevant on what domain probability space the coupling is defined as a random variable.)

The traditional notion of contextuality is confined to consistently connected systems, and it can be rigorously defined in our terminology as follows.

Definition 2. *A consistently connected system $\mathcal{R} \in \mathfrak{R}$ is noncontextual if it has a coupling S in which any connection S_q is the identity coupling of the connection $\mathcal{R}_q$. Otherwise, the system is contextual.*

The class of all possible systems $\mathcal{R}$ in a theory T is denoted $\mathfrak{R}$. For instance, $\mathfrak{R}$ can only contain the systems with finite sets Q and C, or only the systems with dichotomous random variables. By constraining the class $\mathfrak{R}$, one induces constraints on all possible random variables, $R_q^c \in \mathfrak{R}_+^+$, on bunches of random variables, $R^c \in \mathfrak{R}^+$, and on possible connections, $\mathcal{R}_q \in \mathfrak{R}_+$.

In CbD, contextuality of a system $\mathcal{R}$ is defined by considering its couplings S and determining if, in some of them, the couplings S_q of the system's connections $\mathcal{R}_q$ satisfy a certain statement. To generalize this definition to all possible CbD-like theories, all one has to do is to replace this specific statement with one that is (almost) arbitrary. Let C be any statement of the form "the coupling of connection $\mathcal{R}_q$ has the following properties: ...". The only constraints we impose on C are as follows.

Definition 3. C *is considered well-fitting if (1) for any connection $\mathcal{R}_q \in \mathfrak{R}_+$, there is one and only one coupling S_q of $\mathcal{R}_q$ that satisfies C, and (2) if $\mathcal{R}_q$ consists of identically distributed random variables, then the coupling that satisfies C is the identity coupling. We denote such a coupling of $\mathcal{R}_q$ as $\mathsf{C}\left[\mathcal{R}_q\right]$.*

To give an example of a well-fitting statement C: in CbD, if the class $\mathfrak{R}$ of all possible systems is confined to the systems with dichotomous variables, the well-fitting statement is C = "for any two random variables $S_q^{c_1}$ and $S_q^{c_2}$ in the coupling of connection $\mathcal{R}_q$, the probability of $S_q^{c_1} = S_q^{c_2}$ is maximal possible". Another example: if the class $\mathfrak{R}$ of all possible systems is confined to the systems with real-valued (or more generally, linearly ordered) variables, then a well-fitting statement can be C = "for any two random variables $S_q^{c_1}$ and $S_q^{c_2}$ in the coupling of connection $\mathcal{R}_q$, $S_q^{c_1}$ and $S_q^{c_2}$ have the same quantile rank". In Section 5.3, we discuss the possibility of dropping the first of the two defining properties of a well-fitting statement C.

Definition 4. *Given a well-fitting C, a system $\mathcal{R}$ is C-noncontextual if it has a coupling S such that, for any connection $\mathcal{R}_q$ of the system, the connection coupling S_q coincides with $\mathsf{C}\left[\mathcal{R}_q\right]$. Otherwise, the system is C-contextual.*

4. Equivalence and Impossibility Theorems

It follows from the last two definitions that, for a well-fitting C, a consistently connected system is C-noncontextual if and only if it is noncontextual in the traditional sense (i.e., in the sense of Definition 2). In other words, any extension of the notion of contextuality using a well-fitting C properly reduces to the traditional notion when confined to consistently connected systems. This is not, obviously, sufficient to consider the extension of contextuality by means of a well-constructed C. There may be other desiderata for a well-constructed notion of contextuality, and a specific choice of C may not satisfy them. The question we pose now is as follows:

Q*: Is it possible to formulate a set of such desiderata/requirements X for the notion of contextuality that, for some choice of C, (1) X is satisfied for any consistently connected system, but (2) X is not satisfied for some inconsistently connected systems?

Note that we impose no constraints on what X may entail, except for its being related to contextuality. It may, e.g., for some relation B between systems, have the form "if system $\mathcal{R}_1$ is (non)contextual, then any system $\mathcal{R}_2$ related to $\mathcal{R}_1$ by B is (non)contextual" [24].

To answer the question **Q***, we need the following result.

