

Philosophical Problems in Science

**Zagadnienia Filozoficzne
w Nauce**

© Copernicus Center Press, 2022

Except as otherwise noted, the material in this issue is licenced under the Creative Commons BY-NC-ND licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0>.

Editorial Board

dr hab. Paweł Jan Polak (Editor-in-Chief)
dr hab. Janusz Mączka
dr Roman Krzanowski
prof. dr hab. Michał Heller (Honorary Editor)
Piotr Urbańczyk (Editorial Secretary)

Proofreading: dr Roman Krzanowski

Adjustment and correction: Artur Figarski

Cover design: Mariusz Banachowicz

Technical editor: Artur Figarski

Typeset and typographic design: Piotr Urbańczyk

ISSN 0867-8286 (print format)

e-ISSN 2451-0602 (electronic format)

Editorial Office

Zagadnienia Filozoficzne w Nauce

Wydział Filozoficzny UPJPII

ul. Kanonicza 9, 31-002 Kraków

POLAND

e-mail: info@zfn.edu.pl

www.zfn.edu.pl



**Copernicus
Center
PRESS**

Publisher: Copernicus Center Press Sp. z o.o.

pl. Szczepański 8, 31-011 Kraków POLAND

tel. (+48) 12 448 14 12

e-mail: marketing@ccpress.pl

www.ccpress.pl

Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

LXXII ■ 2022

Articles

Artykuły

Jerzy Dajka

Past of a quantum particle: old problem with recent controversies 7

Walter Block

Response to Wysocki on indifference 37

Jacek Rodzeń

*Stanisław Dunin-Borkowski and his views on Einstein's special theory of
relativity 63*

Wojciech P. Grygiel

A critical analysis of the philosophical motivations and development of the concept of the field of rationality as a representation of the fundamental ontology of the physical reality 87

Michał Heller

Józefa Życińskiego zmagania z językiem o Bogu 109

Marcin Miłkowski, Witold Wachowski

Hutchins w obronie interdyscyplinarnych badań nad poznaniem . . 127

Review articles

Artykuły recenzyjne

Marek Kuś

Quantum mechanics of identical particles 169

Roman Krzanowski

Science and its social grounding 179

Ewelina Grądzka

A companion to Kazimierz Twardowski 191

Michał Heller, Janusz Mączka

Idee i ideaty 203

Articles

Artykuły

Past of a quantum particle: old problem with recent controversies

Jerzy Dajka

University of Silesia in Katowice

Abstract

Time-symmetric formulation of quantum mechanics—the two-state vector formalism—is presented as a tool for studying past behaviour of quantum systems. A role of weak measurement and weak values in the Cheshire Cat effect and a nested (Vaidman) three-path interferometer are discussed. Interpretation of a particle's faint trace indicating possibility of discontinuous paths of particles passing the Vaidman interferometer is given. Consistent histories are presented as one of alternative approaches. Multitude of controversial issues is briefly reviewed and discussed.

Keywords

two-state vector formalism, weak values, weak measurement, Cheshire Cat effect.

1. Introduction

Copenhagen orthodoxy limits the predictive role of quantum mechanics to measurement results only and introduces unavoidable indeterminacy at a quantum level. One can prepare (preselect) a quantum system, evolve it (via e.g. unitary time evolution) and then one can measure it and postselect a particular, desired result. The mea-

surement output is known to belong to a certain and well defined set of possible values which occur, as an output, with certain probability. Such an approach is already more than century old and we are used to an uncomfortable fact that at a quantum level not only measurement results but also quantum measurement *per se* require an interpretation.

In a standard von Neumann's scheme a measurement output of an observable is one of eigenvalues a_k of a hermitian operator A representing the observable i.e. $A|a_k\rangle = a_k|a_k\rangle$ with the corresponding eigenvectors $|a_k\rangle$ forming a basis spanning a state space of the system. Let us forewarn that for a simple convenience we use notation suitable for qudits (finite d -dimensional quantum systems) rather than general quantum systems. Such a convention allows one to bound an applied formalism almost solely to linear algebra at a low cost of a little mathematical 'flexibility' required for studying infinite systems. Having a system in a state $|\psi\rangle = \sum_k c_k|a_k\rangle$ a probability of measuring a_k is given by a scalar product of the state and the particular eigenstate $|a_k\rangle$ and reads $prob(k) = |\langle a_k|\psi\rangle|^2$. We humbly acknowledge that one cannot say *before* the measurement is done which among a_k 's is going to be observed and we are left with an expectation only: $\langle A \rangle = \langle \psi|A|\psi\rangle = \sum_k p_k a_k$. Moreover, we know that such a measurement procedure causes a serious disturbance of a system as it results in a jump or collapse of the state $|\psi\rangle \rightarrow |a_k\rangle$ which can be formalised in terms of projection operators $|a_k\rangle\langle a_k|$.

In more formal terms, with an applied notation summarized in Tab.(1), one prepares a composite state of a system and a probe (a pointer) $|\Psi(t_i)\rangle = |\psi(t_i)\rangle|\phi(t_i)\rangle$. To measure a system, the probe must, at least for a certain duration τ , interact with the measured object. Measuring an observable A one assumes that the system–probe interaction is given by a Hamiltonian coupling A and a 'momentum' (an operator canonically conjugated to the probe position)

states of	applied notation
system+probe	$ \Psi\rangle$ (capital)
system preselected	$ \psi\rangle, a\rangle$
system postselected	$ b\rangle$
probe only	$ \phi\rangle$

Table 1: Notation of states applied in the paper with time parameter and subscripts skipped.

$H_{int} = g(t)AP$ with a coupling strength $g = \int_{t_i}^{t_i+\tau} g(t)dt$. Hence, the von Neumann ideal measurement is nothing but a dynamic process

$$\begin{aligned}
 |\Psi(t_i + \tau)\rangle &= e^{-igAP} |\psi(t_i)\rangle |\phi_{x_0}\rangle \\
 &= \sum_k |a_k\rangle \langle a_k | \psi(t_i)\rangle e^{-iga_k P} |\phi_{x_0}\rangle \\
 &= \sum_k |a_k\rangle \langle a_k | \psi(t_i)\rangle |\phi_{x_0 - ga_k}\rangle
 \end{aligned}
 \tag{1}$$

causing shift of a pointer position from x_0 (in the state $|\psi_{x_0}\rangle$ to $x_0 - ga_k$ for $|\phi_{x_0 - ga_k}\rangle$). The last step, i.e. a projective measurement of the pointer, allows one to relate a particular read-out value $x_0 - ga_k$ to a particular value a_k of the observable A .

It is not surprising that there are circumstances where limitations of Copenhagen interpretation (and the celebrated eigenvalue–eigenstate link) become quite bothersome. We present here only one of many problems: a measurement result allows one to gain knowledge about the present of quantum systems (a particle). It is, however, hardly possible to infer upon what was the particle’s *past*, what ‘historic’ properties one can attribute to the just measured systems. Sometimes it is even harder to say anything about quantum systems which are essentially closed (as, for example, the whole universe is) and do

not allow to include an external observer keeping a pointer or these which are simply resistant to measuring attempts. Neutrino, hardly interacting with anything, can serve as a natural example of such a ‘measurement–resistive’ system. Counterfactual reasoning (what would happen if a particular measurement were performed or, more generally, if there were that a , then it would be that b) is related to an another branch of problems where the eigenvalue–eigenstate link induces unwanted constraints.

Our aim is to present a seemingly simple quantum system—a particle in a nested three-path interferometer invited by Lev Vaidman to a modern scientific debate. A particle enters the interferometer, it has three paths at its disposal and then it is measured (postselected) leaving the interferometer by one of three exits. The question which is asked looks innocent and apparently simple: what path took the particle inside the interferometer? What was its trajectory? The questions are natural (particles entering and leaving the interferometer must somehow pass it) and interesting (cf. reference list below) although from the Copenhagen perspective slightly out of place. In this paper we focus on the two-state vector formalism recently applied to answer the above question and leading to hardly acceptable and very disputable possibility: the particle in the Vaidman interferometer follows a *discontinuous* path. Surprisingly, such a weird answer, seemingly contradicting not only our intuition but also common sense, is supported by very recent experiments and as such requires careful analysis not only from a physical but also a more fundamental, philosophical perspective. Although the two-state vector formalism is nothing but a time-symmetric version of a very standard quantum mechanics, it exhibits its true predictive power combined with a notion of *weak measurement* and *weak value* of an observable. Both concepts, exemplified with help of the Cheshire Cat effect, are lacking from the

eigenvalue-eigenstate link and require an interpretation. The paper is organized as follows: in Sec. 2 the two-state vector formalism is presented, a notion of weak measurement and weak values are applied to present the Cheshire Cat effect of a spatial separation of a photon position and polarization. Sec 3. is devoted to the Vaidman's construction of a nested interferometer and a faint trace as a tool verifying the presence of a weakly measured particle inside the interferometer. As the two-state formalism is not the only formalism applied to study particle's past in the Vaidman interferometer, in Sec 4. we very briefly present one of alternatives: an approach based on consistent histories and we indicate probably the most obvious and easiest to formulate objection against this approach. Concluding in Sec 5. we list selected problems which—at least in our opinion—urgently require a professional philosophical analysis.

2. Two–state vector formalism and the Cheshire Cat

Classical dynamics of a particle is a solution of the Newton equation which, to uniquely predict a final state, needs initial conditions (an initial state) to be specified. For quantum system to predict a measurement output at a given time it is not enough to preselect an initial state and then to solve the Schrödinger equation as a resulting final state may be a superposition. To get unique prediction of a measurement output one needs an additional information what state was postselected at the end of evolution. In other words, in classical physics all the results of future measurements are constrained by the results of the past measurements leading to an initial preparation. For quantum systems, however, a measurement output is only partially determined by past measurements and we recognize that

quantum mechanics, contrary to the classical one, is essentially time asymmetric. This asymmetry limits an ability of reproducing the past solely upon the present measurement results. The two-state vector formalism (Aharonov and Vaidman, 2008) (TSVF) is an attempt to make quantum mechanics time-symmetric by adding an information concerning a postselected state. According to the TSVF, to describe a quantum system at a time t one needs a two-state vector

$$(2) \quad \langle \omega(t) | \alpha(t) \rangle$$

where $|\alpha(t)\rangle$ is a ‘usual’ state of a time-evolving quantum system which is a fundamental object of a standard quantum mechanics and $\langle \omega(t) |$ is a state evolving *backward in time* and determined by the results of measurements performed on the system *after* the time t . In other words, the two-state vector is pair of states consisting of a forward evolving preselected state and a backward evolving post-selected state. Let us imagine that at $t = t_i$ our system is in one of eigenstates of an observable A i.e. $|\alpha(t_i)\rangle = |A = a\rangle$ and finally, at $t = t_f$, $\omega(t_f)\rangle = |B = b\rangle$ for a different observable B . According to the TSVF a complete information about the system at $t \in (t_i, t_f)$ is encoded in $\langle \omega(t) | \alpha(t) \rangle$ where

$$(3) \quad \begin{aligned} |\alpha(t)\rangle &= U(t, t_i) |A = a\rangle \\ \langle \omega(t) | &= \langle B = b | U(t, t_f) \end{aligned}$$

and $U(\cdot, \cdot)$ is a unitary time evolution operator. Let us strongly emphasise that essentially the TSVF is a standard quantum mechanics just equipped with an additional information carried by a backward evolving state. As such, the TSVF is not expected to contradict any ‘standard’ predictions of the quantum mechanics. Minimalists may consider TSVF just as a useful tool to analyze otherwise strange and ‘paradoxical’ phenomena.

Let us describe the following scenario to present a rationale behind the TSVF construction and to show a particular usefulness of this formalism in a weak measurement context. The Aharonov-Bergmann-Lebowitz rule (Aharonov, Bergmann and Lebowitz, 1964) is a working horse of many of subsequent techniques. If there is a quantum system preselected in a state $|\psi(t_i)\rangle$ and finally found in a postselected state $|b_f\rangle$ conditioned probability that an observable $A = \sum_n a_n |a_n\rangle\langle a_n|$ measured at some intermediate stage gives a_n reads as follows:

$$(4) \quad \text{prob}(a_n) = \frac{|\langle b_f | a_n \rangle \langle a_n | \psi(t_i) \rangle|}{\sum_s |\langle b_f | a_s \rangle \langle a_s | \psi(t_i) \rangle|^2}.$$

We consider a narrow time window $t_w \pm \tau/2$ when the pointer-system interaction takes place. Out of this window a time evolution of a system+probe composite is unitary and governed by U . Then, a time evolved system-pointer state reads as follows:

$$(5) \quad \begin{aligned} |\Psi(t)\rangle &= U(t, t_w) e^{-gAP} U(t_w, t_i) |\psi(t_i)\rangle |\phi(t_1)\rangle \\ &= U(t, t_w) e^{-igAP} |\psi(t_w)\rangle |\phi(t_1)\rangle \\ &= U(t, t_w) \sum_k e^{-iga_k P} \langle a_k | \psi(t_w) \rangle |a_k\rangle |\phi(t_i)\rangle \end{aligned}$$

Finally, at $t = t_f$, the system becomes postselected i.e. it is subjected to an (ideal) projective measurement of an observable B reducing its state to of eigenstates $|b_f(t_f)\rangle$ of B . Upon the Aharonov-Bergmann-Lebowitz rule the pointer after the system postselection is left in a state

$$(6) \quad |\phi(t_f)\rangle = \sum_k [\langle b_f(t_f) | U(t_f, t_w) | a_k \rangle \langle a_k | \psi(t_w) \rangle] e^{-iga_k P} |\phi(t_i)\rangle$$

Let us note that, at least formally, $\langle b_f(t_w) | = \langle b_f(t_f) | U(t_f, t_w)$ is a bra-vector which is dual to a state $|b_f(t_w)\rangle = [\langle b_f(t_w) |]^\dagger =$

$[\langle b_f(t_f) | U(t_f, t_w) \rangle]^\dagger = U^\dagger(t_f, t_w) | b_f(t_f) \rangle = U(t_w, t_f) | b_f(t_f) \rangle$ which, in turn, is a postselected system state evolving *backward* in time. If in addition the system-probe interaction is weak—i.e. $g \ll 1$ —one can expand $e^{-iga_k P} \approx 1 - iga_k P$ and obtain

$$(7) \quad \begin{aligned} |\phi(t_f)\rangle &= \langle b_f(t_w) | \psi(t_w) \rangle (1 - ig\langle A \rangle^w P) |\phi(t_i)\rangle \\ &= \langle b_f(t_w) | \psi(t_w) \rangle \exp(-ig\langle A \rangle^w P) |\phi(t_i)\rangle \end{aligned}$$

where the quantity

$$(8) \quad \langle A \rangle^w = \frac{\langle b_f(t_w) | A | \psi(t_w) \rangle}{\langle b_f(t_w) | \psi(t_w) \rangle}$$

defines a *weak value* (Aharonov, Albert and Vaidman, 1988) of A . The weak value is thus a number characterizing a shift (or a motion) of a quantum pointer due to a weak interaction with a measured system. Although the weak value is generically complex, it is a measurable quantity (Dressel et al., 2014) and its real part can be interpreted as an amplitude of a measurement induced transition from a preselected into postselected state. Its squared modulus gives the corresponding probability of the process. Such an interpretation settles weak values in a natural context of a measurement driven dynamic process.

Quantum weak values (Aharonov, Albert and Vaidman, 1988) attract non-decreasing attention of physicists who apply them to analyze otherwise difficult problems. Let us mention only two examples: continuous and sometimes controversial discussion what the history of quantum particle really is (Englert et al., 2017; Vaidman, 2013a; 2014) or one of the spectacular counter-intuitive effects—the quantum Cheshire Cat (Aharonov, Popescu et al., 2013). Weak values seem to be of a particular usefulness in all the circumstances which require simultaneous measurement of otherwise non co-measurable observables (Aharonov, Cohen, Waegell et al., 2018; Aharonov, Popescu

et al., 2013; Aharonov and Vaidman, 2008; Dressel et al., 2014; Vaidman, Ben-Israel et al., 2017). Let us emphasise one but crucial feature concerning *null* weak values. In a well established scheme of an ideal measurement Eq.(1) the pointer state couples to a particular eigenvector with a particular eigenvalue a_k . If $a_k = 0$ there is no effect on a pointer state and the probe remains not ‘shifted’. There is however an obvious modification of the system state which is projected on a subspace corresponding to the null eigenvalue. In the weak measurement scheme the situation essentially differs. If a particular *weak value* vanishes $\langle A \rangle^w = 0$ and a postselected state of a system is obtained the property corresponding to A cannot be detected. The weak value indicates a coupling-induced imprint which is left on the pointer, conditioned on the postselection. As a consequence, the vanishing weak value correlates successful postselection with the quantum pointer having been left unchanged despite the interaction with the system. Let us emphasise that for weak values Eq.(8) there is no clear analogy of the eigenvalue-eigenstate link and their relation to quantum properties remain a subject of an intensive investigation (Aharonov, Cohen, Waegell et al., 2018; Matzkin, 2019; Vaidman, Ben-Israel et al., 2017).

To explore further an interpretation of null weak value (further used in a general context of a past of quantum particles (Duprey and Matzkin, 2017; 2018; Sokolovski, 2018)) we discuss one of the first quantum phenomena exhibiting non-intuitive predictions based upon weak values and weak measurement scheme: the Cheshire Cat effect (Aharonov, Popescu et al., 2013; Ashby, Schwarz and Schlosshauer, 2016). For the sake of making our discussion as self-contained as possible we recall a very basic idea of the Cheshire Cat as primarily introduced in (Aharonov, Popescu et al., 2013). The quantum Cheshire Cat effect indicates separation of internal and exter-

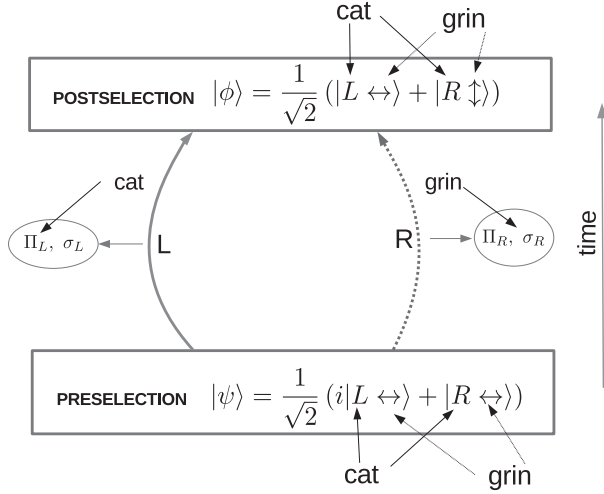


Figure 1: The Cheshire Cat effect. Photon in an interferometer is preselected in a state $|\psi\rangle$. It is then *weakly* measured along left and right arms (possible L, R paths) and then postselected in a state $|\phi\rangle$. The weak values $\langle\Pi_L\rangle^w = 1$, $\langle\Pi_R\rangle^w = 0$, $\langle\sigma_L\rangle^w = 0$ and $\langle\sigma_R\rangle^w = 1$ interpreted as a presence of the Cat in the left arm and its grin in the right one indicate separation of photon polarization (‘grin’) and photon position.

nal degrees freedom of a quantum system corresponding to the Cat’s position and the grin respectively and justifying a direct implementation of terminology adopted from Carroll’s novel *Alice’s Adventures in Wonderland*. The archetype of the Cheshire Cat (Aharonov, Popescu et al., 2013) is a photonic system and an effect takes place in a *two-path* Mach-Zehnder setting with a path ‘chosen’ by the photon—either left L or right R —describing the ‘external’ degree of freedom and photonic polarization—horizontal \leftrightarrow or vertical \updownarrow representing internal degree of freedom as it is presented in Figure 1 above or on Figure 1 in (Aharonov, Popescu et al., 2013). Effectively, both types

of degree of freedom are qubits i.e. a state space of the system is $\mathcal{H} = \mathbf{C}^2 \otimes \mathbf{C}^2 = \text{span}\{|L\rangle, |R\rangle\} \otimes \text{span}\{|\leftrightarrow\rangle, |\updownarrow\rangle\}$ where L, R and $\leftrightarrow, \updownarrow$ label the ‘external’ and internal degrees of freedom respectively. Detection of the Cat’s position corresponds to a measurement related to the projectors: $\Pi_L = |L\rangle\langle L|$ and $\Pi_R = |R\rangle\langle R|$ whereas a measurement of Cat’s grin (an internal degree of freedom) in a given (either left or right) position requires the projectors $\sigma_L = \Pi_L \sigma_z$ and $\sigma_R = \Pi_R \sigma_z$ where $\sigma_z = |+\rangle\langle +| - |-\rangle\langle -|$ for $|\pm\rangle = [|\leftrightarrow\rangle \pm i|\updownarrow\rangle]/\sqrt{2}$. According to the proposal given in (Aharonov, Popescu et al., 2013), the system is preselected (prepared) in a specific but not entangled state $|\psi\rangle = \frac{1}{\sqrt{2}}(i|L\leftrightarrow\rangle + |R\leftrightarrow\rangle)$ and, after passing the interferometer, postselected in a state $|\phi\rangle = \frac{1}{\sqrt{2}}(|L\leftrightarrow\rangle + |R\updownarrow\rangle)$. In the meantime, one measures *weak values* Eq.(8) of the quantities $\Pi_{L,R}$ and $\sigma_{L,R}$ i.e. weak values of the Cat’s position and grin respectively. One obtains (Aharonov, Popescu et al., 2013) $\langle \Pi_L \rangle^w = 1$, $\langle \Pi_R \rangle^w = 0$, $\langle \sigma_L \rangle^w = 0$ and $\langle \sigma_R \rangle^w = 1$. Whenever a weak value is null the corresponding quantum property (either the Cat’s presence or its grin in one of two arm of the interferometer) is absent in a particular arm where the weak measurement was performed. With such an interpretation we infer that a presence of the Cat (indicated by $\langle \Pi_L \rangle^w \neq 0$) in the left L arm of the interferometer is accompanied by a presence of Cat’s grin in the right R interferometer’s arm as indicates $\langle \sigma_R \rangle^w \neq 0$. Clearly, the description provided here is far from being complete. In particular it does not take into account specific experimental circumstances typical for realistic Cheshire Cat measurements (Duprey, Kanjilal et al., 2018), recent experiments (Ashby, Schwarz and Schlosshauer, 2016; Denkmayr et al., 2014) or an effect of decoherence (Schlosshauer, 2007) modifying weak values (Shikano and Hosoya, 2009) and the Cheshire Cat predictions (Dajka, 2020; Richter, Dziewit and Dajka, 2018). Moreover, interpreting null weak value as a hallmark of a ‘non-

presence of something’ is controversial and far from being commonly accepted (Duprey and Matzkin, 2017; 2018; Hance, Rarity and Ladyman, 2021; Sokolovski, 2018; Vaidman, 2017b).

3. Faint trace of a particle in a Vaidman interferometer

Quantum particle can be prepared (preselected) in a given and desired state and may also be postselected in another state with a known at least in principle probability. That what occurs in between, what is the particle’s past, remains problematic due to specific features of quantum measurements, cf. Eq.(1), unavoidably modifying quantum states of measured objects. It is particularly important if one asks about a history of a quantum particle passing through interferometer: the particle enters the device and leaves it (if its outcome is measured), however, inside the interferometer, unless it’s coherence is lost, the particle leaves nothing but a faint trace *defining* its presence, a faint trace which is a result of a *small* change of an amplitude of a component orthogonal to an undisturbed particle’s state (weakly). Such a faint trace is measurable only in experiments operating on an ensemble of particles having the same pre- and postselection. Past of a quantum particle in a nested Mach–Zehnder interferometer (*Vaidman interferometer*)—proposed (Vaidman, 2013a) and presented in Figure (2)—was recently studied in (Vaidman, 2013a) using quantum weak values (Aharonov, Albert and Vaidman, 1988; Aharonov and Vaidman, 2008; Vaidman, Ben-Israel et al., 2017) and the two state vector formalism (TSVF) (Aharonov and Vaidman, 2008). In this approach the faint trace is left by a particle unless a *weak value* of an appropriate projecting operator vanishes. A weak measurement

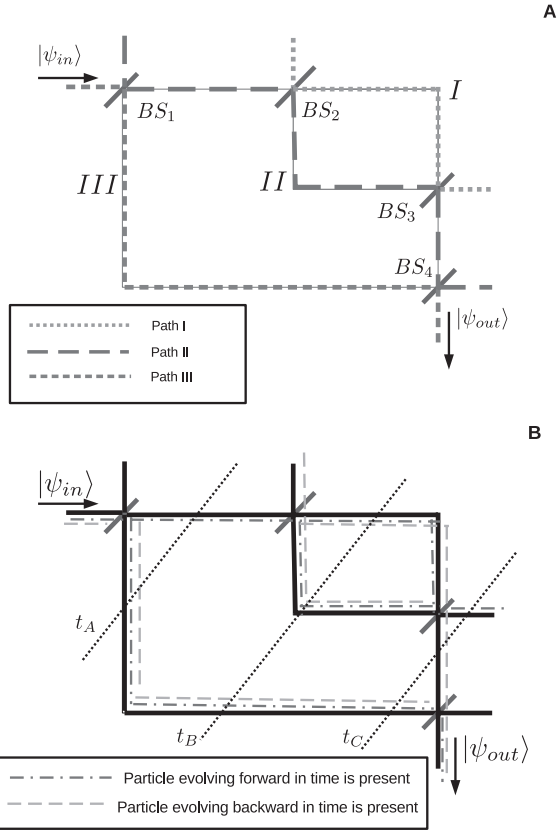


Figure 2: Vaidman interferometer consisting of four beam splitters $BS_{1,2,3,4}$ acting as unitary transformation in Eq.(10). In *panel A* there are three ‘paths’ Eq.(9) indicated *I, II, III* and labelled by different colours. An entrance of an input a particle in a state $|\psi_{in}\rangle$ and an exit of a particle in a state $|\psi_{out}\rangle$ are indicated by arrows. Time instants $t_{A,B,C}$ when weak measurements take place are indicated by dotted lines in *panel B*. The dash-dot and dash lines shown in a legend box of *panel B* indicate a *non-zero* amplitude for meeting a particle evolving forward and (respectively) backward in time. According to the TSVF the particle leaves a faint trace only if *both lines* coincide i.e. on a green path (*III*) and inside the internal interferometer as it is indicated in *panel B*.

of particle traces, contrary to a traditional and collapse assisted one, hardly perturbs the system maintaining its coherence sufficiently for interference effects to occur. The results, however, are highly confounding: particles seem to follow anomalous *discontinuous* path. Such a seemingly weird conclusion results in plethora of controversies (Li, Al-Amri and Zubairy, 2013; Vaidman, 2013b), some of them are quite recent (cf. Hance, Rarity and Ladyman, 2021), and since that time (almost) all works on that problem have come in triads: a paper, commentary inspired by the paper and a reply to the comment (Li, Al-Amri and Zubairy, 2013; Vaidman, 2013b). The main reason is that the TSVF (Aharonov and Vaidman, 2008) applied in (Vaidman, 2013a) is one of few possible approaches to investigate quantum past. The other non-equivalent alternatives are consistent (decoherent) histories (Griffiths, 2016; Vaidman, 2017a) (briefly presented below), standard quantum mechanics (Englert et al., 2017; 2019; Peleg and Vaidman, 2019) and many other other alternative studies (Aharonov, Cohen, Landau et al., 2017; Bartkiewicz et al., 2015; Hashmi et al., 2016; 2018; Potoček and Ferenczi, 2015; Vaidman, 2016a,b; 2018). Moreover, even recent experiments and their detailed analyses fail to resolve all the controversial issues (Aharonov, Cohen, Landau et al., 2017; Danan et al., 2013; Geppert-Kleinrath et al., 2018; Saldanha, 2014; Salih, 2015; Sponar et al., 2019; Vaidman, Danan et al., 2015; Wieśniak, 2018; Zhou et al., 2017). A conception of discontinuous path is clearly counter-intuitive but there are analyses (Aharonov, Cohen, Waegell et al., 2018; Vaidman, 2020) and claims which support the faint-trace anomalous picture as experimentally confirmed.

For a sake of completeness we formalize the Vaidman interferometer and review the controversial features of the faint traces of particles passing it. The original Vaidman interferometer is presented in Fig-

ure (2). It consists of spatial degree of freedom given by three paths denoted by I, II, III respectively and four beam splitters. The Vaidman interferometer can effectively be described (Dajka, 2021; Englert et al., 2017; 2019; Peleg and Vaidman, 2019) as a three level quantum system (a qutrit) with a state space spanned by an orthonormal basis

$$(9) \quad |I\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |II\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } |III\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Let us note an intuitively acceptable rationale behind imposing orthogonality on the set Eq.(9): if a particle collapses to a particular state in Eq.(9), at the same time there is no amplitude to be in an another one. In an ideal setting of a noise-less system (for a noisy dephasing model (cf. Dajka, 2021)), a passage of a particle is described by a unitary transformation composed of four unitaries $U_4 U_3 U_2 U_1$ corresponding to subsequent beam splitters termed as $BS_i, i = 1, \dots, 4$ in Figure 2:

$$(10) \quad U_1 = U_4 = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & -1 & \sqrt{2} \\ 0 & \sqrt{2} & 1 \end{pmatrix} \text{ and } \\ U_2 = U_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}.$$

The strategy applied in (Vaidman, 2013a) to infer the path of a particle entering and leaving Vaidman interferometer in a state $|III\rangle$ was to investigate its *weak trace* at three instants t_A, t_B, t_C indicated in Figure (2). In a framework of the TSVF (Aharonov and Vaidman, 2008) and according to (Vaidman, 2013a) the weak trace is indicated

by a non-vanishing weak value (Aharonov, Albert and Vaidman, 1988; Duprey and Matzkin, 2017) of one of the projectors

$$(11) \quad \langle \Pi_i \rangle_w^q = \frac{\langle \psi_{post}^q | \Pi_i | \psi_{pre}^q \rangle}{\langle \psi_{post}^q | \psi_{pre}^q \rangle}, \text{ where } \Pi_i = |i\rangle\langle i|, \quad i = I, II, III$$

and $q = A, B, C$

where forward in time evolving preselected state (directly prior to the measurement of Π_i) and backward in time postselected state (immediately after the measurement) states compose a *two-state vector* $\langle \psi_{post} | | \psi_{pre} \rangle$.

Most of controversies originate from highly counter-intuitive conclusions provided in (Vaidman, 2013a) indicating possibility of discontinuous trajectories followed by a particle passing through Vaidman interferometer. There are three instants t_A, t_B, t_C where the weak trace is measured: A : just after it passes the interferometer BS_1 but before it arrives to BS_2 , B : where the weak measurement becomes conducted for all three potential paths and C : after the BS_3 beam splitter as presented in Figure (2). The corresponding forward-in-time evolving preselected states read as: $|\psi_{pre}^A\rangle = U_1|III\rangle$, $|\psi_{pre}^B\rangle = U_2U_1|III\rangle$ and $|\psi_{pre}^C\rangle = U_3U_2U_1|III\rangle$. In the time-symmetric TSVF setting the postselected states $|\psi_{post}^A\rangle = U_2^\dagger U_3^\dagger U_4^\dagger|III\rangle$, $|\psi_{post}^B\rangle = U_3^\dagger U_4^\dagger|III\rangle$ and $|\psi_{post}^C\rangle = U_4^\dagger|III\rangle$ describe a hypothetical particle detected at III evolving backward in time. Results of the weak measurement are summarized in Table (2). According to (Vaidman, 2013a) a presence of a particle is defined by its non-vanishing weak trace. The counter-intuitive conclusion which can be inferred upon the results summarized in Table (2) is the following: at A and C the particle is present in III , what upon Figure (2) is intuitively acceptable, but at B it is also present in an internal loop (I, II) of the Vaidman interferometer. The above formal analysis one can support utilizing the

$\langle \Pi_{I,II,III} \rangle_w^{A,B,C}$	I	II	III	$U_{pre}^{A,B,C}$	$U_{post}^{A,B,C}$
A	0	0	1	U_1	$U_2^\dagger U_3^\dagger U_4^\dagger$
B	-1	1	1	$U_2 U_1$	$U_3^\dagger U_4^\dagger$
C	0	0	1	$U_3 U_2 U_1$	U_4^\dagger

Table 2: Weak traces $\langle \Pi_{I,II,III} \rangle_w^{A,B,C}$ Eq.(11) of a particle in Vaidman interferometer in Figure (2) at different instants A, B, C indicated in Figure (2) and corresponding pre- and postselections given by $|\psi_{pre}^{A,B,C}\rangle = U_{pre}^{A,B,C}|III\rangle$ and $|\psi_{post}^{A,B,C}\rangle = U_{post}^{A,B,C}|III\rangle$ respectively. It follows that at t_A and t_C a particle leaves its weak trace on paths I and III respectively whereas at t_B a faint trace is present on each of three paths.

TSVF and a *sine qua non* condition for a presence of particle: to get a non-vanishing *weak value* of a particular projector in a weak measurement scheme both the amplitudes of forward and backward evolving component of the two-vector necessarily must not vanish. The regions where the amplitudes of forward and backward evolving states do not vanish are depicted in Figure (2) with dash-dot and dash lines respectively. The only regions where the lines coincide are the path labeled by III and the internal loop of the Vaidman interferometer. There is a faint trace left by particles on the path II neither between beam splitters BS_1 and BS_2 (potentially used photons entering the internal loop) nor between BS_3 and BS_4 where the photons could exit the internal loop. In simple words, upon the TSVF we conclude that a presence of a particle indicated by a faint trace in an internal loop is not accompanied by any trace of a particle entering or leaving the internal loop and the particle path is *discontinuous*. It is obvious that such a confounding result needs further experimental verification. One can safely assume that any potential experiment,

as it was so far, will be highly subtle and sophisticated (Aharonov, Cohen, Waegell et al., 2018; Dziewior et al., 2019; Rebufello et al., 2021; Vaidman, 2020).

