Violation of contextual generalization of the Leggett–Garg inequality for recognition of ambiguous figures

Masanari Asano¹, Takahisa Hashimoto², Andrei Khrennikov³, Masanori Ohya² and Yoshiharu Tanaka²

¹ Liberal Arts Division, Tokuyama College of Technology, Gakuendai, Shunan, Yamaguchi 745-8585 Japan
² Department of Information Sciences, Tokyo University of Science, Yamasaki 2641, Noda-shi, Chiba, 278-8510 Japan
³ International Center for Mathematical Modeling in Physics, Engineering, Economics, and Cognitive Science Linnaeus University, Växjö-Kalmar, Sweden

E-mail: Andrei.Khrennikov@lnu.se

Received 12 February 2014, revised 6 May 2014
Accepted for publication 2 July 2014
Published 19 December 2014

Abstract

We interpret the Leggett–Garg (LG) inequality as a kind of contextual probabilistic inequality in which one combines data collected in experiments performed for three different contexts. In the original version of the inequality, these contexts have a temporal nature and they are represented by three pairs of instances of time, \((t_1, t_2), (t_2, t_3), (t_3, t_4)\), where \(t_1 < t_2 < t_3\). We generalize LG conditions of macroscopic realism and noninvasive measurability in a general contextual framework. Our formulation is performed in purely probabilistic terms: the existence of the context-independent joint probability distribution \(P\) and the possibility of reconstructing the experimentally found marginal (two-dimensional) probability distributions from \(P\). We derive an analog of the LG inequality, ‘contextual LG inequality’, and use it as a test of ‘quantum-likeness’ of statistical data collected in a series of experiments on the recognition of ambiguous figures. In our experimental study, the figure under recognition is the Schröder stair, which is shown with rotations for different angles. Contexts are encoded by dynamics of rotations: clockwise, anticlockwise and random. Our data demonstrated violation of the contextual LG inequality for some combinations of the aforementioned contexts. Since in quantum theory and experiments with quantum physical systems, this inequality is violated, e.g. in the form of the original LG-inequality, our result can be interpreted as a sign that the quantum-like models can provide a more adequate description of the data generated in the process of recognition of ambiguous figures.

Keywords: contextuality, Leggett–Garg inequality, quantum and cognitive measurements, Kolmogorov probability model, non-Kolmogorovness

1. Introduction

Mathematical modeling of the process of recognition of ambiguous figures is an intriguing problem that still has no completely satisfactory solution. Recently quantum-like models based on the mathematical formalism of quantum mechanics and its generalizations were applied to this problem [1–12]. As in any mathematical modeling project, the output of models has to be compared with results of experiments.

One of the basic intrinsically quantum probabilistic effects is the violation of the formula of total probability, which is experimentally exhibited in the interference effect [6–8]. In a series of papers [1–3] such an effect was found in experimental data collected in sequential recognition of a pair of ambiguous figures. Later it was also found [4] that these data violate Bell’s inequality [5].

Another quantum-like study on the recognition of ambiguous figures was done by Atmanspacher et al [10, 11] a
quantum-like model of bistable perception. (A generalized quantum formalism was in use [12].) One of the important novelties in [10, 11] was the application of temporal Bell inequalities, specifically the Leggett–Garg (LG) inequality [13].

Recently, a quantum-like model of the recognition of ambiguous figures was presented in the paper of Asano et al [14]. The model matches very well with experimental data on the determination of the structure of the Schröder stair, which was shown with rotations for different angles.

In this paper, we consider the possibility of using the inequalities of the LG type to check the ‘quantum-likeness’ of statistical data collected [14] in experimental studies on bistable perception. We interpret the LG inequality as a kind of contextual probabilistic inequality in which one combines data collected in experiments performed for three different contexts; cf [6, 8–10, 15]. In the original version of the inequality, these contexts have a temporal nature and are given by three pairs of instances of time, \((t_1, t_2), (t_2, t_3), (t_3, t_4)\), where \(t_1 < t_2 < t_3\). We generalize the LG conditions of macroscopic realism and noninvasive measurability in the general contextual framework. Our formulation is done in purely probabilistic terms: the existence of the context-independent, joint-probability distribution \(P\) and the possibility of reconstruct the experimentally found marginal (two-dimensional) probability distributions from \(P\).

We derive an analog of the LG inequality, ‘contextual LG inequality’, and use it as a test of ‘quantum-likeness’ of statistical data collected in a series of experiments on the recognition of ambiguous figures. In our experimental study, the figure under recognition is the Schröder stair [16], which is shown with rotations for different angles. Contexts are encoded by dynamics of rotations: clockwise, anticlockwise and random. Our data demonstrated the violation of the contextual LG inequality for some combinations of the aforementioned contexts. Since in quantum theory and experiments with quantum physical systems this inequality is violated, e.g. in the form of the original LG-inequality (see, e.g. [17] and references hereby), our result can be interpreted as a sign that the quantum-like models can provide a more adequate description of the data generated in the process of recognizing ambiguous figures.