Theorem 1. *For any well-fitting* C *and system* $\mathcal{R}$*, there is a consistently connected system* $\mathcal{R}^{\ddagger}$ *that canonically models it (Definition 1), such that* $\mathcal{R}$ *is* C*-contextual (Definition 4) if and only if* $\mathcal{R}^{\ddagger}$ *is contextual in the traditional sense (Definition 2).*

Proof. Let $\mathcal{R}^{\ddagger}$ be a canonically modeling system for $\mathcal{R}$, with

$$\mathcal{R}^{(q,\cdot)} = \left\{ R_{(q,c)}^{(q,\cdot)} : (q,c) \prec^{\ddagger} (q,\cdot) \right\} \overset{d}{=} \mathsf{C}\left[\left\{ R_q^c : q \prec c \right\} \right] = \mathsf{C}\left[\mathcal{R}_q \right]. \tag{*}$$

One can check that $\mathcal{R}^{\ddagger}$ is consistently connected: every connection $\mathcal{R}_{(q,c)}$ of $\mathcal{R}^{\ddagger}$ consists of precisely two variables, $R_{(q,c)}^{(\cdot,c)}$ and $R_{(q,c)}^{(q,\cdot)}$, where $R_{(q,c)}^{(\cdot,c)} \overset{d}{=} R_{(q,c)}^{(q,\cdot)}$. Indeed, $R_{(q,c)}^{(\cdot,c)} \overset{d}{=} R_q^c$, because $R^{(\cdot,c)} \overset{d}{=} R^c$ in any canonically modeling system and $R_{(q,c)}^{(q,\cdot)} \overset{d}{=} R_q^c$ because we know from (*) that $R_{(q,c)}^{(q,\cdot)} \overset{d}{=} S_q^c$, where $S_q^c \in \mathsf{C}\left[\mathcal{R}_q \right]$.

The system $\mathcal{R}^{\ddagger}$ thus constructed is referred to as a *consistification* of $\mathcal{R}$. We can now define the consistification $S^{\ddagger}$ of a coupling S of a system in precisely the same way as for the system itself, except that (*) is replaced with the straightforward

$$S^{(q,\cdot)} = S_q,$$

with the obvious correspondence between the different indexations within the two random vectors. Clearly, $S^{\ddagger}$ is a coupling of $\mathcal{R}^{\ddagger}$.

Assume now that $\mathcal{R}$ is noncontextual. This means that it has a coupling S such that (a) $S^c \overset{d}{=} R^c$ for every $c \in C$ and (b) $S_q = \mathsf{C}\left[\mathcal{R}_q \right]$ for every $q \in Q$. Then, in the coupling $S^{\ddagger}$ of system $\mathcal{R}^{\ddagger}$, we have (a') $S^{(\cdot,c)} \overset{d}{=} R^{(\cdot,c)}$ for every $(\cdot,c) \in C^{\ddagger}$, and (b') $S^{(q,\cdot)} = \mathsf{C}\left[\mathcal{R}_q \right]$ for every $(q,\cdot) \in C^{\ddagger}$. Moreover, since both $S_{(q,c)}^{(\cdot,c)}$ and $S_{(q,c)}^{(q,\cdot)}$ equal S_q^c, we have (c') $S_{(q,c)}^{(\cdot,c)} = S_{(q,c)}^{(q,\cdot)}$. However, (a')-(b')-(c') mean that $\mathcal{R}^{\ddagger}$ is noncontextual in the traditional sense. The implication here is easily seen to be reversible, and we conclude that $\mathcal{R}$ is noncontextual if and only if so is $\mathcal{R}^{\ddagger}$. $\square$

In our example (4), $\mathcal{B}$ is a consistification of $\mathcal{A}$ if we specify the rule for the auxiliary bunches as follows: $R_{(q,c)}^{(q,\cdot)} \overset{d}{=} R_{(q,c)}^{(\cdot,c)}$, and the distribution of $R^{(q,\cdot)}$ is the same as that of $\mathsf{C}\left[\mathcal{R}_q \right]$.