4. Consistent histories as an alternative

Orthodox (Copenhagen orthodox) researchers claim that one cannot talk about a quantum particle between measurements at all. This is an obvious limitation of standard quantum theory radically excluding important questions related to a past behaviour of quantum systems. Previously discussed two-state vector formalism (Aharonov and Vaidman, 2008) and consistent (or decoherent) histories approach (Gell-Mann and Hartle, 1993; 1999; Griffiths, 1984; 2003; Omnès, 1988; 1994) serve as fruitful examples of theoretical extensions going beyond (and sometimes across) the Copenhagen interpretation. In particular, we have recently been witnessing how the two above-mentioned approaches are competitively applied to a problem of a past behaviour of a quantum particle in a Vaidman interferometer (Dajka, 2021; Vaidman, 2013a; 2017a). Using different methods, consistent histories allow one to gain an additional insight if they are applied to problems ranging from a small but fundamental (Griffiths, 2013; 2014; 2015; 2017) to the largest scale (Craig, 2016; Riedel, Zurek and Zwolak, 2016). The consistency of histories allows one to assign probabilities to sequences of suitably defined events for a quantum system. As the quantum events in this perspective do not rely on the notion of measurement *per se*, the consistent histories approach enables one to design a new type of logical approach (Griffiths, 1984) which is essentially different to the Birkhoff and von Neumann quantum logic (Birkhoff and Von Neumann, 1936). There is a particular

practical advantage of using consistent histories formalism to gain information about a system if an external measurement is not available either because of fundamental or simply technical reasons as it is in the case of Vaidman interferometer where an experiment output is highly sensitive to a measurement-induced coherence deficiency. Quantum reasoning based on consistent histories (Griffiths, 1996) uses projective decomposition of identity $PDI = \{P^k\}$ where $P^j P^k = \delta_{jk} P^k$ and $\sum_j P^j = I$ as its cornerstone (Griffiths, 2003). It serves as a quantum-mechanical counterpart of an event algebra used in standard stochastics. There is, however, one crucial yet fundamental requirement which is additionally imposed: *the single framework rule* (Griffiths, 1996; 2003; 2015). According to that rule simultaneous reasoning to physical properties is meaningful if it is limited to ‘events’ which are compatible (Griffiths, 1996; 2003; 2013; 2015). At each time instant all quantum properties of a system correspond to elements of an instantaneous PDI having assigned probabilities and a time evolution of the quantum systems studied with consistent histories model can be considered as a stochastic process. With such an approach neither the future nor the past states of the system must be determined by the present state since, instead, they are (only) related by their probabilities. If the probabilities are 0 or 1 one arrives at a deterministic time evolution. A mathematical stage accommodating sequences of events is a tensor product of ‘instantaneous’ Hilbert spaces. Quantum properties (events) related to time evolving systems are its *histories* forming time-dependent PDIs where (generically) $PDI(t_i) \neq PDI(t_j)$ for different time instants giving one a chance to pose different questions concerning different quantum properties of the systems at different time instants. Assigning probabilities to non-commuting quantum properties is only meaningful if there is no interference between pairs of histories which suppose to be *decoher-*

ent. After extending the celebrated Born rule to multi-time history one can assign a weight to a sequence of events (Griffiths, 1984; 1996; 2003; 2013; 2015) and impose *consistency condition* (Griffiths, 1984; 1996; 2003; 2013; 2015) satisfied by histories which are meaningful. Let us note that a quantum history of a physical system is a sequence of quantum events at successive times, where a quantum event at a particular time can be any quantum property of the system in question (Griffiths, 2003). Such a convenient tool can serve to analyse properties of systems that are very difficult to measure and, in particular, consistent reasoning has already been applied to study history of a particle in a Vaidman interferometer. The result was that the TSVF predictions described above are based upon inconsistent histories and hence are meaningless. Clearly, an associate debate (Griffiths, 2016; Vaidman, 2017a) was surprising to nobody. Consistent histories seem to be an attractive and mathematically sound extension of quantum mechanics (or quantum stochastics) also suitable to study the past of quantum systems or even to dissolve the (in)famous measurement problem (Griffiths, 1996; 2015). The single framework rule, however, crucial for consistent reasoning, is highly more ‘invasive’ for quantum theory in comparison to a simple inclusion of a backward in time evolving postselected state as it is done in the TSVF. To avoid long and technical argumentation to support this statement and to indicate both existing and potential problems with the consistent histories let us invoke Mermin’s opinion from (2013):

[But] I am disconcerted by the reluctance of some consistent historians to acknowledge the utterly radical nature of what they are proposing. The relativity of time was a pretty big pill to swallow, but the relativity of reality itself is to the relativity of time as an elephant is to a gnat.

5. Concluding (yet not conclusive) remarks

Weak values, contrary to standard eigenvalues, are still waiting for a commonly accepted interpretation. In an absence of the eigenstate-eigenvalue link it is not obvious if weak measurements and their outputs can credibly describe ‘elements of reality’ and properties of quantum systems (Matzkin, 2019; Vaidman, 2017b). It is agreed that weak values possess certain non-trivial predictive value and they are related, at least to some extent, to real properties of physical systems. Unfortunately, it is difficult to assign any ‘hard’ limits of their applicability. Life would be definitely simpler if weak values were strongly measurable. In particular, physical meaning of vanishing weak values applied in a current context of a past of quantum systems remains disputable (Duprey and Matzkin, 2017; 2018; Hance, Rarity and Ladyman, 2021; Sokolovski, 2018; Vaidman, Ben-Israel et al., 2017) although there are analyses and experiments (Aharonov, Cohen, Waegell et al., 2018; Dziewior et al., 2019; Rebufello et al., 2021; Vaidman, 2020) which support the faint-trace anomalous picture as experimentally confirmed.

Time-symmetric description of quantum systems utilizing the TSVF (Aharonov, Cohen and Landsberger, 2017; Aharonov and Vaidman, 2008), despite certain problematic issues concerning its adaptation to open quantum systems requiring mixed states for their description (Dajka, 2021; Vaidman, Ben-Israel et al., 2017), seems to be less controversial. At the same time, however, it affects new research areas such as counterfactual reasoning (Vaidman, 1999) or even the hot-forever *free will problem* (Aharonov, Cohen and Shushi, 2016).

Counterfactuality and counterfactual reasoning (‘if it were a , then it would be b ’) (Lewis, 2001; Vaidman, 1999) is a natural extension of an interaction-free measurement, cf. Elitzur and Vaidman’s interaction-

free bomb detector (Elitzur and Vaidman, 1993), where a bomb (a detector, using more pacifistic terminology) indicating a presence of a single photon is put in one of the arms of a Mach–Zehnder interferometer. Even if the bomb does not blow up, its presence affects an interference pattern at the output of the interferometer. Direct application of the Aharonov-Bergman-Lebowitz rule Eq.(4) for counterfactual reasoning is not always sufficient and acceptable (Vaidman, 1999). The idea of using a time-symmetric approach to quantum counterfactuals has recently found a promising application in quantum-based counterfactual communication. Such communication protocols are counterfactual by using quantum effects to send messages without any matter or energy transfer between communicating parties (Vaidman, 2019; Wander, Cohen and Vaidman, 2021)). There are already known crypto protocols (Hance, Ladyman and Rarity, 2021; Kamaruddin, Shaari and Kolenderski, 2020; Noh, 2009; Rao and Srikanth, 2021) based upon that idea. Obviously, all the controversies concerning weak values, past of quantum systems and counterfactual reasoning both mentioned and not mentioned in this work to be resolved need to be supported by further experimental investigations.

One can conclude that with a tremendous development of highly sophisticated experimental techniques we are faced (maybe for the first time) with problems which *simultaneously* require advanced technology, fresh and flexible theory and, last but not least, sound interpretation. Let us take up this challenge.

Bibliography

- Aharonov, Y., Albert, D.Z. and Vaidman, L., 1988. How the Result of a Measurement of a Component of the Spin of a Spin-1/2 Particle Can Turn Out to Be 100. *Physical Review Letters* [Online], 60(14), pp.1351–1354. <https://doi.org/10.1103/PhysRevLett.60.1351>.
- Aharonov, Y., Bergmann, P.G. and Lebowitz, J.L., 1964. Time Symmetry in the Quantum Process of Measurement. *Physical Review* [Online], 134(6B), B1410–B1416. <https://doi.org/10.1103/PhysRev.134.B1410>.
- Aharonov, Y., Cohen, E., Landau, A. and Elitzur, A.C., 2017. The Case of the Disappearing (and Re-Appearing) Particle. *Scientific Reports* [Online], 7, p.531. <https://doi.org/10.1038/s41598-017-00274-w>.
- Aharonov, Y., Cohen, E. and Landsberger, T., 2017. The Two-Time Interpretation and Macroscopic Time-Reversibility. *Entropy* [Online], 19(3). <https://doi.org/10.3390/e19030111>.
- Aharonov, Y., Cohen, E. and Shushi, T., 2016. Accommodating Retrocausality with Free Will. *Quanta* [Online], 5(1), pp.53–60. <https://doi.org/10.12743/quanta.v5i1.44>.
- Aharonov, Y., Cohen, E., Waegell, M. and Elitzur, A.C., 2018. The Weak Reality That Makes Quantum Phenomena More Natural: Novel Insights and Experiments. *Entropy* [Online], 20(11). <https://doi.org/10.3390/e20110854>.
- Aharonov, Y., Popescu, S., Rohrlich, D. and Skrzypczyk, P., 2013. Quantum Cheshire Cats. *New Journal of Physics* [Online], 15(11), p.113015. Available at: <<http://stacks.iop.org/1367-2630/15/i=11/a=113015>>.
- Aharonov, Y. and Vaidman, L., 2008. The Two-State Vector Formalism: An Updated Review. In: J. Muga, R.S. Mayato and Í. Egusquiza, eds. *Time in Quantum Mechanics* [Online]. Berlin, Heidelberg: Springer, pp.399–447. https://doi.org/10.1007/978-3-540-73473-4_13.
- Ashby, J.M., Schwarz, P.D. and Schlosshauer, M., 2016. Observation of the Quantum Paradox of Separation of a Single Photon from One of Its Properties. *Physical Review A* [Online], 94(1), p.012102. <https://doi.org/10.1103/PhysRevA.94.012102>.

- Bartkiewicz, K. et al., 2015. One-State Vector Formalism for the Evolution of a Quantum State Through Nested Mach-Zehnder Interferometers. *Physical Review A* [Online], 91(1), p.012103. <https://doi.org/10.1103/PhysRevA.91.012103>.
- Birkhoff, G. and Von Neumann, J., 1936. The Logic of Quantum Mechanics. *Annals of Mathematics* [Online], 37(4), pp.823–843. <https://doi.org/10.2307/1968621>.
- Craig, D.A., 2016. The Consistent Histories Approach to Loop Quantum Cosmology. *International Journal of Modern Physics D* [Online], 25(08), p.1642009. <https://doi.org/10.1142/S0218271816420098>.
- Dajka, J., 2020. Scattering–Like Control of the Cheshire Cat Effect in Open Quantum Systems. *Quantum Reports* [Online], 2(1), pp.1–11. <https://doi.org/10.3390/quantum2010001>.
- Dajka, J., 2021. Faint Trace of a Particle in a Noisy Vaidman Three-Path Interferometer. *Scientific Reports* [Online], 11(1), p.1123. <https://doi.org/10.1038/s41598-020-80806-z>.
- Danan, A., Farfurnik, D., Bar-Ad, S. and Vaidman, L., 2013. Asking Photons Where They Have Been. *Physical Review Letters* [Online], 111(24), p.240402. <https://doi.org/10.1103/PhysRevLett.111.240402>.
- Denkmayr, T. et al., 2014. Observation of a Quantum Cheshire Cat in a Matter-Wave Interferometer Experiment. *Nature Communications* [Online], 5, p.4492. Available at: <<http://dx.doi.org/10.1038/ncomms5492>>.
- Dressel, J. et al., 2014. Colloquium. *Reviews of Modern Physics* [Online], 86(1), pp.307–316. <https://doi.org/10.1103/RevModPhys.86.307>.
- Duprey, Q., Kanjilal, S. et al., 2018. The Quantum Cheshire Cat Effect: Theoretical Basis and Observational Implications. *Annals of Physics* [Online], 391, pp.1–15. <https://doi.org/10.1016/j.aop.2018.01.011>.
- Duprey, Q. and Matzkin, A., 2017. Null Weak Values and the Past of a Quantum Particle. *Physical Review A* [Online], 95(3), p.032110. <https://doi.org/10.1103/PhysRevA.95.032110>.
- Duprey, Q. and Matzkin, A., 2018. Reply to “Comment on ‘null Weak Values and the Past of a Quantum Particle’”. *Physical Review A* [Online], 97(4), p.046103. <https://doi.org/10.1103/PhysRevA.97.046103>.

- Dziewior, J. et al., 2019. Universality of Local Weak Interactions and Its Application for Interferometric Alignment. *Proceedings of the National Academy of Sciences* [Online], 116(8), pp.2881–2890. <https://doi.org/10.1073/pnas.1812970116>.
- Elitzur, A.C. and Vaidman, L., 1993. Quantum Mechanical Interaction-Free Measurements. *Foundations of Physics* [Online], 23(7), pp.987–997. <https://doi.org/10.1007/BF00736012>.
- Englert, B.-G. et al., 2017. Past of a Quantum Particle Revisited. *Physical Review A* [Online], 96(2), p.022126. <https://doi.org/10.1103/PhysRevA.96.022126>.
- Englert, B.-G. et al., 2019. Reply to “Comment on ‘past of a Quantum Particle Revisited’ ”. *Physical Review A* [Online], 99(2), p.026104. <https://doi.org/10.1103/PhysRevA.99.026104>.
- Gell-Mann, M. and Hartle, J.B., 1993. Classical Equations for Quantum Systems. *Physical Review D* [Online], 47(8), pp.3345–3382. <https://doi.org/10.1103/PhysRevD.47.3345>.
- Gell-Mann, M. and Hartle, J.B., 1999. Innovation and Intellectual Property Rights. In: J. Fagerberg, D.C. Mowery and R.R. Nelson, eds. *Complexity, Entropy, and the Physics of Information*. Addison-Wesley, pp.266–290.
- Geppert-Kleinrath, H. et al., 2018. Multifold Paths of Neutrons in the Three-Beam Interferometer Detected by a Tiny Energy Kick. *Physical Review A* [Online], 97(5), p.052111. <https://doi.org/10.1103/PhysRevA.97.052111>.
- Griffiths, R.B., 1984. Consistent Histories and the Interpretation of Quantum Mechanics. *Journal of Statistical Physics* [Online], 36(1), pp.219–272. <https://doi.org/10.1007/BF01015734>.
- Griffiths, R.B., 1996. Consistent Histories and Quantum Reasoning. *Physical Review A* [Online], 54(4), pp.2759–2774. <https://doi.org/10.1103/PhysRevA.54.2759>.
- Griffiths, R.B., 2003. *Consistent Quantum Theory*. Cambridge: Cambridge University Press.

- Griffiths, R.B., 2013. A Consistent Quantum Ontology. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* [Online], 44(2), pp.93–114. <https://doi.org/10.1016/j.shpsb.2012.12.002>.
- Griffiths, R.B., 2014. The New Quantum Logic. *Foundations of Physics* [Online], 44(6), pp.610–640. <https://doi.org/10.1007/s10701-014-9802-4>.
- Griffiths, R.B., 2015. Consistent Quantum Measurements. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* [Online], 52, pp.188–197. <https://doi.org/10.1016/j.shpsb.2015.07.002>.
- Griffiths, R.B., 2016. Particle Path Through a Nested Mach-Zehnder Interferometer. *Physical Review A* [Online], 94(3), p.032115. <https://doi.org/10.1103/PhysRevA.94.032115>.
- Griffiths, R.B., 2017. What Quantum Measurements Measure. *Physical Review A* [Online], 96(3), p.032110. <https://doi.org/10.1103/PhysRevA.96.032110>.
- Hance, J.R., Ladyman, J. and Rarity, J., 2021. How Quantum Is Quantum Counterfactual Communication? *Foundations of Physics* [Online], 51(1), p.12. <https://doi.org/10.1007/s10701-021-00412-5>.
- Hance, J.R., Rarity, J. and Ladyman, J., 2021. *Does the Weak Trace Show the Past of a Quantum Particle in an Unperturbed System?* [Online]. Available at: <<https://arxiv.org/abs/2109.14060>>.
- Hashmi, F.A., Li, F., Zhu, S.-Y. and Zubairy, M.S., 2016. Two-State Vector Formalism and Quantum Interference. *Journal of Physics A: Mathematical and Theoretical* [Online], 49(34), p.345302. <https://doi.org/10.1088/1751-8113/49/34/345302>.
- Hashmi, F.A., Li, F., Zhu, S.-Y. and Zubairy, M.S., 2018. Reply to the Comment on ‘two-State Vector Formalism and Quantum Interference’. *Journal of Physics A: Mathematical and Theoretical* [Online], 51(6), p.068001. <https://doi.org/10.1088/1751-8121/aa8d23>.

- Kamaruddin, S., Shaari, J.S. and Kolenderski, P., 2020. Counterfactual Protocol Within Device Independent Framework and Its Insecurity. *Scientific Reports* [Online], 10(1), p.5847. <https://doi.org/10.1038/s41598-020-62812-3>.
- Lewis, D., 2001. *Counterfactuals*. 2nd ed. Oxford: Blakwell.
- Li, Z.-H., Al-Amri, M. and Zubairy, M.S., 2013. Comment on “Past of a Quantum Particle”. *Physical Review A* [Online], 88(4), p.046102. <https://doi.org/10.1103/PhysRevA.88.046102>.
- Matzkin, A., 2019. Weak Values and Quantum Properties. *Foundations of Physics* [Online], 49(3), pp.298–316. <https://doi.org/10.1007/s10701-019-00245-3>.
- Mermin, N.D., 2013. *Annotated Interview with a Qbist in the Making* [Online]. Available at: <<https://arxiv.org/abs/1301.6551>> [visited on 9 September 2022].
- Noh, T.-G., 2009. Counterfactual Quantum Cryptography. *Physical Review Letters* [Online], 103(23), p.230501. <https://doi.org/10.1103/PhysRevLett.103.230501>.
- Omnès, R., 1988. Logical Reformulation of Quantum Mechanics. I. Foundations. *Journal of Statistical Physics* [Online], 53(3), pp.893–932. <https://doi.org/10.1007/BF01014230>.
- Omnès, R., 1994. *The Interpretation of Quantum Mechanics*. Princeton: Princeton University Press.
- Peleg, U. and Vaidman, L., 2019. Comment on “Past of a Quantum Particle Revisited”. *Physical Review A* [Online], 99(2), p.026103. <https://doi.org/10.1103/PhysRevA.99.026103>.
- Potoček, V. and Ferenczi, G., 2015. Which-Way Information in a Nested Mach-Zehnder Interferometer. *Physical Review A* [Online], 92(2), p.023829. <https://doi.org/10.1103/PhysRevA.92.023829>.
- Rao, V.N. and Srikanth, R., 2021. Noiseless Attack and Counterfactual Security of Quantum Key Distribution. *Physical Review A* [Online], 104(2), p.022424. <https://doi.org/10.1103/PhysRevA.104.022424>.
- Rebuffleo, E. et al., 2021. Anomalous Weak Values Via a Single Photon Detection. *Light: Science & Applications* [Online], 10(1), p.106. <https://doi.org/10.1038/s41377-021-00539-0>.

- Richter, M., Dziewit, B. and Dajka, J., 2018. The quantum cheshire cat effect in the presence of decoherence. *Advances in Mathematical Physics* [Online], 2018(1), p.7060586. <https://doi.org/10.1155/2018/7060586>.
- Riedel, C.J., Zurek, W.H. and Zvolak, M., 2016. Objective Past of a Quantum Universe: Redundant Records of Consistent Histories. *Physical Review A* [Online], 93(3), p.032126. <https://doi.org/10.1103/PhysRevA.93.032126>.
- Saldanha, P.L., 2014. Interpreting a Nested Mach-Zehnder Interferometer with Classical Optics. *Physical Review A* [Online], 89(3), p.033825. <https://doi.org/10.1103/PhysRevA.89.033825>.
- Salih, H., 2015. Commentary: “Asking Photons Where They Have Been” - Without Telling Them What to Say. *Frontiers in Physics* [Online], 3, p.47. <https://doi.org/10.3389/fphy.2015.00047>.
- Schlosshauer, M., 2007. *Decoherence and the Quantum-to-Classical Transition* [Online], *Frontiers Collection*. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-540-35775-9>.
- Shikano, Y. and Hosoya, A., 2009. Weak Values with Decoherence. *Journal of Physics A: Mathematical and Theoretical* [Online], 43(2), p.025304. <https://doi.org/10.1088/1751-8113/43/2/025304>.
- Sokolovski, D., 2018. Comment on “Null Weak Values and the Past of a Quantum Particle”. *Physical Review A* [Online], 97(4), p.046102. <https://doi.org/10.1103/PhysRevA.97.046102>.
- Sponar, S. et al., 2019. Asking Neutrons Where They Have Been. *Journal of Physics: Conference Series* [Online], 1316, p.012002. <https://doi.org/10.1088/1742-6596/1316/1/012002>.
- Vaidman, L., 1999. Time-Symmetrized Counterfactuals in Quantum Theory. *Foundations of Physics* [Online], 29(5), pp.755–765. <https://doi.org/10.1023/A:1018826607583>.
- Vaidman, L., 2013a. Past of a Quantum Particle. *Physical Review A* [Online], 87(5), p.052104. <https://doi.org/10.1103/PhysRevA.87.052104>.
- Vaidman, L., 2013b. Reply to “Comment on ‘past of a Quantum Particle’”. *Physical Review A* [Online], 88(4), p.046103. <https://doi.org/10.1103/PhysRevA.88.046103>.

- Vaidman, L., 2014. Tracing the Past of a Quantum Particle. *Physical Review A* [Online], 89(2), p.024102. <https://doi.org/10.1103/PhysRevA.89.024102>.
- Vaidman, L., 2016a. Comment on “One-State Vector Formalism for the Evolution of a Quantum State Through Nested Mach-Zehnder Interferometers”. *Physical Review A* [Online], 93(3), p.036103. <https://doi.org/10.1103/PhysRevA.93.036103>.
- Vaidman, L., 2016b. Comment on “Which-Way Information in a Nested Mach-Zehnder Interferometer”. *Physical Review A* [Online], 93(1), p.017801. <https://doi.org/10.1103/PhysRevA.93.017801>.
- Vaidman, L., 2017a. Comment on “Particle Path Through a Nested Mach-Zehnder Interferometer”. *Physical Review A* [Online], 95(6), p.066101. <https://doi.org/10.1103/PhysRevA.95.066101>.
- Vaidman, L., 2017b. Weak Value Controversy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* [Online], 375(2106), p.20160395. <https://doi.org/10.1098/rsta.2016.0395>.
- Vaidman, L., 2018. Comment on ‘two-State Vector Formalism and Quantum Interference’. *Journal of Physics A: Mathematical and Theoretical* [Online], 51(6), p.068002. <https://doi.org/10.1088/1751-8121/aa8d24>.
- Vaidman, L., 2019. Analysis of Counterfactuality of Counterfactual Communication Protocols. *Physical Review A* [Online], 99(5), p.052127. <https://doi.org/10.1103/PhysRevA.99.052127>.
- Vaidman, L., 2020. Neutrons and Photons Inside a Nested Mach-Zehnder Interferometer. *Physical Review A* [Online], 101(5), p.052119. <https://doi.org/10.1103/PhysRevA.101.052119>.
- Vaidman, L., Ben-Israel, A. et al., 2017. Weak Value Beyond Conditional Expectation Value of the Pointer Readings. *Physical Review A* [Online], 96(3), p.032114. <https://doi.org/10.1103/PhysRevA.96.032114>.
- Vaidman, L., Danan, A., Farfurnik, D. and Bar-Ad, S., 2015. Response: Commentary: “Asking Photons Where They Have Been” - Without Telling Them What to Say. *Frontiers in Physics* [Online], 3, p.48. <https://doi.org/10.3389/fphy.2015.00048>.

- Wander, A., Cohen, E. and Vaidman, L., 2021. Three Approaches for Analyzing the Counterfactuality of Counterfactual Protocols. *Physical Review A* [Online], 104(1), p.012610. <https://doi.org/10.1103/PhysRevA.104.012610>.
- Wieśniak, M., 2018. Spectra in Nested Mach–Zehnder Interferometer Experiments. *Physics Letters A* [Online], 382(36), pp.2565–2568. <https://doi.org/10.1016/j.physleta.2018.06.032>.
- Zhou, Z.-Q. et al., 2017. Experimental Observation of Anomalous Trajectories of Single Photons. *Physical Review A* [Online], 95(4), p.042121. <https://doi.org/10.1103/PhysRevA.95.042121>.

Response to Wysocki on indifference¹

Walter Block

Loyola University New Orleans

Abstract

Nozick (1977) was a critique of the view of Austrian economics which rejected the notion of indifference in human action. This author claimed that this stance was incompatible with the notion of the supply of a good, and, also, with diminishing marginal utility, both of which were strongly supported by this praxeological school of thought. Block (1980) was an attempt to rescue the Austrian school from this brilliant intellectual challenge. Hoppe (2005; 2009) rejected Nozick's challenge, and, also, Block's (1980) response. Block (2009a) and Block and Barnett (2010), defended Block's (1980) analysis of indifference. The latest contribution to this ongoing discussion is Wysocki (2021) who maintains that Hoppe was correct in his rejection of Nozick, while Block was not. The present paper is a rejoinder to Wysocki (2021).

Keywords

indifference, supply, diminishing marginal utility, praxeology.

¹I wish to thank two referees of This Journal for helpful comments on an earlier version of this paper which when incorporated greatly improved it. The usual caveats of course always apply.

1. Introduction

Wysocki (2021) is written in support of Hoppe (2005; 2009) in his debate with Block (2009a) and Block and Barnett (2010). This dispute concerns the best way to counter Nozick's (1977) rejection of Austrian praxeological economics which involves indifference, the supply of a good, and the law of diminishing marginal utility.

I will be quoting widely from Wysocki (2021)² and then responding to his many, and important points. Let us begin.

According to Wysocki (2021):

If he were forced to give up a unit of apple juice or the (sic) one of mineral water, he would be indifferent between the two.

Before responding to this quote, let me try to give the reader the context. I'm going to be putting "words into the mouth" of this author, but, hopefully, I will be accurately transmitting his overall viewpoint. In my understanding of what he is about, he is supporting Hoppe's criticism of my critique of Nozick. All three of us, Hoppe, Wysocki and me reject Nozick's criticism of Austrian economics; but Hoppe and Wysocki maintain that my rejection of Nozick is a failure. The issue turns on indifference and the supply curve. All economists, Austrian or not, maintain that there is such a thing as a supply curve. If this is to exist, avers Nozick, then the economic actor such see every element of the supply as equivalent. If there is to be a supply curve of apples, for example, the consumer must see all the apples as the same; he must be indifferent between them. If he is not, if he sees some apples as different from others, there cannot be one supply curve of apples; there must be two or more, depending upon how many types

² All references to Wysocki, unless otherwise indicated, will be to this one article of his.

of apples are perceived. However, asserts Nozick, correctly, Austrians reject the concept of indifference; they are thus logically inconsistent in acquiescing as to the supply curve.

Let me now respond to this opening quote from Wysocki, who rejects Nozick's critique of Austrianism, and supports Hoppe's criticism of my critique of Nozick. My claim is that Austrians can have our cake and eat it too; we may acknowledge the legitimacy of the supply curve (the demand curve too), and still not buy into the concept of indifference, at least not when choices are being made.

I just do not see how anything like this could be true. For if this economic actor gave up one of these economic goods, that would, surely, demonstrate (Rothbard, 2011) that he valued it less than the other; and vice versa. The point is, there is no way he could reveal, in action, that he was indifferent between these two items.

If he were truly "indifferent between the two" then of course, *a la* Wysocki, he could not choose to set aside either. However, this contradicts this author's premise. He states that the economic actor was *forced* to give up one unit of either the juice or the H₂O. We much assume he complied. If so, he might well have tossed a coin to determine which one he gave up, but, in going along with the coin toss, he is not acting in an indifferent manner. Logically, he cannot do any such thing.

Our author continues:

Specifically, indifference has such a propositional content (believing that x and y are economically identical) that it cannot motivate an actor to act on it.

But this appears to be a direct contradiction of his previous statement. If indifference "cannot motivate" someone to undertake an action

of picking and choosing and setting aside, how can the person in question objectively reveal that he saw the apple juice and mineral water as equivalents?

Wysocki then avers as follows:

Specifically, indifference has such a propositional content (believing that x and y are economically identical) that it cannot motivate an actor to act on it. By contrast, homogeneity is a relation holding between economic goods. However, remember that economic goods are not mere physical goods.

Whenever a person engages in human action, giving up a gallon of water or purchasing an extra bottle containing this amount of H_2O , he cannot be engaging in an act of indifference between his present stock and the additional unit. Indeed, there is no such thing as an “act of indifference.” There can only be acts of preference and of setting aside. People can only be preferring or setting aside; e.g., treating these units as *different* from one another. A supply curve, then, consists of goods that are seen as physically indistinguishable one from the other in the absence of any human action that takes place with regard to them.³

In the view of Wysocki:

When asked how we should understand the concept of the same good presupposed by the universal law of time preference, an Austrian economist might reply in a similar fashion: we can easily learn whether x_1 (some economic good at t_1) and x_2 (some economic good at t_2) are the units of the same good. We would do so by checking whether an actor would now necessarily prefer x_1 to x_2 .

³ That would be a necessary condition. But not sufficient, for we must incorporate Machaj's (2007) insightful example of the same physical ring on the hand of one's fiancé, and another physically identical one in the jewelry shop (see Block, 2009b; also Block, 2012).

But how are we to “check” whether that is true or not? All we can observe is human action: people choosing amongst alternatives. It is logic alone, not any presumably empirical “checking” that can make any such determination. But Wysocki rejects this as “tautologies.” He states in this regard:

But these two apodictically true statements come at a price. For the consequence of the lack of the independent (of the laws in question) notion of the same good, would turn those laws into concealed tautologies.

This author continues:

Incidentally, similar remarks would apply to Austrian formulation of the universal law of time preference. Austrians hold that for one (and the same!) end, each actor would prefer to achieve it sooner rather than later. Note, this law also presupposes the notion of the same good—but this time in a sort of a temporal way for it is the same economic good that is carried over time (we may obtain it at t_1 or at t_2). When asked how we should understand the concept of the same good presupposed by the universal law of time preference, an Austrian economist might reply in a similar fashion: we can easily learn whether x_1 (some economic good at t_1) and x_2 (some economic good at t_2) are the units of the same good. We would do so by checking whether an actor would now necessarily prefer x_1 to x_2 . But these two apodictically true statements come at a price. For the consequence of the lack of the independent (of the laws in question) notion of the same good, would turn those laws into concealed tautologies.

Our author continues in his footnote 13:

In fact, it was Rothbard (2011, pp.15–16) himself who resorted to this tautologous defense of the universal law of time preference, which is evidenced by the following passage: ‘Time

preference may be called the preference for present satisfaction over future satisfaction or present good over future good, provided it is remembered that it is the same satisfaction (or “good”) that is being compared over the periods of time. Thus, a common type of objection to the assertion of universal time preference is that, in the wintertime, a man will prefer the delivery of ice the next summer (future) to delivery of ice in the present. This, however, confuses the concept “good” with the material properties of a thing, whereas it actually refers to subjective satisfactions. Since ice-in-the-summer provides different (and greater) satisfaction than ice-in-the-winter, they are not the same, but different goods. In this case, it is different satisfactions that are being compared, despite the fact that physical property of the thing may be the same. Whereas in the body of the text we considered the possible Austrian rejoinder in the form of saying that the same good can be conceptualized as the one that obeys the universal law of time preference, Rothbard merely contraposes by saying that if there is an apparent preference for a future good (ice cream in summer) over a present good (ice-cream in winter now), these two cannot constitute one and the same economic good. So, not only is the Rothbardian solution clearly circular, but also it gives the impression of fudging the notion of the same good. It may seem that whatever counterexamples to the law of time preference one may possibly come up with, Rothbard would rebut it by claiming that his critic invokes two distinct economic goods. This is yet another indication that an independent concept of the same good is logically required.

There are problems with this analysis. A tautology gives us no information about the real world. Rather, it merely indicates definitions. For example, “bachelor” is an “unmarried man.” This insight offered by Rothbard and rejected by Wysocki constitutes much more: it is a synthetic a priori statement which is both necessarily true and,

also, give us profound insight as to how the real world operates. As to “fudging” this seems an erroneous characterization. Let us consider a different case. It is alleged that “two things cannot be in the same place at the same time.” Someone objects: but more than two people can both be in New York City at the same time. The “fudger” responds: that’s too big a “place.” Whereupon the critic says: but more than two people can both be in Manhattan at the same time, and that borough is smaller than the entire city.” What is the “fudger” to do but to “fudge” once again? And to keep on doing so until the critic gives up. But “fudging” is not the correct description of this process. “Clarification” would do much better.

Wysocki states:

First and foremost, Nozick’s challenge may be construed as a purely logical objection to Austrian repudiation of indifference. After all, remember, the gist of Nozick’s objection was that without the concept of indifference, Austrians would be unable to formulate the law of diminishing marginal utility. Indubitably, in this respect Nozick is right. The law of diminishing marginal utility has it that when we deal with a supply of economically same units, each additional unit we value less; or, in other words, each additional unit is of lower utility.

To my way of thinking, this all depends upon precisely what is meant by “when we deal with.” If “dealing with” means, merely, contemplating a stock of goods, then “indifference” is quite acceptable. After all, it is a perfectly good word in the English language. Mere contemplation is not a human action, at least not vis a vis the object of the contemplation.⁴ But if “dealing with” is taken as acting with regard

⁴ It is a human action insofar as it indicates that the person prefers to contemplate at this moment in time when he could have been doing something else, which he sees as a lesser value to him.

to, say, a stock of goods, then, it depends upon the precise action. For example, suppose someone sells the entire stock to someone else. Then, again, indifference vis a vis elements of the stock with each other, is again acceptable.⁵ But, if he sells only one of these units, for example the 72nd one of them out of a holding of 100, then he demonstrates that he prefers the other elements to this one.

Here is our author once again:

Second, we must also concede to Nozick that pricing of the commodity also seems to rest on the notion of indifference. Then, if Austrians fail to somehow accommodate indifference into their theory, this would have disastrous consequences for their entire conceptual edifice. After all, it must be borne in mind that the market (equilibrium) price of a given product is a function of supply and demand. And the demand curve is but a reflection of the diminishing marginal utility of a given product. That is, the demand curve—rather unsurprisingly—slopes downwards because the more we have (of a given product), the less we value marginal units. And it is precisely why we are ready to buy more (of pretty much anything) only when the successive units of the product in question cost less and less. Therefore, it is clear to see that the demand curve reflects the logic of diminishing marginal utility. Hence, Nozick is right.

I do not agree with Nozick's attack on Austrian economics. A supply curve consists of a locus of points indicating prices and quantities. We refer here to quantities of goods, of course, such as apples, or shoes. Are we indifferent amongst all of these objects? Of course we are. We have not yet acted upon any of them. When and if we do, we can no longer be indifferent between them. If we choose one of them, we will have demonstrated preference, not indifference. Says

⁵ Here, there would be no question of indifference between the entire stock and the money the vendor receives in payment for it. Clearly he prefers the latter.

Nozick (paraphrase on my part): “Aha! You Austrians are logically inconsistent. You admit that indifference is required if the supply of a good is to make any sense.” But there is no inconsistency. Of course, indifference is all around us.⁶ It only breaks up when we actually *do* anything with regard to this supply, such as buy or sell any of it. In like manner, the same holds true for diminishing marginal (ordinal!) utility. We have the proverbial three bottles of water. We are indifferent between them all. They are now just sitting there. They all have precisely the same chemical properties (H_2O) and we have no psychic preferences amongst any of them. We cannot so much as tell the difference between them. But then, for some reason we have to give up one of them. A robber demands this of us, and we decide to comply. We are still indifferent between them. So, we grab one, at random. Behold, we are no longer indifferent! By our own action, we reveal that we prefer the other two bottles to this one; that is why we are giving up this one, and none other, to the thief. Has Nozick upset any Austrian theory? I cannot see how he has. Is this mode of refuting him fallacious as Hoppe contends, and, now Wysocki? I cannot see my way clear to agreeing with that, either. I fear that Wysocki is giving away too much of the Austrian store to Nozick in his concession: “Hence, Nozick is right.”

Rather, we can have our cake and eat it too. When indifference is required, in the absence of human action, in order to undergird diminishing marginal utility, or the supply and demand curves, we can retain it. But when human action takes place, we can jettison this concept.

We need not concede anything whatsoever to Nozick (1977) in this regard. He asserts that indifference is required to demonstrate diminishing marginal utility. It is not. Before action, yes, there is

⁶ In the television series Mary Tyler Moore, “love was all around us.”

indifference. There are stocks, supplies of goods, after all. But, when action occurs, when a unit of the good is given up, or added, there is preference, not indifference.

We now arrive, more specifically, at Wysocki's dissatisfaction with my (Block, 1980) attempted refutation of Nozick (1977). Wysocki is kind enough to credit me (and my friend and colleague Hoppe) with "eloquence" and for having "contributed most to the entire debate in question" and I am grateful to him for that compliment.