In probabilistic terms, the context-dependence of (probabilistically determined) mental states implies that the conventional model of probability theory, the Kolmogorov measure-theoretic model [18], cannot be applied to describe statistics concerning the recognition of ambiguous figures. Thus our result on the violation of the contextual LG inequality restricts the domain of applications of the Kolmogorov model [4]. More general probabilistic models have to be tested, e.g. quantum probability and its generalizations [6, 12, 19]. Since neither physicists nor psychologists have a proper education in probability theory, at least in the axiomatic approach, we complete this paper with an extended appendix presenting both a brief introduction to the classical Kolmogorov model and a discussion on possible non-Kolmogorovian generalizations as well as Bell-type inequalities.

The contextuality of mental representations is one of the fundamental features of cognition. In particular, mental contextuality is one of the main motivations for applications of the quantum formalism to modeling of cognition [6, 14, 19–24] and more generally biological information processing [25, 26], since quantum mechanics is also fundamentally contextual. Typically, quantum contextuality is expressed in the form of the Kochen–Specker theorem. However, contextuality has recently been represented with the aid of Bell-type inequalities, e.g. [27]. These recent theoretical and experimental results on the Bell-type expression of contextuality match well with the contextual approach to the problem of violation Bell’s inequalities that was developed by one of the authors of this paper [9, 28–30], see also [31–39]. Originally, Bell mixed in one method nonlocality and realism. The standard conclusion from violation of the inequalities of the CHSH-type is that ‘local realism’ is incompatible with quantum behavior is not easy to interpret. What is the problem? Nonlocality? Realism? Both? The contextual viewpoint regarding the violation of Bell-type (that is, spatial) inequalities provides much help towards clarifying this problem. The contextual (spatial) Bell inequality is violated for a single particle, e.g. a neutron [27], therefore, the problem of (non)locality can be automatically excluded from consideration. The same happens in the case of temporal contextual Bell inequalities, including the LG inequality. We shall come back to this discussion and extend it to the problem of the inter-relation of mental contextuality and mental realism in sections 2.1 and 7. (We also remark that T Nieuwenhuizen invented the terminology Contextuality Loophole summarizing outputs of studies [9, 28–36].)

We point out that contextually dependent systems (in physics, biology, economics, finance and social science) and non-Kolmogorov probability theory can be described mathematically [8, 19] by the adaptive dynamics and the operation of lifting (the latter is widely used in quantum information theory [40]).

It is important to remark that we consider contextuality in the most general form, as N Bohr [41] did: the whole experimental arrangement has to be taken into account. J Bell [5] considered only a very special form of contextuality: dependence of the result of measurement of some observable \(A\) on joint measurement of another observable \(B\) compatible with \(A\). In cognitive science, the situation is even trickier. Measurements are typically self-observations that the brain performs on itself. Therefore, ‘the whole experimental arrangement’ includes not only the ‘external experimental arrangement’, e.g. prepared by researchers in cognitive psychology, but also ‘internal arrangement’ including the brain state.

We remark that this paper has nothing to do with the study of quantum physical processes in the brain. We proceed with the operational approach to quantum theory as a formalism describing in probabilistic terms measurements of, in general, incompatible observables. Such observables can be of any nature, physical, mental or biological [6, 19].

\footnote{In fact, the real situation is more complicated; see section A.4.}
2. Leggett–Garg inequality

2.1. Conditions of derivation

At the beginning of the discussion in the paper [13], Leggett and Garg (LG) postulated the following two assumptions:

- (A1) **Macroscopic realism**: a macroscopic system with two or more macroscopically distinct states available to it will at all times be in one or the other of these states.
- (A2) **Noninvasive measurability**: it is possible, in principle, to determine the state of the system without arbitrarily small perturbation on its subsequent dynamics.

Under these assumptions, the correlation functions must satisfy the LG inequality that will be presented in the next section. However, quantum mechanics violates the LG inequality as well as the same analogue of Bell’s inequality or the CHSH inequality. Therefore, this violation means that at least one of the two assumptions fails for quantum systems.

Although in the derivation of the LG inequality, section 3, both conditions play important roles, their foundational value is different. The main issue is realism, whether one can still proceed with (A1), macroscopic realism, in the quantum world. Therefore, the main part of the LG paper [13] is devoted to a discussion about possible physical experimental schemes that could lead to noninvasive measurements or at least measurements in which invasiveness is small compared with the degree of violation of the LG inequality. There are claims, e.g. the experiment in [17], that such negligibly invasive measurements were performed experimentally and the LG inequality was violated. This is often seen as an important argument in favor of the nonobjectivity of quantum observables.

However, the LG approach plays an important foundational role even if the possibility that measurements are non-negligibly invasive cannot be excluded. We know that classical systems and measurements on such systems satisfy conditions (A1) and (A2); e.g. an airplane’s trajectory. Therefore, by violating LG we at least know that a phenomenon under study cannot be described classically.