If C is chosen as in CbD, the consistification of the system $\mathcal{R}$ in (2) is the system below (omitting, for simplicity, the parentheses and commas in $R^{(\cdot,c)}_{(q,c)}$ and $R^{(q,\cdot)}_{(q,c)}$):

$$
\begin{array}{|c|c|c|c|c|c|c|c||c|}
\hline
R^{\cdot 1}_{11} & R^{\cdot 1}_{21} & & & & & & & c = \cdot 1 \\
\hline
& & R^{\cdot 2}_{22} & R^{\cdot 2}_{32} & & & & & c = \cdot 2 \\
\hline
& & & & R^{\cdot 3}_{13} & R^{\cdot 3}_{33} & & & c = \cdot 3 \\
\hline
& & & & & & R^{\cdot 4}_{14} & R^{\cdot 4}_{24} & R^{\cdot 4}_{34} & c = \cdot 4 \\
\hline
R^{1\cdot}_{11} & & & & R^{1\cdot}_{13} & & R^{1\cdot}_{14} & & c = 1\cdot \\
\hline
& R^{2\cdot}_{21} & R^{2\cdot}_{22} & & & & & R^{2\cdot}_{24} & c = 2\cdot \\
\hline
& & & R^{3\cdot}_{32} & & R^{3\cdot}_{33} & & & R^{3\cdot}_{34} & c = 3\cdot \\
\hline
q = 11 & q = 21 & q = 22 & q = 32 & q = 13 & q = 33 & q = 14 & q = 24 & q = 34 & \mathcal{R}^{\ddagger} \\
\hline
\end{array}
\quad , \quad (11)
$$

where all variables are assumed to be dichotomous, and in each of the auxiliary bunches, the variables are pairwise equal with maximal possible probability.

For the purposes of contextuality analysis, $\mathcal{R}^{\ddagger}$ can be viewed as a mere reformulation of $\mathcal{R}$, a different form of the same substance. We express this fact by saying that $\mathcal{R}$ and $\mathcal{R}^{\ddagger}$ are *contextually equivalent*. (In Refs. [24,26], contextual equivalence is defined more narrowly, requiring also the numerical coincidence of certain measures of contextuality, such as contextual fraction [32]. In this paper, however, the level of abstraction is higher, and we only consider the notion of contextuality rather than its quantifications.)

Consider now a theory of (generally, extended) contextuality $\mathsf{T} = \mathsf{T}(\mathsf{C}, \mathfrak{R})$. In accordance with Theorem 1, we can form the class $\mathfrak{R}^{\ddagger}$ of the consistifications of the elements of $\mathfrak{R}$ in a bijective correspondence with $\mathfrak{R}$. By extension of the term, we can say that T and $\mathsf{T}' = \mathsf{T}\left(\mathsf{C}_0, \mathfrak{R}^{\ddagger}\right)$ are *contextually equivalent*. C_0 here denotes the statement "the connection R_q has an identity coupling" that underlies the traditional notion of contextuality, because by definition, it can be viewed as a special case of any well-fitting statement C. We have now everything we need to demonstrate our main conclusion. Let there be a set of requirements X of the notion of contextuality that are satisfied by all consistently connected systems (using the traditional contextuality) and contravened by some inconsistently connected ones, using some version of C-contextuality. Let T include some of the inconsistently connected systems contravening X. Clearly then, requirements X contradict theory T, but they are satisfied by the contextually equivalent theory $\mathsf{T}' = \mathsf{T}\left(\mathsf{C}_0, \mathfrak{R}^{\ddagger}\right)$. Therefore, X is not a set of substantive requirements. We can summarize this as a formal theorem.

Theorem 2. *For any well-fitting* C, *there can be no set of substantive requirements* X *of the notion of contextuality that are satisfied by all consistently connected systems (using the traditional contextuality) and contravened by some inconsistently connected ones, using* C-*contextuality.*

Of course, a set of requirements X satisfied by T' but not T can be readily formulated. The theorem says, however, that all it can do is lead one to prefer one of two equivalent representations of contextuality, without affecting the substance of the notion.

Note also that in the theorem just formulated, we assume no relationship between the set of requirements X and the bijective correspondence relating $\mathfrak{R}$ to $\mathfrak{R}^{\ddagger}$. In particular, let X have the form "if system $\mathcal{R}_1$ is contextual, then any system $\mathcal{R}_2$ related to $\mathcal{R}_1$ by relation B is contextual." It is not necessary then, although not excluded either, that $\mathcal{R}_2^{\ddagger}$ is also related to $\mathcal{R}_1^{\ddagger}$ by relation B. All that is stated in the theorem above is that if one wishes to use this X as a substantive principle in testing competing theories, then the failure of a theory to satisfy it cannot be selectively attributed to the fact that its $\mathfrak{R}$ contains inconsistently connected systems.