But then he offers his critique:

Block's account⁷ is doomed to conclude for it is inherently unable to square the two apparent facts: 1) that those units of butter are really (ex hypothesi) 'equally useful, desirable, serviceable' and 2) that 72nd one was 'picked up'. As we are about to see, the whole problem trades on the concept of 'picking up'. Block seems to be lured into thinking that the imagined actor does pick up the 72nd unit where he says: "For if the person didn't really prefer to give up this (72nd) one, why did he pick it to be given". Fair enough, if we assume that he did pick up this very unit, he must have preferred giving up this one to giving up any other, which simply logically follows from the concept of 'picking up' employed herein. And yet, why should we beg any questions? It is to be established first that the actor does indeed pick up the 72nd unit. For settling this issue has a bearing on whether he prefers giving this unit to any other or he does not. And this in turn determines whether the actor conceives of the 72nd unit as the unit of the same supply (with all the other units of butter) or he conceives of the stock before him as consisting of two distinct classes: a) a homogeneous class of 99 units (still intact) and b) a singleton containing the very pound of butter given up. Therefore, it

⁷ In Block (1980) I mentioned 100 pounds of butter, among and between which the owner was initially indifferent, but then for some reason had to give up one of these units. He chose the 72nd one.

seems that something has to give here: either 100 units were not in fact perceived as equally useful or they indeed were but the actor did not (in a relevant sense) choose to give up the 72nd unit. So, Block cannot have it both ways.

It seems to me that I can indeed “have it both ways.” Not, of course, necessarily, with the 72nd unit, but with any pound of butter. In the example, the owner of the butter is just sitting there, perhaps, contemplating his butter stock, and congratulating himself upon his ownership of it. If he is of a philosophical bent, and contemplates which of these butter units he likes best, he could readily admit that he sees them as a homogeneous blob, and is indeed indifferent between them all. But then comes a robber who tells him, at the point of a gun, to select one of these pounds of butter and give it to him. Or perhaps, he faces a customer, who wants to buy one of them. Now, the situation is going to hit the fan. The grocer must choose one of these one-pound packages, to give to the thief/customer. In my example, I chose the 72nd unit. But I am hardly wedded to that number. It could be any unit. But, in the event, based on the example, he must choose one of them. Let us say he chooses the first one, because he closes his eyes, and just happens upon that one. Well, now, he is no longer indifferent. Even Wysocki admits the truth of this when he says: “Fair enough, if we assume that he did pick up this very unit, he must have preferred giving up this one to giving up any other, which simply logically follows from the concept of ‘picking up’ employed herein.” I insist I can indeed “have it both ways.” There are two separate scenarios here: one, before the demand on the part of this other person for some butter, and the other after this takes place. In the first place there is

indifference, plenty of it.⁸ In the second place, there is no such thing. Wysocki in my view places insufficient attention to this two-stage situation. The two scenarios are very, very different.

Wysocki continues:

The problem with this contention is that Block must either invalidate his assumption that they were homogeneous before the choice in order to explain why the choice (i.e. picking up the least preferred unit of butter, as opposed to the remaining ones) took place. Alternatively, if he maintains that the units in question are indeed equally useful, then he cannot explain why this particular unit of butter was picked up because they were assumed to be equally valuable in the first place. Nozick's challenge comes with vengeance to Block and the reason is precisely that the latter author has a distorted idea of choice.

This author then touches upon various subjects: the essence of choice, the infinite varieties thereof, probability, numerous types of bread, many routes to super markets, the fact that I am "left with no hope of intelligibly conceiving of the same commodity," why the 72nd unit as opposed to any old "a" unit? My answer might be overly simplistic; let me try it on for size in any case. The choice of the 72nd unit was random. It was totally and completely arbitrary, used for illustration purposes only. I fear that Wysocki does not give full credence to the very different thoughts and behavior of the butter owner before and after he is called upon to give up one unit of butter. Before? Homogeneity, to be sure. After? Heterogeneity is the order of the day. Why this should be shocking, why impossible, why positing this to take place should be logically inconsistent, is beyond me. Time changes all sorts of things; why not this too?

⁸ Don't ask me how much. There is no such thing as a unit of indifference.

We then arrive at Wysocki's second interpretation of my response to Nozick; it is true and correct, but merely trivial: "On the second reading, Block's position may be rendered true but then it would amount to the mere restatement of the law of marginal utility." My first reaction to this is "Please don't help me so much," or "With friends like this, who needs (intellectual) enemies" (in this regard see Riley, 2016). On a more serious note, I fail to see how my distinction between before and after human action amounts, merely, to restating the law of diminishing marginal utility. It cannot be denied, of course, that with the 72nd unit of butter out of the way, all the others are now of greater value, but this hardly is what my refutation of Nozick is all about.

It is more than passing curious that in the next section of his paper, where Wysocki claims that Hoppe's refutation of Nozick is vastly superior to my own feeble attempt, that we read this incisive and correct statement of his: "[...] indifference cannot be demonstrated in action, as is usually reiterated by Austrians (Block, 2009a; Block and Barnett II, 2010; Rothbard, 2011). By no means can we deduce from any actual choice whether we were confronted with the units of the same good or with the ones of distinct goods." But this is precisely the point I employed to launch my critique of Nozick, to which Wysocki gives the back of his hand.

In section 5 ("Hoppe's account as a remedy for Block's shortcomings") of his paper, Wysocki will demonstrate that while my disproof of Nozick's criticism of Austrian economics fails, Hoppe's succeeds. Given that this is his goal, he gets off on the wrong foot. He opens with this statement:

First and foremost, it must be noted that—unlike Block's solution—[...] (Hoppe's) [...] involves both doing justice to indifference (at least admitting that a man can be genuinely

indifferent between some options) and barring it steadfastly from the realm of choice. Briefly speaking, Hoppe (2005) maintains that one cannot make a choice under indifference.

But I too, do this. There is not a dime's worth of difference between Hoppe and me on this issue. We both acknowledge that choice, or human action, is incompatible with a state of indifference! We both, Hoppe and I, agree with, in Wysocki's words: "the truth of the proposition that a man cannot choose when indifferent."

Wysocki's account gets curiouser and curiouser. He continues: "Specifically, Hoppe defines choice in such a way that it entails the lack of indifference. That is, if man chooses x over y, he is not (and, logically speaking, cannot) indifferent between the two." Well, Hoppe is not the only one to point to this insight. That is precisely my position, too.

However, Wysocki does not continue down this path. Instead of failing to see that Hoppe and I are on the same page, he now, correctly, notes a deviation between the two of us, and error number two on his part, supports Hoppe's incorrect position; writes our author:

Likewise, a mother who sees her equally loved sons Peter and Paul drown and who can only rescue one does not demonstrate that she loves Peter more than Paul if she rescues the former. Instead, she demonstrates that she prefers a (one) rescued child to none. On the other hand, if the correct (preferred) description is that she rescued Peter, then she was not indifferent as regards her sons.

No, no, no. While it is of course true that this poor mother⁹ "prefers one rescued child to none" she also places a higher value on Peter than Paul. She did rescue the former, when she could have chosen

⁹ The movie "Sophie's choice" subjects a poor woman to this excruciating choice.

differently, and selected the latter for retrieval, did she not? I simply cannot understand how it can be denied that she favored Peter over Paul, in the face of her action in rescuing Peter, not Paul.

Wysocki's account violates the spirit of praxeology. Now, it is quite alright to do so, at least in terms of congruence with mainstream economics. Our author is here on firm ground in this regard. My argument is that both Wysocki and most economists are in error here. I simply do not see how it is possible to deny that since the mother chose Peter, not Paul, she was not indifferent between them. Even if she did this with her eyes closed, and just grabbed the nearest son, this I think is the correct view.

Suppose she was offered the choice between a chicken dinner and a fish dinner. And, she selects the former. If Wysocki and Hoppe stick to their "logic" they would of course be correct in concluding that the woman preferred to eat dinner rather than go hungry. But they would be logically precluded from, also, maintaining that she preferred chicken to fish. The two of them, on the one hand, and I, on the other, live on different planets. Or, perhaps Mises was incorrect in his castigation of polylogism; there are two different "logics" at work in this case, mine and theirs.¹⁰

Wysocki misconstrues Buridan's ass in the same manner. This beast, let us say, chooses the bale of hay to the right. The correct interpretation of this is two fold: one, this creature preferred life to death, and, two, he favored the hay on the right to the hay on the left. In Wysocki's correct interpretation of Hoppe, and his own, only the first is true. The second, amazingly, is not. But, but, but, the donkey

¹⁰ I am of course joking about this. Mises was correct, as am I in this case. The same cannot be said for Hoppe and Wysocki.

moved to his right, not his left! If this is not evidence that he preferred the right to the left bale, there can be no such thing as evidence, at least not in cases like this.

At least Wysocki is logically consistent. Let us give credit where it is due: he makes the same identical mistake when it comes to the butter example. He writes:

[...] since the actor did indeed give up the 72nd pound of butter (while holding all of them equally serviceable), he must have preferred giving up a unit of butter for some pecuniary equivalent. In other words, the actor preferred one unit of butter less, but some increment of money to retaining his entire stock of butter but depriving himself of an opportunity to earn this money. The second possibility is that the correct description of an action is that the actor really dispreferred that 72nd unit that he actually gave up. If so, that unit was not the same economic good as all the other units in the first place and therefore, trivially, the original supply of 100 units was heterogeneous.

Again, I say, no, no, no. Why is it so hard to realize that one and the same thing can be altogether different, given the passage of time? At time t_1 , before any choice was made, yes, all units of butter were “equally serviceable.” Their owner was indifferent between all of them. They were homogeneous as far as he was concerned.¹¹ But then, at time t_2 , when push came to shove, the grocer had a decision to make: he valued the money more than a unit of butter, any unit of this product, to be sure. In the event, he was compelled to choose one of them if he wanted to make the sale. If this does not establish that he valued this particular one, the 72nd unit, less than the others, then there is no such thing as choice, utility, economic theory, common sense.

¹¹ If he thought about it at all, which he almost certainly did not.

Wysocki characterizes the Peter and Paul example as “the celebrated Hoppean thought.” Sure.

Here we come to the very core of why our author favors Hoppe’s analysis vis a vis mine. Says Wysocki:

Block’s point against Hoppe would be decisive if the act of saving a particular child (under this description) instead of another were inherently preference-implying. That is, Block would succeed if we can infer from the fact of saving a particular child (or from bringing about the event of a particular child being saved) that this particular child was preferred to the other. Yet, there is a deep distinction favored by Davidson (2001) between what an actor does (including his primitive action consisting in his bodily movements up to everything they cause) and what he does intentionally. As Davidson (2001, p.45) put it: “[. . .] although intentionality implies agency, the converse does not hold”. Therefore, it would simply beg the question to say that by the act of saving Peter the mother demonstrated her preference for Peter over Paul. As established above by alluding to the Davidsonian insight, from the event that the mother authored, we cannot infer which aspects thereof were informed by her preference. Therefore, not to beg any questions, we should treat the act of saving Peter in the non-choice- (and hence also non-preference) -implying sense. Alternatively, just to remain neutral on whether the mother did actually choose to save Peter or chose to save a child, we could say that what the mother in fact did was to save Peter. After all, to say that the mother saved Peter is only to attribute her agency to this event (in other words, it is to say that she authored the event of Peter having been saved), which does not imply that she saved Peter intentionally. And this is the key insight which, in our view, counts in favor of the Hoppean account.

I have no problem with what Davidson writes. My difficulty concerns the manner in which Wysocki utilizes that insight in behalf of his own views and those of Hoppe. Why? Davidson's perception is a matter of psychology, not praxeology. But we are now in the latter realm, not the former. Remember, this is Austrian economics we three are defending against Nozick's critique. As psychologists, we are undoubtedly interested in whether something is intentional or not. But I insist, this is not the case in our role as praxeologists. Suppose someone shoves two plates of ice cream at you and demands you choose one. You are very busy with something else. You reach out and grab one. You are not aware of which flavor you now have at your disposal. You may be so concerned with these other matters that the actual taste of this treat does not even register with you. As psychologists, there is no objection, at least not stemming from the present quarter, in saying that you were totally unaware of what you were doing, at least ice cream wise. The other matter took all of your intention. But, in sharp contrast, we as praxeologists must note that you actually reached out and grabbed one of them, not the other. This is the essence of Hoppe's error, with support from Wysocki. That mother reached out and saved Peter not Paul. What might well have been on her mind had nothing to do with Peter nor Paul. It might well have been as Hoppe opined, she was just preferring to save one of her sons, rather than none. Who knows, she might have been thinking about ice cream, as far as we praxeologists are concerned. This does not matter in the slightest for the praxeologist. We see her grabbing Peter, not Paul, to safety, and we are compelled by praxis logic, e.g., praxeology, to note that she was not indifferent between her sons, she could not have been indifferent between them, given that she chose the

one, not the other. In not seeing this, in maintaining the very opposite, Hoppe and Wysocki are guilty of what Rothbard (2011) labeled the error of psychologizing:

Psychologizing is a common error in utility analysis. It is based on the assumption that utility analysis is a kind of ‘psychology,’ and that, therefore, economics must enter into psychological analysis in laying the foundations of its theoretical structure. Praxeology, the basis of economic theory, differs from psychology, however. Psychology analyzes the how and the why of people forming values. It treats the concrete content of ends and values. Economics, on the other hand, rests simply on the assumption of the existence of ends, and then deduces its valid theory from such a general assumption. It therefore has nothing to do with the content of ends or with the internal operations of the mind of the acting man.

Consider an additional illustrative example. A man chooses women who do not respect him. He has a long trail of such psychologically unsatisfying and debilitating relationships. He knows he should select a different kind of women the next time he is free to do so. But when the occasion again arises, he conforms to the same old sick pattern. A psychologist might analyze this series of events by saying that the man full well knows what he is doing is counter-productive, and that he really wants a better relationship with a nice woman. But the Austrian praxeologist is precluded from making any such determination. He is required to look, only, at human action, actual behavior, and conclude that the man really prefers that type of woman.

Like a good psychologist, Wysocki places great weight on “what the mother intended to do” regarding her two sons. But as praxeologists, we are forbidden to even take this into account to the slightest

degree. Rather, we are (logically) required to ignore that entirely, and focus, instead, on the decision she actually made. She was *not* indifferent between her two sons, as matters unfolded. What went through her mind at the actual point of choosing is entirely irrelevant. We praxeologists must focus, only, on the decision she actually made.

Nor will Parfit's (2011, p.289) analysis save Wysocki's nor Hoppe's bacon. This philosopher points to

Two merchants [...] who [...] may both act on the policy 'Never cheat my customers'. But these merchants act on different maxims if one of them never cheats his customers because he believes this to be his duty, while the other's motive is to preserve his reputation and his profits.

This is indeed, again, grist for the mill of the psychologist; but for the praxeologist it makes a no never mind. We praxeologists could not be less interested in motivation. For us, the only, sole, issue is, what did the merchants actually do? They did exactly the same thing here. End of praxeological story!

Opines Wysocki: "Finally, let us note that Nozick's challenge leaves the Hoppean position unscathed." Not so, not so, say I. Wysocki continues: "when two—economically identical—goods are at stake, our goal (maxim) is satisfied to the same degree regardless of whether one good or the other is employed." I repeat myself, not so, not so. Hoppe does not lay a glove on Nozick in this instance.

Why did the grocer choose the 72nd unit? He could have picked any of the other 99 pounds of butter. But he chose that one. That necessarily, logically, implies that at least at the moment of choice,¹² he preferred to jettison that one, not any of the others. He was indeed

¹² Not before, of course, when he was indifferent to the entire lot of them.

indifferent between them at time t_1 , but not at all at time t_2 . In the latter epoch, he demonstrated preference, not indifference, *a la* Hoppe and Wysocki.

So far in this paper I have been highly critical of Wysocki.¹³ I have stressed the difference between the two of us. However, perhaps, there is hope for a reconciliation, at least partially. For in his section 6 “Extending the Hoppean framework: stating the law of diminishing marginal utility,” we seem to be more in tandem with one another. He now acknowledges the importance of time; this is something I have been emphasizing ever since Block 1980.

Wysocki avers: “that a given stock of units may be considered by an economic actor as a supply of the same commodity only relative to a given moment. Strictly speaking, it is a matter of course that human action is sequential (in a temporal sense) by nature; yet, an actor at t_1 may envisage the way he is going to employ consecutive units at later times. This double time indexation—one standing for a given moment in which an actor envisages the employment of his successive means and the other standing for the actual time at which they are employed—is necessary.”

Perhaps, who knows, Wysocki may one day, give him time, come to the realization that one and the same stock at time t_1 may be totally homogeneous for a grocer,¹⁴ and yet, later on, when push comes to shove and a decision has to be made, this will cease, and one of these units, maybe even the 72nd one just for illustration purposes, may seem to be more expendable than the others, in which case indifference between all of them no longer holds.

Here is yet another inkling of a coming together of Wysocki and myself. He says of the “Hoppean account” that it “accommodates

¹³ And, *en passant*, of Hoppe.

¹⁴ Thus, he is indifferent between all of the units.

indifference and keeps it steadfastly from the realm of choice—very much in line with the demands of praxeology itself.” This, too, is very much congruent with what I have been saying since 1980.¹⁵

It is time to conclude this paper. I do so by quoting and responding to a statement made by a referee of *This Journal* to an earlier version of this paper:

- the author does not seem to provide a satisfactory explanation (or reference, perhaps) of why indifference breaks up at a moment of choice/action, especially if that action is not internally motivated but externally enforced; if it is enforced, it is not necessarily well reflected upon (especially under time-pressure), it could well be automatic and mean nothing for the agent’s individual assessment of the value of the thing chosen;
- what is also missing is some explanation of why passing of time in itself is enough for an agent to move from a state of indifference to a state of defined preference; from the examples used by Wysocki, and then commented on in this text, it seems that the apparent ‘preference’ is enforced by external (and sometimes extreme) circumstances, e.g. the rescuing of one son only; what is puzzling here is the lack of interest in the role of intentionality in the act of ‘picking’ as ‘choosing’ (self-motivated) which is missing in an act of ‘picking’ under someone’s enforcement; there seems to be a marked difference between internally and externally motivated choice; the author explains towards the end of the paper that intentionality (sic) or motivation does not interest praxeologists who are merely concerned with the end result of a decision, but that seems to be the core of the misunderstanding between Wysocki and the author of the rejoinder.

I thank this referee for giving me this opportunity to further clarify.

Indifference breaks up at a moment of choice/action because that is the only precise time when it logically cannot exist. We can all be indifferent between the dozens if not scores or even hundreds of

¹⁵ Other hopes for reconciliation; the two of us are several times co-authors on similar matters: (Wysocki and Block, 2018; 2019).

cans of Coca-Cola in the grocery store before selecting one of them. If not, then the word “indifference” has no meaning in the English language. But it clearly does. We all use this word upon occasion.¹⁶ But when push comes to shove, we select *this* can of coke, and *not* any of the others. How are we social scientists to account for this? Why, by saying that indifference is an accurate account for how we see large aspects of reality, with the exception of when choice/action takes place.

Does this apply, also, when the choice is made under duress? Yes, indeed, it does. It is certainly “externally enforced” when the mother can save only one of her children. It might well then be “automatic.” But this does not, cannot, change the primordial fact that even upon this occasion she chose this boy to save, and not the other. Can we really say, then, that she was indifferent between the two of them? Yes, of course we can: from a psychological point of view. She loves them both; equally. However, Nozick, Hoppe and Wysocki do not properly occupy the psychological or ordinary world. My debate with these three scholars has nothing whatsoever to do with that realm. Rather, this discussion is a matter of praxeology, or technical economics. My major criticism of these three in this debate is that they fail to make this distinction, and stick to it no matter what.

The “passing of time” is crucially important since it demarcates between the situation in which choice, or human action in Mises’ terminology does not occur, and when, later on, it does take place. When looking out upon a horde of coke cans, it is easy to be “indifferent”

¹⁶ Economics is not the only discipline to utilize a technical language for ordinary words. In physics, “work” means something very different than how this word is ordinarily employed.

between them all. Not so, later on, when one has to grab one. But which *one*? When one is selected, and the other is put aside, there can be no room for indifference in the analysis.

“Intentionality” is a psychological concept, not an (Austrian) economic one. We as economists are not privy to the internal thoughts of other people. We can only observe their “human action,” their actual behavior, the choices they make. We see someone selecting *this* can of coke, not any of the others; we observe the mother saving *this* son, not the other one. What are we to say about this? That indifference is now occurring? Not at all. Rather, that there is *preference* taking place.

What of the claim that “praxeologists [...] are merely concerned with the end result of a decision.” This is not so. Rather, Austrian economists are concerned with the decision itself. As Mises (1998, p.97) wrote: Human “Action is an attempt to substitute a more satisfactory state of affairs for a less satisfactory one. . .” This is what choice is all about. There is simply no room in such behavior for indifference.

Bibliography

- Block, W.E., 1980. On robert nozick’s ‘on austrian methodology’. *Inquiry* [Online], 23(4), pp.397–444. <https://doi.org/10.1080/00201748008601918>.
- Block, W.E., 2009a. Rejoinder to Hoppe on Indifference. *The Quarterly Journal of Austrian Economics* [Online], 12(1), pp.52–59. Available at: <http://mises.org/journals/qjae/pdf/qjae12_1_4.pdf> [visited on 15 September 2022].
- Block, W.E., 2009b. Rejoinder to Machaj on Indifference. *New Perspectives on Political Economy*, 5(1), pp.65–71.

- Block, W.E., 2012. Response to Ben O'Neill on indifference. *Dialogue* [Online], (2), pp.76–93. Available at: <https://dlib.uni-svishtov.bg/bitstream/handle/10610/2380/DialogueBook2eng2012_76_93.pdf> [visited on 15 September 2022].
- Block, W.E. and Barnett II, W., 2010. Rejoinder to Hoppe on indifference, once again. *Reason Papers* [Online], 32, pp.141–154. Available at: <http://reasonpapers.com/pdf/32/rp_32_9.pdf> [visited on 15 September 2022].
- Davidson, D., 2001. Agency. *Essays on Actions and Events*. Second edition. Oxford: Clarendon Press, pp.43–61.
- Hoppe, H.-H., 2005. Must Austrians embrace indifference? *The Quarterly Journal of Austrian Economics* [Online], 8(4), pp.87–91. Available at: <<https://mises.org/library/must-austrians-embrace-indifference>> [visited on 15 September 2022].
- Hoppe, H.-H., 2009. Further Notes on Preference and Indifference: Rejoinder to Block. *The Quarterly Journal of Austrian Economics* [Online], 12(1), pp.60–64. Available at: <https://cdn.mises.org/qjae12_1_5.pdf> [visited on 15 September 2022].
- Machaj, M., 2007. A praxeological case for homogeneity and indifference. *New Perspectives on Political Economy*, 3(2), pp.231–238.
- von Mises, L., 1998. *Human Action: A Treatise on Economics*. Scholar's ed. Auburn AL: Ludwig von Mises Institute.
- Nozick, R., 1977. On Austrian Methodology. *Synthese* [Online], 36(3), pp.353–392. <https://doi.org/10.1007/bf00486025>.
- Parfit, D., 2011. *On What Matters. Vol. 1*. Oxford: Oxford University Press.
- Riley, J.L., 2016. *Please Stop Helping Us: How Liberals Make It Harder For Blacks To Succeed*. New York: Encounter Books.
- Rothbard, M.N., 2011. Toward a reconstruction of utility and welfare economics. *Economic Controversies* [Online]. Auburn AL: Ludwig von Mises Institute, pp.289–333. Available at: <<https://mises.org/library/toward-reconstruction-utility-and-welfare-economics-1>> [visited on 15 September 2022].
- Wysocki, I., 2021. The problem of indifference and homogeneity in Austrian economics: Nozick's challenge revisited. *Philosophical Problems in*

Science (Zagadnienia Filozoficzne w Nauce) [Online], (71), pp.9–44. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/554>> [visited on 24 January 2022].

Wysocki, I. and Block, W.E., 2018. An analysis of the supply curve: does it depict homogeneity among its constituent elements? Another rejoinder to Nozick. *MEST Journal* [Online], 6(1). Ed. by Z. Cekerevac, pp.132–143. <https://doi.org/10.12709/mest.06.06.01.14>.

Wysocki, I. and Block, W.E., 2019. Homogeneity, Heterogeneity, the Supply Curve, and Consumer Theory. *Quarterly Journal of Austrian Economics* [Online], 21(4), pp.398–416. <https://doi.org/10.35297/qja.010004>.

Stanisław Dunin–Borkowski and his views on Einstein’s special theory of relativity

Jacek Rodzeń

Jan Kochanowski University in Kielce

Abstract

The main purpose of this article is to discuss the views of the Jesuit Stanisław Dunin–Borkowski (1864–1934) about Albert Einstein’s theory of relativity. These days, Dunin–Borkowski is a rather obscure figure despite rising to fame in the interwar period as an outstanding expert in the philosophy of Baruch Spinoza. Thus, the secondary aim of this article is to remind ourselves of this somewhat forgotten scholar. As a researcher, writer, and pedagogue, Dunin–Borkowski was interested in numerous fields of knowledge. Among these were the natural sciences, including physics and the influence that new physical theories had on philosophical thought. This present study therefore fills a gap in the existing research about how Polish philosophers received Einstein’s theories. The example of Dunin–Borkowski also serves as a basis for discussing some of the fundamental problems of neo-scholasticism in receiving new mathematicised scientific theories.

Keywords

Stanisław Dunin-Borkowski, Spinoza, special theory of relativity, neo-scholastic philosophy, philosophy of nature, science and theology.

1. Introduction

In the history of contemporary philosophy, Stanisław Dunin-Borkowski (1864–1934) is rarely mentioned, and there is scant recognition of his writings. For example, two monumental works on the history of philosophy, one by F. Copleston and another by W. Tatarkiewicz merely mention him (Copleston, 1994, p.209; Tatarkiewicz, 1998, p.365). Moreover, the recently published voluminous *History of Polish Philosophy*, as edited by J. Skoczylski and J. Woleński (2010), contains not even a single word about Dunin-Borkowski. Only among the works of Jesuit biographers can one encounter descriptions of his life and academic achievements (cf. Siwek, 1935; Pummerer, 1935; Darowski, 2001, pp.103–104).¹

Stanisław Dunin-Borkowski deserves a secure place in the history of philosophy for at least one crucial reason, namely his seminal works on Baruch (Benedictus) Spinoza (1632–1677). Indeed, Dunin-Borkowski devoted four hefty volumes totaling around 2,000 pages to the life and works of this thinker from Amsterdam. In 1937, in his review of them, Richard P. McKeon (1900–1985), a preeminent American philosopher and historian who is regarded as a representative of *The Chicago School* of literary criticism, wrote the following in *The Journal of Philosophy*:

Data and explications important to the study of the seventeenth century and its intellectual background, bibliographies, recondite information concerning men, movements, and places, brief histories of concept, problems, and methods, frequently carried far beyond the immediate scope of Spinoza's usage or

¹ Dunin-Borkowski and the first part of his comprehensive work on Spinoza were mentioned at the 1911 meeting of the Philosophical Society in Cracow by Rev. Stefan Pawlicki (1912, pp.13–14).

his knowledge, make Dunin Borkowski's work indispensable in the study of Spinoza or of the seventeenth century (McKeon, 1937, p.381).

Dunin-Borkowski presented the results of his studies of Spinoza at prestigious international conferences. After one such event where he had delivered a paper on Spinoza's physics (Dunin-Borkowski, 1933), León Brunschvicg (1869–1944) referred to him as “le grand historien de Spinoza” (Brunschvicg, 1934, p.427).

However, Spinoza was not Dunin-Borkowski's only interest. He left a sizeable legacy in the form of scientific and popular science literature on pedagogy, the history of religion (e.g., Buddhism), theology, and philosophy, but he was also familiar with issues connected with the contemporary natural sciences, including physics, and their relation to philosophy and theology. Several of his semi-popular research works bear significant traces of this interest (cf. Dunin-Borkowski, 1898; 1911). Although Dunin-Borkowski was associated with Thomism (Dunin-Borkowski, 1921a,b; also Siwek, 1935, p.141; 1938, pp.306–307), aside from some several-page articles, he did not leave any thorough work that represented this thought movement. His association with Thomism, or more generally with neo-scholasticism, should come as no surprise, however, given that he belonged to the Society of Jesus.

This present article is devoted to just one aspect of Dunin-Borkowski's work, however, namely this Polish thinker's appearance as a commentator on the emergence of Albert Einstein's theory of relativity.² It is an episode that can be followed through two relatively short research works by Dunin-Borkowski (1921a,b),

² Dunin-Borkowski's comments concern only the special theory of relativity.

yet it appears these works have some relevance for broadening our perspective of how this physical theory was received in neo-scholastic circles in the early decades of the 20th century.

There is a moderate amount of literature on the early reception of relativity by neo-scholastics. In an article, S.L. ten Hagen (2020, pp.238–239) briefly discussed the views of some Belgian adherents (mainly in Louvain) along this line of thought, including, among others, the Jesuit H. Dopp, P. Drumaux and D. Nys. In a few paragraphs of his work, A.C. Flipse (2010, pp.1148–1149) introduced the attitudes of P. Hoenen, a Dutch Jesuit, toward relativity. T. Glick (1987, pp.240–242), in turn, discussed the views of neo-scholastics in Spain. Works have also been devoted to J. Maritain's views on Einstein's theory (Kłósak, 1980, pp.161–182; Wolak, 1991). The views of some Polish Catholic philosophers about this theory have also been discussed by P. Polak (2016).

Before we present Stanisław Dunin-Borkowski's views about the theory of relativity, let us first briefly discuss the life of this relatively unknown Polish philosopher, theologian, and pedagogue. Moreover, because he devoted so much of his activity as a philosopher and historian to the philosophy of Spinoza, we will focus a little more on this interest of the Polish Jesuit.

2. A count who became a Jesuit and a Spinoza specialist

Stanisław Dunin-Borkowski, who is known chiefly as Stanislaus von Dunin-Borkowski in the literature, was the son of Count Witold Dunin-Borkowski, an owner of landed property in the village of Winiczki near Lvov, and Countess Kazimiera, née Fredro. Stanisław's

grandfather, Aleksander, was an envoy to the Diet of Galicia and Lodomeria and a founder and president of the Society of Fine Arts Enthusiasts in Lvov. His grandmother, Henryka, was a head of the Lvov Educational Care Centre for the Blind. Suffused with the spirit of humanism and respect for art *sensu largo*, a family tradition of public activity left an undeniable imprint on Count Stanisław's character and his later interests.

Initially, Dunin-Borkowski's education was intended to prepare him to perform state administrative functions later in life, so the young count was sent to the famous *Theresianische Akademie* (*Theresianum*) in Vienna. This episode was short-lived, however, and Stanisław soon moved to the Jesuit school *Stella Matutina* in the Austrian town of Feldkirch. Sometime later, he entered the Jesuit novitiate before taking courses in classical languages and philosophy at the Jesuit study houses in the Netherlands. He completed a four-year theology course among the English community at Ditton Hall, which had been established in the 1870s by German Jesuits who had fled Bismarck's *Kulturkampf*. In 1889, between his philosophical and theological studies, Dunin-Borkowski began a pastoral and teaching job at *Stella Matutina*.

After taking holy orders in 1896, Dunin-Borkowski moved to the Jesuit writers' house in Limpertsberg (Lampertsbierg), a district of Luxembourg City. According to the account of Rev. Paweł Siwek, S.J., he there fell under the friendly guidance of Rev. Erich Wasmann, S.J. (1859–1931), and the young Polish count and Jesuit was “[...] trained for skillful use of the most effective weapon of today—the pen” (Siwek, 1935, p.137). Dunin-Borkowski was in daily contact with Wasmann, who was a famous Jesuit entomologist, Catholic polemicist,

and an advocate of the theistic interpretation of the theory of evolution, being dubbed “the father of the ants” (*Ameisenpater*) (for more on Wasmann’s activity see Baranzke, 1999; Polak, 2007).³

In 1920, Rev. Dunin–Borkowski, S.J. was appointed the spiritual director of the priestly monastery school in Wrocław, where he worked until 1931. He then moved to Koblenz and later Munich, where he died in 1934. The fruits of his activity can now only be found in his publications, because his unpublished manuscripts, notes, and other materials were destroyed when the Gestapo took control of the Munich Jesuit house in April 1941 (Stasiewski, 1959).

To this day, Dunin–Borkowski’s pastoral activity and pedagogical thought is best known in Germany and Austria. In the European and global dimensions, mainly in the interwar period, this Polish Jesuit was respected as an expert in the life and works of Baruch Spinoza (cf. Siwek, 1935, p.139). Dunin–Borkowski often tried to combine these areas of knowledge and practice, and he was convinced that “[...] pedagogical theory that did not include the philosophy and psychology of love (*die Philosophie und Psychologie der Liebe*) in the core of its content would never be able to find the road to pastoral deeds and success” (Dunin–Borkowski, 1926, p.43).

A rather obvious question arises, though: Why did such a well-educated Jesuit, educator, and theologian become interested in a Jewish–Dutch thinker who created an extremely complex philosophical system? There is no simple answer to this question, because none of his works unequivocally explains Dunin–Borkowski’s motivations for this interest. It is worth noting that he did not consider himself an advocate of Spinoza’s views. In an early article about this thinker from

³ For the subsequent years, Dunin–Borkowski’s and Wasmann’s publications would appear “side by side” in Jesuit journals *Stimmen aus Maria Laach* and *Stimmen der Zeit* (published by Herder).

Amsterdam, he declared that his mission was “[...] just to understand Spinoza, but not admire or condemn him” (Dunin-Borkowski, 1902, p.126).

On reading some of the remarks scattered among his works on Spinoza, one could conclude that Dunin-Borkowski was intrigued by the manifold interpretations of the oeuvre of this author, which included *Ethics* (Dunin-Borkowski, 1910, p.30). Moreover, interest in Spinoza was especially intense in the German cultural area in which Dunin-Borkowski lived and was active.⁴ Of the interpretations that developed in this area, two are particularly notable while also being essentially opposed to each other. The first reaches back to the 18th-century German Romantic movement (e.g., F.H. Jacobi, J.W. Goethe) and then to the great idealist systems of F.W.J. Schelling and G.W.F. Hegel. F. Copleston wrote that the German Romantics “found or thought they found in Spinoza a kindred soul” (Copleston, 1994, p.261). The second well-known pattern of thought that selected Spinoza as its guide was the monistic and materialistic movement of L. Büchner and K. Vogt (cf. Pawlicki, 1912).

It is hard to clearly establish which of the above two interpretations spurred Dunin-Borkowski into researching Spinoza’s work. In his first volume of analyses (dated 1910), references to Jacobi and Goethe are as numerous as references to the monistic concept of psychophysical parallelism. At the same time, Dunin-Borkowski appeared to share the beliefs held by the materialistically oriented monists who saw Spinoza as their patron:⁵

⁴ There was also an interest in Spinoza among Jewish-German circles in the Weimar Republic (see, for example, Wertheim, 2011).

⁵ It is also noteworthy that the references to the author of *Ethics* in the context of the discussion of the soul-body problem and the parallelistic concepts came to feature

You will only have half of Spinoza if you emphasize his anti-Christian monism, but omit his silent struggle against the pestering, purely materialistic tendencies that undermined morality and religion back then (Dunin-Borkowski, 1910, p.XVI).