In cognitive science, it is not easy (if possible at all) to come with an experimental scheme that would lead to (at least approximately) noninvasive measurements. The brain is a kind of self-measurement device; by giving an answer to a question the brain definitely perturbs its mental state and non-negligibly. And introspective measurements have definitely the lowest degree of noninvasiveness. Therefore, it seems that violations of the LG-type inequalities for data collected, e.g. in cognitive psychology, cannot lead to the conclusion that mental realism is questionable. Here, realism is understood in the sense of the objectivity of mental observables, that their values can be assigned, say to the brain, a priori, i.e. before measurements. Nevertheless, such violations show that the data under consideration is nonclassical, i.e. it is not similar to data collected, e.g. from an ensemble of moving airplanes.

However, our main point is that in relation to the problem of cognition, the standard physical viewpoint on conditions for derivation of the LG inequality, namely, the mixture of macrorealism and noninvasiveness, does not match so well with the mental situation. As was emphasized, the very notion of (non)invasive measurement loses its clearness for self-measuring devices, and the brain is one such device. This paper advertises the contextual viewpoint on the mental phenomena developed in the series of works [6, 14, 19–24]. It seems that Bell-type inequalities, including the temporal ones, can be used to distinguish contextual and noncontextual realism and more generally (since mental processes are fundamentally random) contextual and noncontextual probabilistic representations. As we will see in the following presentation, the noncontextuality of representations of probabilistic data implies constraints on such data, in the form of various inequalities. By using contextual representation, a system (including the brain) can violate such constraints. We shall come back to this discussion in section 7.

3. Contextual viewpoint on the proof of LG inequality

To provide to a reader the possibility to compare the original LG inequality with our contextual generalization, see section 4. Moreover, to add contextual flavor to the LG approach, we present the original LG derivation by considering time as a contextual parameter.

Let \( Q \) be an observable quantity that takes either +1 or −1. In the original discussion by LG, \( Q \) is the observable position of a particle in the two potential wells. However, we can discuss another two-level system, e.g. spin\( -\frac{1}{2} \) system.

The measurement of the two-level system is performed on a single system at different times \( t_k \) \( (k = 1, 2, 3) \). By repeating a series of three measurements, we can estimate the value of correlation functions by

\[
C_{ij} = \frac{1}{N} \sum_{n=1}^{N} q_{i}^{(n)} q_{j}^{(n)},
\]

where \( q_{i}^{(n)} \) (or \( q_{j}^{(n)} \)) is a result of the \( n \)th measurement of \( Q_i \) (or \( Q_j \)). Note that the correlation between \( Q_i \) and \( Q_j \) takes the maximum value \( C_{ij} = 1 \) when \( q_{i}^{(n)} q_{j}^{(n)} \) equals to 1 for all the repeated trials. Here, consider the assumption (A1), when the state of the system is determined at all times, even when the measurement does not perform on the system. Therefore, the values of joint probabilities of \( Q_i \), \( Q_j \) and \( Q_k \) are determined a priori at initial time \( t_0 \). We denote it by

\[
q_{i}^{(n)} q_{j}^{(n)}
\]
the symbol \( P_{ij}(Q_i, Q_j, Q_k) \). Notice that the pairs of indexes \( i, j \) encode the situation that only two observables \( Q_i \) and \( Q_j \) are measured. In other words, the joint probability depends on the situations where pairs of observables are measured. (We can consider pairs of indexes, instances of time, as parameters encoding three temporal contexts, \( C_{12}, C_{23}, C_{31} \); cf section 4.) However, if one considers (A2), then the joint probabilities do not depend on temporal contexts \(^6\):

\[
P_{ij}(Q_i, Q_j, Q_k) = P(Q_i, Q_j, Q_k) \quad \forall i, j
\]

Then we have the following equalities:

\[
P(Q_1, Q_2) = \sum_{Q_i = \pm 1} P(Q_1, Q_2, Q_3),
\]

\[
P(Q_2, Q_3) = \sum_{Q_i = \pm 1} P(Q_1, Q_2, Q_3),
\]

\[
P(Q_1, Q_3) = \sum_{Q_i = \pm 1} P(Q_1, Q_2, Q_3)
\]

which are consequences of the additivity of classical (Kolmogorov) probability. Thus pairwise joint probability distributions are context independent (as a consequence of (A2)). We also have

\[
P(Q_i) = \sum_{Q_i = \pm 1} P(Q_1, Q_2) = \sum_{Q_i = \pm 1} P(Q_1, Q_i); \quad (1)
\]

\[
P(Q_2) = \sum_{Q_i = \pm 1} P(Q_2, Q_1) = \sum_{Q_i = \pm 1} P(Q_2, Q_i); \quad (2)
\]

\[
P(Q_3) = \sum_{Q_i = \pm 1} P(Q_3, Q_1) = \sum_{Q_i = \pm 1} P(Q_3, Q_i); \quad (3)
\]

Thus, for each observable, its probability distribution is also context-independent (as a consequence of (A2)). Violation of these equalities is interpreted as an exhibition of contextuality. In psychology and cognitive science the equalities (1)–(3) represent the special case of so-called marginal selectivity [37–39]. It is clear that if at least one of these equalities is violated then one cannot assume existence of context independent joint probability distribution.