5. Miscellaneous Remarks

Here, we consider a few issues related to the main point of this paper.

5.1. Interpretation of Contents and Contexts

Dealing with consistified systems $R^{\ddagger}$, one needs to get used to a new interpretation of contents and contexts of the random variables: as mentioned previously, in $R^{\ddagger}$, contents are "contextualized," with (q, c) in place of just q, and the contexts are simply marginalized contents, $(\cdot, c)$ and $(q, \cdot)$. Consider as an example the EPR/Bohm experiment, the most widely investigated paradigm in contextuality/nonlocality research [1,33,34]. In the usual CbD notation, the system representing it is

R_1^1	R_2^1			$c = 1$	
	R_2^2	R_3^2		$c = 2$	
		R_3^3	R_4^3	$c = 3$	,
R_1^4			R_4^4	$c = 4$	
$q = 1$	$q = 2$	$q = 3$	$q = 4$	$\mathcal{A}$	

(12)

where $q = 1$ and $q = 3$ denote two settings (axes) to be chosen between by Alice, $q = 2$ and $q = 4$ are settings to choose between by Bob, c indicates the combination of their choices, and R_q^c is the dichotomous (spin-up/spin-down) variables. The consistified representation of the same experiment is (again, omitting the parentheses and commas in the indexation)

$R_{11}^{\cdot 1}$	$R_{21}^{\cdot 1}$							$c = \cdot 1$	
		$R_{22}^{\cdot 2}$	$R_{32}^{\cdot 2}$					$c = \cdot 2$	
				$R_{33}^{\cdot 3}$	$R_{43}^{\cdot 3}$			$c = \cdot 3$	
						$R_{44}^{\cdot 4}$	$R_{14}^{\cdot 4}$	$c = \cdot 4$	
$R_{11}^{1\cdot}$							$R_{14}^{1\cdot}$	$c = 1\cdot$	.
	$R_{21}^{2\cdot}$	$R_{22}^{2\cdot}$						$c = 2\cdot$	
			$R_{32}^{3\cdot}$	$R_{33}^{3\cdot}$				$c = 3\cdot$	
					$R_{43}^{4\cdot}$	$R_{44}^{4\cdot}$		$c = 4\cdot$	
$q = 11$	$q = 21$	$q = 22$	$q = 32$	$q = 33$	$q = 43$	$q = 44$	$q = 14$	$\mathcal{A}^{\ddagger}$	

(13)

The interpretation of, say, the content $q = (3, 2)$ here is as follows: it is the choice of axis 3 (that we know to be made by Alice) when Bob's choice of his axis forms combination 2 with Alice's choice (which we know to mean that Bob chooses axis 2). The interpretation of context $c = (\cdot, 2)$ is that it is simply the set of contents whose second component is 2. Similarly, $c = (3, \cdot)$ is the set of contents whose first component is 3. The random variables within context $c = (\cdot, 2)$ are jointly distributed by observation, whereas the random variables within context $c = (3, \cdot)$ are jointly distributed by computation (that, in turn, is uniquely determined by the observations). If C is defined in accordance with CbD, $\left(R_{21}^{2\cdot}, R_{22}^{2\cdot} \right)$ is computed so that $R_{21}^{2\cdot} \overset{d}{=} R_{21}^{\cdot 1}$, $R_{22}^{2\cdot} \overset{d}{=} R_{22}^{\cdot 2}$ (consistent connectedness), and the probability of $R_{21}^{2\cdot} = R_{22}^{2\cdot}$ is maximal possible. In particular, if $R_{21}^{\cdot 1} \overset{d}{=} R_{22}^{\cdot 2}$, then $R_{21}^{2\cdot} = R_{22}^{2\cdot}$.