It seems reasonable to assume that these prevalent anti-Christian interpretations of Spinoza to some extent motivated Dunin-Bukowski's studies of Spinoza, but it would go too far beyond the scope of this article to elaborate on this assumption. Anyway, in the face of the multiple interpretations of Spinoza's system, Dunin-Borkowski resolved to perform a historical and critical analysis of the available early editions of Spinoza's works, as well as manuscripts and other materials related to this Dutch thinker. Employing philological and historical methods, he set out to capture a proper sense of Spinoza's statements. He even sometimes tried to correct Spinoza's thoughts, believing that this would be what the author of *Ethics* would have done to maintain consistent adherence to the premises of his own system (Siwek, 1938, p.308). Indeed, the Polish Jesuit was perfectly prepared for such tasks. In addition to his fluency in Latin, he was also proficient in Dutch, among other things. For example, he had studied 17th-century philosophical concepts and the specificity of the Dutch culture of that era. In Dunin-Borkowski's opinion, it was only from such a research perspective that one could furnish explications that would enable a proper reading of Spinoza's work (Dunin-Borkowski, 1910, p.XI–XXIII).

quite prominently in the works by other Polish Jesuits active in the days of Dunin-Borkowski and beyond, e.g. Rev. Fryderyk Klimke (1878–1924) and Rev. Paweł Siwek (1893–1986) (cf. 1906, pp.6; 1911, pp.312–332; 2020, pp.1310–1314).

Having discussed Dunin-Borkowski's research interests, particularly his passion for analyzing the life and work of Spinoza, we will now focus on the Polish Jesuit's own views about the special theory of relativity proposed by Albert Einstein (1879–1955).⁶

3. The theory of relativity as merely a mathematical tool

In 1921, the Jesuit journal *Stimmen der Zeit* published an article of Dunin-Borkowski entitled “Neue philosophische Strömungen [New philosophical currents]” (1921b). In reality, he did not discuss any emergent thought currents but rather the philosophical repercussions associated with scientific ideas, such as the physical concept of force (in light of the classical concept of substance), the problem of the so-called psychophysical parallelism, and Einstein's special theory of relativity. Dunin-Borkowski paid most attention to the last of these matters. In his own words, he was interested in “a certain philosophical line of thought” that he believed was among “the most important tasks of the current philosophy of nature (*Naturphilosophie*).” This is why he chose to “take up several philosophical questions that this new theory [of relativity—J.R.] prompts” (1921b, p.211).

The special theory of relativity (STR) gives rise to conclusions about the contraction of object lengths and time dilation in moving systems (uniformly and rectilinearly) at high speeds (i.e., comparable with the speed of light) in relation to a selected system. Even

⁶ The convergence of Dunin-Borkowski's extensive studies on Spinoza and the widely known fascination on the part of Einstein with the life and works of this Dutch thinker, it seems, was entirely coincidental. There is no information either that the Polish Jesuit knew Einstein's interests in Spinoza, or that Einstein read Dunin-Borkowski's work *Der junge De Spinoza* (1910).

at the beginning of his commentary on these aspects of the STR, Dunin-Borkowski made a surprising statement: “[...] of course these differences cannot be proven experimentally” (1921b, p.211). For him, the conclusions that followed from the STR were “purely mathematical, drawn thanks to the application of the Lorentz equations” (1921b, p.212).

A similar opinion was expressed by the well-known Belgian neo-scholasticist Rev. Desiré Nys (1859–1927): “The theory of relativity appears to us a purely mathematical conception [...]. It would be reckless to regard the theory as a representation of reality” (1922, p.321). The view expressed by Rev. Konstantin Gutberlet (1837–1928)—a German theologian, philosopher, and friend of G. Cantor’s—went in a similar vein. In a 1913 work, he wrote that while the STR could be expressed in a form of “an even more adventurous fiction” (*noch abenteuerlichen Fiktion*) as a four-dimensional spacetime, “[...] this has nothing in common with reality” (1913, p.334).

Dunin-Borkowski was, however, inclined to admit that the mathematical results related to the STR could significantly explain (*erklären*) a number of physical and astronomical facts in a “simple and certain” way (1921b, p.212). Nevertheless, he believed achieving this kind of agreement provides explication (*Erklärung*) but not comprehension (*Erkenntnis*) of objective reality. Thus, the Polish thinker treated the STR as a mere mathematical tool and physical hypothesis.

Interpreting the STR solely as a mathematical tool and an experimentally unverifiable physical concept is reminiscent of the reaction by Catholic dignitaries and theologians to any questioning of the geocentric world system in the 16th and 17th centuries. Admittedly, as a mathematical theory and a physical hypothesis that afforded simple explanations and served certain practical purposes, Einstein’s “calculi” were accepted to some extent by authors like Dunin-Borkowski, just

like Copernicus' "hypotheses" in the past, but they were rejected outright as experimentally confirmed statements about reality (cf. Benk, 2000, p.121). An analogous situation involved the famous foreword (*Ad Lectorem*) that unbeknownst to Nicolaus Copernicus had been included in the first edition of *De revolutionibus*. In this, Andreas Osiander (1498–1552) tried to demonstrate that "these hypotheses need not be true nor even probable. On the contrary, if they provide a calculus consistent with the observations, that alone is enough" (Osiander, 1978, p.XVI). Similar advice was presented to Galileo by Cardinal Robert Bellarmine (1542–1621), where the motion of the planets in the solar system was presented as merely a hypothesis in order to avoid conflicting with Catholic doctrine and its closely related geocentric cosmology⁷ (on the various ways of understanding the concept of "hypothesis," from antiquity to Newton, see the work of Koyré, 1968, pp.23–52).

In addition, in light of the above seemingly surprising statement by the Polish Jesuit, whereby the conclusions that follow from the STR "cannot of course be proven experimentally," the situation becomes clearer. In this case, Dunin-Borkowski was not referring to a hypothesis that is empirically tested based on mathematical and natural methods but rather a hypothesis as a purely mathematical construct that reveals the existing knowledge (here: Lorentz's theory) in a new way, thus enabling calculations. A hypothesis understood this way cannot be confirmed (proven) experimentally, because in principle and in itself, it tells us nothing about reality. Such a mindset clearly shows that Dunin-Borkowski had problems understanding

⁷ The fact that the Copernican system was treated by Catholic (and Protestant) theologians as a hypothesis did not, however, stand in the way of the system being used as the foundation for the Gregorian calendar reform of 1582 (cf. Crombie, 1953, p.326).

the role of mathematics in modern natural sciences, and these likely resulted from his adoption of the current-specific Aristotelian concept of mathematics combined with the principle of *metabasis*.⁸

With regard to the motion of physical systems in relation to one another, Dunin–Borkowski wrote the following in his 1921 article:

But according to Einstein, it is not motion in itself (*die Bewegung an sich*) which exerts specific influence. The difference rather lies in the influence of purely **relative** motion—the only one he allows. [...] In other words, relative motion influences (*beeinflusst*) the length and time in exactly the same manner as every absolute does (*wie eine etwaige absolute*) [emphasis—Dunin–Borkowski] (Dunin–Borkowski, 1921b, p.212).

What is notable in this passage is the distinction between “purely relative motion” and “motion in itself,” as well as the words about the “influence” of relative motion on length. The latter statement, along with its association with the notion of something “absolute,” may imply the presence of some real force that changes the length of an object as it moves. *Nota bene*, the concept of the “influence of motion” appears very often in Dunin–Borkowski’s article.

The mathematical structure of the STR does not require the introduction of the above distinctions in relation to motion. The phrase “motion in itself” rather shows that Dunin–Borkowski imposed a conceptual pattern that was alien to the physical theory, and at the same time, he suggests an interpretation that attempts to impose a specific philosophical character on the theory. The phrase “the influence of

⁸ As formulated by Aristotle, this prohibited the use of methods and concepts from one area of knowledge (e.g., geometry) in another (e.g., the philosophy of nature). The areas of knowledge were classified by the Stagirite, with them having an immutable basis in an ontologically understood reality.

motion” for the change in length is essentially erroneous, because according to the STR, this length contraction is kinematic (cf. Heller, 1992, p.216; 1995, p.40).

For Dunin-Borkowski, the STR was nothing more than “a science of the system of relative motion” (1921b, p.213). Indeed, the Polish thinker wrote about “the unacceptable **objectification of pure relation** [emphasis—Dunin-Borkowski]” (1921b, pp.214–215).⁹ As he criticized such an approach, he concluded that what was being “[...] introduced is the concept of objective changes instead of only relative shifts (*relative Verschiebungen*)” (1921b, p.214).

Dunin-Borkowski by no means denied the fundamental possibility of acquiring objective knowledge about reality—he merely challenged the belief that such knowledge might be gleaned through the STR. “You mustn’t misunderstand Einstein,” wrote the Polish Jesuit, “he does not presuppose that the length of a rigid rod has become **objectively** shorter through motion [emphasis—Dunin-Borkowski]” (1921b, p.212). Nevertheless, Dunin-Borkowski levelled the following accusation at Einstein: “[...] he and some of his supporters [...] appear to be drawing philosophical conclusions that do not spring from [the STR—J.R.] by necessity” (ibidem), and “under closer scrutiny the surprising philosophical consequences¹⁰ do not take place” (1921b, p.215).

The Polish Jesuit disagreed with Einstein wherever he perceived the threat of “the possibility of absolute definitions (*Bestimmungen*)” (Dunin-Borkowski, 1921b, p.214) in relation to the material outside (i.e., objective) world. In particular, he favored the notion of “absolute

⁹ His opposition to the “objectification of pure relation” might have been caused by Dunin-Borkowski’s adoption of the Aristotelian concept of categories in which a relation is only one of the nine accidental categories in relation to the overriding category of substance/object.

¹⁰ Among which Dunin-Borkowski reckoned also the so-called twin paradox.

motion” (1921b, p.213). According to A. Benk, who briefly referred to Dunin–Borkowski’s views about the theory of relativity from the physical viewpoint, similar interpretative attempts can still, to some extent, be likened to the efforts of Hendrik A. Lorentz (1853–1928), the Dutch physicist who attempted to explain light propagation after the classical fashion. As he changed the frame of reference, apart from motion, he also transformed time, but he maintained the “true” or “absolute” time in relation to the local time in other frames of reference (Benk, 2000, p.123; cf. also Wróblewski, 2006, p.440). Only the STR made it clear that from the physical aspect, no inertial frame of reference is favored, so by extension, this new physics does not justify the need to refer to absolute motion or absolute time. Still, in his 1921 article, Dunin–Borkowski unpretentiously stated that unlike Einstein and the STR, he “did not see it necessary to be in favor of such conclusions” (Dunin–Borkowski, 1921b, p.214).

In the same issue of *Stimmen der Zeit* in which Dunin–Borkowski published his article, another Jesuit, Rev. Theodor Wulf (1868–1946), wrote a text on the STR, expressing that philosophy “certainly [teaches—J.R.] about the existence of absolute motion” (Wulf, 1921, p.115). This statement is interesting for at least one reason, namely that Wulf was an educated physicist who had studied under Walther Nernst (1864–1941) in Göttingen and authored some groundbreaking works on cosmic radiation. Neither Dunin–Borkowski nor Wulf provided a more explicit justification for their assertion that in terms of philosophy, one can speak of absolute movement. However, at least one explanation can be given for this attitude, such that both authors understood movement as a change that is, in the scholastic sense, the transition of being from a potentiality to an act. The change in this case is objective and absolute, so accordingly, we can speak of

absolute movements.¹¹ However, this is an ontological understanding of a movement expressed in the context of a specific philosophy. Such terms cannot be compared with those of physics, which are empirically operational, and one cannot judge physical theories based on a specific philosophical theory. This has unfortunately been attempted by many neo-scholastic thinkers in relation to Einstein's theories.

4. Dunin-Borkowski, Cardinal Bellarmine, and fictionalism

In the literature, Stanisław Dunin-Borkowski's works, aside from those devoted to Spinoza, are associated with the *sensu largo* neo-scholastic current, which chiefly refers to the intellectual legacy of Aristotle and Thomas Aquinas (Bruehl, 1921, p.305; Benk, 2000, pp.119–123). This current grew in strength in the 1880s following a number of initiatives by Catholic Church authorities (mainly Pope Leo XIII) aimed at recommending a revival in the studies and teachings of Aquinas' thoughts. A crucial element of the neo-scholastic program was opening the research of Catholic scholars up to the contemporary currents of philosophy and science, including the natural sciences. Apparently, the neo-scholastic current could include Dunin-Borkowski's work "Neue philosophische Strömungen,"¹² which more comprehensively addresses Einstein's special theory of relativity.

One might conclude from Dunin-Borkowski's arguments that to some extent, the Polish Jesuit accepted the special theory of relativity

¹¹ I would like to thank one of the reviewers for this suggestion.

¹² As well as Dunin-Borkowski's article (1921a), in which the author clearly confronts contemporary philosophical views and scientific concepts (and the theory of relativity, but only in one sentence) with the neo-scholastic viewpoint.

merely as a mathematicised hypothesis or theory (i.e., he writes about the application of Lorentz equations) that allows the “explication” (i.e., achieving an agreement between the conclusions drawn and the observations) of certain physical and astronomical facts. At the same time, he levelled criticism against “a certain philosophical line of thought” that in his opinion accompanied the special theory of relativity as a kind of interpretation.

Given Dunin–Borkowski’s article, it is difficult to establish what source materials about the theory of relativity were used by the author. The semi-popular character of the work probably explains why there are no referenced quotations. The form of discourse used resembles descriptions of the physical consequences of this theory (e.g., length contraction, time dilation, relativity of motion) that are frequently presented, albeit for popularization purposes, by physicists themselves in everyday language. However, the Polish thinker intersperses his descriptions with analyses involving philosophical terminology (e.g., “motion in itself,” “an objective absence of change,” “philosophical contradiction”). This seems unacceptable, especially in a situation where one is trying to render the proper physical meaning of a scientific theory’s consequences.

Any presentation of a physical theory like the special theory of relativity should, it seems, not ignore the properties of the mathematical apparatus in which it is being expressed or the methodological character of its references to physical reality, which serves as the testing ground for its reliability. Unfortunately, Dunin–Borkowski did not consider these basic elements of the cognitive character of mathematical physics. In his philosophical discourse, he engaged in polemics with paraphrases of popular (which, however, is not to be understood as erroneous!) descriptions of the consequences of the theory of relativity rather than referring directly to the mathematical

presentation and sense. (For more on this subject see (Heller, 1999, pp.101–102)). This is particularly evident in the Polish author's comparison of Einstein's alleged "speculations" about the relativity of time and space to, in his opinion, "similar ideas" by Jaime Balmes (1810–1848),¹³ a 19th-century Spanish Thomist (Dunin-Borkowski, 1921a, p.275). Being experienced in the milieu-related and literary analyses of Spinoza's works, Dunin-Borkowski probably did not fully realize that his skill set was not applicable to contemporary, mathematicised physical theories. His method for the exegesis of texts on theory rather than the exegesis of the structures of mathematical theories unfortunately meant that him and many other neo-scholastic thinkers were not in a place where they could have a responsible discussion about the philosophical aspects of new physical theories.

After reading this present paper, one may question whether Dunin-Borkowski's article should be placed within the interpretative tradition of factionalism, which was developed by Jesuits as early as the 16th and 17th centuries. According to William B. Ashworth, Jr. this tradition dates back to the works of Rev. Christoph Clavius, S.J., and its most prominent manifestation was the famous words of Cardinal R. Bellarmine, S.J. that were directed at Carmelite Fr. Paolo A. Foscarini, OCD in a 1615 letter:

I say that it seems to me that Your Paternity and Mr. Galileo are proceeding prudently by limiting yourselves to speaking suppositionally (*ex suppositione*) and not absolutely, as I have always believed that Copernicus spoke (cit. by Finocchiaro, 1989, p.67).

¹³ Balmes developed a completely non-physical and thoroughly philosophical concept of a relation in which time is a factor secondary to changes taking place in reality. In particular, the relation was concerned with changes in the sequence of "be" after "not be" (*non être*) (cf. Balmes, 1852). For more on this concept see (Wojciechowski, 1955, pp.678–683).

Both Clavius and Bellarmine expressed the opinion that all astronomical systems are concepts (hypotheses) designed only to “save phenomena,” with them having nothing in common with reality. Reality is therefore the exclusive domain of philosophers and not astronomers (Ashworth, 1986, p.158). The remarkable thing about the above quotation is the striking similarity between Bellarmine’s suggestion that Foscarini and Galileo treat the Copernican system as a hypothesis rather than an absolute formulation and Dunin–Borkowski’s accusation that Einstein treated the concepts of relative motion, relative time, and relative length as something absolute (i.e., as if they referred to reality).¹⁴

Stanisław Dunin–Borkowski’s articles are not isolated examples given the number of works by other neo-scholastic (neo-Thomistic) authors who grappled intellectually with Einstein’s theory of relativity in the interwar period (cf. Polak, 2016; and beyond the above-mentioned works Glick, 1987; Flipse, 2010; Hagen, 2020). Of course, criticism of, and even opposition to, the new concepts of physics (they were also concerned with quantum mechanics) did not exert any influence on the further course of its history.¹⁵ On the other hand, part of the writing activity from this Polish count, Jesuit, and pedagogue revealed the crucial systemic difficulties faced by representatives of

¹⁴ Interestingly, at the beginning of his article, as he briefly analyses the multitude of positions concerning the cognitive status of contemporary scientific concepts, Dunin–Borkowski repeatedly refers to Friedrich Lipsius (1873–1934), an advocate of Hans Vaihinger’s views. Vaihinger (1852–1933) created a thought current underpinned by contemporary positivism, and referred to as fictionalism (*Philosophie des Als Ob*), which states that, *inter alia*, mathematical and physical concepts are fictions.

¹⁵ Georges Lemaître can be considered an exception in this circle. Although he was thoroughly educated in mathematics, physics, and also neo-scholastic philosophy (i.a. at the Cardinal Mercier’s Higher Institute for Philosophy), he was able to take his thoughts above the narrow interpretations of Einstein’s theory and appreciate the importance of mathematics in understanding of nature (Hagen, 2020, pp.241–242).

the neo-scholastic movement when engaging with the methods and language of the mathematicised natural sciences.¹⁶ It should be added that Dunin-Borkowski, despite his Polish origin, was associated with a German-speaking (mainly Austrian) area of culture for most of his life, and he was largely shaped by this area. In any comparisons between his views and the works of other neo-scholastic authors, research should start with the relationship between his thoughts and the intellectual tendencies that were characteristic of circles in the mainly Austrian and German Catholic philosophy of that time.

Acknowledgments The author wishes to express his sincere gratitude to Paweł Polak and the anonymous reviewers for the immensely valuable remarks they made when this paper was being prepared.

Bibliography

- Ashworth Jr., W., 1986. Catholicism and Early Modern Science. In: R.L. Numbers and D.C. Lindberg, eds. *God and Nature: Historical Essays on the Encounter between Christianity and Science*. Berkeley; Los Angeles: University of California Press, pp.136–166.
- Balmes, J.L., 1852. *Philosophie fondamentale* (É. Manec, Trans.). Vol. 1. Paris: Auguste Vaton, Libraire-Éditeur.

¹⁶ This state of affairs concerned not only neo-scholastic thinkers who were active in the area of interwar Germany and Austria, but also numerous representatives of this current of thought in France, Poland, and even Belgium (along with the still living tradition of studies originated by Cardinal D.-J. Mercier). In Poland the situation took a qualitative turn only in the 1970s, which saw the beginning of collaboration between physicists, philosophers and theologians, most of whom were associated with the Cracow academic milieu (cf. Wolak, 1991; Polak and Rodzeń, 2021; Trombik, 2021).

- Baranzke, H., 1999. Erich Wasmann (29.5.1859–27.2.1931) – Jesuit und Zoologe in Personalunion. *Jahrbuch für Geschichte und Theorie der Biologie* [Online], 6, pp.77–140. Available at: <<http://www.moraltheologie.uni-wuppertal.de/wp-content/uploads/2015/01/Baranzke-Wasmann-JGTG-VI-1999.pdf>> [visited on 10 September 2022].
- Benk, A., 2000. *Moderne Physik und Theologie: Voraussetzungen und Perspektiven eines Dialogs*. Mainz: Matthias-Grünwald-Verlag.
- Bremer, J. and Poznański, J., 2020. Philosophy and Psychology in the Service of the Catholic Faith: Paweł Siwek, SJ and His Legacy. *Revista Portuguesa de Filosofia* [Online], 76(4), pp.1297–1330. Available at: <<https://www.jstor.org/stable/26986576>> [visited on 10 September 2022].
- Bruehl, C., 1921. Some Implications of the New Realism. *The Ecclesiastical Review*, 64, pp.300–305.
- Brunschvicg, L., 1934. Septimana Spinozana – Review of Acta Conventus æcumenici in memoriam Benedicti de Spinoza, diei natalis trecentiesimi, Hagae comitis habiti, curis Societatis Spinozanæ edita. *Revue Philosophique de la France et de l'Étranger* [Online], 118(11/12), pp.426–427. Available at: <<https://www.jstor.org/stable/41084115>> [visited on 10 September 2022].
- Copleston, F., 1994. *A History of Philosophy. Vol. 4, Modern Philosophy: from Descartes to Leibniz*. New York: Doubleday.
- Crombie, A., 1953. *Augustine to Galileo. The History of Science A.D. 400–1650*. Cambridge, MA: Harvard University Press.
- Darowski, R., 2001. *Filozofia Jezuitów w Polsce w XX wieku: próba syntezy - słownik autorów*. Kraków: Wyższa Szkoła Filozoficzno-Pedagogiczna Ignatianum; Wydawnictwo WAM.
- Dunin-Borkowski, S.v., 1898. Populärer Materialismus und Wissenschaft. *Stimmen aus Maria Laach*, 55, pp.486–506.
- Dunin-Borkowski, S.v., 1902. Leben und Lehre bei Bened. de Spinoza. *Stimmen aus Maria Laach*, 62, pp.121–133.
- Dunin-Borkowski, S.v., 1910. *Der junge De Spinoza. Leben und Werdegang im Lichte der Weltphilosophie*. Münster: Aschendorffschen Verlagsbuchhandlung.

- Dunin-Borkowski, S.v., 1911. Wissenschaft und Übernatur. *Stimmen aus Maria Laach*, 81, pp.477–486.
- Dunin-Borkowski, S.v., 1921a. Auf dem Weg zur Weltphilosophie. *Stimmen aus Maria Laach*, 100, pp.271–281.
- Dunin-Borkowski, S.v., 1921b. Neue philosophische Strömungen. *Stimmen aus Maria Laach*, 100, pp.209–221.
- Dunin-Borkowski, S.v., 1926. Autobiographie. In: E. Hahn, ed. *Die Pädagogik der Gegenwart in Selbstdarstellungen*. Vol. 1. Leipzig: Felix Meiner, p.43.
- Dunin-Borkowski, S.v., 1933. *Aus den Tagen Spinozas. Geschehnisse, Gestalten, Gedankenwelt*, vol. 1: *Das Entscheidungsjahr 1657*. Münster: Aschendorffschen Verlagsbuchhandlung.
- Finocchiaro, M.A., ed., 1989. *The Galileo Affair: A Documentary History*, *California Studies in the History of Science*. Berkeley; Los Angeles; London: University of California Press.
- Flipse, A., 2010. Between Neo-Thomist Natural Philosophy and Secular Science: Roman Catholic Scientists in the Netherlands. In: H. Hunger, F. Seebacher and G. Holzer, eds. *Styles of Thinking in Science and Technology: Proceedings of the 3rd International Conference of the European Society for the History of Science, Vienna, September 10-12, 2008*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, pp.1148–1149.
- Glick, T.F., 1987. Relativity in Spain. In: T. Glick, ed. *The Comparative Reception of Relativity* [Online], *Boston Studies in the Philosophy and History of Science*, 103. Dordrecht; Boston; Lancaster; Tokyo: D. Reidel Publishing Company, pp.231–263. <https://doi.org/10.1007/978-94-009-3875-5>.
- Gutberlet, C., 1913. Der Streit um die Relativitätstheorie. *Philosophisches Jahrbuch*, 26, pp.328–335.
- Hagen, S.L.t., 2020. The Local Versus the Global in the History of Relativity: The Case of Belgium. *Science in Context* [Online], 33(3), pp.227–250. <https://doi.org/10.1017/S0269889721000028>.

- Heller, M., 1992. Materia – geometria. In: *Zagadnienia filozoficzne współczesnej nauki: wstęp do filozofii przyrody*. 3rd ed. Warszawa: Akademia Teologii Katolickiej, pp.155–281.
- Heller, M., 1995. Henri Bergson i szczególna teoria względności. *Znak*, 47(6), pp.34–43.
- Heller, M., 1999. Physics and Philosophy in the 20th Century. *Organon*, 28-30, pp.95–110.
- Klimke, F.A., 1906. *Teorya parallelizmu psychofizycznego* [Online]. Vol. 4, *Broszury o chwili obecnej*. Kraków: nakł. Red. "Przeglądu Powszechnego"; (druk W. L. Anczyca i Spółki). Available at: <<https://rcin.org.pl/dlibra/doccontent?id=31373>> [visited on 28 January 2022].
- Klimke, F.A., 1911. *Der Monismus und seine philosophischen Grundlagen: Beiträge zu einer Kritik moderner Geistesströmungen* [Online]. Freiburg im Breisgau: Herdersche Verlagshandlung; (Buchdruckerei der Herderschen Verlagshandlung). Available at: <<https://rcin.org.pl/dlibra/doccontent?id=29490>> [visited on 28 January 2022].
- Kłósak, K., 1980. *Z teorii i metodologii filozofii przyrody*. Poznań: Księgarnia św. Wojciecha.
- Koyré, A., 1968. *Newtonian Studies*. Chicago: The University of Chicago Press.
- McKeon, R., 1937. A review of S. von Dunin-Borkowski's 'Aus den Tagen Spinozas. Geschehnisse, Gestalten, Gedankenwelt, vol. 3&4'. *The Journal of Philosophy*, 34(14), pp.380–383.
- Nys, D., 1922. *La notion d'espace* [Online]. Bruxelles: R. Sand; [etc., etc.]. Available at: <<http://archive.org/details/lanotiondespace00nysd>> [visited on 5 October 2021].
- Osiander, A., 1978. Foreword. *Nicolas Copernicus on the Revolutions, vol. 2* (E. Rosen and J. Dobrzycki, Trans.). Warsaw: Polish Scientific Publishers, p.XVI.
- Pawlicki, S.Z., 1912. *Spinoza i dzisiejszy monizm* [Online], *Wydawnictwa Towarzystwa Filozoficznego w Krakowie*, 5. Kraków: Towarzystwo Filozoficzne. Available at: <<http://rcin.org.pl/dlibra/doccontent?id=16416>>.
- Polak, P., 2007. Spór wokół teorii ewolucji przed stu laty. *Zagadnienia Filozoficzne w Nauce*, (41), pp.56–90.

- Polak, P., 2016. Zmagania polskich filozofów z ogólną teorią względności: przypadek neoscholastycznej recepcji teorii Einsteina przed II wojną światową. In: P. Polak and J. Mączka, eds. *Ogólna teoria względności a filozofia: sto lat interakcji*. Kraków: Copernicus Center Press, pp.29–64.
- Polak, P. and Rodzeń, J., 2021. The science-religion relationship in the academic debate in Poland, 1945-1998. *European Journal of Science and Theology*, 17(6), pp.1–17.
- Pummerer, A., 1935. P. Stanislaus von Dunin-Borkowski. *Mitteilungen aus den deutschen Provinzen der Gesellschaft Jesu*, 103, pp.549–573.
- Siwek, P., 1935. Stanisław Dunin Borkowski. *Przegląd Filozoficzny*, 38, pp.135–141.
- Siwek, P., 1938. Spinoza w świetle badań ks. St. Dunin-Borkowskiego. *Przegląd Filozoficzny*, 41(3), pp.305–314.
- Skoczyński, J. and Woleński, J., 2010. *Historia filozofii polskiej, Myśl Filozoficzna*. Kraków: Wydawnictwo WAM.
- Stasiewski, B., 1959. Dunin-Borkowski, Zbigniew Stanislaus Graf. *Neue Deutsche Biographie*, vol. 4. Dittel–Falck. Herausgegeben von der historischen Kommission bei der Bayerischen Akademie der Wissenschaften [Online]. Berlin: Duncker&Humblot, p.198. Available at: <<https://www.deutsche-biographie.de/gnd101302150.html>> [visited on 16 September 2022].
- Tatarkiewicz, W., 1998. *Historia filozofii. T. 3, Filozofia XIX wieku i współczesna*. 15th ed. Warszawa: Wydawnictwo Naukowe PWN.
- Trombik, K., 2021. *Koncepcje filozofii przyrody w Papieskiej Akademii Teologicznej w Krakowie w latach 1978-1993: studium historyczno-filozoficzne [Concepts of Philosophy of Nature at the Pontifical Academy of Theology in Krakow 1978-1993: A Historical and Philosophical Study]*. I. Kraków: Wydawnictwo „scriptum”.
- Wojciechowski, T., 1955. Teorie czasu scholastyczne a Einsteinowska. *Collectanea Theologica*, 26(4), pp.663–776.
- Wolak, Z., 1991. Filozofia przyrody Maritaina a szczególna teoria względności. *Zagadnienia Filozoficzne w Nauce*, (13), pp.33–42.

Wróblewski, A.K., 2006. *Historia fizyki od czasów najdawniejszych do współczesności*. 1st ed. Warszawa: Wydawnictwo Naukowe PWN.

Wulf, T., 1921. *Einsteins Relativitätstheorie*. Innsbruck [etc.]: Verlagsanstalt Tyrolia.

A critical analysis of the philosophical motivations and development of the concept of the field of rationality as a representation of the fundamental ontology of the physical reality

Wojciech P. Grygiel

Pontifical University of John Paul II in Krakow

Abstract

The unusual applicability of mathematics to the description of the physical reality still remains a major investigative task for philosophers, physicists, mathematicians and cognitive scientists. The presented article offers a critical analysis of the philosophical motivations and development of a major attempt to resolve this task put forward by two prominent Polish philosophers: Józef Życiński and Michał Heller. In order to explain this particular property of mathematics Życiński has first introduced the concept of the field of rationality together with the field of potentiality to be followed by Heller's formal field and the field of categories. It turns out that these concepts are fully intelligible once located within philosophical stances on the relations between mathematics and physical reality. It will be argued that in order to achieve more extended conceptual clarification of the precise meaning of the field of rationality, further advancements in the understanding of the nature of the human mind are required.

Keywords

ontology, mathematics, platonism, category theory, Roger Penrose, Alfred North Whitehead.

1. Introduction

The problem of the ontology implied by the contemporary physical theories ties up closely with the philosophical issues pertaining to the nature of their fundamental language, that is, mathematics. The nature of purely mathematical objects and structures as well of how they can be known has been subject of lively debates among philosophers and mathematicians since the times of antiquity and it continues to inspire both ontological and epistemological reflection until the present day (e.g. Shapiro, 2000). In order to provide philosophical justification of how the mathematical object exist and what the reasons for the great effectiveness of mathematics in the physical description of nature are, a prominent Polish philosopher of science, Józef Życiński, coined out in 1987 the concept of the *field of rationality* and its close correlate the *field of potentiality* (Życiński, 1987; 1988). His most direct motivation to introduce the field of rationality as constitutive to the ontology of the Universe flows from a careful analysis of the practice of science which shows that the generalization of the theoretical description of the physical reality effects a radical shift from concrete things to abstract mathematical structures.

A concept similar to that of the field of rationality has emerged from a different approach assumed by another prominent Polish philosopher of science, cosmologist theologian and a co-worker of Życiński, Michał Heller. Contrary to Życiński, however, Heller ven-

tures out in his inquiry not from the nature of the physical world but from the nature of mathematics itself to arrive at the concept of *the formal field*. Consequently, he considers the field of rationality to be an ontological interpretation of the formal field or, more properly, of its subfield (Heller, 2014, p.442). Heller insists that the idea of the field of rationality is still very “fuzzy” and is in need of further elaboration. In order to provide the necessary insight, he turns to the category theory as one of the most abstract and general theories of the contemporary mathematics often considered as a firmer foundation of mathematics in comparison to set theory. Heller’s key claim in the effort of “unfuzzifying” the concept of the field of rationality is to equate it with the *field of categories* (Heller, 2014, p.453).

Despite of Heller’s quite sophisticated ways of unpacking and sharpening of the conceptual content of the field of rationality much remains to be done especially in view of the relations between the meaning of the field of rationality with the formal field. The main scope and motivation of this article is to bring in the desired clarity into the meaning of the field of rationality as well as all its derivatives mentioned above. The inquiry will proceed in four stages. Firstly, a detailed critical review of the justification of the concept of the field of rationality as proposed by Życiński will be carried out. Secondly, special attention will be devoted to some inconsistencies in how he handles the platonic doctrine of the ideal forms. In that perspective, the subsequent introduction of the concept of the *noumenal structure* will be addressed. Thirdly, the alternative approach launched by Heller leading to the introduction of the field of categories will be surveyed. Fourthly, it will be concluded that although much better clarification of the concept of the field of rationality and its derivatives and their

mutual relations has been achieved in the course of this study, further investigative efforts are still needed so that these concepts may achieve satisfactory account and clarity in their meaning.

2. The Birth of the Field

The development of the concept of the field of rationality as well as its derivatives is a rather lengthy process which was initiated by Życiński in 1985 and developed by him through a series of subsequent articles and books (Życiński, 1985; 1987; 1988; 1991; 1995; 2006; 2013b). Although Życiński's *post mortem* published platonic manifesto entitled *Świat matematyki i jej matematycznych cieni* (Życiński, 2011; 2013b) evidently aspires to give full expression to the concept of the field of rationality, many arguments presented in this manifesto are at most extended elaborations on what had already appeared in the preceding articles. He declares expressly that the ontological assumptions on which the concept of the formal field rests are based on the philosophy of Plato and Alfred North Whitehead (Życiński, 1987, p.171). These philosophers exert their marked influence on other areas of Życiński's philosophical reflection as well. Moreover, Życiński fell under considerable influence of a renowned British theoretical physicist, mathematician, philosopher and the 2020 laureate of the Nobel Prize in physics, Roger Penrose and especially by his famous work entitled *The Emperor's New Mind* (1989). Although heavily criticized by experts from a wide range of disciplines¹, Penrose made a substantial effort to justify his platonic view of reality in which mathematical objects and structures exist in the world of ideal platonic forms.

¹ For a review of this critics see (Grygiel and Hohol, 2009).

The term “field of rationality” was formally introduced by Życiński with all due detail in his 1987 work. However, an important conceptual prelude appeared in his incisive treatment of the method of theology launched yet in 1985 in the first volume of the work entitled *Teizm i filozofia analityczna* (Życiński, 1985, pp.187–207). In this prelude, Życiński takes up the issue of the *ontic* and *epistemic rationality* of nature with special emphasis on the mathematicity of nature as a constitutive element of the ontic rationality. Consequently, Życiński’s main reason for introducing the concept of the field of rationality is to provide the precise philosophical explanation of why the language of mathematics developed independently of the investigations of nature applies so accurately to its physical description (Życiński, 1987). He argues that this applicability compels to assume that the fundamental mathematical structures are ontologically prior to their observable physical instantiations as well as to the development of mathematics itself revealing the existence of a great number of structures that do not find their use in the study of nature. In short, “the nature is mathematical because the level of the field of rationality is the fundamental level in its ontic structure” (Życiński, 1987, p.176) [author’s translation]. Also, Życiński highlights the fact that while the human psychological disposition favors the treatment of concrete things as the fundamental constituents of reality, the development of science forces radical departure from common sense perceptions of this kind towards abstract mathematical structures. These perceptions are conditioned by the phylogenetic conditions of the evolutionary development of the human cognitive apparatus proper to the level of reality at which this development had occurred.