Under the assumption of the existence of the joint (triple) probability distribution, the correlation functions are written with the joint probabilities \( P(Q, Q) \) as

\[
C_{ij} = P(Q_i = 1, Q_j = 1) + P(Q_i = -1, Q_j = -1)
\]

\[
- P(Q_i = -1, Q_j = 1) - P(Q_i = 1, Q_j = -1)
\]

\[
= 2 \left( P(Q_i = 1, Q_j = 1) + P(Q_i = -1, Q_j = -1) \right) - 1.
\]

\(^6\) We notice that conditions (A1) and (A2) were formulated in a physical framework. Therefore any study about the LG-inequality performed at the mathematical level of rigorousness has to present some mathematical formalization of these conditions. In our study, (A1) and (A2) imply that there exists the joint probability distribution that does not depend on experimental contexts. In particular, we identify (A2) with noncontextuality. We understand well that this is not the only possible probabilistic interpretation of (A1) and (A2).
The Schröder stair, which is leaning at a certain angle $\theta$, is an ambiguous figure that induces an optical illusion (see figure 1). We show the subjects the picture of Schröder's stair leaning at a certain angle $\theta$ (see figure 1). We prepare the 11 pictures that are leaning at different angles: $\theta = 0, 10, 20, 30, 40, 45, 50, 60, 70, 80, 90$. A subject must answer either (i) ‘L is front’ or (ii) ‘R is front’ for every picture. We arrange the computer experiment to change the pictures and to record their answers.

Before the experiment, we divided the subjects into three groups: (A) 55 persons, (B) 48 persons, (C) 48 persons. For the first group (A), the order of showing is randomly selected for each person. To assume statistically uniform randomness of this selection, we use a computer-implemented function (e.g. java.rand). For the second group (B), the angle $\theta$ is increased from a small value: 0, 10, ..., 90. Inversely, for the third group (C), the angle $\theta$ is decreased from a large value: 90, 80, ..., 0.

Thus we have the three kinds of experimental data: (A) the angle of Schröder stair changes randomly, (B) from 0 to 90 and (C) from 90 to 0. These experimental contexts are denoted by $C_A$, $C_B$ and $C_C$. Let $X_0$ be a random variable with a value of ±1. The event that a subject says ‘left side is front’ corresponds to the result $X_0 = +1$. Then, from the repeated trials for each experimental context, we have the experimentally obtained values of joint probabilities:

\[ P_{C_A}(X_0, X_{10}, \ldots, X_{90}), P_{C_B}(X_0, X_{10}, \ldots, X_{90}), P_{C_C}(X_0, X_{10}, \ldots, X_{90}). \]

The correlation functions are given by

\[ C_{12} = 2 \left\{ P_X(X_{\theta_1} = 1, X_{\theta_2} = 1) + P_X(X_{\theta_1} = -1, X_{\theta_2} = -1) \right\} - 1, \]
\[ C_{23} = 2 \left\{ P_Y(X_{\theta_2} = 1, X_{\theta_3} = 1) + P_Y(X_{\theta_2} = -1, X_{\theta_3} = -1) \right\} - 1, \]
\[ C_{13} = 2 \left\{ P_Z(X_{\theta_1} = 1, X_{\theta_3} = 1) + P_Z(X_{\theta_1} = -1, X_{\theta_3} = -1) \right\} - 1. \]

Here, the triple $(X, Y, Z)$ is given by a combination of the contexts $C_A$, $C_B$, and $C_C$. We show the values of $C_{12}$, $C_{23}$ and $C_{13}$ in table 1. We estimate the LHS of the inequality:

\[ K(\theta_1, \theta_2, \theta_3) = C_{12} + C_{23} - C_{13}. \]

Tables 2 and 3 show that the value of $K$ with respect to the
Table 1. Experimental correlations.

|      | \(C_{12}\) | \(C_{23}\) | \(C_{13}\) | \(C_{12}\) | \(C_{23}\) | \(C_{13}\) |
|------|------------|------------|------------|------------|------------|------------|
| \(C_A\) | 1.000      | 0.964      | 0.964      | 0.091      | 0.091      | 0.127      |
| \(C_B\) | 0.917      | 0.833      | 0.750      | 0.375      | 0.625      | 0.083      |
| \(C_C\) | 0.917      | 1.000      | 1.000      | 0.625      | 0.375      | 0.167      |

(Left) \((\theta_1, \theta_2, \theta_3) = (0, 10, 20)\) (Right) \((\theta_1, \theta_2, \theta_3) = (40, 45, 50)\)

Table 2. The triple of angles \((0, 10, 20)\). The values of \(K\) for various combinations of contexts. For the contexts \((C_A, C_C, C_B)\), \(K\) approaches its maximal value.