5.2. Hidden Variable Models

One possible argument against contextuality in inconsistently connected systems is that it is not distinguishable from inconsistent connectedness itself in the language of hidden variable models (HVMs). If, the argument goes, a consistently connected system $\mathcal{R}$ in (1) is noncontextual, it has a coupling S in which all random variables can be presented as

$$S_q^c = F(q, \Lambda),$$

(14)

where Λ is a "hidden" random variable [35]. If $\mathcal{R}$ is contextual, then all its couplings can only be presented as

$$S_q^c = F(q, c, \Lambda),$$

(15)

with ineliminable c. However, the latter HVM representation is also required for all inconsistently connected systems, irrespective of whether they are C-contextual or C-

noncontextual. We would argue in response that this only means that on this general level (merely showing the arguments of the functions), the language of HVMs is too crude to capture the subtler properties of the couplings, such as contextuality under inconsistent connectedness. However, even if one takes this issue as a matter of concern, it is eliminated by the consistification procedure. The system $\mathcal{R}^{\ddagger}$ corresponding to $\mathcal{R}$ is noncontextual if and only if it has a coupling $S^{\ddagger}$ such that, for all $(q,c) \in Q^{\ddagger}$,

$$S^{(\cdot,c)}_{(q,c)} = G((q,c),\Lambda) = S^{(q,\cdot)}_{(q,c)}, \tag{16}$$

for some random variable Λ. If $\mathcal{R}^{\ddagger}$ is contextual, then in all its couplings, for some $(q,c) \in Q^{\ddagger}, S^{(\cdot,c)}_{(q,c)} \neq S^{(q,\cdot)}_{(q,c)}$, which means that their HVM representations can only be different functions,

$$S^{(\cdot,c)}_{(q,c)} = G_1((q,c),\Lambda), \qquad S^{(q,\cdot)}_{(q,c)} = G_2((q,c),\Lambda), \tag{17}$$

or, equivalently, the same function but with differently distributed hidden variables,

$$S^{(\cdot,c)}_{(q,c)} = H((q,c),\Lambda_1), \qquad S^{(q,\cdot)}_{(q,c)} = H((q,c),\Lambda_2). \tag{18}$$

It is instructive to apply this to the EPR/Bohm systems $\mathcal{A}$ and $\mathcal{A}^{\ddagger}$ in (12) and (13). Here, contextuality is traditionally referred to as nonlocality because for the contextual system $\mathcal{A}$, all its couplings are represented in the form of (15): the ineliminable dependence on c here is interpreted as the dependence of a measurement on a remote setting. However, if one models the EPR/Bohm experiment by system $\mathcal{A}^{\ddagger}$ instead, the HVM representations (16) and (17) both contain the contextualized content (q,c) as an argument. Following the logic above, they should both be considered nonlocal, even though one of them represents a noncontextual system and is equivalent to (14), while the other represents a contextual system and is equivalent to (15). It seems to us, in agreement with other authors [36], that this demonstration speaks against a naturalistic interpretation of the HVMs in terms of physical dependences.

5.3. The Existence and Uniqueness Constraint

In the definition of C-couplings, their reducibility to identity couplings when applied to identically distributed variables is indispensable because without it, the C-contextuality will not be an extension of traditional contextuality. How critical, however, is the second constraint imposed on well-fitting C, that the $\mathsf{C}\left[\mathcal{R}_q\right]$-coupling always exists and is unique? What if one considers statement C for which $\mathsf{C}\left[\mathcal{R}_q\right]$ is a set that may be empty or contain more than one coupling? This complicates the matters conceptually because then, in the consistification procedure, the $(q,\cdot)$-type bunches, those filled with the $\mathsf{C}\left[\mathcal{R}_q\right]$-couplings, cannot be formed unquely or cannot be formed at all. However, the main point of this paper can still be made, with some qualifications.

We can agree that the consistification of an inconsistently connected system R is not a single system $R^{\ddagger}$ but a cluster of systems $\left\{R^{\ddagger}_i : i \in \mathbb{I}\right\}$, the elements of which are obtained by filling the $(q,\cdot)$-type bunches in the consistification of R by all possible couplings of R's connections. We can further agree that the cluster $\left\{R^{\ddagger}_i : i \in \mathbb{I}\right\}$ is considered noncontextual if it contains a noncontextual system $R^{\ddagger}_i$. In particular, if $\left\{R^{\ddagger}_i : i \in \mathbb{I}\right\}$ is empty (which means that $\mathsf{C}\left[\mathcal{R}_q\right]$ does not exist for at least one of the connections of R), the latter definition is not satisfied, and the cluster should be considered contextual. Once again, we have a theory dealing with consistently connected systems only, except that the empirical or theoretical situations they depict are represented by clusters of systems sharing a format and the $(\cdot,c)$-bunches.