While one can easily agree with Życiński that accepting the field of rationality as the fundamental level of reality implies “certain version of platonism” (Życiński, 1987, p.174), considerable amount

of further analysis will be necessary in order to establish to what degree this stance corresponds with the original ontological views of Plato. The central claim that Życiński builds into the concept of the field of rationality is that the abstract mathematical structures inherent in this field contain in themselves potentially all possible concrete objects and structures which can be actualized into existence at a later stage of the development of the Universe. He also asserts that the field of rationality imposes certain restrictions on the ontology of the Universe which is particularly evident in very specific kinds of symmetries that strictly define the form of the fundamental equations describing the dynamics of particles and interactions (Życiński, 1987, p.180).

As the primary example illustrating this claim Życiński refers to the theoretical description of the physical fields and in particular to the concept of the quantum vacuum which is a physically existing field with the lowest energy with no particles in it. The creation of a particle occurs when a creation operator acts on the wave function proper to the wave function of the vacuum thereby revealing obvious dependence of the reality of concrete physical particles on the mathematically abstract structure of the vacuum. Thus the designation of the field of rationality as the field of potentiality receives its proper explanation. This dependence in turn is intelligible only when the field of rationality represented by the vacuum is ontologically prior to concrete particles which are actualizations of the potentialities inherent in this vacuum through excitation to states of higher energy (Życiński, 1987, pp.176–178). In this context Życiński provides justification as to why this ontologically primitive realm of abstract mathematical structures should be called a field and not a matrix or a network. The

physical field serves as a suitable metaphor to illustrate the dynamic character of a field out of which new entities can emerge as opposed to the static and invariant status of matrices and networks.

Some general references to the philosophy of Plato or more generally understood platonism are made in Życiński's 1987 article. Nevertheless in the same essay Życiński devotes more attention to show the importance of the legacy of Alfred North Whitehead (Życiński, 1987, pp.181–185). He states that although Whitehead does not explicitly mention this concept in his writings, his main tenets in the interpretation of the mathematical character of nature coincide what is implied by the field of rationality. In particular, Życiński highlights that this interpretation bears markedly platonic character and what Whitehead has in mind is a certain matrix of abstract structures which constitute the most fundamental ontology of the Universe and from which all concrete physical objects may emerge. Lastly, Życiński points to Whitehead's theological interpretation of this matrix by associating it with a certain mode of the Divine presence in the Universe. This issue will receive its further clarification in Życiński's subsequent publications on the field of rationality.

It turns out that even on the contemporary scene Życiński is not at all alone in his insistence on the fundamental ontology of abstract mathematical structures. For instance, Penrose from whom he draws much inspiration in regards to the concept of the field of rationality promotes a unique belief that the complex numbers and abstract mathematical structures based on them bearing the name of holomorphic structures underlie the physical reality at its most fundamental level. Penrose goes as far as to assert that "we shall see something of the remarkable way in which complex numbers and holomorphic functions can exert their magic from behind the scenes" (Penrose, 2005, p.151). This magic is well evidenced in

the example of quantum mechanics which Penrose brings forth on numerous occasions. He takes every effort to emphasize that it is the complex character of the weightings in the linear combination of states composing an entangled state which is directly responsible for the quantum interference manifesting itself in the double slit experiment (Penrose, 1989, pp.236–242; 1997, pp.50–92; 2005, pp.553–559).

3. How Platonic is the Field?

A closer survey of Życiński's articles on the field of rationality following the original one discussed above (Życiński, 1987) reveals that in this article Życiński presented the majority of his key arguments to justify the meaningfulness of this concept. The articles following that of 1987 are mainly devoted to the search for the philosophical interpretation of the field of rationality in agreement with the declaration that a "certain version of platonism" must be sought for this purpose. It must be remembered, however, that in reference to the ontological status of the abstract mathematical formalisms of the contemporary physical theories the term platonism is used in a much broader sense as originally intended by Plato. The main difficulty consists in that the abstractness of the mathematical structures known today greatly exceeds the ancient mathematics of numbers and simple geometrical figures. As a result, the references to the original platonic thought are never straightforward and demand particular care in relating the meanings of concepts developed in quite different intellectual environments.

In article (Życiński, 1991) on the concept of field of rationality Życiński quite rightly locates its justification within the classical philosophical problem of the existence of abstract entities, that is, the

problem of universals in which as a radical form of realism platonism occupies an important position. In the following text on the field of rationality (Życiński, 1995), he evidently seeks further support for this concept by referring to the famous debate between two influential Polish philosophers Tadeusz Kotarbiński and Roman Ingarden on the fundamental ontology whether these are things (reism) or abstract objects or structures, respectively (Kotarbiński, 1920; Ingarden, 1972, pp.483–507). Unfortunately, Życiński addresses this debate in general terms only so that it is difficult to see how the argumentation for the field of rationality developed so far and merely restated here ties with the complexity of the debate in question.

As a direct reference to the works of Plato Życiński picks the interpretation of Plato's statement from *Phaedrus* 247 C offered by G.M.A. Grube in light of which Plato should be read as implying that these are ideas that dwell above the heavens. Grube's claim (1958, pp.30–35) that the statement must be taken metaphorically serves as a basis for Życiński to infer that the mathematical objects and structures contained in the field of rationality bear transcendent character in relation the realm of concrete physical objects because they exist beyond the spatio-temporal regime proper to what is termed as physical. Keeping in mind that the openness of metaphors allows for a variety of interpretations, the interpretative path of the platonic thought assumed by Życiński shows inconsistencies with the basic tenets of Plato's doctrine on the ideal forms.

First of all, Życiński incorrectly places mathematical forms directly in the world of the platonic ideas. According to Plato, mathematics lies below the ideas in the hierarchy of being and, for instance, the idea of a number can never enter into any computation (Shapiro, 2000, pp.52–60). Next, (Życiński, 1991) makes a lot of effort to demonstrate the existence of radical gap between the abstract realm of the field

of rationality and the concrete physical objects. In addition to the aforementioned example of the theory of physical fields, Życiński proposes three new illustrations of this gap which include the Kepler laws, the DNA code and the algorithms. In particular, the example of algorithms attests to the inspirations that Życiński has drawn from the works of Penrose. For instance, in case of the Kepler's laws he suggests that although not physically instantiated before the birth of stars and galaxies, they "somehow existed in the structures of the early Universe" (Życiński, 1991, p.71). This means that despite of the radical separation on which Życiński insists, the abstract structures of the field of rationality may enter into causal relationships with other structures of the field to produce concrete physical objects. In other words, the fact that the abstract structures do not occupy the spatiotemporal realm does not imply that they must be relegated to the acausal and eternal realm of the platonic forms. However, the force with which Życiński suggests this separation does make an impression that such a relegation is indeed intended. This would be still consistent with the interpretational freedom of the proposed metaphorical reading of Plato and would bring Życiński closer to the thought of Penrose who asserts that "I might baulk at actually attempting to identify physical reality within the abstract reality of the Plato's world" (Penrose, 2005, p.1029). Interestingly enough, Penrose reveals a slightly different understanding of physicality as compared to Życiński because in his ontology of the three worlds he evidently designates as physical the entire realm of what is constitutive to the structure of the Universe being a subset of the platonic world of mathematical forms. Życiński's use of the term "physical" only in reference to concrete spatiotemporal instantiations makes his reading of Plato all the more difficult.

Another serious concern that arises on the grounds of Życiński's justification of the concept of the field of rationality is the exact

meaning of the term “concrete object” or “concrete thing”. Aside from standing in marked opposition to the abstract objects or structures and most likely occupying the spatiotemporal realm not much more can be asserted. For instance, when he states that the motion of stars follows the patterns determined by Kepler laws and he considers the stars to be concrete and the laws abstract, what are the stars made out of? There is no doubt that this is an extremely intuitive and imprecise account which does not seem to differ much from the pre-scientific understanding of matter as a chunk of stuff contained in a given volume. By doing this, Życiński evidently falls into contradiction with himself as on one hand he heavily criticizes the use of the intuitive common sense concepts in science and, on the other, he makes them a foundational concept in the process of defining the field of rationality. If the concrete structures and objects are actualized into existence from the field of rationality, from the point of view of the contemporary physical theory of matter it must be regarded as a highly complex combination of fields and their excitations representing elementary particles that the molecules of the stuff are made out of. This is often referred to as the problem of the emergence of the classical world from the quantum domain. For instance, it receives a detailed treatment in the works of Penrose who considers the reduction of the wave vector as a real physical process induced by gravitational interactions (e.g. Penrose, 2005, pp.816–868). Unfortunately, Życiński does not mention these issues in any of his works.

The interpretative difficulties associated with Życiński’s attempt to locate the field of rationality in the original thought of Plato find at least partial solution in two of his final works on the subject published in 2005 and *post mortem* in 2013. As Heller rightly points out, Życiński has eventually abandoned platonic metaphysical view of the field of rationality and switched to its ontological interpretation by calling

the field of rationality the *noumenal structure* of the Universe (Heller, 2014, p.442). Although Życiński still mentions Grube's metaphorical reading of Plato, he admits of the multiplicity of possible interpretations of Plato's thought and shifts his emphasis to different texts of Plato, namely *Parmenides 132 D* and *Philebus*, where the participation of the concrete physical objects in the abstract structures is clearly admitted (Życiński, 2006, p.58). Evidently Życiński begins to withdraw from his former stance of the radical separation between the two in favor of treating the field of rationality as the constitutive element of the fundamental ontology of the Universe. Hence comes the term "noumenal structure" which reflects close relationship of the field of rationality with the laws that govern the Universe (Życiński, 2006, pp.58–59). Życiński reaffirms his ontological approach to the field of rationality in his 2013 publication by asserting that the Kepler's laws had existed in the structures of the early Universe before they emerged with the formation of stars and galaxies (Życiński, 2013b, p.161). This remains in good agreement with the theological interpretation of the field of rationality in which he refers to this field as the immanence of Logos in the Universe (Życiński, 2006; 2013b, pp.170–172) and that the actualization of potentialities inherent in this field does not require a separate act of the Divine creation. It is worth mentioning that Życiński has also argued for the key role of the field of rationality in the origin and development of life in the Universe (Życiński, 2009).

In his introduction to Życiński's *The World of Mathematics and Its Material Shades* Heller suggests a novel interpretational approach to Plato's perspective on mathematics which in his opinion provides suitable philosophical setting for Życiński's ontological views expressed by the concept of the field of rationality (Życiński, 2013b, pp.5–15). This approach is marshaled by a Polish philosopher Bogdan Dembiński who maintains that in its late period the platonic School

was influenced by the Pythagoreans and led to the reshaping of the platonic understanding of the nature of mathematics (Dembiński, 2003; 2010; 2015; 2017; 2019). This change was prompted mainly by the disciples of Plato: Speusippus (410–339 BC), Xenocrates (396–314 BC) and Eudoxos (408–355 BC) who turned mathematics into the main topic of discussions in the academy. Ultimately, these discussions resulted in the belief that it is mathematics that constitutes the fundamental stuff of the Universe. While Speusippus assigned all the characteristics of the ideal numbers to the mathematical numbers, that is, the separate existence, eternity, unchangeability and objectivity (Dembiński, 2010, pp.109–138), Xenocrates turned mathematics into ontology (Dembiński, 2010, pp.139–170). According to Dembiński (2010, p.158), Xenocrates is rightly called the forerunner of the concept of the mathematicity of the Universe. Undoubtedly, his ontological stance provides the most consistent philosophical environment for the proper foundation of Życiński's mature understanding of the concept of the field of rationality.

4. Starting with mathematics

It turns out that the path to the concept of the field of rationality does not have to commence with the physical reality as is the case of Życiński. As a skilled mathematician, Heller approaches the field of rationality in abstraction to any physical instantiations by referring directly to the nature of mathematics itself (Heller, 1997, pp.216–238). Heller's focal point in this regard is the famous Gödel theorem which stipulates that if any axiomatic system rich enough to contain arithmetic is complete then it must be contradictory (e.g. Penrose, 1994, pp.66–116). The direct consequence of this theorem is that if one

selects a non-contradictory axiomatic system then in there will be theorems whose truth will not be provable within this system. To put things in short, provability cannot be equated with truth and mathematics cannot be reduced to an axiomatic system. Much rather one should think of axiomatisation as mere human means of capturing the complexity of all possible mathematical structures.

This is precisely the point where Heller clarifies his usage of the term “field”. In doing so, he wishes to purposely avoid referring to all these structures as a set of objects in the strict set-theoretic sense. What he rather has in mind is an ensemble of structures linked together with all possible paths of inference (Heller, 1997, pp.236–238). Why then a field? First of all, the concept of a field takes into account all relations between the structures and secondly, as Heller insists, this concept conveys the idea of potentiality, that is, the field contains not only already known mathematical structures both those who still await their discovery in the future. Since these are not only the structures but all possible relations of inference among them Heller proposes to qualify this field as the *formal field*. This field provides a good framework for the interpretation of the Gödel theorem because to construct an axiomatic system means to select a certain small area in the formal field which is always too small to deductively grasp all theorems lying within this system.

Interestingly enough, while Heller does not invoke his famous distinction between two kinds of mathematics in this context: mathematics with the “small m” that evidently stands for axiomatised mathematics as a human creative activity written down in academic books and mathematics with the “capital M” as the universe of objectively existing mathematical structures which the former purports to describe (e.g. Heller, 2010). Moreover, it is worth stressing that Heller leaves the meaning of the potentiality of a mathematical structure

somewhat vague as he merely asserts that “these are the structures that have not been discovered yet or structures that will never be discovered but are in some sense possible” (Heller, 1997, pp.236–237). Since the existence of any mathematical structure is guaranteed only by its inner non-contradiction, there are no reasons to think of a possible mathematical structure as non-actualized from the point of view of mathematics alone. Much rather the problem lies in their not being known yet.

Ultimately, however, Heller shifts his attention to the ontological application of the formal field to explain the effectiveness of mathematics in modeling the physical world (Heller, 1997, p.237). As he rightly notices, this demands the existence of an intimate connection between the formal field and the fabric of the Universe. On such an interpretation, however, Heller suggests replacing the term “formal field” with “field of rationality”. And this is precisely the point where Heller’s thought meets with the thought of Życiński. The concept of the field of rationality carries with it much more philosophical implications for it directly refers to the idea of the rationality of the Universe which occupies the central position in the works of both of these thinkers (e.g. Heller, 2006; Życiński, 2013a).

The concept of the field of rationality finds its further development in the thought of Heller as he turns his scientific interest to a highly abstract mathematical theory known as the *category theory* (Heller, 2014). He is aware that a mathematical theory that aspires to better capture the concept of the field of rationality must bear some marks of fundamentality for mathematics as a whole. In this regard he relies on the opinion of one of the authors of the category theory, Saunders Mac Lane who claims that this theory has foundational significance (Mac Lane, 1992). This goal of Heller finds its somewhat humorous expression as he sees the need to make the concept of the field of

rationality “less fuzzy”. Quite surprisingly, as the interpretational perspective for this task Heller selects the ontology in the sense Willard V.O. Quine “which does not aspire to establish what exists, but rather what a given theory or doctrine assumes there exists”. The import of this choice will become more evident in the critical conclusion of this study (Heller, 2014, p.442).

Since a rigorous presentation of the category theory cannot be offered in this article (for a suitable source see for example (Simmons, 2011)), several general points will suffice to illuminate the desired philosophical import. The category theory is not just another branch of mathematics like calculus, linear algebra or set theory, for instance. Instead, the theory sees these and other branches as separate categories whereby it provides an overview “from above” and reveals possible connections among them. A category is a collection of objects connected by means of arrows which are called *morphism*. The most important aspect of the application of the category theory to the study of the structure of the Universe is that each major physical theory such as quantum mechanics or general relativity supervenes on a certain area of mathematical discourse and, consequently, on a certain category. In other words, a separate category may be selected to represent a section of the field of rationality that constitutes a matrix for the functioning of a given region of physical reality. Moreover, since different logics govern different categories, different logics may apply to different physical theories, as it is seen in the case of quantum mechanics. Also, Heller gives several other arguments of more formal nature in support of matching the field of rationality with *the field of categories*. Since their complexity would require lengthy explanations reaching beyond the scope of this paper, only the hierarchical structure of the category theory is worth mentioning here. In other words, mathematics cannot be captured into one axiomatic system but

it should be looked upon as an ensemble of structures and structures of these structures, part of which may constitute the matrix of the functioning of the Universe.

5. Knowing the Field: A Critical Conclusion

As the inquiry into the philosophical motivations and development of the concept of the field of rationality nears its conclusion, it is fitting to voice some closing critical remarks as well as to indicate possible areas of further research. The suitable point of departure is the seemingly uncontroversial matching of the formal field with the field of rationality suggested both by Życiński and by Heller. With this match in force, two fundamental difficulties arise: (1) how to explain the nature of mathematics as a necessary and *a priori* knowledge which can be acquired entirely independently from the empirical method and (2) how to explicate the “excessive” status of mathematics evident in the predictive power of the laws of physics such as the Einstein field equation to contain information on phenomena not imagined by their authors. In short, this issue focuses on the central problem in philosophy of mathematics—how the mathematical structures exist and how they can be known. Both Życiński and Heller directly express their awareness of the problem of how the human mind gets to know mathematics and the solutions they propose center on the evolutionary scenarios responsible for the acquisition of the mathematical knowledge (Heller, 2010; Życiński, 2010). Of course, this is fully consistent with the ontological views of Xenocrates that they both adhere to in which mathematics is matched with ontology and

no ontologically distinct platonic world is explicitly proposed. On such reading, however, the two difficulties mentioned above remain without explanation.

Interestingly enough, Życiński mentions the alternative solution in which the acceptance of the ontologically distinct world of mathematics is proposed and which is boldly marshaled by Penrose. In other words, the formal field differs from the field of rationality and the field of rationality constitutes but a subsection of the formal field which governs the behavior of the Universe. While this swiftly explains the nature of mathematics, it generates two new problems: (1) how the human mind gains the quasi-mystic access to the platonic world of mathematical objects and structures and, most importantly, (2) it turns the mathematicity of the Universe into a mystery because it is entirely unclear how the contingent physical world should emerge out of the acausal and eternal realm of mathematical forms. Moreover, Życiński not only devotes a considerable portion of *The World of Mathematics and Its Material Shades* to the survey of the typical arguments brought forth in support of the mathematical platonism but declares his open opposition to the purely naturalistic metaphorical approach to mathematics promoted by Lakoff and Núñez (2000, see also Życiński, 2013b, pp.48–49). Paradoxically, this approach would better correspond with the ontology of the formal field which he confesses.

At this point one can get an impression that a lot of effort have been devoted into making the concept of the field of rationality “less fuzzy”, much fuzziness still remains. In a way this fuzziness is enhanced by Heller himself who by considering the field of categories from the point of view of the ontology of Quine eschews the fundamental question of what objects and structures are actualized into existence and what mechanisms are responsible for this process by reducing this question to what is warranted by the discourse of the

category theory only. Although the field of categories may indeed capture some of the hierarchical relations implied by the category theory, it works in this capacity at best as a model and does not exhaust the full complexity of how a given section of the field of rationality finds its application to a given area of physical reality. However, the dependence of the distinction between the formal field and the field of rationality on the assumed philosophical position, namely that of Plato or Xenocrates, as well as the difficulties implied by both of them reveal a deeper source of “fuzziness” of these concepts which at this point must remain a mystery. After all, it is Penrose himself who saw mystery behind his ontology of the three worlds and claimed that now it is the time to study the nature of the human mind to shed some new light on what so far seems to resist our intellectual insight (Brożek and Hohol, 2014).

Bibliography

- Brożek, B. and Hohol, M., 2014. *Umysł matematyczny*. Kraków: Copernicus Center Press.
- Dembiński, B., 2003. *Późna nauka Platona: związki ontologii i matematyki, Prace Naukowe Uniwersytetu Śląskiego w Katowicach*, nr 2143. Katowice: Wydaw. Uniwersytetu Śląskiego.
- Dembiński, B., 2010. *Późny Platon i Stara Akademia, Fundamenta: studia z historii filozofii*, t. 63. Kęty: Wydawnictwo Marek Derewiecki.
- Dembiński, B., 2015. O niektórych aspektach platońskiej filozofii matematyki. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (58), pp.45–61. Available at: <<http://zfn.edu.pl/index.php/zfn/article/view/7>>.

- Dembiński, B., 2017. U źródeł pytania o matematyczność świata: Pitagorejczycy, Platon i Arystoteles. In: P. Polak, J. Mączka and W.P. Grygiel, eds. *Oblicza filozofii w nauce: księga pamiątkowa z okazji 80. urodzin Michała Hellera*. Kraków: Copernicus Center Press, pp.11–27.
- Dembiński, B., 2019. The theory of ideas and Plato's philosophy of mathematics. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), pp.95–108. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/468>> [visited on 23 September 2019].
- Grube, G.M.A., 1958. *Plato's Thought* [Online]. Boston: Beacon Press. Available at: <<http://archive.org/details/platosthought>> [visited on 14 September 2022].
- Grygiel, W.P. and Hohol, M., 2009. Rogera Penrose'a kwantowanie umysłu. *Filozofia nauki*, 17(3(67)), pp.5–31.
- Heller, M., 2014. The field of rationality and category theory. In: M. Eckstein, M. Heller and S.J. Szybka, eds. *Mathematical Structures of the Universe* [Online]. Kraków: Copernicus Center Press, pp.441–457. Available at: <<http://adsabs.harvard.edu/abs/2014msu..book..441H>> [visited on 28 January 2020].
- Heller, M., 1997. *Uchwycić przemijanie*. Kraków: Znak.
- Heller, M., 2006. Czy świat jest matematyczny? *Filozofia i wszechświat: wybór pism*. Kraków: TAIWPN UNIVERSITAS, pp.48–57.
- Heller, M., 2010. Co to znaczy, że przyroda jest matematyczna? In: J. Życiński and M. Heller, eds. *Matematyczność przyrody*. Kraków: Wydawnictwo Petrus, pp.7–18.
- Ingarden, R., 1972. *Z teorii języka i filozoficznych podstaw logiki, Dzieła filozoficzne / Roman Ingarden*. Warszawa: PWN.
- Kotarbiński, T., 1920. Sprawa istnienia przedmiotów idealnych. *Przegląd Filozoficzny*, 23, pp.149–170.
- Lakoff, G. and Núñez, R.E., 2000. *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics Into Being*. New York: New York.
- Mac Lane, S., 1992. The protean character of mathematics. In: J. Echeverria, A. Ibarra and T. Mormann, eds. *The Space of Mathematics* [Online]. Berlin: De Gruyter, pp.3–13. <https://doi.org/10.1515/9783110870299.3>.

- Penrose, R., 1989. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford [etc.]: Oxford University Press.
- Penrose, R., 1994. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford; New York: Oxford University Press.
- Penrose, R., 1997. *The Large, the Small, and the Human Mind*. Ed. by M.S. Longair. Cambridge: Cambridge University Press.
- Penrose, R., 2005. *The Road to Reality: A Complete Guide to the Laws of the Universe*. 1st American ed. New York: Alfred A. Knopf.
- Shapiro, S., 2000. *Thinking About Mathematics: The Philosophy of Mathematics*. Oxford: Oxford University Press.
- Simmons, H., 2011. *An Introduction to Category Theory*. Cambridge [etc.]: Cambridge University Press.
- Życiński, J., 1985. *Teizm i filozofia analityczna. T. 1*. Kraków: Znak.
- Życiński, J., 1987. Filozoficzne aspekty matematyczności przyrody. In: M. Heller, A. Michalik and J. Życiński, eds. *Filozofować w kontekście nauki*. Kraków: Polskie Towarzystwo Teologiczne, pp.170–185.
- Życiński, J., 1988. *Teizm i filozofia analityczna. T. 2*. Kraków: Znak.
- Życiński, J., 1991. Poza granicami konkretności. Spór o powszechność w kontekście rozwoju nauki nowożytnej. In: M. Heller, W. Skoczny and J. Życiński, eds. *Spór o uniwersalia a nauka współczesna: symposium, Kraków 11–12 maja 1990*. Kraków: Ośrodek Badań Interdyscyplinarnych przy Wydziale Filozofii Papieskiej Akademii Teologicznej, pp.55–80.
- Życiński, J., 1995. Status przedmiotów idealnych a ontologia współczesnej fizyki. In: W. Stróżewski and A. Węgrzecki, eds. *W kręgu filozofii Romana Ingardena: materiały z konferencji naukowej, Kraków 1985: praca zbiorowa*. Warszawa; Kraków: Wydawnictwo Naukowe PWN, pp.97–110.
- Życiński, J., 2006. Pole potencjalności a ewolucja Wszechświata. In: J.A. Janik, ed. *Nauka, religia, dzieje: materia i forma, potencja i akt, czyn: XIII seminarium w serii seminariów z Castel Gandolfo, 25–27 września 2005, Lublin*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, pp.49–60.
- Życiński, J., 2009. *Wszechświat emergentny: Bóg w ewolucji przyrody, Filozofia Przyrody i Nauk Przyrodniczych, 4*. Lublin: Wydawnictwo KUL.

- Życiński, J., 2010. Jak rozumieć matematyczność przyrody? *Matematyczność przyrody*. Kraków: Wydawnictwo Petrus, pp.19–36.
- Życiński, J., 2011. *Świat matematyki i jej materialnych cieni: elementy platonizmu w podstawach matematyki*. Kraków: Konsorcjum Akademickie – Wydawnictwo WSE w Krakowie, WSliZ w Rzeszowie i WSZiA w Zamościu.
- Życiński, J., 2013a. *Granice racjonalności*. Kraków: Wydawnictwo Petrus.
- Życiński, J., 2013b. *Świat matematyki i jej materialnych cieni*. 2nd ed. Kraków: Copernicus Center Press.

Józefa Życińskiego zmagania z językiem o Bogu

Michał Heller

Centrum Kopernika Badań Interdyscyplinarnych

Joseph Życiński's struggle with the language about God

Abstract

In the two-volume work *Theism and the Analytical Philosophy* (1985; 1988a) Joseph Życiński took up the challenge of renewing Christian metaphysics so that it could appear as a full-fledged partner in the dialogue with other streams of contemporary philosophy. This renewal should use two sources: the methodological principles of analytic philosophy, especially its philosophy of language, and certain elements of Whitehead's process philosophy. This study presents a critical reconstruction of Życiński's arguments contained in the first two chapters of (1985), which are devoted to the problem of language. Main results of this part of Życiński's work are negative, that is, they refute the arguments and interpretations of those analytical philosophers who show the meaninglessness of the theistic language or try to assimilate it to other standard languages, depriving it of a reference to the transcendent reality.

How can a positive part of the Życiński program be developed? It seems that only by formulating specific problems in the field of philosophy of God, or even theology, and choosing the right linguistic

tools to drill down on a given problem and seek its solution. This is in line with Wittgenstein's concept of language games. Życiński tries to do this in (1988a).

Życiński turned out to be a precursor of nowadays increasingly developing analytical theology.

Keywords

philosophy of language, process philosophy, analytical theology, language game, Józef Życiński, Ludwig Wittgenstein, Alfred North Whitehead, Willard Van Orman Quine.

1. Wstęp

W twórczości Józefa Życińskiego trudno jest nakreślić ostrą linię graniczną pomiędzy artykułem naukowym, esejem i prasową polemiką. Gdy autora ponosi pisarski temperament (a zdarza się to często), elementy tych stylów mieszają się ze sobą, tworząc całość z daleka rozpoznawalną jako charakterystyczny dla niego „rodzaj literacki”. Życiński jest jednak także głębokim filozofem, choć nie zawsze łatwo jest spod jego stylu wydobyć oryginalną myśl. Zapewne z tego powodu jest on bardziej znany jako pisarz-polemista i działacz na polu społeczno-kościelnym niż jako filozof, a jego praktycznie jedyna monografia naukowa¹, dwutomowy *Teizm i filozofia analityczna* (Życiński, 1985; 1988a), przeszła właściwie niezauważona. Myślę, że

¹ Może obok *Struktury rewolucji metanaukowej* (Życiński, 2013). Oryginał w języku angielskim (Życiński, 1988b).

jesteśmy winni jego pamięci, by dzieło to „ocalić od zapomnienia”, tym bardziej, że kryją się w nim nowatorskie myśli, które nie tylko warto przypomnieć, ale także twórczo rozwijać.

Zamiarem Życińskiego, gdy pisał to dzieło, było – ni mniej, ni więcej – tylko odnowienie, czy uwspółcześnienie, metafizyki, na której mogłaby się oprzeć chrześcijańska filozofia i teologia. Odnowienie to, jego zdaniem, powinno wykorzystać dwa źródła: dorobek współczesnej filozofii analitycznej i pewne elementy filozofii procesu Whiteheada. Pierwsze z tych źródeł miałoby wprowadzić do myśli chrześcijańskiej element ścisłości i krytycyzmu (temu poświęcony jest pierwszy tom *Teizmu*); drugie wypełnić metodą analityczną niezbędną treścią ontologiczną. *Teizm i filozofia analityczna* zawiera dość szczegółowo zarysowany projekt i wizję całości. Niestety, Życiński nigdy bardziej systematycznie nie wrócił do tego tematu. Obowiązki biskupie pozwalały mu na pracę filozoficzną jedynie „z doskoku”.

Wobec ciężącego na nas obowiązku przypomnienia filozoficznego dorobku Józefa Życińskiego, cel niniejszego opracowania jest raczej skromny. Pragnę wziąć na warsztat jedynie dwa podrozdziały pierwszego rozdziału pierwszego tomu *Teizmu* (Życiński, 1985, s. 13–36), poświęcone problemowi języka, i dokonać rekonstrukcji przebiegu rozumowań (argumentacji) autora, a tym samym wyraźnie sformułować wnioski, do jakich jest uprawniony. Nie jest to zadanie trywialne, ponieważ i tu Życiński niekiedy daje się ponieść literackiej swadzie, a jego własne argumenty są wplatane w liczne odwoływania się do poglądów innych autorów i polemik między nimi. W efekcie czytelnik otrzymuje efektowny pejzaż różnych opinii oraz autorskich komentarzy i trzeba niemałego wysiłku i uwagi, by wyłowić nić logicznych powiązań, która prawdopodobnie autorowi wydawała się oczywistą.

2. Problem

Problem, z jakim Życiński pragnie się zmierzyć w tych dwu podrozdziałach, można ująć w postaci następującego pytania: Czy język teizmu może być sensowny? W związku z tym pytaniem nasuwają się dwie uwagi:

Pierwsza: z przebiegu dyskusji, jaką w dalszym ciągu prowadzi Życiński, wynika, że chodzi również o język metafizyki (nie tylko teizmu), ponieważ teizm – jak zobaczymy – jakąś metafizykę zakłada.

Druga: należy zapytać, o jaką sensowność chodzi. Odpowiedź na to pytanie nie jest trudna – taką, o jakiej mówi się we współczesnej filozofii języka. W pierwszej połowie lat osiemdziesiątych zeszłego stulecia, wtedy, gdy książka była pisana, filozofię języka kształtowała filozofia analityczna, wywodząca się z empiryzmu logicznego Koła Wiedeńskiego. Wpływy samego Koła Wiedeńskiego były wówczas znacznie mocniejsze niż obecnie. W kontekście tak rozumianej filozofii języka pytanie o sensowność języka teizmu (czy ogólniej – metafizyki, a w dalszej perspektywie również teologii) jest szczególnie doniosłe, gdyż bezpośrednie stosowanie standardowych kryteriów sensowności, wyznawanych przez empiryzm logiczny, do języka teizmu prowadzi wprost do odmówienia mu sensowności. Celem dzieła Życińskiego jest więc stworzenie takiego języka, który mógłby być językiem teizmu, ale równocześnie korzystałby w jak największym stopniu z narzędzi analizy językowej wypracowanych przez współczesną filozofię języka – stworzenie takiego języka, lub przynajmniej ukazania kierunku, w jakim należy zmierzać, by taki język stworzyć. Życińskiemu przyświeca więc cel apologetyczny – dostarczenie filozofii Boga takiego wyrazu językowego, w którym mogłaby się ona stać pełnoprawnym partnerem dialogu na scenie współczesnych kierunków filozoficznych.

3. Język dyskursu

Życiński rozpoczyna swoje analizy od wyróżnienia w nurcie analitycznym dwu głównych koncepcji języka:

1. Koncepcja empiryzmu logicznego – według niej istnieje „radikalne cięcie między naukowymi i sensownymi zdaniami, które można sprawdzić empirycznie i między pozbawionymi sensu czysto ekspresywnymi zdaniami poezji, metafizyki czy filozofii Boga” (Życiński, 1985, s. 13). Koncepcja ta nawiązuje do poglądów „pierwszego Wittgensteina”, wyrażonych w jego *Traktacie logiczno-filozoficznym*. Chcąc tworzyć sensowny język teizmu, nie można oczywiście podążać tą drogą.
2. Konkurencyjna w stosunku do poprzedniej, koncepcja „gier językowych”, pochodząca od „drugiego Wittgensteina”, głównie z jego *Dociekań filozoficznych*. Zdaniem Wittgensteina w języku należy wyróżnić strukturę powierzchniową i strukturę głęboką. Struktury powierzchniowej można nauczyć się ze słownika i podręcznika gramatyki danego języka. Strukturę głęboką można opanować jedynie „grając danym językiem”, czyli posługując się nim w codziennych sytuacjach. Każdy język ma swoją własną „grę językową”. Życiński w tym miejscu (ani gdzie indziej w tej książce) nie wyjaśnia dokładniej, na czym polega koncepcja gier językowych. Najwidoczniej zakłada, że czytelnik powinien to wiedzieć. Zaznacza tylko, swoim dziennikarskim stylem, że „w interpretacji tej za bezsensowne uznane zostały próby poszukiwania idealnego języka. Uznano je za nierealistyczne i bezcelowe, choćby z tej racji,

iż nawet w różnych działach fizyki możliwości empirycznej weryfikacji poszczególnych twierdzeń są bardzo zróżnicowane” (Życiński, 1985, s. 13).

Istnieje oczywiście wiele wariantów tych dwu koncepcji, niektóre z nich pojawiają się w dalszych rozważaniach, zwłaszcza przy referowaniu poglądów różnych autorów.

Już na tym etapie rozważań łatwo się domyślić, że myśl Życińskiego będzie zmierzać w kierunku języka teizmu jako właściwej mu gry językowej. Jednakże poprzestanie na tym stwierdzeniu oznaczałoby właściwie zaniechanie tego tematu i pozostawienie go w stanie, w jakim obecnie się znajduje, wzbogaconym tylko o formułę „gry językowe”. Posługując się tak ogólnikowym rozumieniem koncepcji gier językowych, można by uzasadniać dowolny żargon jako język jakiejś „nauki”. Życiński podkreśla: „W perspektywie takiej teorii języka wiary można by szukać usprawiedliwienia również dla irracjonalno-fideistycznych postaw wobec Absolutu [...]” (Życiński, 1985, s. 17). Należy więc dążyć do takiej koncepcji gry językowej, w której określone byłyby „specyficzne reguły sensowności dyskursu filozofii Boga i precyzowane warunki metodologiczne jego poprawności” (Życiński, 1985, s. 14).

Następuje teraz dłuższy wywód na temat, czy sam Wittgenstein uznawał istnienie Boga i ewentualnie, w jakim sensie. Swoim zwyczajem Życiński rozgrywa ten problem, przytaczając opinie różnych autorów. Faworyzuje przy tym opinie, które przychylają się do teistycznych interpretacji poglądów Wittgensteina. Oczywiście – dodajmy od siebie – to, czy Wittgenstein był teistą, czy nie, nie ma istotnego znaczenia dla interesującego nas problemu, to znaczy po-

szukiwania właściwego języka o Bogu. Nie znaczy to jednak, że taka dyskusja jest nie na miejscu. Podczas niej padają bowiem argumenty, które już takie znaczenie mają.