| \(X, Y, Z\) | \(C_A, C_A\) | \(C_A, C_B\) | \(C_A, C_C\) | \(C_B, C_A\) | \(C_B, C_B\) | \(C_B, C_C\) | \(C_C, C_A\) | \(C_C, C_B\) | \(C_C, C_C\) |
|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| \(C_A\)     | 1.000      | 1.214      | 1.047      | 0.870      | 1.083      | 0.917      | 1.036      | 1.250      | 1.083      |
| \(C_B\)     | 0.917      | 1.130      | 0.964      | 0.786      | 1.000      | 0.833      | 0.953      | 1.167      | 1.000      |
| \(C_C\)     | 0.917      | 1.130      | 0.964      | 0.786      | 1.000      | 0.833      | 0.953      | 1.167      | 1.000      |

Table 3. The triple of angles \((40, 45, 50)\). The values of \(K\) for various combinations of contexts. For the contexts \((C_C, C_B, C_A)\), \(K\) approaches its maximal value.

| \(X, Y, Z\) | \(C_A, C_A\) | \(C_A, C_B\) | \(C_A, C_C\) | \(C_B, C_A\) | \(C_B, C_B\) | \(C_B, C_C\) | \(C_C, C_A\) | \(C_C, C_B\) | \(C_C, C_C\) |
|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| \(C_A\)     | 0.055      | 0.099      | 0.015      | 0.589      | 0.633      | 0.549      | 0.339      | 0.383      | 0.299      |
| \(C_B\)     | 0.339      | 0.383      | 0.299      | 0.873      | 0.917      | 0.833      | 0.623      | 0.667      | 0.583      |
| \(C_C\)     | 0.589      | 0.633      | 0.549      | 1.123      | 1.167      | 1.083      | 0.873      | 0.917      | 0.833      |

\((\theta_1, \theta_2, \theta_3) = (0, 10, 20)\) and \((40, 45, 50)\). The value of \(K\) exceeding 1 is seen in several cases.

6. Statistical analysis

We start from the random variable:

\[ K = Q_1Q_2 + Q_2Q_3 - Q_1Q_3 \]

Here, \(K\) takes \(-4, -2, 0\) or \(+3\) since \(Q_i\) takes \(+1\) or \(-1\). The probability distribution of \(K\) is not known, but it has mean value \(\mu\) and variance \(\sigma^2\) and their statistical estimates can be found. To find the confidence interval in such a situation, we apply the simplest method of nonparametric statistics, namely, the method based on the Chebyshev inequality. (However, from the very beginning, we recalled that this method gives us only a rough estimate for the confidence interval.) This method was recently used [43] for the analysis of statistical data from the Vienna-test for the Bell-type inequality and the Eberhard inequality, which finally closed the fair sampling loophole. In this test, because of the presence of slight drift depending on the experimental setting, one cannot assume Gaussianity of data. It seems that using the Chebyshev inequality is the simplest way to resolve this problem.

We can apply the Chebyshev inequality to the sample mean of \(K\)

\[ P(|m - \mu| > c) \leq \frac{\sigma^2}{nc^2} \]

with positive constant \(c\). Here, \(m\) is a sample mean of independent random variables \(K_1, ..., K_n\):

\[ m = \frac{K_1 + K_2 + ... + K_n}{n} \]

Although we do not know the value of \(\sigma^2\), we can estimate \(\sigma^2\) with unbiased sample variance. Then \(\mu\) is estimated by \(m\) with a confidence interval \([m - c, m + c]\).

We take the 80% confidence level. In the case that the order of the contexts is \((X, Y, Z) = (C_A, C_C, C_B)\), and the angle \(\theta = 0, 10, 20\) (This case has a maximum value of \(K\)), we estimate the value of \(K\) as follows:

\[ K = 1.250 \pm 0.213. \]

Statistical analysis shows that the violation of the LG-inequality is statistically significant. However, it is clear that one has to perform better experiments to get a higher level of violation. The main problem of the present experiment is that the sample is not large enough. It seems that this will be the problem of all Bell-type tests in cognitive science. All such inequalities are based on correlations, so samples of students have to be sufficiently large in order calculate correlations (this needs human recourses) and get statistically significant violation (additional human recourse) of the corresponding inequality for such correlations.

7. Concluding remarks

A violation of the contextual LG inequality by statistical data collected for observations of the Schröder stair rotated for different angles supports the contextual cognition paradigm
presented in the series of works [6, 14, 19–24]. Our experimental statistical data is fundamentally contextual⁹. The brain does not have a priori prepared ‘answers’ to the question about the R/L structure of the Schröder stair for the fixed angle θ. Answers are generated depending on the mental context. Thus, mental realism is a kind of contextual realism; cf [30]. There are practically no (at least not so many), so to say, ‘absolute mental quantities’, ‘answers’ to the same question vary essentially depending on context. This conclusion is not surprising in the framework of cognitive science and psychology, where various framing effects are well known. Thus, the main contribution of this paper is the demonstration of the applicability of a statistical test of contextuality borrowed from quantum physics. We also can consider this study as a step towards the creation of a unified mathematical picture of the world: physical and mental phenomena can be described by the same equations; cf [6].