It might seem that dealing with an infinity of possible couplings $\mathsf{C}\left[\mathcal{R}_q\right]$ or proving that $\mathsf{C}\left[\mathcal{R}_q\right]$ is empty is a significantly more difficult mathematical task than when C is well-fitting. This is not the case, however, as the complication is not necessarily major. Mathematically,

the problem of finding whether a system R is contextual consists in determining whether the variable S having the same format as R can be assigned a probability measure subject to certain constraints on its marginals. The constraints are imposed by the distributions of the bunches R^c (that S^c have to match) and by the statement C that has to be satisfied by the couplings S_q of the connections $\mathcal{R}_q$. For discrete random variables and finite sets Q and C, this is a linear programming task, provided that the compliance with C can be presented in terms of linear inequalities of the probabilities in the distribution of S_q. For the consistification $R^\ddagger$ the problem is precisely the same, except that in place of connection couplings, one deals with $(q, \cdot)$-type bunches.

References

1.   Bell, J. On the Einstein-Podolsky-Rosen paradox. *Phys. Phys. Fiz.* **1964**, *1*, 195–200. [CrossRef]
2.   Bell, J.S. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.* **1966**, *38*, 447–453. [CrossRef]
3.   Kochen, S.; Specker, E.P. The problem of hidden variables in quantum mechanics. *J. Math. Mech.* **1967**, *17*, 59–87. [CrossRef]
4.   Suppes, P.; Zanotti, M. When are probabilistic explanations possible? *Synthese* **1981**, *48*, 191–199. [CrossRef]
5.   Cabello, A. Experimentally testable state-independent quantum contextuality. *Phys. Rev. Lett.* **2008**, *101*, 210401. [CrossRef] [PubMed]
6.   Abramsky, S.; Brandenburger, A. The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.* **2011**, *13*, 113036–113075. [CrossRef]
7.   Kurzynski, P.; Ramanathan, R.; Kaszlikowski, D. Entropic test of quantum contextuality. *Phys. Rev. Lett.* **2012**, *109*, 020404. [CrossRef]
8.   Budroni, C.; Cabello, A.; Gühne, O.; Kleinmann, M.; Larsson, J.Å. Kochen-Specker contextuality. *Rev. Mod. Phys.* **2022**, *94*, 045007. [CrossRef]
9.   Dzhafarov, E.N.; Cervantes, V.H.; Kujala, J.V. Contextuality in canonical systems of random variables. *Philos. Trans. R. Soc. A* **2017**, *375*, 20160389. [CrossRef]
10. Kujala, J.V.; Dzhafarov, E.N. Contextuality and dichotomizations of random variables. *Found. Phys.* **2021**, *52*, 13. [CrossRef]
11. Bacciagaluppi, G. Leggett-Garg inequalities, pilot waves and contextuality. *Int. J. Quant. Found.* **2017**, *1*, 1–17.
12. Zhan, X.; Kurzyński, P.; Kaszlikowski, D.; Wang, K.; Bian, Z.; Zhang, Y.; Xue, P. Experimental detection of information deficit in a photonic contextuality scenario. *Phys. Rev. Lett.* **2017**, *119*, 220403. [CrossRef] [PubMed]
13. Amaral, B.; Duarte, C.; Oliveira, R.I. Necessary conditions for extended noncontextuality in general sets of random variables. *J. Math. Phys.* **2018**, *59*, 072202. [CrossRef]
14. Flühmann, C.; Negnevitsky, V.; Marinelli, M.; Home, J.P. Sequential modular position and momentum measurements of a trapped ion mechanical oscillator. *Phys, Rev. X* **2018**, *8*, 021001. [CrossRef]
15. Malinowski, M.; Zhang, C.; Leupold, F.M.; Cabello, A.; Alonso, J.; Home, J.P. Probing the limits of correlations in an indivisible quantum system. *Phys. Rev. A* **2018**, *98*, 050102. [CrossRef]
16. Kupczynski, M. Contextuality-by-Default description of Bell tests: Contextuality as the rule and not as an exception. *Entropy* **2021**, *23*, 1104. [CrossRef]
17. Khrennikov, A. Contextuality, complementarity, signaling, and Bell tests. *Entropy* **2022**, *24*, 1380. [CrossRef]
18. Wang, P.; Zhang, J.; Luan, C.-Y.; Um, M.; Wang, Y.; Qiao, M.; Xie, T.; Zhang, J.-N.; Cabello, A.; Kim, K. Significant loophole-free test of Kochen-Specker contextuality using two species of atomic ions. *Sci. Adv.* **2022**, *8*, eabk1660. [CrossRef]
19. Workshop Quantum Contextuality in Quantum Mechanics and Beyond (QCQMB), December 2022, Prague, Czech Republic. Available online: http://www.psych.purdue.edu/~ehtibar/QCQMB2022 (accessed on 21 March 2023).
20. Abramsky, S.; Barbosa, R.S.; Searle, A. Combining Contextuality and Causality I. Paper Presented at QCQMB 2022. 17–18 December 2022. Available online: http://www.psych.purdue.edu/~ehtibar/QCQMB2022/abstracts/Abramsky.html (accessed on 21 March 2023).
21. Searle, A.; Abramsky, S.; Barbosa, R.S. Combining Contextuality and Causality II. Paper Presented at QCQMB 2022. 17–18 December 2022. Available online: http://www.psych.purdue.edu/~ehtibar/QCQMB2022/abstracts/Searle.html (accessed on 21 March 2023).