I tak z dyskusji tej wynika „zasadnicza kwestia, którą trzeba rozstrzygnąć w teorii gier językowych”. Można ją sformułować w postaci następującego pytania: „Czy poza pojętym szeroko językiem wiary, językiem, którego wyrazem są zarówno słowa modlitw, gesty kulturowe, jak i postawy etyczne, istnieje nieredukowalna do tych elementów rzeczywistość transcendentna opisywana w dyskursie religii?” (Życiński, 1985, s. 16). Takie sformułowanie „zasadniczej kwestii” wymaga komentarza. Analizując język, nie można wyjść poza język i stwierdzić, czy to, o czym język mówi, istnieje, czy nie. A więc nie można tego pytania rozstrzygnąć „w teorii gier językowych”². Można co najwyżej, wiedząc skądinąd, że dany przedmiot istnieje, *przyporządkować* mu pewne wyrażenie językowe. Jest to oczywiście wyjście poza język do problemu istnienia, a więc do problemu ontologicznego. Życiński jest tego świadomy, chociaż w jego wypowiedziach granica między ontologią a filozofią języka niekiedy się zaciera. Przykładem, w którym tej nieostrości nie ma, jest następujący cytat: „[...] trzeba odpowiedzieć na pytanie, do jakiej rzeczywistości pozajęzykowej odnosi się jej [danej teorii gier językowych] dyskurs i jakie *pozajęzykowe implikacje ontologiczne* można mu przyporządkować” (Życiński, 1985, s. 18)[podkreślenia Życińskiego].

To rozróżnienie warstwy językowej i pozajęzykowej nie przeszkadza jednak Życińskiemu mówić o „ontologicznych implikacjach teorii *language-game*” (Życiński, 1985, s. 18). Wobec niemożności wyjścia poza język, jak więc należy rozumieć „ontologiczne implikacje

² Język matematyki może bardzo precyzyjnie opisywać jakieś zjawisko fizyczne, na przykład kwanty pewnego pola, ale aby stwierdzić, czy kwanty te istnieją jako cząstki fizyczne, niezbędny jest eksperyment, a więc wyjście poza język.

języka”? Istnieje tylko jedna odpowiedź na to pytanie – może tu być mowa jedynie o ontologii w sensie Quine’a. Willard Van Orman Quine w słynnym eseju „O tym, co istnieje” (Quine, 1969b) rozpatrywał kwestię „ontologicznych zaangażowań języka”. Zaangażowania takie mają miejsce, gdy badając język, przeprowadzamy analizę nie po to, „by dowiedzieć się, co istnieje, lecz po to, by dowiedzieć się, co dana wypowiedź lub teoria – nasza, czy też sformułowana przez kogoś innego – *uznaje* za istniejące” (Quine, 1969b, s. 29). Innymi słowy, tego rodzaju ontologiczna analiza języka zmierza do tego, by wyodrębnić minimum bytów, założenie istnienia których jest niezbędne, by zagwarantować sensowność danej wypowiedzi. Na przykład wiem, że Pegaz nie istnieje³, ale jeżeli mówię coś o Pegazie i jeżeli moja wypowiedź ma mieć sens, to musi ona traktować Pegaza jakby istniał naprawdę. Dla niej (dla wypowiedzi), nie dla mnie, Pegaz istnieje. Moja wypowiedź jest zaangażowana w istnienie Pegaza.

W tym kontekście Życiński nie wspomina Quine’a, ale z całego przebiegu jego argumentacji jasno wynika, że od języka teizmu wymaga on czegoś więcej niż zaangażowania tego języka (w sensie Quine’a) w istnienie Boga, a więc przyjęcia istnienia Boga jedynie „na potrzeby tego języka”; chodzi mu o istnienie Boga jako rzeczywistości pozajęzykowej. Kilka stron dalej Życiński polemizuje z poglądem W. Zuurdeega, który utrzymywał, że „w filozofii analitycznej niemożliwe jest postawienie kwestii prawdziwości sformułowań tego języka [chodzi o język teizmu]. Co najwyżej można tylko stwierdzić, iż sformułowania te odnoszą się do swoistej ‘rzeczywistości’ uznawanej przez grupę posługującą się tym językiem” (Życiński, 1985, s. 28). Zauważmy, że jest tu mowa o „rzeczywistości” zrelatywizowanej do języka, jakim posługuje się dana grupa. Jest to więc w gruncie

³ Mam oczywiście na myśli Pegaza – mitycznego konia, a nie szpiegowskie urządzenie elektroniczne o tej nazwie, które na pewno istnieje (w sensie ontologicznym).

rzeczy ontologia w sensie Quine'a. „W ujęciu takim – pisze Życiński – uderza przede wszystkim *zrównanie statusu wszystkich przekonań* i nadanie im charakteru pozaracjonalnego” (Życiński, 1985, s. 28). Jest to wymowny dowód tego, że w dyskursie religijnym Życiński nie zadowala się ontologią w sensie Quine'a.

A więc, według niego, w rozważaniach o języku teizmu analizy językowe (w duchu filozofii analitycznej) należy uzupełnić o składową (autentycznie, nie w duchu Quine'a) metafizyczną. Chcąc jakoś usprawiedliwić ten istotny „dodatek”, Życiński polemizuje z autorami, którzy Wittgensteinowskie gry językowe, związane z językiem o Bogu, interpretują ściśle wedle reguł sensu logicznego pozytywizmu.

* * *

Wnioski, jakie Życiński wyprowadza z analiz, przeprowadzonych w tym podrozdziale, są oszczędne:

1. „[...] odmienne koncepcje teizmu można łączyć z Wittgensteinowską teorią gier językowych” (Życiński, 1985, s. 22). Ściślej byłoby: „z różnymi wersjami Wittgensteinowskich teorii gier językowych”.
2. Czynnikiem decydującym o wyborze którejś z koncepcji teizmu „jest podzielana koncepcja języka wiary oraz koncepcja metafizyki” (Życiński, 1985, s. 22). Tu konieczne jest następujące uściślenie.

W tym podrozdziale kilkakrotnie pojawia się określenie „język wiary”. Częściej jednak Życiński mówi po prostu o „języku teizmu” albo o „języku o Bogu”. Proponowałbym, aby uwypuklić to rozróżnienie:

z jednej strony „język dyskursu o Bogu” lub „język filozofii Boga” i z drugiej „język wiary”. Na początku następnego podrozdziału Życiński wyjaśnia:

Wyrażenie „język wiary” ma szerszy zakres niż wyrażenie „język filozofii Boga”. Jakkolwiek Bóg filozofii i Bóg wiary jest tym samym Bogiem, to jednak w pierwszym przypadku jest On poznawany wyłącznie za pośrednictwem inferencji logicznych. W przypadku drugim jawi się On jako Osoba, w stosunku do której możliwa jest nie tylko refleksja filozoficzna, lecz także ułatwiane przez prawdy objawione przeżycie fascynacji, miłości, czci, czy zależności egzystencjalnej (Życiński, 1985, s. 22).

W analizowanym podrozdziale Życiński mówi w zasadzie o języku filozofii Boga, choć określenie „język wiary” pojawia się kilkakrotnie i nie zawsze granica pomiędzy nimi jest wyraźna. Powyższy wniosek (2) należy odnieść do języka dyskursu o Bogu, a nie do języka wiary. Nie znaczy to, że podobnego wniosku nie można wyciągać również dla języka wiary, ale wydaje się, że w tym miejscu nie było to intencją autora.

Nasuwa się wreszcie uwaga co do sposobu, w jaki Życiński wykorzystał poglądy Wittgensteina. Zrobił to w sposób mocno wybiórczy, czerpiąc z nich tylko to, co stanowiło podatny materiał do budowania zrębów własnego systemu (taki jest niewątpliwie cel obydwu tomów *Teizmu*). Oczywiście, każdy autor ma prawo korzystać z dorobku innych myślicieli w sposób, jaki mu odpowiada, ale pewnego rodzaju lojalność wobec swojego „źródła” wymagałaby, by nie sprawiać wrażenia, że „źródło” we wszystkim zgadza się z autorem.

4. Język wiary

Drugi podrozdział nosi tytuł „Specyfika języka wiary”. Życiński omawia w nim i poddaje krytyce „współczesne dyskusje z zakresu teorii języka”, które „znalazły swe odbicie w nowych propozycjach dotyczących zasadniczej reinterpretacji języka wiary” (Życiński, 1985, s. 22–23). Zważywszy antyreligijne nastawienie większości ówczesnych filozofów analitycznych, nie dziwi fakt, iż na ogół propozycje te zmierzają do takich interpretacji, które by ukazywały bądź bezsensowność, bądź jedynie ekspresywność języka wiary.

Wśród tych prób są i takie, które usiłują tak przykroić język religijny, aby w jakimś stopniu spełniał on neopozytywistyczne kryteria. Środkiem do tego celu byłoby tego rodzaju przeformułowanie zasady weryfikacji, które by odnosiło się do eksperymentalnych przejawów religijności, na przykład do „specyficznego ujmowania świata innych ludzi i osobistego życia na podstawie przesłanek religijnych” (Życiński, 1985, s. 23), lub wręcz „redukcja języka wiary do języka obserwowanych zachowań” (Życiński, 1985, s. 25) (P. van Buren).

Niejako kontrprzykładem w stosunku do tego rodzaju redukcjonizmu jest, zdaniem Życińskiego, fideizm, przejawiający się w poglądach D.Z. Phillipsa, który będąc człowiekiem religijnym, utrzymywał, że „filozofia nie jest ani za, ani przeciw wierzeniom religijnym. Jej rola kończy się na wprowadzaniu porządku do gramatyki tych wierzeń” (Życiński, 1985, s. 26).

Zauważmy, że Życiński stosuje tu swoją ulubioną strategię: chcąc zneutralizować poglądy jakiegoś myśliciela (van Buren), przeciwstawia mu poglądy innego autora (Phillips), umiejętnie przy tym rozkłada akcenty w ten sposób, żeby opinia, której nie sprzyja, okazała się fałszywa lub przynajmniej nieatrakcyjna.

Nie wnikając w szczegóły (jest ich naprawdę dużo), dokonajmy pobieżnego zestawienia poglądów, z którymi Życiński polemizuje. Są to:

- Wspomniane wyżej poglądy van Burena: modyfikacja zasady weryfikacji, redukcja języka wiary do języka obserwowalnych zachowań.
- Koncepcja „weryfikacji eschatologicznej”, „według której empiryczna weryfikacja teizmu jest możliwa i dokonuje się po biologicznej śmierci” (Życiński, 1985, s. 27) (J. Hick).
- Wspomniane wyżej poglądy W. Zuurdeega, w zasadzie powtarzające klasyczne zarzuty neopozytywistów pod adresem języka wiary, związane z kryterium sensu.
- „Interpretacja języka wiary w kategoriach autorytetu i posłuszeństwa” (A. MacIntyre). „U podstaw tej koncepcji znajduje się założenie, iż dowolne wierzenia religijne usprawiedliwiamy jako całość, ‘odnosząc je do autorytetu’” (Życiński, 1985, s. 29).
- Redukowanie języka wiary do języka etyki (R.B. Braithwaite).
- „Próby redukowania funkcji języka wiary do ekspresji postawy czci, kultu i szacunku, opisu doświadczeń numinotycznych czy przeżywania *sacrum* jako wartości” (Życiński, 1985, s. 31).

Ogólne konkluzje, jakie Życiński wyciąga z tej obszernej dyskusji, są następujące:

1. „[...] trzeba definitywnie odrzucić możliwość zredukowania [języka wiary] do jego innych poziomów bez uwzględniania jego składowej ontycznej odnoszącej się do rzeczywistości transcendentnej” (Życiński, 1985, s. 32).

2. „Ten specyficzny wymiar [ontyczny] dyskursu teizmu sprawia, iż w języku wiary pojawiają się nowe uwarunkowania epistemologiczne. Zarówno status empiryczny, jak i kryteria racjonalności tego języka są częściowo różne od analogicznych cech innych języków” (Życiński, 1985, s. 32).
3. „Odmienność ta nie przesądza *a priori* niczego o wartości dyskursu teistycznego”, także „nie implikuje niemożności stosowania do teizmu zwykłych kryteriów racjonalności” (Życiński, 1985, s. 32).

Jak wiadomo, neopozytywiści uznawali tylko dwa rodzaje sensownych wypowiedzi: tautologie i zdania empiryczne. Co więcej, podział ten jest, według nich, dychotomiczny. Zdania, które nie są ani tautologiami, ani zdaniami empirycznymi, tworzą zbiór zdań bezsensownych, „zawierający jako podzbiór właściwy klasę zdań metafizyki i teologii” (Życiński, 1985, s. 34). Odpowiadając na ten zarzut, Życiński, całkiem słusznie, przywołuje pracę Quine’a, który podważył tę dychotomiczność (Quine, 1969a), ale powołuje się także na drugi argument, a mianowicie na „brak ścisłej i przyjmowanej powszechnie definicji zdań analitycznych” (Życiński, 1985, s. 34). Na dowód tego przytacza listę aż dziewięciu różnych tego rodzaju definicji. Jego zdaniem, ukazuje to „nierealistyczny charakter propozycji Koła Wiedeńskiego” (Życiński, 1985, s. 34). Jest to jednak argument nietrafny, ponieważ zdanie „Bóg istnieje” nie jest ani zdaniem empirycznym, ani tautologią w żadnym z tych dziewięciu znaczeń. Co więcej, wobec trafności argumentacji Quine’a, jest to argument zbyteczny.

5. Wnioski i dalszy rozwój

Pamiętajmy, jest rok 1985 (rok wydania pierwszego tomu *Teizmu*). Życiński proponuje reformę języka o Bogu, a w perspektywie całej teologii, w ten sposób, by mogła ona przemówić do tej części społeczeństwa, która mówi językiem kształtowanym przez nauki ścisłe. Jest to krok nowatorski i prekursorski. Teologia katolicka już raz niedawno zdobyła się na językową transformację: po Soborze Watykańskim II porzuciła język scholastyki na rzecz języka kształtowanego przez różne filozofie typu egzystencjalno-fenomenologicznego i nauki humanistyczne, ale pozostała całkowicie obojętna na falę przemian wywołaną stopniową ewolucją empiryzmu logicznego w różne odmiany filozofii analitycznej. Wkrótce ta fala ogarnęła cały świat anglojęzyczny i zaczęła sobie torować drogę w Europie. Życiński, jako jeden z pierwszych, i na pewno pierwszy w Polsce, dostrzegł wagę problemu. Dziś, po trzydziestu pięciu latach, na naszych oczach rodzi się nowy kierunek w teologii – teologia analityczna. Wprawdzie dzieje się to niezależnie od prekursorskich projektów Życińskiego, ale jego praca nie zdezaktualizowała się, mieści się w niej bowiem szereg elementów, na które do dziś nie zwrócono baczniejszej uwagi. Żeby je dostrzec, spójrzmy jeszcze raz krytycznie na wyniki naszych analiz.

Jak widzieliśmy, Życiński wyróżnił język dyskursu teistycznego i język wiary, choć w trakcie swoich rozważań nie zawsze dość rygorystycznie przestrzegał tego rozróżnienia. W teologii praktyka religijna odgrywa ważną rolę: jako *locus theologicus* może być ona źródłem teologicznej argumentacji. Nie należy więc pomniejszać znaczenia analiz języka wiary, ale trzeba mieć również na uwadze, że dla projektu reformy języka teologicznego pierwszorzędne znaczenie ma przede wszystkim laboratoryjna praca nad językiem dyskursu.

Główna część tej pracy, jaką wykonał Życiński, ma charakter negatywny, to znaczy obala ona argumenty i interpretacje tych analityków, którzy wykazują bezsensowność języka teistycznego lub starają się „uanalitycznić” ten język, pozbawiają go odniesienia do rzeczywistości transcendentnej. Nie jest to robota błaha: sama w sobie ma charakter analityczny i toruje drogę do bardziej konstruktywnej części projektu.

Powstaje pytanie: w jaki sposób można wypracować pozytywną część programu Życińskiego? Wydaje się, że jedynie w konkretnej pracy wykonywanej tym językiem, to znaczy formułując konkretne problemy z zakresu filozofii Boga, lub wręcz teologii i dobierając właściwe narzędzia językowe, drążyć dany problem i poszukiwać jego rozwiązania. Zgodnie z tezą Wittgensteina, trzeba tak długo grać danym językiem, aż się on wykrystalizuje. Życiński próbuje to robić w drugim tomie *Teizmu* (Życiński, 1988a), w którym stara się wypracować zręby chrześcijańskiej metafizyki opartej na filozofii Whiteheada. Inna sprawa, czy mu się to udało i w jaki sposób ewentualnie kontynuować jego przedsięwzięcie.

Takie postawienie sprawy jest w pełni zgodne z innym, bardzo doniosłym wnioskiem, jaki wyłania się z negatywnej części programu Życińskiego (sam ten wniosek już nie jest negatywny): wprowadzie analizy czysto językowe nie pozwalają wyjść do rzeczywistości pozajęzykowej, ale język dyskursu teistycznego musi być tego rodzaju, żeby nie wykluczał jego odniesienia do rzeczywistości transcendentnej czyli Boga. „Do rozstrzygnięcia tej kwestii konieczne jest przeanalizowanie epistemologicznego statusu metafizyki i filozofii Boga” (Życiński, 1985, s. 36). Ale to jest już przedmiotem następnego rozdziału pierwszego tomu *Teizmu*.

Życiński pod wieloma względami okazał się prekursorem, zwłaszcza na terenie myśli polskiej, gdzie – o ile mi wiadomo – nie było

przed nim prób „uanalitycznienia” filozoficznego systemu Whiteheada i wykorzystania go w filozofii Boga. Z czasem teologia dostrzegła możliwości, jakie kryją się w metodach analitycznych. Zrobiła to na innej drodze niż proponował Życiński, ale efekt zmierza w tym samym kierunku, jaki on proponował. Życiński zaczął od sformułowania projektu i drażenia pytania, w jakim stopniu metody analityczne są w stanie sprostać wymaganiom teistycznego języka. Dzisiejsi teologowie po prostu biorą na warsztat konkretne problemy teologiczne i starając się je rozwiązać, naśladują to, co analitycy robią „u siebie” (por. np. Crisp i Rea, 2009; pojawiła się także polska „jaskółka”: Hołda, 2021). Jeżeli po drodze pojawiają się jakieś problemy metafizyczne – a są one nieuniknione – to trzeba sobie z nimi także jakoś radzić. Oczywiście, nie brakuje również czynionej przy okazji refleksji metodologicznej. Jest to na pewno strategia owocna, gdyż metoda zawsze najlepiej wyodróżnia się w działaniu.

Nie można jednak zapominać, że projekt Życińskiego był znacznie bardziej ambitny: najpierw przeorać fundamenty – dokładnie rozróżnić to, co analitycy już w tej dziedzinie zrobili, także ich krytykę języka filozofii Boga i teologii, wybrać z tego i rozwinąć to, co można odpowiednio przystosować do konstrukcji odnowionego języka teizmu, a dopiero potem, przy pomocy takiego narzędzia, budować chrześcijańską metafizykę i w dalszej perspektywie teologię. Projekt Życińskiego jest nadal wielkim wyzwaniem. Czy ktoś je podejmie?

Bibliografia

Crisp, O. i Rea, M.C., red., 2009. *Analytic theology: new essays in the philosophy of theology*. Oxford; New York: Oxford University Press.

- Hołda, M., 2021. Kochać Boga na sposób analityczny. O tym, co teologia dogmatyczna może zyskać w spotkaniu z filozofią analityczną. W: R.J. Woźniak, red. *Dogmat i metoda: wprowadzenie do badań interdyscyplinarnych w teologii dogmatycznej*. Kraków: Wydawnictwo WAM, ss. 453–475.
- Quine, W.V.O., 1969a. Dwa dogmaty empiryzmu. W: *Z punktu widzenia logiki: eseje logiczno-filozoficzne* (B. Stanosz, tłum.). Warszawa: Państwowe Wydawnictwo Naukowe, ss. 35–70.
- Quine, W.V.O., 1969b. O tym, co istnieje. W: *Z punktu widzenia logiki: eseje logiczno-filozoficzne* (B. Stanosz, tłum.). Warszawa: Państwowe Wydawnictwo Naukowe, ss. 9–34.
- Życiński, J., 1985. *Teizm i filozofia analityczna. T. 1*. Kraków: Znak.
- Życiński, J., 1988a. *Teizm i filozofia analityczna. T. 2*. Kraków: Znak.
- Życiński, J., 1988b. *The Structure of the Metascientific Revolution: An Essay on the Growth of Modern Science*, seria: *Philosophy in Science Library*, 3. Tucson: Pachart Publishing House.
- Życiński, J., 2013. *Struktura rewolucji metanaukowej: studium rozwoju współczesnej nauki* (M. Furman, tłum.). Kraków: Copernicus Center Press.

Hutchins w obronie interdyscyplinarnych badań nad poznaniem

Marcin Miłkowski

Instytut Filozofii i Socjologii PAN

Witold Wachowski

Uniwersytet Marii Curie-Skłodowskiej

Edwin Hutchins: In defense of interdisciplinary research on cognition

Abstract

The article presents the interdisciplinary approach of Edwin Hutchins, analyzing his conception of distributed cognition as probably the most important and lasting contribution of anthropology to the repertoire of theoretical tools in cognitive science. At the same time, this conception resulted in one of the most interesting relationships between cognitive science and social sciences. These relationships are made possible by the assumptions of Hutchins' conception, which directly contribute to interdisciplinary collaboration. His account of distributed cognition has enormous potential, allowing the integration of research into cognitive and social processes. This is also because it breaks with methodological individualism.

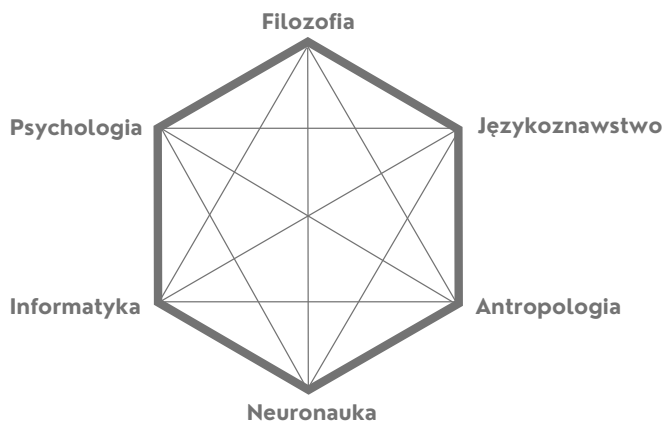
Keywords

cognitive anthropology, ethnography, methodological individualism, cognitive science, distributed cognition, representational states, interdisciplinarity.

1. Wprowadzenie

Antropologia jako jeden z głównych filarów kognitywistyki podejmuje zagadnienia nauk poznawczych i społecznych. Jeżeli więc szukać zależności i oddziaływań między nimi, nie sposób tej dziedziny nie uwzględnić: bodaj to głównie ona powinna łączyć kognitywistykę (czyli interdyscyplinarny konglomerat badań nad poznaniem) z prowadzonymi przez antropologów poznawczych czy socjologów wiedzy badaniami nad społecznymi procesami poznawczymi. A jednak historia decydującego udziału antropologii w kognitywistyce jest stosunkowo krótka. Wydaje się, że pozostał po niej jeden wyróżniający się i rozwijany po dziś dzień nurt, mianowicie badania nad poznaniem rozproszonym; ich najbardziej zasłużonym i rozpoznawalnym przedstawicielem jest Edwin Hutchins (2022).

Przyjrzyjmy się jednak historycznym perturbacjom poprzedzającym te badania. Wymowną pamiątką po początkowym równorzędnym udziale antropologii w powstawaniu kognitywistyki jest klasyczny heksagram z 1978 roku (zob. np. Miller, 2003), obrazujący udział sześciu dyscyplin: językoznawstwa, neurobiologii, sztucznej inteligencji, filozofii, psychologii oraz właśnie antropologii (patrz Ilustracja 1). Wydaje się, że kognitywiści tamtych czasów doceniali społeczno-kulturowe aspekty procesów poznawczych, toteż antropologia pozostawała wówczas w dość dobrych relacjach z kognitywistyką (Bender, Hutchins i Medin, 2010). Allen Newell (1980; 1990) wymieniał społeczne aspekty poznania wśród swoich 13 kryteriów jednolitej teorii poznawczej. Donald Norman (1981), będący jednocześnie klasykiem badań nad poznaniem oraz przedstawicielem koncepcji poznania rozproszonego, wskazywał na systemy wiedzy kulturowej wśród 12 kluczowych zagadnień w kognitywistyce. Z kolei Howard Gardner (1985) wymieniał emocje, kontekst, kulturę i historię jako istotne elementy



Ilustracja 1: Heksagram obrazujący udział sześciu dyscyplin: językoznawstwa, neurobiologii, sztucznej inteligencji, filozofii, psychologii oraz antropologii w powstawaniu kognitywistyki.

aktywności poznawczej. W ten sposób udział antropologii otwierał przed kognitywistyką perspektywę systematycznych dociekań nad kulturowymi mechanizmami systemów wiedzy, z uwzględnieniem ograniczeń poznawczych i ekologicznych.

Co się jednak stało, że ostatecznie wykluczono rolę społeczeństwa i kultury z głównego nurtu kognitywistyki, doprowadzając przy tym do rozejścia się dróg antropologów i kognitywistów? Za jedną z przyczyn uważa się rozłam w samej antropologii na obszar przyrodniczo-ściśły i humanistyczny, gdzie antropologia poznawcza, wykazująca cechy obu, okazała się „niekochanym dzieckiem”, którym nie opiekowało się żadne z „rodziców”. Również sama współpraca antropologów z kognitywistami pozostawiała sporo do życzenia. Tych pierwszych zniechęcały preferowane przez drugich ściśle kontrolowane eksperymenty, prowadzone w sztucznych warunkach laboratoryjnych,

przy niedocenianiu roli kontekstu i realnych warunków. Kognytywiści z kolei zarzucali antropologom brak odpowiedniego rygoru badawczego i nadmierne skupienie na danych jakościowych. James Boster podkreślał, że antropologia, w odróżnieniu od psychologii w kognitywistyce, „skupia się na treści (nie na procesie), na społecznościach i kontekstach społecznych (nie na jednostkach), na warunkach naturalnych (nie laboratoryjnych), na uchwytowywaniu zjawisk w świecie rzeczywistym, nawet jeśli wymaga to pewnego złagodzenia rygorów, i troszczy się o to, czy zebrane dane faktycznie dowodzą tego, co może się wydawać na pierwszy rzut oka” (za Bender, Hutchins i Medin, 2010, s. 377, przekład własny). Dodajmy, że do podtrzymania tego stanu rzeczy przyczynili się również kontynuatorzy Newella, eliminując rolę społeczeństwa i kultury z projektu formowania jednolitej teorii poznania (Hutchins, 1995a; Bender, Hutchins i Medin, 2010; Anderson i Lebiere, 2003; Kronenfeld i in., 2011).

Jak wskazaliśmy na początku artykułu, bezpośredni udział antropologii w kognitywistyce utrzymał się zasadniczo w jednym przypadku: w rozwijanym od lat osiemdziesiątych XX w. nurcie badań nad poznaniem rozproszonym, którym przewodził Hutchins. Badacz ten, związany z Wydziałem Kognitywistyki na Uniwersytecie Kalifornijskim w San Diego, od pierwszych prac zdradzał zainteresowanie związkami między myśleniem, językiem i kulturą. W połowie lat siedemdziesiątych XX w. prowadził badania etnograficzne na Wyspach Trobriandzkich w Papui-Nowej Gwinei, koncentrując się na argumentacji w sporach publicznych. W ramach stażu podoktorskiego opracował model tradycyjnej nawigacji mikronezyjskiej, głównie na podstawie dostępnych opracowań praktyk nawigacyjnych. Zatrudniony później przez marynarkę wojenną Stanów Zjednoczonych, wykorzystał spostrzeżenia z bezpośrednich badań etnograficznych do tworzenia komputerowych systemów szkoleniowych dla układów

o napędzie parowym oraz nawigacji radarowej. Swoje obserwacje poszerzył o nawigację morską, a następnie wykorzystał je w swoim pierwszym, obszernym opracowaniu podstaw teorii poznania rozproszonego, czyli najbardziej docenianej książce *Cognition in the Wild* (1995a). Pod koniec lat osiemdziesiątych XX w. zaczął prowadzić obserwacje w lotnictwie cywilnym (sam jest dyplomowanym pilotem). Ponadto badania Hutchinsa objęły takie aktywności poznawcze, jak interakcje człowiek–komputer, obliczeniowe symulacje procesów kulturowych, praktyki naukowców, społeczności małych człokształtnych, a ostatnio również – społeczności delfinów¹.

Będziemy argumentować, że pod względem konceptualizacji systemów i procesów poznawczych teorię poznania rozproszonego można traktować jako swego rodzaju konkurencję dla rozpowszechnionej w badaniach społecznych teorii aktora-sieci (ANT) Brunona Latoura. Ta ostatnia oferuje ujęcie relacji poznawczych, które zupełnie zrywa z indywidualizmem metodologicznym – czego w kognitywistyce nie zaoferowało żadne inne podejście poza teorią poznania rozproszonego (co wykazano w: Wachowski, 2022). W przeciwieństwie jednak do ANT, teoria Hutchinsa sprzyja współpracy interdyscyplinarnej, którą Latour dogmatycznie odrzuca. Stanowisko Latoura wydaje nam się o tyle symptomatyczne, że może świadczyć o pewnym braku porozumienia między badaczami procesów społecznych i poznawczych, które prowadzi do stosunkowo małej obecności prac antropologów w kognitywistyce (Núñez i in., 2019) i zbyt ograniczonej, naszym zdaniem, recepcji prac spod znaku poznania rozproszonego w samej antropologii. Przy tym wszystkim należy podkreślić, że mimo wszystko teoria poznania rozproszonego została w wyraźnej mierze zainspirowana propozycją Latoura (zob. np. Hutchins, 2001).

¹ Tę ewolucję potwierdza również sam badacz w korespondencji prywatnej, zaznaczając, że wycofywał się (zwłaszcza w ciągu ostatnich 15 lat) z agnostycyzmu co do natury procesów mózgowych pod wpływem postępów w naukach poznawczych – co powinno być istotne dla tych, którzy traktują *Cognition in the Wild* (1995a) jako obowiązującą w całości po dziś dzień biblię teorii poznania rozproszonego.

Interdyscyplinarne badanie procesów poznawczych i społecznych w ich wzajemnym powiązaniu jest niewątpliwie o tyle trudne, że wymaga kontrolowania bardzo wielu czynników jednocześnie, co jest kłopotliwe szczególnie w badaniach obserwacyjnych, a nie eksperymentalnych. Antropologowie i badacze społecznych procesów wiedzy chcą je zaś często obserwować w warunkach naturalnych. Propozycja Hutchinsa jednak pozwala połączyć takie obserwacyjne podejście z rygorem typowym dla kognitywistyki. Daje to szansę na rozwój badań zarówno społecznych, jak i poznawczych. W tym przypadku interdyscyplinarność nie tylko wspiera się na argumentach pojęciowych – takich jak (1) możliwość ugruntowania procesów społecznych w dobrze rozumianych procesach poznawczych, (2) możliwość teoretycznej unifikacji nauk społecznych przez odwołanie do poznawczych wyjaśnień zachowań racjonalnych, (3) konieczność poprawnego ujęcia poznawczych podstaw procesów społecznych, czy (4) uzupełnianie się kognitywistyki i socjologii poznawczej (por. Kaidesoja, Sarkia i Hyryläinen, 2019) – lecz także na dobrze opracowanej metodologii badawczej. Hutchins jest więc, naszym zdaniem, znakomitym ambasadorem interdyscyplinarnej pracy badawczej na styku nauk społecznych i poznawczych. Pokazuje nieodzowność i sposób uwzględniania społecznych determinant rozproszonych procesów poznawczych, a także kognitywistyczną metodę modelowania tych procesów społecznych.

2. Teoria, metoda i przedmiot badań

Hutchins forsuje projekt włączenia „kultury, kontekstu i historii” w sam rdzeń badań nad poznaniem (Hutchins, 2001, s. 2072). Nie neguje klasycznej metodologii i dorobku nauk poznawczych, propo-

nuje jednak ich istotne rozszerzenie: metodologiczne i konceptualne. Przedmiotem jego badań najczęściej są struktury złożone z ludzi oraz rzeczy (przeważnie artefaktów). Autor *Cognition in the Wild* traktuje ludzkie myślenie jako działalność kulturową. W tym kontekście można wyróżnić teorię poznania rozproszonego oraz ekologię poznawczą, dominującą w ostatnich pracach Hutchinsa. Ta ostatnia stanowi jednak rodzaj szerszej ramy pojęciowej, w której plasuje między innymi własną teorię.

Celem ekologii poznawczej jest badanie zjawisk poznawczych w kontekście kulturowo-społecznym. Opisuje ona system zależności między procesami poznawczymi a strukturami grup. Badane zależności dotyczą procesów umysłowych, społecznych i materialnych oraz ciał. Elementy ekologii poznawczej od dawna bywały sporadycznie obecne w kognitywistyce. Wraz z dowartościowaniem kontekstu kulturowo-społecznego poznanie przestaje być postrzegane jako proces racjonalnego wnioskowania logicznego, a zaczyna być traktowane też jako zjawisko biologiczne. W ekologii poznawczej wskazuje się na sieć wzajemnych zależności między elementami systemu poznawczego. Znaczej wagi nabiera pytanie o jednostkę analizy. Hutchins wychodzi od banalnej konstatacji, że wszystko się ze sobą wiąże. Jednak nie każde połączenie jest jednakowo istotne, dzięki czemu badania naukowe pozostają możliwe. Wytyczenie granicy zawsze ułatwia spostrzeżenie jednego, a utrudnia czy wręcz uniemożliwia – ujrzenie czegoś innego. Hutchins powołuje się na Gregory'ego Batesona, który podkreślał, że wytyczanie granic jednostek analizy nie powinno pozostawiać ważnych kwestii niewyjaśnionych ani niewyjaśnialnych (Hutchins, 2008; 2010a).

Hutchins wskazuje na trzy ujęcia historycznie ważne dla ekologii poznawczej. Jednym z nich jest zainicjowana przez Jamesa Gibsona psychologia ekologiczna z jej teorią afordancji. Gibson „afordancjami”

nazywał właściwości relacyjne ustrukturyzowanego środowiska, prowokujące organizm do określonego, niewymagającego namysłu zachowania. Wbrew Gibsonowi, który odrzuca wyjaśniania odwołujące się do procesów reprezentowania czy obliczania, Hutchins przyjmuje typowe w kognitywistyce założenia o obliczeniowo-reprezentacyjnym charakterze poznania. Kolejne ważne tutaj ujęcie to Batesonowska ekologia umysłu, odwołująca się do cybernetycznych sprzężeń i teorii systemów. Trzecie zaś to kulturowo-historyczna teoria czynności Lwa Wygotskiego. Hutchins widzi ich kontynuację we współczesnych kognitywistycznych teoriach ucieleśnienia, rozszerzenia, ogólniej – usytuowania poznania, jak również we własnej teorii poznania rozproszonego (Hutchins, 2010a, s. 707–712).