Such studies on the usage of standard quantum mechanical tests in other domains of science also contribute to the foundations of quantum physics, since they can be used as at least indirect arguments supporting some interpretations of outputs of these tests for physical systems. In our case, the viewpoint that contextuality (and not nonlocality) is the basic source of the nonclassical probabilistic behavior of the cognitive systems (see also [44]), can be used to support local contextual models of quantum physical phenomena.

Acknowledgments

We would like to thank H Atmanspacher, I Basieva, N Watanabe, and I Yamato for discussions. This work was partially supported by MPNS COST Action MP1006 (Fundamental Problems in Quantum Physics) and visiting fellowships (AK) to the Center of Quantum BioInformatics, Tokyo University of Science, and the Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences.

⁹ In fact, it violates even the conditions of the marginal selectivity. We interpret violation of marginal selectivity as one of the signs of contextuality. Of course, it is important to approach violation of the LG inequality in combination with marginal selectivity. And this is a delicate issue. Up to authors’ knowledge, even in physics the experimental situation is not completely clear. Practically all publications on the violation of the Bell-type inequalities represent only the values of correlations (or even only the values of outputs of these tests for physical systems. However, the degree of such dependence can be considered as ‘sufficiently small comparing with the degree of violation of the corresponding inequality.’ Of course, the latter statement has to be presented in terms of statistical analysis.

Appendix. Kolmogorov probability model and Bell-type inequalities

A.1. Kolmogorov axiomatics, probability space

We start this section with a rather long introduction to the classical measure-theoretic model of probability theory [18]. Our aim is to show that the problem of constructing a common probability measure was one of the basic problems even in classical probability theory, namely, the invention of stochastic processes was based on such a special construction (Kolmogorov’s theorem [18]). Finding a positive solution to this problem in the case of stochastic processes played an important role in the formation of the present ideology of classical probability: observables have to be represented by random variables on common probability space. However, in the case of quantum observables, this is not true.

Classical probability theory is based on the model of A. N. Kolmogorov [18]. Its basic notion is probability space, a triple $\mathcal{P} = (\Omega, \mathcal{F}, P)$, where $\Omega$ is a set, $P$ is a probability measure, and $\mathcal{F}$ is a collection of subsets of $\Omega$ on which probability is defined¹⁰. In the Kolmogorovean model [18] an observable, say $a$, is represented by a random variable¹¹: a map $\alpha: \Omega \to \mathbb{R}$ such that, for each interval, its pre-image, $\{\omega \in \Omega : a(\omega) \in [\alpha, \beta]\}$, belongs to $\mathcal{F}$. Its probability distribution is defined as $p_a(\omega) = P(\omega : a(\omega) \in A)$. A system of observables is represented by a vector of random variables $a = (a_1, ..., a_n)$. Its probability distribution is defined as

$$
p_a(A_1 \times \cdots \times A_n) = P(\omega : a_1(\omega) \in A_1, \cdots, a_n(\omega) \in A_n). \tag{A.1}
$$

The notion of a random vector is generalized to the notion of a stochastic process. Suppose that the set of indexes is infinite; for example, $a_i, i \in [0, +\infty)$. Suppose that, for each finite set $(t_1, ..., t_k)$, the vector $(a_{t_1}, ..., a_{t_k})$ can be observed and its probability distribution $p_{t_1, ..., t_k}$ is given. By selecting $\Omega_{t_1, ..., t_k} = \mathbb{R}^k$, $p_{t_1, ..., t_k} = p_{t_1, ..., t_k}$, and $\mathcal{F}$ as the Borel $\sigma$-algebra, we obtain the probability space $\mathcal{P}_{t_1, ..., t_k}$ describing measurements at points $t_1, ..., t_k$. At the beginning of the 20th century the main mathematical question of probability theory was whether it is possible to find a single probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$ such that all $a_i$ be represented as random variables on this space and all probability distributions $p_{t_1, ..., t_k}$ are induced by

¹⁰ This is a $\sigma$-algebra (‘$\sigma$-field’): a system of sets that is closed with respect to countable unions, intersections and the operation of a complement.

¹¹ We note that the Kolmogorovean model is not simply a mathematical theory. In the same way as the Euclidean model provided for the mathematical formalization of the geometry of physical space, the Kolmogorovean model provided mathematical formalization of the theory of measurements of random variables. Euclid formalized such heuristic notions as point, straight line, plane, angle, etc.; Kolmogorov formalized such notions as event, probability, random observable (= random variable). In section A.3, we shall come back to a comparison of the roles played by the Euclidean model of geometry of space and the Kolmogorovean model of random measurement in physics.
the same $P$:
\[ p_{\lambda_{n+1}}(A_1 \times \cdots \times A_k) = P(\omega \in \Omega; \alpha_n(\omega) \in A_1, \alpha_{n+1}(\omega) \in A_k). \]

Kolmogorov found natural conditions for the system of measures $p_{\lambda_{n+1}}$ which guarantee the existence of such a probability space; see [18]\(^{12}\). And during the next 80 years, analysis of properties of (finite and infinite) families of random variables defined on one fixed probability space was the main activity in probability theory.