22. Mansfield, S. Addressing the Compatibility Loophole in the Absence of Space-like Separation. Paper Presented at QCQMB 2022. 17–18 December 2022. Available online: http://www.psych.purdue.edu/~ehtibar/QCQMB2022/abstracts/Mansfield.html (accessed on 21 March 2023).

23. Tezzin, A.; Jones, M.; Wolfe, E.; Amaral, B. Impossibility Theorem for Extending Contextuality to Disturbing Systems. Paper Presented at QCQMB 2022. 17–18 December 2022. Available online: http://www.psych.purdue.edu/~ehtibar/QCQMB2022/abstracts/Tezzin.html (accessed on 21 March 2023).

24. Dzhafarov, E.N.; Cervantes, V.H.; Kujala, J.V. Contextuality and informational redundancy. *Entropy* **2023**, *25*, 6. [CrossRef]

25. Thorisson, H. *Coupling, Stationarity, and Regeneration*; Springer: New York, NY, USA, 2000.

26. Dzhafarov, E.N. The Contextuality-by-Default view of the Sheaf-Theoretic approach to contextuality. In *Samson Abramsky on Logic and Structure in Computer Science and Beyond, in Series Outstanding Contributions to Logic*; Palmigiano, A., Sadrzadeh, M., Eds.; Springer Nature: Berlin/Heidelberg, Germany, 2023.

27. Dzhafarov, E.N. Replacing nothing with something special: Contextuality-by-Default and dummy measurements. In *Quantum Foundations, Probability and Information*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 39–44.

28. Pfanzagl, J. *Theory of Measurement*; Wiley: New York, NY, USA, 1968.

29. Narens, L. *Abstract Measurement Theory*; MIT Press: Cambridge, MA, USA, 1985.

30. Moore, D.W. Measuring new types of question-order effects. *Pub. Opin. Quart.* **2002**, *66*, 80–91. [CrossRef]

31. Wang, Z.; Busemeyer, J.R. A quantum question order model supported by empirical tests of an a priori and precise prediction. *Top. Cog. Sci.* **2013**, *5*, 689–710.

32. Abramsky, S.; Barbosa, R.S.; Mansfield, S. The contextual fraction as a measure of contextuality. *Phys. Rev. Lett.* **2017**, *119*, 050504. [CrossRef]

33. Clauser, J.F.; Horne, M.A.; Shimony, A.; Holt, R.A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **1969**, *23*, 880–884. [CrossRef]

34. Fine, A. Joint distributions, quantum correlations, and commuting observables. *J. Math. Phys.* **1982**, *23*, 1306–1310. [CrossRef]

35. Dzhafarov, E.N. On joint distributions, counterfactual values, and hidden variables in understanding contextuality. *Philos. Trans. R. Soc. A* **2019**, *377*, 20190144. [CrossRef]

36. Khrennikov, A. Bell-Boole inequality: Nonlocality or probabilistic incompatibility of random variables? *Entropy* **2008**, *10*, 19–32. [CrossRef]