Kultura nie stanowi dla tego badacza zbioru rzeczy, materialnych czy abstrakcyjnych, ale proces – zachodzący wewnątrz i na zewnątrz ludzkich umysłów – na który składają się przede wszystkim procesy poznawcze. Jest ona procesem adaptacyjnym, który akumuluje rozwiązania powtarzających się problemów, a badanie go wymaga oderwania się od myślenia w kategoriach przyjętych z góry granic systemów poznawczych. Na ludzki ekosystem składają się różnorodne zasoby poznawcze, do których należą obiekty fizyczne, modele umysłowe, a także praktyki kulturowe. Zarówno artefakty, jak i praktyki mają swoje historyczne i kulturowe ścieżki rozwojowe. Ludzkie myślenie Hutchins opisuje jako rodzaj praktyki kulturowej, nabywanej w toku ucieleśnionych interakcji i zgodnie z powszechnymi społecznymi i materialnymi prawidłowościami oraz znaczeniami obecnymi w świecie, a ściślej biorąc, w kulturowym ekosystemie jednostki ludzkiej, uczącej się nie przez obserwację, lecz poprzez współuczestniczenie. Dla Hutchinsa ów kulturowy ekosystem poznawczy staje się przedmiotem nowej nauki, dzięki której można go lepiej poznać (Hutchins, 1995a, s. 353–355; 2006; 2014).

Istotnym kulturowo procesem jest materialne osadzanie modeli pojęciowych. Hutchins podaje przykład kolejki w punkcie usługowym. Praktyka ta stwarza rodzaj przestrzennej pamięci, wprowadzając liniowy porządek umożliwiający zrealizowanie usługi. Nie każdy taki porządek jest kolejką – mimo zewnętrznego podobieństwa. Praktykę korzystania z kolejki można postrzegać więc jako poznawczą, bo bez niej obsługa w kolejności jest trudna. Ma ona w istocie ponadindywidualne funkcje poznawcze: przede wszystkim umożliwia rejestrowanie i utrwalanie kolejności przybywających klientów, jak również zarządzanie zapominaniem, gdy ludzie opuszczają kolejkę przed zrealizowaniem swojej potrzeby lub po jej realizacji (Hutchins, 2005, s. 1559–1560; 2014, s. 39–40).

Wszystkie założenia ekologii poznawczej w ramach kognitywistyki ma uwzględniać teoria poznania rozproszonego. Przyjrzymy się jej, uwzględniając przemiany stanowiska Hutchinsa w kontekście rozwoju nauk poznawczych, jego kolejne analizy założeń teoretycznych, by wreszcie skonfrontować ją z koncepcją Brunona Latoura.

Według teorii poznania rozproszonego właściwa jednostka analizy – przedmiot badań kognitywistycznych i społecznych – powinna być ustalana nie *a priori*, lecz z uwagi na naturę badanego zjawiska. Owa jednostka, czyli system poznawczy, niekiedy może mieścić się w granicach czaszki lub skóry człowieka czy innego zwierzęcia, może jednak również wykraczać poza jednostkę, obejmując zarówno żywe, jak i nieorganiczne składniki, a więc funkcjonując w pewnej makroskali (Hutchins, 1995a, s. 128–129; 2001, s. 2068; 2010b, s. 426). Jest to modelowy wręcz przykład zawieszenia indywidualizmu metodologicznego, tak jak rozumie się go w kognitywistyce, ale i w filozofii umysłu.

Pojęcie „indywidualizm metodologiczny” jest dość dobrze ugruntowane w socjologii (np. Schumpeter, 1908; Weber, 2002), gdzie

wyjaśnianie zjawisk społecznych prowadzić ma do ukazania ich jako skutków indywidualnych działań. W filozofii umysłu i kognitywistyce jednak pojmowane jest nieco inaczej. Jerry Fodor odróżnił pojęcie indywidualizmu metodologicznego od pojęcia solipsyzmu metodologicznego. Jak twierdzi, według indywidualizmu metodologicznego stany psychiczne indywiduuje się ze względu na ich moce przyczynowe. Zgodnie z indywidualizmem można więc mówić o relacjach przyczynowych między podmiotem a środowiskiem, jeśli są istotne dla stanów psychicznych. Według solipsyzmu metodologicznego zaś stany psychiczne indywiduuje się niezależnie od wartościowania semantycznego (tego, do czego stany te mają się faktycznie odnosić); a więc pomijane są relacje przyczynowe między podmiotem a środowiskiem (Fodor, 1980; por. też Heath, 2020).

W kontekście kognitywistyki przyjmujemy rozumienie indywidualizmu metodologicznego odbiegające nieco od stanowiska bronionego przez Fodora. W wersji radykalnej oznacza on założenie, że wiedza o podmiocie poznania (procesach zachodzących w jego umyśle czy mózgu) jest konieczna i wystarczająca do zrozumienia procesów poznawczych, wobec czego ignoruje się środowisko podmiotu oraz interakcje z nim, gdyż nie są one własnościami podmiotu. W takiej postaci indywidualizm metodologiczny nie jest już może stanowiskiem dominującym w kognitywistyce. Można wskazać różne stopnie zerwania z nim czy też zawieszenia go (patrz Ilustracja 2), od założenia istotnej roli ciała pozaneuronalnego w poznaniu, przez rolę środowiska, w którym podmiot jest poznawczo zakorzeniony, po koncepcje umysłu rozszerzonego czy też szerokiego systemu, w którym indywidualny podmiot ludzki jest jednym z komponentów, i to obok artefaktów².

² Więcej na temat indywidualizmu metodologicznego i przypadków zrywania z nim: (Wachowski, 2022).



Ilustracja 2: Różne stopnie odejścia od indywidualizmu metodologicznego.

Dla Hutchinsa każdy taki system jest rozproszonym systemem poznawczym, niezależnie od jego skali – ta bowiem może ulegać zmianie w działaniu, a przez to również przesuwając granice systemu. Poznanie rozproszone nie stanowi żadnego określonego rodzaju poznania, lecz ujęcie bądź sposób badania poznania każdego rodzaju: „Dla każdego procesu poznawczego musi istnieć sposób ujęcia go jako rozproszonego”. Dlatego „podejście to nie wnosi żadnych twierdzeń o naturze świata. Jest raczej sposobem patrzenia na ten świat” (Hutchins, 2014, s. 36). W omawianej teorii właściwe pytania nie dotyczą więc tego, czy dane poznanie jest, czy nie jest rozproszone, ale składników danego systemu i relacji między nimi, wyłaniania się procesów poznawczych z interakcji w systemie itp. Jak podkreśla badacz, perspektywa poznania rozproszonego nie ma żadnych założeń empirycznych. Nie można jej potwierdzić ani obalić. Można natomiast w jej obrębie stawiać hipotezy empiryczne, między innymi hipotezę umysłu rozszerzonego, którą powinno się móc uzasadnić bądź odrzucić w trakcie badań empirycznych (Hutchins, 2014, s. 36–37).

Warto zatem porównać teorię Hutchinsa z koncepcją umysłu rozszerzonego (Clark i Chalmers, 2008). Koncepcja ta głosi, że granice umysłu nie kończą się na głowie: częścią (a nie tylko narzędziem) umysłu mogą być zewnętrzne wobec układu nerwowego procesy

i przedmioty fizyczne, np. implanty stanowiące dodatkowe receptory zmysłowe, urządzenia obliczeniowe, systemy symboliczne (np. notacja matematyczna) lub nawet całe współpracujące ze sobą grupy. Badacz potwierdza pewne podobieństwa między nimi; podkreśla jednak dwie ważne różnice. Po pierwsze, koncepcja Clarka i Chalmersa dotyczy tylko pewnego rodzaju poznania: w jej ujęciu umysł czasami rozszerza się poza granice mózgu, a czasami nie. Po drugie, w systemie umysłu rozszerzonego daje się wyróżnić jego centrum (zwykle mózg lub też cały organizm). Dlatego też koncepcję tę można traktować jako dotyczącą szczególnego przypadku funkcjonowania umysłu – czyli może ona być rozwijana na podstawie szerszych założeń teorii poznania rozproszonego (Hutchins, 2011, s. 437–446; 2014, s. 35–36).

Hutchins, odrzucając indywidualizm metodologiczny, nie odchodzi od założeń komputacjonistycznych i reprezentacjonistycznych, gdy określa poznanie jako obliczanie realizowane poprzez wytwarzanie, przetwarzanie i przekazywanie stanów reprezentacyjnych (Hutchins, 1995a, s. 49). Przez „stany reprezentacyjne” rozumie się stany nośników materialnych (zewnętrznych lub np. mózgowych), które oznaczają lub opisują przedmioty procesów poznawczych. Jednak zakres, jaki Hutchins przypisuje poznawczym reprezentacjom i poznawczym procesom obliczeniowym, zdecydowanie wybiega – na mocy głównego założenia jego teorii – poza granice tradycyjnie rozumianych systemów poznawczych. Przedstawiana przez niego perspektywa przypomina nieco punkt widzenia Guliwera, którym byłby psycholog poznawczy, w krainie olbrzymów: oto otrzymuje on wręcz wymarzoną możliwość bezpośredniej obserwacji procesów przebiegających w systemie poznawczym w makroskali, a więc nie „w głowie”, lecz w otoczeniu, w które może wejść. Reprezentacje poznawcze stają się więc widzialne i namacalne: mogą być nimi artefakty. Przez

„artefakty poznawcze” teoretycy poznania rozproszonego rozumieją obiekty, które służą do „przechowywania, udostępniania lub przetwarzania informacji na potrzeby funkcji reprezentowania, wpływając na poziom wykonania procesów poznawczych u ludzi” (Norman, 1991, s. 11). Warto pamiętać, że artefaktami poznawczymi są również gesty czy wypowiedzi. Podobnie dostępne i obserwowalne jest ich wytwarzanie, przetwarzanie i przekazywanie, niekiedy dosłownie z rąk do rąk (Hutchins, 1995a, s. 128–129). Tym samym taki makrosystem poznawczy ma własne systemy pamięci zewnętrznej, jak i mechanizmy oceny i kontroli sytuacji, z których żadne nie daje się sprowadzić do pamięci ani aktów oceny i kontroli indywidualnych podmiotów poznania. Są one natomiast „rozproszone”, czy też rozdzielone między ludźmi w kooperującej grupie lub artefaktami.

Perspektywa poznania rozproszonego to nie po prostu rozwinięcie kognitywistycznej historii o sprzężeniu podmiotu poznawczego z otoczeniem. Niemniej jednak powtarzające się odwołania Hutchinsa do obiektów na różnych skalach czasoprzestrzennych, takich jak mózg, układ nerwowy czy ciało jednostkowego podmiotu ludzkiego, dobitnie pokazują, że stanowiska badacza nie można sprowadzać do eksternalistycznych, a niekiedy wręcz „neurofobicznych”, ujęć organizacji systemów rozwiązywania problemów, którego modelowym przykładem miałyby być klasyczne studium nawigacji morskiej. Do takich wniosków mógł prowokować deklarowany przezeń we wcześniejszych pracach częściowy agnostycyzm co do możliwości badania wewnętrznych procesów umysłowych, z biegiem czasu coraz bardziej zredukowany³ – choć trudno posądzać Hutchinsa o ignorowanie wagi sprzężenia poznawczego między wewnętrznymi i zewnętrznymi procesami poznawczymi, o czym świadczy chociażby wnikliwa analiza

³ Szczegółowa rekonstrukcja w (Afeltowicz, 2012, s. 178–188; Afeltowicz i Wachowski, 2015).

funkcjonowania systemu kokpitu samolotu (zob. Hutchins, 1995b; Hutchins i Klausen, 1996). Z drugiej strony nie brakowało wyważonych ujęć jego teorii również we wcześniejszym okresie, jak chociażby autorstwa Yvonne Rogers (1997) czy też takich współpracowników Hutchinsa, jak David Kirsh (1999). Jednak nie bez znaczenia są też odniesienia do koncepcji Latoura, przeciwnika uwzględniania jednostkowej psychologii w wyjaśnieniach funkcjonowania kolektywnych struktur poznawczych, jak również twórcy podejścia metodologicznego, kojarzonego z podejściem reprezentantów nurtu społecznych badań nad nauką i techniką (zob. np. Hess, 2001).

Główna metoda badawcza Hutchinsa jest etnograficzna. W podejściach etnograficznych docenia Hutchins możliwość badań nad rzeczywistym poznaniem w rzeczywistym świecie, pozwalającą na wyprowadzenie ich poza ograniczoną przestrzeń laboratorium (Hutchins, 2014, s. 43). Takie podejście, zastosowane do procesów poznawczych, wydaje się w pełni legitymizować odejście od indywidualizmu metodologicznego, nie tak oczywiste w kognitywistyce. Jakościowa metoda etnograficznych badań poznawczych wykorzystuje obserwacje funkcjonowania danej wspólnoty do analiz jej aktywności poznawczej.

Różnicę między etnografią tradycyjną a poznawczą można za Robertem Williamsem określić następująco: podczas gdy pierwsza zajmuje się znaczeniami wytwarzanymi przez członków danej wspólnoty kulturowej, druga koncentruje się na tym, jak te znaczenia są przez nich wytwarzane. Etnografowie tradycyjni wskazują materialne i konceptualne zasoby składające się na lokalną rzeczywistość danej grupy, a etnografowie kognitywni badają sposób, w jaki te zasoby są wykorzystywane w aktywności kulturowej. Pierwsi analizują myślenie ludzi definiujące ich grupę kulturową, natomiast drudzy obserwują

sam ten proces; pierwsi opisują wiedzę, a drudzy – procesy konstrukcji i wykorzystywania wiedzy. Oczywiście różnice te nie świadczą o istnieniu między nimi przepaści (Williams, 2006, s. 838).

Co tutaj istotne, zainteresowanie procesami poznawczymi funkcjonującymi w środowisku naturalnym (tytułowe *in the wild*) nie oznacza, że dla Hutchinsa procesy przebiegające w warunkach laboratoryjnych są nienaturalne i niegodne badania lub że wręcz stoją poza kulturą. Nic bardziej błędnego: poznanie nie zna próżni kulturowej, więc miejsca, w których przeprowadza się wyizolowane eksperymenty, nie mogą stanowić tutaj wyjątku. Obserwacje etnograficzne w laboratoriach są po prostu trudniejsze, bo aspekty kulturowe pozostają tam niejako przezroczyste, co wymaga po prostu szczególnego podejścia badawczego (Hutchins, 1996, s. 66–67). Że są możliwe i interesujące, sam Hutchins udowadnia między innymi we współautorskim studium etnograficznym aktywności neurobadaczy (Alač i Hutchins, 2004).

Przedmiotem badań nad poznaniem rozproszonym ma być „natura poznania” w różnego rodzaju systemach działalności. W badaniach Hutchinsa były one dość różnorodne, a więc kultura Wysp Trobriandzkich, systemy nawigacji i awiacji, społeczności małych członkowskich oraz delfinów, systemy interakcji człowiek-komputer (HCI), wspólnoty naukowców, jak również grupy realizujące tak codzienne praktyki kulturowe, jak opisana już kolejka. Poza obserwacjami *in the wild* badacz ten przeprowadzał również obliczeniowe symulacje procesów kulturowych. Wskazując przy tym na różne skale realnych rozproszonych systemów poznawczych, plasuje on umysł rozszerzony na średnim poziomie skali, podkreślając, że w systemie tego typu mamy do czynienia z centralizacją poznawczą (Hutchins, 2014, s. 37). Systemami pozbawionymi takiej centralizacji byłyby już takie funkcjonalne struktury człowiek-artefakt, jak kokpity pojazdów,

laboratoria, instytucje finansowe czy internet. Istnieją także rozległe i różnorodne przestrzenie inteligencji zbiorowej, gdzie jednostkami systemu są samodzielne, aktywne podmioty poznania. Tutaj przykładem mogą być społeczności owadów, rynki ekonomiczne oraz media społecznościowe (Hutchins, 2014, s. 37).

Przyjrzyjmy się najbardziej znanemu Hutchinsowskiemu studium przypadku, jakim były czteromiesięczne badania etnograficzne na początku lat dziewięćdziesiątych XX w. nad nawigacją morską na pokładzie amerykańskiego lotniskowca⁴. Zespół marynarzy i oficerów musi odpowiedzieć na prosto brzmiące pytanie: „gdzie jesteśmy i dokąd zmierzamy?”, stawiane oczywiście z perspektywy osób przebywających na pokładzie statku. Zapomnijmy tutaj na moment o GPS i systemach satelitarnych, ponieważ dopiero zaczynano je wdrażać i były jeszcze dość zawodne.

Aby dokładnie wytyczyć położenie statku, nawigatorzy muszą dysponować co najmniej trzema łukami pozycyjnymi (AOP), wyznaczanymi na podstawie odległości do określonego punktu orientacyjnego, bądź liniami pozycyjnymi (LOP), wyznaczanymi na podstawie mierzonego w stopniach położenia okrętu względem widocznego punktu orientacyjnego. Owe łuki i linie oznaczane są na mapie cyrklem oraz specjalnym kątomierzem. Natomiast wyznaczanie LOP odbywa się z wykorzystaniem pelorusów, czyli specjalnych urządzeń przypominających kompasy bez igły, umieszczonych na burcie lub pokładzie, dzięki którym obsługujący je marynarze wyszukują punkt obserwacyjny i w żądanym momencie podają kwatermistrzowi odczyt z tarczy. Najważniejszym przyrządem – określanym przez

⁴ Literatura przedmiotu obejmuje również prace poświęcone poznaniu rozproszonemu społeczeństwu, które można by traktować jako pewnego rodzaju formę przejściową w stosunku do analiz heterogenicznych struktur dokonywanych przez Hutchinsa (Roberts, 1964; Salomon, 1993).

Hutchinsa jako analogowy komputer – jest mapa rozłożona w kabinie nawigacyjnej, gdyż to wokół niej koncentrują się działania nawigatorów. Reprezentując przestrzeń geograficzną, daje ona nawigatorom unikatową perspektywę. Ponadto jednak mapa nawigacyjna stanowi urządzenie obliczeniowe: dzięki niej zachodzi zewnętrzne obliczanie, niemające odpowiednika w niczyjej obecnej tam głowie, nawet oficerów. Odpowiedzi na zadawane pytania nawigatorzy otrzymują, wykonując proste czynności, takie jak obsługa cyrkla i kątomierza, rysowanie linii czy przepisywanie cyfr. Cała procedura stanowi, zdaniem Hutchinsa, doskonałą ilustrację Latourowego myślenia za pomocą rąk i oczu (Hutchins, 1995a, s. 142–143). W tym kontekście Hutchins przytacza słowa Herberta Simona: „rozwiązanie problemu to po prostu przedstawienie go w taki sposób, aby rozwiązanie stało się oczywiste” (Hutchins, 1995a, s. 117).

Mapa nawigacyjna pełni kilka ważnych funkcji. Stanowi podstawową reprezentację zewnętrzną, umożliwiającą przeprowadzanie obliczeń. Pozwala na integrowanie i porównywanie informacji pochodzących z różnych źródeł. Służy do koordynacji ludzkich działań. Stanowi także środek monitorowania jakości pracy nawigatorów. Przy tym wszystkim jest zewnętrznym systemem pamięci (zresztą przechowywanym pieczołowicie i w kilku kopiach), obejmującym wiedzę (wraz z jej modyfikacjami) zarówno członków danej załogi okrętu, jak i różnych pokoleń marynarzy, badaczy, geografów, kartografów (Afeltowicz i Wachowski, 2015, s. 93–98; Hutchins, 1995a, s. 61–143).

3. Korzenie, inspiracje, interpretacje

Do jakiego stopnia mamy tutaj do czynienia z nową propozycją teoretyczną? Jak się mają te badania nad poznaniem do reszty nauk

poznawczych? Odpowiedzi na oba te ważne pytania wiążą się ze sobą, a wstępną odpowiedź naszkicowaliśmy już wcześniej, omawiając ekologię poznawczą.

Osbeck i Nersessian (2014), podejmując własną próbę historycznego usytuowania teorii poznania rozproszonego, wskazują na analogie z psychologią funkcjonalną – w tym stanowiskami Williama Jamesa i Johna Deweya – jednocześnie sugerując związki teoretyczne z pragmatyzmem filozoficznym. Funkcjonałści podejmowali „problem kontekstu”, ujmując działający podmiot w „pełnej” sytuacji, jego życie umysłowe jako proces adaptacyjny, a procesy poznawcze raczej jako pewną aktywność niż nabywanie wiedzy. Porównując psychologię funkcjonalną i teorię poznania rozproszonego, Osbeck i Nersessian wskazują na zbieżności: perspektywa systemowa i dynamiczna, nacisk na rozwiązywanie problemów, sprzeciw wobec czysto intelektualnej wizji poznania, oddzielonego od świata rzeczy i zjawisk społecznych, preferowanie badań nad zjawiskami zachodzącymi w naturalnym (a nie sztucznym, laboratoryjnym) środowisku.

Na uwagę zasługują wpływy i zbieżności z perspektywą radzieckiej szkoły kulturowo-historycznej Lwa Wygotskiego i Aleksandra Łurii. Ten ostatni, komentując stanowisko Wygotskiego (np. Wygotskij, 1971; warto także porównać z: Mead, 1975), wywodzącego indywidualne akty psychiczne z zachowań grupowych, stwierdzał, że jeśli mamy wyjaśnić złożone formy ludzkiej świadomości, powinniśmy w analizach wyjść poza granice organizmu, w sferę zewnętrznych warunków jego życia, uwzględniając tak społeczny, jak i historyczny ich wymiar (Luriâ, 1979). Michael Cole i Yrjö Engeström wymieniają następujące elementy podejścia kulturowo-historycznego, które można traktować jako początki teorii poznania rozproszonego: zasady rządzące „naturalnymi” psychicznymi czynnościami człowieka są różne od zasad rządzących jego czynnościami „kulturowymi” –

upośrednionymi przez narzędzia i normy społeczne; to kulturowe upośrednienie kształtuje szczególnego rodzaju strukturę ludzkiego umysłu i zachowania, modyfikując podmiot i jednocześnie jego otoczenie; interakcję między nimi regulują kulturowe artefakty, zarówno materialne, jak i symboliczne. Niezwykle ważnym narzędziem tego upośrednienia jest język. W danym otoczeniu kulturowym skumulowana jest również wiedza poprzednich pokoleń, dlatego w rozwoju osobniczym ważną rolę odgrywa środowisko społeczne. Stąd też naturalną jednostką analizy w badaniach nad ludzkim zachowaniem stają się systemy czynności, rozumiane jako historycznie uwarunkowane systemy relacji między jednostkami ludzkimi a ich najbliższym, kulturowo ustrukturyzowanym otoczeniem. Autorzy zestawiają wyniki prac radzieckich psychologów nurtu kulturowo-historycznego z wnioskami wspomnianego już Deweya, ale także Wilhelma Wundta i Hugo Münsterberga, dostrzegając w nich zbliżone sposoby konceptualizacji poznania jako zjawiska rozproszonego, w ramach których dochodzi do zjednoczenia nauk przyrodniczych i kulturowych (Cole i Engeström, 1993).

Dysponujemy już pewną wiedzą o stosunkowym nowatorstwie teorii Hutchinsa, powróćmy więc do pytania o to, jak ma się ona do szerszego dorobku nauk poznawczych. O stopniu osadzenia teorii poznania rozproszonego w klasycznej kognitywistyce dobre pojęcie daje książka *Cognition in the Wild*. Uwidacznia się to między innymi w specyficznym rozumieniu obliczania i reprezentacji: choć sama przyjmowana przez tego autora definicja poznania – czyli (przypomnijmy) obliczania realizowanego poprzez wytwarzanie, przetwarzanie i przekazywanie stanów reprezentacyjnych – bezkolizyjnie łączy omawianą perspektywę z klasycznymi podejściami w naukach poznawczych, to wątpliwości pojawiają się co do zakresu, jaki Hutchins przypisuje pojęciom reprezentacji oraz procesom obliczeniowym.

Trudno bowiem uznać reprezentacje zewnętrzne (znaki, ilustracje czy gesty) za po prostu równoważne wewnętrznym reprezentacjom umysłowym. Podobnie brakuje powszechnej zgody na to, by uznać zbiorową czynność kolektywu współpracującego przy rozwiązywaniu wspólnego problemu za czynność poznawczą.

Autor *Cognition in the Wild* wskazuje na pewne nieporozumienia związane z pojęciami stosowanymi w kognitywistyce, a także postuluje reinterpretację jej historii. U podstaw nieporozumień – historycznie wzmocnionych – leży, jego zdaniem, podstawowe niezrozumienie podobieństw i różnic między komputerami a umysłami.

Dla Hutchinsa idea procesu obliczeniowego jest podstawowa nie tylko dla poznania, lecz także kultury; opanowanie systemów formalnych to wręcz klucz do rozumienia nowoczesnej cywilizacji (Hutchins, 1995a, s. 359–360). Jednak to nie tyle umysł ludzki, ile większy wycinek ludzkiego świata kulturowego miałby być komputerem: „Architektura systemu fizycznych symboli nie stanowi modelu indywidualnego poznania, ale model działania systemu społeczno-kulturowego, z którego usunięto indywidualny ludzki podmiot” (Hutchins, 1995a, s. 363). Analogicznie Hutchins traktuje system Searle’owskiego chińskiego pokoju jako społeczno-kulturowy system poznawczy, którego właściwości wykraczają poza możliwości poznawcze ludzkiej jednostki. Searle przedstawia następujący eksperyment myślowy (Searle, 1995). W pewnym pomieszczeniu – zwanym „chińskim pokojem” – znajduje się człowiek, posługujący się skomplikowanym podręcznikiem przekładu z języka chińskiego na angielski. Sam jednak w ogóle nie zna chińskiego. Przez otwór do pokoju dostaje pytanie w języku chińskim. Korzystając z angielskiego podręcznika i wertując go, odpowiada na pytanie, też po chińsku. Czy powiedzielibyśmy, że rozumie po chińsku? Nie! Dlatego też, powiada Searle, nie ma podstaw, aby przypisywać komputerom rozumienie języka naturalnego. Searle jed-

nak nie rozpoznaje złożoności rozproszonego systemu poznawczego, który sam opisuje w swoim eksperymencie myślowym. Podobnie, zgodnie z fałszywą (zdaniem Hutchinsa) ideą, że komputer wykonany jest na obraz i podobieństwo człowieka, dziesiątki lat historii kognitywistyki można postrzegać jako dążenie do przerobienia obrazu człowieka na obraz komputera (Hutchins, 1995a, s. 361–363). Przy takim podejściu ciało, emocje, kultura musiały stać się elementami dodanymi, uwzględnianymi dopiero na późniejszym etapie dociekań.

Autor *Cognition in the Wild* zakłada, że ludzie rzeczywiście przetwarzają wewnętrzne reprezentacje symboli. Ale nie wierzy w to, by to manipulacja symbolami stanowiła istotę indywidualnego poznania. Pylyshynowi, który opisuje manipulowanie symbolami zakodowanymi w koralikach liczydła (Pylyshyn, 1989, s. 56), Hutchins zarzuca zignorowanie tego, co robi liczący, abstrahowanie od funkcji jego rąk i oczu, od istoty samego uczenia się, na rzecz skupienia się na właściwościach systemu, który wyłania się dzięki aktywności osoby manipulującej koralikami. Oczywiście Pylyshyn dobrze opisuje obliczeniowe właściwości systemu społeczno-kulturowego, lecz nie aktywność poznawczą podmiotu ludzkiego. Hutchins domaga się podejścia ucieleśnionego, to znaczy uwzględniającego konkretny, fizyczny udział osobnika w czynnościach liczenia i uczenia się, gdyż dopiero wtedy można mówić o jego użytkowaniu przyrządu do rozwiązywania problemów oraz o nabywaniu określonych umiejętności. W tych dociekaniaх konieczne jest usytuowanie podmiotu poznającego w realnym świecie. Hutchins przyznaje, że klasyczni zwolennicy hipotezy *systemu fizycznych symboli* są tego świadomi. Allen Newell i Herbert Simon potwierdzają istotność tak zwanych pamięci zewnętrznych – a więc materialnych zasobów dostępnych w otoczeniu podmiotu rozwiązującego dany problem poznawczy i wykorzystywanych przezeń równolegle z pamięcią wewnętrzną (Newell i Simon, 1972,

s. 800–803). Badacz dostrzega dobry punkt wyjścia, zauważa jednak, że przez dziesiątki lat od wydania przytaczanej tu książki Newella i Simona wykorzystanie materialnych struktur otoczenia w procesach poznawczych jakoś nie interesowało głównego nurtu kognitywistyki (Dahlbäck i Kristiansson, 2016; Hutchins, 1995a, s. 364–370).

Badania Hutchinsa wpisują się w dosyć swoiście pojmowaną naturalizację badań społecznych – a mianowicie pozwalają spełnić wymóg naturalizmu metodologicznego, głoszący, że w badaniach empirycznych dopuszczalne są wszystkie i tylko te metody, które są typowe dla nauk przyrodniczych. Za taką naukę przyrodniczą uchodzić ma, rzecz jasna, antropologia kognitywna, integrująca się również z pozostałymi elementami konglomeratu kognitywistycznego, w tym także do pewnego stopnia z neuronauką w ramach neuroantropologii (Lende i Downey, 2012). Jest tak dlatego, że Hutchins bardzo swoiście rozumie też antropologię poznawczą: obejmuje ona badanie kultury metodami, które proponowano do wyjaśniania procesów rozwiązywania problemów w psychologii poznawczej. Te wyjaśnienia odwołują się m.in. do operowania symbolami.

Hutchins odwołuje się do bardzo wpływowej propozycji metodologicznej, sformułowanej przez Davida Marra (1982), a powszechnie przyjmowanej w badaniach kognitywistycznych. Marr uważał, że systemy obliczeniowe należy wyjaśniać na trzech poziomach ich realizacji, ponieważ dopiero wyjaśnienie wszystkich poziomów jest kompletne i satysfakcjonujące. Przede wszystkim należy wiedzieć, jakie zadanie – opisywane w kategoriach przetwarzania informacji – realizuje dany system i dlaczego realizacja tego zadania jest właściwa w danym otoczeniu. Ten poziom Marr zwał „obliczeniowym”, lecz jest to termin nieco mylący; inni stosowali określenia „poziom wiedzy” (Newell, 1981) „poziom semantyczny” (Pylyshyn, 1984) lub „ekologiczny” (Sterelny, 1990). W przypadku określenia położenia

statku wejściowe informacje dostępne są z wielu, często dających rozbieżne wyniki, przyrządów pomiarowych, których obsługa wymaga współdziałania wielu osób. Rozwiązaniem ma być niezbędna w żegludze para współrzędnych nałożona na mapę.

Na drugim poziomie – kluczowym dla Marra – opisuje się algorytmy i reprezentacje, na których owe algorytmy pracują. Innymi słowy, opisuje się system obliczeniowy jako urzeczywistniający precyzyjnie opisane procesy przetwarzania informacji, których postać jest dokładnie określona. Tu Hutchins abstrakcyjnie opisuje strukturę systemu – odzwierciedlającą podział pracy poznawczej – i poszczególne składowe reprezentacje oraz kanały przesyłania informacji (np. telefon na pokładzie).

Na trzecim zaś poziomie Marr domaga się wskazania realizacji sprzętowej danego systemu obliczeniowego. Hutchins wskazuje w szczególności na ograniczenia materialne interakcji społecznej, dzięki którym działania są lepiej koordynowane, a sekwencja działań mająca prowadzić do ustalenia położenia w danym momencie jest wykonywana bez zakłóceń. Pozostaje on w dużej mierze neutralny w sprawie tego, jak mają się działania poszczególnych osób, będących częściami dużego, rozproszonego systemu poznawczego, do ich indywidualnych zdolności poznawczych. Uważa, że właściwe wyjaśnienie systemu rozproszonego nie musi schodzić na taki niższy poziom, a zjawisko można wyjaśnić trafnie na bardziej abstrakcyjnym poziomie opisu struktury społeczno-kulturowej wraz z jej materialną realizacją.

Przy takim podejściu widać od razu, że Hutchins korzysta z narzędzia, które pozwala mu modelować podział pracy poznawczej, i to także z uwzględnieniem skuteczności tej pracy. W tym miejscu tematyka jego prac wprost łączy się z klasycznym wątkiem badań socjologicznych, a metoda badawcza – opis antropologiczny – zbliża

go do badań społecznych prowadzonych w mikroskali, lecz jednocześnie metodami umożliwiającymi matematyzację. Tu Hutchins różni się też od innych antropologów poznawczych, którzy, w odróżnieniu od większości kognitywistów korzystających w ogromnej mierze z modelowania obliczeniowego, rzadko odwołują się do symulacji komputerowych, a prawie nigdy nie przeprowadzają symulacji realnych procesów społeczno-poznawczych w celu ich wyjaśniania i przewidywania.

Hutchins między innymi opracował symulacje zjawiska zwanego „błędem potwierdzania”. Polega ono na tym, że skłonni jesteśmy uznawać za prawdziwą tę informację, która potwierdzałyby wcześniejsze założenia lub oczekiwania. Badacz pokazał, że zjawisko to może zachodzić w grupie w formie procesu rozproszonego; wykorzystał w tym celu sieć koneksjonistyczną zwaną „siecią spełniania ograniczeń”, uzyskując uproszczony model grupowej interakcji. Przy okazji zaobserwował, że błąd potwierdzania zachodzi tym słabiej, im większa jest komunikacja (interakcja) między jednostkami w sieci (Hutchins, 1995a, s. 239–255).

Podział pracy poznawczej stanowi istotę działania wielu organizacji i instytucji społecznych, a próby udoskonalenia ich działania – jeśli same nie adaptują się do otoczenia – mogą z pewnością skorzystać z bardziej kognitywnego podejścia. Dzięki zastosowanemu tutaj podejściu Hutchinsa można opisać przebieg procesu przetwarzania informacji i przeanalizować jego efektywność ze względu na założony cel poznawczy. Możliwa jest więc poznawcza analiza społecznych procesów podejmowania decyzji, tak politycznych, jak i naukowych czy potocznych.

Nie oznacza to wszelako, że metoda obrana przez Hutchinsa jest idealna. Metodologia Marra została opisana przezeń jedynie we wstępie metodologicznym do jego wpływowej książki i nie odzwierciedla

zacierania się granic między kognitywistyką a neurokognitywistyką, a także między badaniem procesów kulturowych i poznawczych. Badacze interesują się bowiem często tym, jakie mechanizmy odpowiadają za procesy psychiczne, i starają się zawęzić swoje hipotezy, odwołując się do świadectw neurobiologicznych czy kulturowych. Marr, *notabene*, też tak czynił, kiedy zwracał uwagę na rozłączne pojawianie się pewnych zaburzeń poznawczych przy ogniskowych uszkodzeniach mózgu. Jednak co do zasady uważał, że nie jest to bynajmniej konieczny element metody wyjaśniania, a jego opis poziomu realizacji sprzętowej (neuronalnej) pozostaje, podobnie nieco jak i u Hutchinsa, osobliwie nieprecyzyjny.

Skupienie się Hutchinsa na budowie i morfologii systemów poznawczych na różnych poziomach ich podsystemów, a przez to na udziale artefaktów w ludzkich procesach poznawczych oraz ogólnie funkcjonowaniu ekosystemów kulturowych, wykazuje duży związek z pracami Donalda Normana, zwłaszcza z książką *Things that Make us Smart* (1993). Pojawia się w niej zresztą pojęcie i wyjaśnienie poznania rozproszonego oraz wyników badań Hutchinsa jeszcze przed publikacją *Cognition in the Wild*.

Miejsce i recepcja Normana w kognitywistyce wymaga kilku słów wyjaśnienia. Ten amerykański kognitywista i teoretyk wzornictwa przemysłowego, a jednocześnie najbardziej chyba wpływowy badacz interakcji człowiek-komputer (HCI), jak i szerzej: człowiek-artefakt, uprawia swoistą inżynierię poznawczą. Zainteresowany jest nie tyle rozwijaniem teorii znanych już interakcji człowiek-artefakt, ile projektowaniem tych interakcji, ich kulturowym zastosowaniem (i tym samym badaniem przez projektowanie). To nastawienie bardzo dobrze współgra z równie praktyczną orientacją Hutchinsa, konsekwentnie odwołującego się do kulturowego wykorzystywania rzeczy.