A.2. Kolmogorovian formalization of Bell’s argument

Although J. Bell did not formulate his argument in the terms of a probability space, the problem of local realism for quantum observables was formalized in complete accordance with classical probability theory. In the Bell framework, $\Omega$ is selected as the set of hidden variables $\lambda$; local realism is equivalent to mathematical presentation of observables by random variables $A \rightarrow a_\lambda(\lambda)$. Here, $a_\lambda$, $A = a_\lambda, a_\lambda$ are observables depending on the parameter $\alpha$, experimental settings, at 'Alice’s lab' and $b_\beta$, $\beta = b_1, b_2$, are observables depending on the parameter $\beta$, experimental settings, at 'Bob’s lab'. In the framework of the Kolmogorov probability model, the Bell inequality is a theorem. A. Fine rigorously proved [47] that the Bell inequality is satisfied if the common probability space for random variables representing quantum observables does exist.

In the LG inequality (which is a special form of contextual Bell inequalities) there is a single observable $a_\lambda$ depending on the parameter $\alpha$. Kolmogorovness means that all these observables can be represented by random variables, $\lambda \rightarrow a_\lambda(\lambda)$, on common Kolmogorov probability space.

A.3. Models: (non-)Euclidean geometry and (non-)Kolmogorovian probability

Although the main stream in classical probability was, so to speak, the ‘common probability space stream’, we can point to a few attempts to swim against this stream, e.g. [9, 15, 20, 28–39].

Now we come back to a comparison of mathematical formalizations of the geometry of physical space and random observations. Since the work of Lobachevsky (and Gauss and Boyai), mathematicians have understood that there exist different mathematical possibilities for the representation of space geometry. We remark that already Lobachevsky and Gauss studied the problem of adequacy of the Euclidean model to physical reality. Lobachevsky proposed some astronomical tests, Euclidean contra Lobachevsky geometries; in Germany, Gauss (who had some administrative obligations regarding the measurement of land) measured the angles of a huge triangle formed by three mountains. The latter test confirmed that at least locally we live in Euclidean space. Later Riemann formulated the general principles of geometry that played the fundamental role in the mathematical representation of Einstein’s general relativity. (Lobachesky geometry was used in special relativity).

We emphasize that mathematicians understood long before physicists (who were at that time completely busy with Newtonian physics based on Euclidean geometry) that Euclidean geometry is one of the possible models of space. On the basis of such an experience collected in the mathematical community, physicists were not astonished by the appearance of Minkowski space in special relativity and then (pseudo-)Riemann space in general relativity. Here, the general approach of D. Hilbert on the axiomatization of physics (also known as Hilbert’s sixth problem) was respected.

Development of the mathematical formalization of probability was very different. Mathematicians (with a few exceptions) did not question the Kolmogorov axiomatics. The first non-Kolmogorov model was elaborated in physics as a part of a new physical theory—quantum mechanics. And in the classical probability community, quantum probability is still not recognized as a probability theory, but as some exercises in noncommutative algebra. Therefore, in probability it is more difficult than it was in the case of geometry: any mathematical model, including the Kolmogorovian model, has a restricted domain of application. Quantum phenomena simply showed that one special model of probability cannot be applied. From the viewpoint of Hilbert, the axiomatization of physics is the end of the story, i.e. one need not search for additional ‘explanations’ of non-Kolmogorovness; one simply must find a new appropriate mathematical model of probability and proceed with that model. Thus, from this viewpoint, the Bell argument is not about the locality and realism, but about inadequacy of the Kolmogorov model\(^{13}\). (The mathematical foundation of non-Kolmogorov probability theory was discussed in the books [8, 9, 40].)

Let us again make a comparison with geometry. Did Einstein try to ‘explain’ the appearance of (pseudo-)Riemannian geometries in general relativity, instead of one special geometry (Euclidean)? Not at all. He simply identified (pseudo-)Riemannian geometry with physical space.

Thus, experimental tests of Bell’s inequality can be considered as tests of adequacy of the traditional Kolmogorov model for quantum physical phenomena (cf with the aforementioned Gauss test of the Euclidean model).

Typically, adherents of the non-Kolmogorovness viewpoint on the violation of the Bell inequality consider physics (in the spirit of the Hilbert program of axiomatization of physics) as a collection of mathematical models formalizing

\(^{12}\) We just remark that the $\Omega$ is selected as the set of all trajectories $t \rightarrow \omega(t)$. The random variable $a_\lambda$ is defined as $a_\lambda(\omega) = \omega(t)$. Construction of the probability measure $P$ serving for all finite random vectors is mathematically advanced and going back to the construction of the Wiener measure on the space of continuous functions.

\(^{13}\) We point out that one of the problems slowing clarification of Bell’s argument is that here, probability (both classical and quantum) is typically not treated in an axiomatic framework. There is a prejudice that ‘probability is probability’ and it can be understood heuristically without going into mathematical axiomatization. However, the positive experience of the physical applications of axiomatic mathematical models of geometry tells us that this is the most fruitful way even for probabilistic applications. (Nowadays in physics nobody would work in the framework of ‘heuristically understandable geometry’.)
various natural phenomena. From the very beginning, it is assumed that any such model has a restricted domain of applications. A physical experiment of which the output cannot be described by a model under this application is considered as a signal for the creation of a new mathematical model; e.g. the Euclidean model could not be used for special relativity, and new non-Euclidean models were explored (the Lobachevsky, in special relativity and pseudo-Riemannian geometries in general relativity.)

**A.4. Reconstructing Kolmogorovness from contextuality**

The main message of this paper is that experimental statistical data collected in cognitive science and psychology are fundamentally contextual. In general, for each experimental context, data is described by its own Kolmogorov probability space.

> **Is it possible to unify these spaces in some way?**

One possibility is to use quantum probability and the formalism of complex Hilbert space. This is, so to speak, nonclassical unification. Surprisingly it is even possible to unify these probability spaces classically, i.e. to embed them into a single ‘big Kolmogorov space’. Such classical unification is based on taking into account randomness of realizations of contexts. In this ‘big probability space’ the original probabilities appear as conditional probabilities with respect to various contexts.

The first version of such unification was presented in the paper [48] (see [49] for a better-structured presentation), where the probabilistic data collected in experiments to determine the EPR–Bohm correlations were embedded into single Kolmogorov probability space. (This construction is evidently generalized to statistical data collected for any family of experimental contexts.)

The reader may ask: why do you emphasize non-Kolmogorovness in such a situation? Although a ‘big Kolmogorov space’ unifying data collected for different contexts exists, it cannot be used as successfully as quantum probability. The latter describes all possible experimental contexts homogeneously. In the classical approach based on the reconstruction of Kolmogorovness from contextuality, one has to take into account the randomness of realizations of concrete contexts. Of course, one can construct a huge Kolmogorov space unifying all possible contexts and all possible types of randomness for them. However, it would be practically impossible to work with such space.

Another problem of the Kolmogorovian unification of contextual experimental data is that by taking into account the randomness of realizations of contexts (e.g. for the Bell-type experiments, how often each pair of orientations of polarization beam splitters is realized in the concrete experiment), we lost the internal description of data: an experimenter’s ‘free will’ (to use this or that experimental context for the next trial) also has to be taken into account. This is a complex interpretational problem related to the construction of ‘big Kolmogorov spaces’; see [49] for a discussion. Quantum mechanics provides a description of experimental probabilities without taking into account the randomness of realizations of experimental contexts; in this sense, the quantum description can be treated as a kind of intrinsic description. And data can be ‘intrinsically’ Kolmogorovian or non-Kolmogorovian (although, as we emphasized in this section, it is always possible to make these data Kolmogorovian ‘externally’).

### References

[1] Conte E, Todarello O, Federici A, Vitelli F, Lopane M and Khrennikov A 2004 A preliminary evidence of quantum-like behaviour in measurements of mental states Proc. Int. Conf. on Quantum Theory: Reconsideration of Foundations (Ser. Math. Modelling in Phys. Cogn. Sci. vol 10) (Växjö: Växjö University Press) pp 679–702

[2] Conte E, Todarello O, Federici A, Vitelli F, Lopane M, Khrennikov A and Zbilut J P 2007 Some remarks on an experiment suggesting quantum-like behavior of cognitive entities and formulation of an abstract quantum mechanical formalism to describe cognitive entity and its dynamics Chaos Solitons Fractals 31 1076–88

[3] Conte E, Khrennikov A, Todarello O, Federici A, Mendolicchio L and Zbilut J P 2009 Mental state follow quantum mechanics during perception and cognition of ambiguous figures Open Syst. Info. Dyn. 16 1–17

[4] Conte E et al 2008 A preliminary experimental verification on the possibility of Bell inequality violation in mental states Neuroquantontology 6 214–21

[5] Bell J 1987 Speakable and Unspeakable in Quantum Mechanics (Cambridge: Cambridge University Press)

[6] Khrennikov A 2010 Ubiquitous Quantum Structure: From Psychology to Finances (Berlin: Springer)

[7] Busch J R and Bruza P D 2012 Quantum Models of Cognition and Decision (Cambridge: Cambridge University Press)

[8] Asano M, Khrennikov A, Ohya M, Tanaka Y and Yamato I 2014 Quantum Adaptivity in Biology: From Genetics to Psychology to Finances (Berlin: Springer)

[9] Khrennikov A 2004 A preliminary experimental verification on the possibility of Bell inequality violation in mental states Neuroquantontology 6 214–21

[10] Atmanspacher H, Flik Th and Römer H 2004 Quantum zeno behaviour in measurements of bistable perception Biol. Cybern. 90 33–40

[11] Atmanspacher H and Flik Th 2012 Temporal nonlocality in bistable perception Quantum Theory: Reconstructions of Foundations–6 (Quantum-Like Decision Making: From