4. Szkodliwe moratorium

Perspektywą najbardziej zbliżoną do perspektywy Hutchinsa jest nurt społecznych badań nad nauką i techniką, reprezentowany między innymi przez Brunona Latoura, do którego prac Hutchins zresztą wyraźnie się odwoływał.

W klasycznych badaniach nad poznaniem i mózgiem zagadnienia społeczne w poznaniu sprowadza się zasadniczo do procesu społecznego wiążanego z nabywaniem i przetwarzaniem informacji o innych ludziach, komunikowaniem się, empatią. Ścieżka analizy wiedzy tutaj od jednostkowych umysłów czy mózgów do społeczeństwa (zob. np. Fiske i Taylor, 2008).

Tymczasem – jak pokazują Ronald Giere i Barton Moffatt – perspektywa naukoznawcza ukazuje wspomnianą relację jako opozycję, w pewnych przypadkach wskazując na dominację wymiaru społecznego, co byłoby pokłosiem socjologii wiedzy (zob. także Brown, 2011; Giere i Moffatt, 2003, s. 301–302). Wyzywającą manifestacją tej dominacji było dziesięcioletnie moratorium na odwoływanie się do umysłu w wyjaśnianiu sukcesu nauki, które zaproponował Latour ze Steve'em Woolgarem. Rzucali oni wyzwanie kognitywistycie wyjaśniającej efektywność nauki przez odwołanie do umysłu, ujmowanego jeszcze w sposób odcieleśniony i oderwany od kontekstu społecznego i materialnego: „My zaś w tym miejscu obiecujemy, że jeśli pod koniec tego okresu zostanie cokolwiek [związanego z poznaniem naukowym] do wyjaśnienia, to my także zwrócimy się w stronę umysłu!” (Latour i Woolgar, 2020, s. 366).

W badaniach nad poznaniem naukowym Hutchins wydawałby się doskonałym partnerem Latoura; wszak analizował czynności badaczy w laboratorium neuronaukowym (Alač i Hutchins, 2004). Jego uwagi

dotyczące wizualizacji naukowych wpisują się w perspektywę Latoura (np. Latour, 2012). Prześledźmy ich pouczający spór (Latour, 1996; Hutchins i Klausen, 1996).

Latour jest badaczem społecznym i filozofem nauki, kojarzonym z konstruktywistyczną teorią aktora-sieci (ANT), która przypisuje sprawczość czynnikom ludzkim i pozaludzkim, znajdującym się w zmiennych (czasowo wzmacnianych lub też słabnących) sieciach relacji (zob. np. Latour, 2010). „W szesnastym wieku nie pojawił się żaden nowy człowiek, a we współczesnych laboratoriach nie pracują wcale mutanci o rozrośniętych mózgach” – stwierdza, namawiając do szukania przyczyn spektakularnych sukcesów nauki w praktykach i narzędziach badaczy (Latour, 2012, s. 208). Analogicznie Hutchins – co nie uszło uwadze Latoura – każe szukać źródeł ludzkiej mocy poznawczych w przetworzonym już otoczeniu i użytkowanych artefaktach, krytykując popularne przekonanie o zaawansowanej technice jako konsekwencji zaawansowanych zdolności poznawczych czy umysłów (Hutchins, 1995a, s. 169, 355).

Latour dostrzega w *Cognition in the Wild* kompleksowe ujęcie praktyk poznawczych w ich naturalnym kontekście. Hutchins pokazuje, jak owe praktyki mogą być dobrze skoordynowane, precyzyjne i zdyscyplinowane, co pozwala na sformalizowaną analizę, mimo iż przeprowadzaną poza przestrzenią laboratorium. Badacz jawi się tutaj jako jeden z antropologów (obok Jean Lave, Lucy Suchman czy Chucka Goodwina), którzy w swojej karierze badali zarówno społeczności tradycyjne o nisko zaawansowanej technice, jak i silnie zmodernizowane. Przewaga Hutchinsa nad nim polega na wykorzystaniu formalnego rozumowania, metrologii i matematyki. Latour docenia wagę i radykalność głównej tezy książki, która w jego ujęciu określa opisywane tam zjawisko poznania jako niemające „nic wspólnego z umysłami czy jednostkami ludzkimi, lecz z rozprzestrze-

nianiem reprezentacji przez różnorodne ośrodki, koordynowane przez dość słabo wyposażony ludzki podmiot działający w grupie, w ramach jakiejś kultury, z wykorzystaniem artefaktów, mogący internalizować pewną część tego procesu” (Latour, 1996, s. 56). Według Latoura teza ta mogłaby zreformować całość nauk poznawczych. Jedną z radykalnych zmian to przekierowanie uwagi z umysłowych czy indywidualnych czynności na przebieg zmian reprezentacji. W takich badaniach tracą znaczenie sformułowania typu „myśle” lub „reprezentuję”; nie ma sensu pytać, co się dzieje w umyśle kreślarza na statku podczas wykonywania zbiorowej pracy nawigacyjnej. Latour podkreśla konsekwencje tego odwrócenia od zdarzeń umysłowych i zwrotu ku ośrodkom: docenienie roli techniki i zapośredniczających artefaktów, które nie służą wzmacnianiu ludzkich zdolności poznawczych, lecz odkrywaniu nowych problemów poznawczych.

Zdaniem Latoura, „pod skórą” indywiduum przebiegają te same procesy, co na zewnątrz. Myślący podmiot stanowi dla Hutchinsa szczególnego rodzaju ośrodek umożliwiający koordynację wielu innych ośrodków, zewnętrznych, wewnętrznych, w postaci artefaktów, idei czy relacji społecznych. W stwierdzeniu tym Latour widzi ostateczny rozpad psychologii. Myślący podmiot przypomina tutaj coś w rodzaju białego biurka dobrze zorganizowanego menedżera: jest pusty, odkąd wszystko inne zostało oddelegowane na zewnątrz do czegoś lub kogoś innego (Latour, 1996, s. 59). Nic więc dziwnego, że Latour z aprobatą powtarza przekonanie Hutchinsa, że komputery, „ukochany model umysłu” przedstawicieli klasycznej perspektywy, mogą podsuwać dobry opis funkcjonowania systemów społeczno-kulturowych, ale nie ludzkich procesów poznawczych.

Na tym aprobatą Latoura się jednak kończy. Ku jego rozczarowaniu Hutchins okazuje się niekonsekwentny, liczy bowiem na psychologię lepiej ugruntowaną empirycznie – podczas gdy, zda-

niem Latoura, stara się ona jedynie wślaczać kategorie poznawcze do indywidualnych umysłów obdarzonych świadomością i odpowiedzialnością. Równie niepokojącą dla niego ideą przewijającą się przez książkę jest założenie, że po jednej stronie (tego, co dane) mamy świat, a po drugiej zdolności poznawcze, co niekoniecznie zgadza się z ideą rozproszenia (wedle Latoura). Hutchins pisze o procedurze bezpośredniego uzgadniania mapy ze światem – a przecież operatorzy pelorusa uzgadniają nie świat z mapą, tylko odczyty z kompasu z istniejącymi już punktami orientacyjnymi na mapie, w czym istotną rolę odgrywają zapośredniczenia, o których autor najwyraźniej tutaj zapomina (Latour, 1996, s. 60).

Latour walczy z mitem wyższych zdolności poznawczych i lepiej rozwiniętych cywilizacyjnie umysłów: „Myślą laboratoria, odkrywają wspólnoty, rozwijają się dziedziny, spostrzegają przyrządy, a nie jednostkowe umysły” (Latour, 1996, s. 61). Jego zdaniem, pogląd ten został już należycie wyrażony i sprecyzowany przez naukowców, z czego Hutchins nie zdaje sobie sprawy i dlatego nie usuwa indywiduum z pola widzenia w poznaniu rozproszonym.

Hutchins w odpowiedzi zauważa, że Latour chciałby rozpuścić indywidualny podmiot i psychologię indywiduum. Sprzątnąć, wymieść to, co indywidualne, z podmiotu, niczym z biurka, jak w przytoczonym wcześniej porównaniu. Twórca teorii aktora-sieci chciałby oddelegować wszystko „na zewnątrz do czegoś lub kogoś innego” – tylko kogo, skoro według niego żadnego podmiotu w ogóle nie ma? Wbrew Latourowi ambicją Hutchinsa nie jest wyprzątnięcie wszystkiego z podmiotu ludzkiego, tylko zespolenie tego, co w nim, z jego otoczeniem. Zdarzenia poznawcze muszą gdzieś się odbywać, a częściowo odbywają się wewnątrz podmiotu. Choć fragment książki o myślącym podmiocie jako specjalnym ośrodku koordynującym Latour przytacza na poparcie swojej wizji rozpuszczenia psychologii

indywidualizm, Hutchins stawia w nim pytanie, jakiego rodzaju ośrodkiem jest myślące indywiduum – na które odpowiedzią byłaby nowa, adekwatna do aktualnych wyzwań badawczych teoria psychologiczna. Dzięki niej umysł ujawni się nam nie jako „uprzątnięty”, lecz jako „umeblowany” inaczej, niż wcześniej sądzono. A kwestia, jak dobrze opisywać oraz modelować takie systemy, pozostaje trudna (Hutchins, 1996, s. 64–66, 68).

Jaki jest stosunek Latoura do interdyscyplinarności, czy też szerzej – „wzajemnego oświecania się” dyscyplin? Z jednej strony, uderza on w mit odrębnych, autonomicznych dyscyplin. Z drugiej strony, nie najlepiej ocenia dotychczasowe (a więc podejmowane przed Hutchinsem) próby współpracy psychologii i nauk społecznych, określając na przykład psychologię społeczną jako monstrum, które skupia to, co najgorsze, z obu dyscyplin (Latour, 1996, s. 58). Nie można mu rzeczywiście odmówić, że na różne sposoby wyraża wsparcie dla zasypywania granic czy też likwidacji napięcia w podejściach antropologii i kognitywistyki, prowadzające się do dylematu, czy poznanie jest konstytuowane przez kulturę, czy też jest od kultury niezależne. I w tym kontekście zasadniczo komplementuje właśnie Hutchinsa za wyczulenie na ten problem i pracę na rzecz uzgodnienia nauk przyrodniczo-ścisłych z badaniami nad poznaniem „in the wild”. Recenzent *Cognition in the Wild* chciałby bardzo, jak twierdzi, aby odnowiona antropologia poznawcza „współpracowała blisko z tymi przedstawicielami nauk przyrodniczo-ścisłych, którzy mają takie same zainteresowania i wielu tych samych wrogów” (Latour, 1996, s. 61). Cóż z tego jednak, skoro Latour wydaje się nie akceptować samego przyjmowanego w kognitywistyce rozumienia tego, co poznawcze. Jego moratorium zaś miało być także wyrazem sprzeciwu wobec „luźnego i łatwego” podejścia epistemologów do wyjaśnień poznawczych (Latour, 1996, s. 62). Innymi słowy, nie ma zgody

Latoura na to, jak Hutchins rozumie dalej to, co jest przetwarzane obliczeniowo, z jednego medium reprezentacyjnego do innego. Po tym, jak Hutchins zredefiniował poznanie w kategoriach koordynacji mediów reprezentacyjnych, jego obowiązkiem byłoby określenie różnych trybów koordynacji. Zamiast tego używa on bardzo słabych i wieloznacznych metafor, wyczuwając trudność, ale unikając jej: określa ludzi jako „klej spajający sprzęt”, a związek przyczynowy dostrzega w „tkance relacji międzyludzkich”, w której osoby zaangażowane w czynności nawigacyjne „ograniczają” nawzajem swoje działania na rzecz przetwarzania stanów reprezentacyjnych (Latour, 1996, s. 61; zob. Hutchins, 1995a, s. 202).

Różnice między Latourem a Hutchinsem można do pewnego stopnia sprowadzić do różnic między etnografią klasyczną a poznawczą, które wyznaczają nie tyle stosowane metody badawcze – bo te nierzadko są identyczne – ile ich konceptualizacja (Williams, 2006, s. 838–839).

Powyższa konfrontacja pomaga uzmysłwić ryzyko kreślenia zbyt dużych analogii między teorią rozproszonych systemów poznawczych a teorią aktora-sieci czy też bezmyślnego ich wiązania. Spór z Latourem ma pomóc w stwierdzeniu, czy i jak nauki społeczne i kognitywistyka mogą współpracować. Giere i Moffatt, analizując z kolei prace Latoura w ramach teorii poznania rozproszonego, postulują łączenie aspektu poznawczego i społecznego w badaniach dlatego, że – jak twierdzą – nie możemy zrozumieć, jak naukowcy wspólnie realizują dane zadanie poznawcze bez scharakteryzowania interakcji społecznych między nimi (Giere i Moffatt, 2003, s. 4–8). I może o taką komplementarność celów – bardziej niż o cel wspólny – we współczesnej nauce chodzi.

5. Podsumowanie

W artykule skupiliśmy się na najbardziej wpływowym okresie działalności badawczej Hutchinsa, w którym mocno uwidoczniły się wpływy nauk społecznych. Przede wszystkim zmienna jest polemika z Latourem, naszym zdaniem doskonale ujmująca najważniejsze zgodności i różnice między autorem *Cognition in the Wild* a twórcą ANT. Był to okres formułowania i efekownego testowania koncepcji poznania szerokiego, nie bez związków z dużą w tym czasie aktywnością Normana w badaniach nad artefaktami i projektowaniem. Przedstawiliśmy w skrócie również etap późniejszy prac Hutchinsa, znaczony pojęciem ekologii poznawczej. Całokształt tych prac wymownie oddaje – charakterystyczne dla badań nad poznaniem rozproszonym i w pełni tutaj konsekwentne – odejście od indywidualizmu metodologicznego.

Hutchins w swoich badaniach wyraźnie pokazuje, że rozwiązywanie złożonych problemów wymaga podziału pracy poznawczej: rozproszone procesy poznawcze obejmują wytwarzane i przekazywane dalej materialne nośniki reprezentacji zewnętrznych, ich przekształcenia, często wiele artefaktów o różnych rolach obliczeniowych oraz żywe organizmy. Procesy poznawcze bada się w naturalnym ich otoczeniu, a nie tylko w uproszczonych warunkach laboratorium.

Dzięki pracom Hutchinsa, przy wydatnym wpływie Normana, Kirsha, Nersessian oraz późniejszych pokoleń badaczy tego nurtu, w repertuarze pojęciowym kognitywistyki zaistniały na trwałe lub nabrały znaczenia takie kategorie, jak „poznanie rozproszone” „szeroki [czy rozproszony] system poznawczy”, „artefakt poznawczy”, „reprezentacja zewnętrzna”, „interakcja człowiek–artefakt”. Co istotne, pojęcia te nie wykluczają się z siatką pojęciową klasycznego nurtu

kognitywistyki, czego nie można powiedzieć o wszystkich pozostałych nurtach badań poznawczych zrywających z indywidualizmem metodologicznym.

Poznanie rozproszone to więc bodaj najbliższa naukom społecznym koncepcja kognitywistyczna. Podczas gdy wcześniej badania w psychologii poznawczej i w kognitywistyce w ogóle koncentrowały się w dużej mierze na procesach poznawczych zachodzących u pojedynczych osób, biernie poddających się manipulacjom eksperymentalnym, obecnie coraz silniej do głosu dochodzą nurty, w których uwzględnia się nie tylko rolę materialnego otoczenia, ale także cielesności i społeczno-kulturowych determinant poznania. Ułatwia to znacznie dialog kognitywistyki z socjologią, nie mówiąc o antropologią, która ma szansę płynnego łączenia się z innymi badaniami poznawczymi. Krótko mówiąc, Hutchinsa obrona interdyscyplinarności – wbrew Latourowemu moratorium i wbrew pesymizmowi Núñeza – skutecznie przyczynia się do współpracy ponad podziałami między tzw. dyscyplinami naukowymi.

Bibliografia

- Afeltowicz, Ł., 2012. *Modele, artefakty, kolektywy: praktyka badawcza w perspektywie współczesnych studiów nad nauką*, seria: *Monografie Fundacji na Rzecz Nauki Polskiej*. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- Afeltowicz, Ł. i Wachowski, W., 2015. How far we can go without looking under the skin: the bounds of cognitive science. *Studies in Logic, Grammar and Rhetoric* [Online], 40(1), ss. 91–109. <https://doi.org/10.1515/slgr-2015-0005>.

- Alač, M. i Hutchins, E., 2004. I see what you are saying: action as cognition in RMRI brain mapping practice. *Journal of Cognition and Culture* [Online], 4(3-4), ss. 629–661. <https://doi.org/10.1163/1568537042484977>.
- Anderson, J.R. i Lebiere, C., 2003. The Newell Test for a theory of cognition. *Behavioral and Brain Sciences* [Online], 26(5), ss. 587–601. <https://doi.org/10.1017/S0140525X0300013X>.
- Bender, A., Hutchins, E. i Medin, D., 2010. Anthropology in cognitive science. *Topics in Cognitive Science* [Online], 2(3), ss. 374–385. <https://doi.org/10.1111/j.1756-8765.2010.01082.x>.
- Brown, M.J., 2011. Science as socially distributed cognition: Bridging philosophy and sociology of science. W: B.v. Kerkhove, K. François, B. Löwe i T.W. Müller, red. *Foundations of the Formal Sciences VII: Bringing Together Philosophy and Sociology of Science*, seria: *Studies in Logic*, 32. London: College Publications, ss. 17–31.
- Clark, A. i Chalmers, D.J., 2008. Umysł rozszerzony. W: M. Miłkowski i R. Poczubut, red. *Analityczna metafizyka umysłu: najnowsze kontrowersje* (M. Miłkowski, tłum.), seria: *Umysł: prace z filozofii i kognitywistyki*. Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii Polskiej Akademii Nauk, ss. 342–357.
- Cole, M. i Engeström, Y., 1993. A cultural-historical approach to distributed cognition. W: G. Salomon, red. *Distributed Cognitions: Psychological and Educational Considerations*. 1 wyd., seria: *Learning in doing*. Cambridge; New York: Cambridge University Press, ss. 1–46.
- Dahlbäck, N. i Kristiansson, M., 2016. A perspective on all cognition? A study of everyday environments from the perspective of distributed cognition. W: A. Papafragou, D. Grodner, D. Mirman i J. Trueswell, red. *CogSci 2016 Proceedings* [Online]. Philadelphia, Pennsylvania: Cognitive Science Society, Inc., ss. 734–739. Dostępne na: <<https://cogsci.mindmodeling.org/2016/papers/0137/paper0137.pdf>> [ostatni dostęp: 19 września 2022].
- Fiske, S.T. i Taylor, S.E., 2008. *Social Cognition: From Brains to Culture*. Boston [etc.]: McGraw-Hill Higher Education.

- Fodor, J.A., 1980. Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences* [Online], 3(1), ss. 63–73. <https://doi.org/10.1017/S0140525X00001771>.
- Gardner, H., 1985. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.
- Giere, R.N. i Moffatt, B., 2003. Distributed Cognition: Where the Cognitive and the Social Merge. *Social Studies of Science* [Online], 33(2), ss. 301–310. <https://doi.org/10.1177/03063127030332017>.
- Heath, J., 2020. Methodological Individualism. W: E.N. Zalta, red. *The Stanford Encyclopedia of Philosophy* [Online]. Summer 2020. Stanford, Calif.: Metaphysics Research Lab, Stanford University. Dostępne na: <<https://plato.stanford.edu/archives/sum2020/entries/methodological-individualism/>> [ostatni dostęp: 19 września 2022].
- Hess, D., 2001. Ethnography and the Development of Science and Technology Studies. W: P. Atkinson i in., red. *Handbook of Ethnography* [Online]. London: SAGE Publications Ltd, ss. 234–245. <https://doi.org/10.4135/9781848608337.n16>.
- Hutchins, E., 1995a. *Cognition in the Wild*. Cambridge, MA; London: The MIT Press.
- Hutchins, E., 1995b. How a cockpit remembers its speeds. *Cognitive Science* [Online], 19(3), ss. 265–288. https://doi.org/10.1207/s15516709cog1903_1.
- Hutchins, E., 1996. Response to Reviewers. *Mind, Culture, and Activity* [Online], 3(1), ss. 64–68. https://doi.org/10.1207/s15327884mca0301_6.
- Hutchins, E., 2001. Cognition, Distributed. W: N.J. Smelser i P.B. Baltes, red. *International Encyclopedia of the Social & Behavioral Sciences* [Online]. Oxford: Pergamon, ss. 2068–2072. <https://doi.org/10.1016/B0-08-043076-7/01636-3>.
- Hutchins, E., 2005. Material anchors for conceptual blends. *Journal of Pragmatics* [Online], 37(10), ss. 1555–1577. <https://doi.org/10.1016/j.pragma.2004.06.008>.
- Hutchins, E., 2006. The Distributed Cognition Perspective on Human Interaction. W: S.C. Levinson i N.J. Enfield, red. *Roots of Human Sociality*. 1 wyd. Oxford: Berg Publishers, ss. 375–398.

- Hutchins, E., 2008. Cognitive ecology. *CogSci 2008. 30th Annual Meeting of the Cognitive Science Society* [Online]. Dostępne na: <<http://csjarchive.cogsci.rpi.edu/proceedings/2008/AnniversarySymposia/Hutchins.pdf>> [ostatni dostęp: 7 stycznia 2017].
- Hutchins, E., 2010a. Cognitive ecology. *Topics in Cognitive Science* [Online], 2(4), ss. 705–715. <https://doi.org/10.1111/j.1756-8765.2010.01089.x>.
- Hutchins, E., 2010b. Enaction, imagination, and insight. W: J. Stewart, O. Gappenne i E.A. Di Paolo, red. *Enaction: Toward a New Paradigm for Cognitive Science* [Online]. Cambridge, MA: The MIT Press, ss. 425–450. <https://doi.org/10.7551/mitpress/9780262014601.003.0016>.
- Hutchins, E., 2011. Enculturating the supersized mind. *Philosophical Studies* [Online], 152(3), ss. 437–446. <https://doi.org/10.1007/s11098-010-9599-8>.
- Hutchins, E., 2014. The cultural ecosystem of human cognition. *Philosophical Psychology* [Online], 27(1), ss. 34–49. <https://doi.org/10.1080/09515089.2013.830548>.
- Hutchins, E., 2022. *Hutchins Home Page* [Online]. Dostępne na: <<https://pages.ucsd.edu/~ehutchins/>> [ostatni dostęp: 19 września 2022].
- Hutchins, E. i Klausen, T., 1996. Distributed Cognition in an Airline Cockpit. W: Y. Engeström i D.J. Middleton, red. *Cognition and Communication at Work*. 1 wyd. Cambridge; New York; Melbourne: Cambridge University Press, ss. 15–34.
- Kaidesoja, T., Sarkia, M. i Hyryläinen, M., 2019. Arguments for the cognitive social sciences. *Journal for the Theory of Social Behaviour* [Online], 49(4), ss. 480–498. <https://doi.org/10.1111/jtsb.12226>.
- Kirsh, D., 1999. Distributed Cognition, Coordination and Environment Design. *Proceedings of the European Cognitive Science Society* [Online], ss. 1–11. Dostępne na: <<https://pdfs.semanticscholar.org/97df/530eeb9c0133e1bd1282658cbd7c8b905949.pdf>> [ostatni dostęp: 19 września 2022].
- Kronenfeld, D.B., Bennardo, G., Munck, V.C.d. i Fischer, M.D., red., 2011. *A Companion to Cognitive Anthropology*, seria: *Blackwell companions to anthropology*, 16. Chichester, West Sussex; Malden, MA: Wiley-Blackwell.

- Latour, B., 1996. Cogito ergo sumus! or psychology swept inside out by the fresh air of the upper deck... *Mind, Culture, and Activity*, 3(1), ss. 54–63.
- Latour, B., 2010. *Splatając na nowo to, co społeczne: wprowadzenie do teorii aktora-sieci* (A. Derra i K. Abriszewski, tłum.), seria: *HN Horyzonty Nowoczesności*, 72. Kraków: Universitas.
- Latour, B., 2012. Wizualizacja i poznanie: zrysowywanie rzeczy razem (A. Derra i M. Frąckowiak, tłum.). *Avant*, 3(T) (A. Derra i M. Frąckowiak, tłum.), ss. 207–257.
- Latour, B. i Woolgar, S., 2020. *Życie laboratoryjne: konstruowanie faktów naukowych* (K. Abriszewski, P. Gąska, M. Smoczyński i A. Zabielski, tłum.). 1 wyd., seria: *Myśl o Kulturze*. Warszawa: Narodowe Centrum Kultury.
- Lende, D.H. i Downey, G.J., red., 2012. *The Encultured Brain: An Introduction to Neuroanthropology*. Cambridge; London: MIT Press.
- Luriâ, A.R., 1979. *Âzyk i soznanie*. Moskva: Izdatel'stvo Moskovskogo Universiteta.
- Marr, D., 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W. H. Freeman i Company.
- Mead, G.H., 1975. *Umysł, osobowość i społeczeństwo* (Z. Wolińska, tłum.), seria: *Biblioteka Socjologiczna*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Miller, G.A., 2003. The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences* [Online], 7(3), ss. 141–144. [https://doi.org/10.1016/S1364-6613\(03\)00029-9](https://doi.org/10.1016/S1364-6613(03)00029-9).
- Newell, A., 1980. Physical Symbol Systems. *Cognitive Science* [Online], 4(2), ss. 135–183. https://doi.org/10.1207/s15516709cog0402_2.
- Newell, A., 1981. The Knowledge Level: Presidential Address. *AI Magazine* [Online], 2(2), ss. 1–20, 33. <https://doi.org/10.1609/aimag.v2i2.99>.
- Newell, A., 1990. *Unified Theories of Cognition*. 3rd printing. Cambridge, MA; London: Harvard University Press.
- Newell, A. i Simon, H.A., 1972. *Human Problem Solving*. Englewood Cliffs: Prentice-Hall.

- Norman, D.A., 1981. Twelve issues for cognitive science. *Cognitive Science* [Online], 4(1), ss. 1–32. [https://doi.org/10.1016/S0364-0213\(81\)80002-X](https://doi.org/10.1016/S0364-0213(81)80002-X).
- Norman, D.A., 1991. Cognitive Artifacts. W: J.M. Carroll, red. *Designing Interaction: Psychology at the Human-Computer Interface*, seria: *Cambridge series on human-computer interaction*, 4. Cambridge; New York: Cambridge University Press, ss. 17–38.
- Núñez, R. i in., 2019. What happened to cognitive science? *Nature Human Behaviour* [Online], 3(8), ss. 782–791. <https://doi.org/10.1038/s41562-019-0626-2>.
- Osbeck, L.M. i Nersessian, N.J., 2014. Situating distributed cognition. *Philosophical Psychology* [Online], 27(1), ss. 82–97. <https://doi.org/10.1080/09515089.2013.829384>.
- Pylyshyn, Z.W., 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. 1 wyd. Cambridge, MA; London: The MIT Press.
- Pylyshyn, Z.W., 1989. Computing in cognitive science. *Foundations of Cognitive Science*. Cambridge, MA: The MIT Press, ss. 51–91.
- Roberts, J., 1964. The self-management of cultures. W: W.H. Goodenough, red. *Explorations in Cultural Anthropology; Essays in Honor of George Peter Murdock*. New York, McGraw-Hill Book Co, ss. 433–454.
- Rogers, Y., 1997. *A Brief Introduction to Distributed Cognition* [Online]. Dostępne na: <<http://yvonnerogers.com/wp-content/uploads/2014/07/dcoq-brief-intro.pdf>> [ostatni dostęp: 19 września 2022].
- Salomon, G., red., 1993. *Distributed Cognitions: Psychological and Educational Considerations*. 1 wyd. Cambridge; New York: Cambridge University Press.
- Schumpeter, J.A., 1908. *Das Wesen und der Hauptinhalt der theoretischen Nationalökonomie*. Leipzig: Verlag von Duncker & Humblot.
- Searle, J.R., 1995. Umysły, mózgi i programy. W: B. Chwedeńczuk, red. *Filozofia umysłu*, seria: *Fragmenty Filozofii Analitycznej*, t. 2. Warszawa: Fundacja Aletheia; Wydawnictwo Spacja, ss. 301–324.
- Sterelny, K., 1990. *The Representational Theory of Mind: An Introduction*. Oxford; Cambridge: Basil Blackwell.

- Vygotskij, L.S., 1971. *Wybrane prace psychologiczne* (E. Fleszner i J. Fleszner, tłum.). Warszawa: Państwowe Wydawnictwo Naukowe.
- Wachowski, W., 2022. *Poznanie rozproszone: od heurystyk do mechanizmów*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- Weber, M., 2002. *Gospodarka i społeczeństwo: zarys socjologii rozumiejącej* (D. Lachowska, tłum.), seria: *Biblioteka Socjologiczna*. Warszawa: Wydawnictwo Naukowe PWN.
- Williams, R.F., 2006. Using Cognitive Ethnography to Study Instruction. *Proceedings of the 7th international conference on Learning sciences, ICLS '06*. Bloomington, Indiana: International Society of the Learning Sciences, ss. 838–844.

Review articles

Artykuły recenzyjne

Quantum mechanics of identical particles

Marek Kuś

Center for Theoretical Physics, Polish Academy of Sciences

Tomasz Bigaj, *Identity and Indiscernibility in Quantum Mechanics*, Palgrave MacMillan, Cham 2022, pp.XV+262.

Systems of identical particles pose a major ontological problem in quantum physics. In the classical world, the situation is much simpler. Here Leibniz's *principium identitatis indiscernibilium* states that two objects having exactly the same features are identical, so they are the same object. With its contraposition, stating that two identical objects have exactly the same features, they form one of the most important metaphysical principles, indispensable e.g. in the discussion about individuals (Guay and Pradeu, 2016). Although one can contest the universality of the Leibniz Principle even for macroscopic objects (Black, 1952), the real danger is posed by quantum mechanics. Here the existence of objects identical in every respect, nevertheless, in some sense, different is not only possible but even postulated. Thus all electrons are identical in every respect, and yet their number may be greater than one. It should be thus clear that the Leibniz Principle is violated in the quantum world (Cortes, 1976; Ginsberg, 1981; French and Redhead, 1988; Castellani and Mittelstaedt, 2000).

Let us be more precise. States of physical systems, such as elementary particles (electrons, protons, mesons) are described by vectors in a Hilbert space (in fact, to achieve unambiguity of such representation we should go to the appropriate projective space, however, it is not

important here). Physical quantities, such as energy, spin, etc., are represented by Hermitian operators acting in a given Hilbert space. Quantum mechanics is a statistical theory and to relate the formalism to reality we use average values. There is a precise recipe how, knowing the state of the system, to find the mean value of an observable to compare it with experimental result obtained by performing many runs of an experiment.

In case of more particles (or, in general, more components of the whole system) the appropriate Hilbert space is the tensor product of the single-particle spaces. Thus, to each of the particles there corresponds the number of its Hilbert space in the product, according to the assumed order of numbering the spaces. Thus, e.g., in the case of two particles the appropriate Hilbert space H_S of the system equals $H_1 \otimes H_2$. To refer to a single (say, first) component and measure its properties, we apply an observable of the form $A \otimes I$, where A is an operator acting in H_1 . This recipe works well when we deal with distinguishable particles possessing different physical properties, e.g. electron and proton creating a binary system in a hydrogen atom. But when all particles are of the same type¹ (e.g. they are all electrons), $H_1 = H_2 = H$ and, consequently, $H_S = H \otimes H$. However, quantum mechanics requires that the average value of each observable does not change with any exchange (permutation) of the components. This is the so-called “indistinguishability postulate”. Its requirements will be met by adopting the so-called “symmetry postulate”—the states of the whole system must be completely symmetric or completely anti-symmetric under the exchange of particles. Consequently, the set of states of a complex system is not the whole “available”

¹ This is the term used by the author of the book to avoid the terms “identical particles” or “indistinguishable particles” commonly used in physics, supposing in advance properties of these objects.

Hilbert space—the product of the spaces of individual particles, but its symmetric (antisymmetric) subspace. Whatever happens to such a complex system leaves it in this subspace. It follows that also observables have to be permutation invariant, otherwise the action may produce a state not belonging to the appropriate symmetric (antisymmetric) subspace. Still, at least in the orthodox approach to quantum mechanics, the numbers of spaces correspond to the numbers of individual particles, in other words, each particle is “attached” to its own Hilbert space. Such a postulate is called the “factorism postulate” (Caulton, 2014). A particular constituent (particle) is distinguished and individuated by applying a measurement that acts only in its space, but it is not allowed in light of the previous discussion; such an operator, like $A \otimes I$ in the two-particle case is not permutationally invariant.

Compatible with the Leibniz’s principle, would be thus a view that particles such as electrons cannot be treated as individuals in the sense to which we are attached in the physics of macroscopic systems described by classical mechanics, where all objects can be individuated by the totality of their properties (Redhead and Teller, 1992). However, for a physicist, treating quantum systems of particles of the same type in this way is not attractive. Results of many experiments concern “single particles”, or e.g. objects like “all electrons occupying a given energy shell in an atom”. It is far from clear in what sense such particles could not be individuals.

Tomasz Bigaj’s book (Bigaj, 2022) discusses the problems described above and does it in a masterly way. The main thread of the book is a discussion of whether particles of the same type, like electrons, protons, neutrons, etc., can be distinguished from one another using their physical properties. And if so, how to do it.

Chapter 2 of the book is a very good introduction to the problems discussed further. It contains a clear formulation of the quantum mechanical formalism of many-particle systems, in particular systems of particles of the same type. The most important parts are devoted to the permutation symmetry and its consequences. The role played by the symmetrization postulate stating that states of such systems are either fully symmetric or fully antisymmetric under permutations, (i.e. the appropriate Hilbert spaces are symmetric or antisymmetric subspaces of the tensor product) is thoroughly discussed, with a special emphasis put on its consequence that observables have to be permutationally invariant. The above outlined ensuing problems concerning identity and discernibility in the context of the Leibniz Principle are analyzed.

In Chapter 3 the author presents a deep analysis of the origins of the symmetrization postulate, in particular its relation to the indistinguishability postulate demanding that all that expectation values of all empirically accessible observables are invariant under permutations of the indices of particles.

From the philosopher's point of view, what is important are the ontological and epistemological consequences of the formal results described above. The author discusses in depth the relations between the above-mentioned postulates and the Leibniz Principle and their sources (primarily empirical). The discussion in this context of Leibniz's Principle concerns the features to be taken into account in establishing the indiscernibility of particles—so, for example, assuming that such a feature is *haecceity*, i.e. what makes each particular individual to be this very individual, trivializes Leibniz's Principle by reducing it to logical truth. Since within the interpretation of quantum mechanics of systems of particles of the same type described above, the Leibniz Principle is violated, the basic question remains whether

it is necessary and sufficient for empirical purposes: the existence of the possibility of pointing to individual particles and establishing their numerical identity (e.g. to determine their number).

The most important result is that to secure the possibility of making reference to individual objects and grounding numerical identity in qualitative facts, one should consider “relation-based” types of discernibility. How to do it is the main topic of the rest of the book.

The story begins, in chapter 4, with a discussion of various variants of discernibility. As it seems, the most important from the point of view of quantum mechanics is the so-called “weak discernibility” proposed by Simon Saunders (2003). This concept has gained some popularity. Although it was clear that it does not lead to any possibility of referring to individual objects in the case of indistinguishability, it gave hope for establishing numerical differentiation of objects (e.g. to determine their number). Formally, this principle postulates that in the language we use in discourse about quantum particles of the same type, a symmetric ($R(a, b) = R(b, a)$) and non-reflexive ($\sim R(a, a)$) relation can be constructed. Such a relation (weakly) distinguishes objects if it is satisfied by objects a and b . This way of distinguishing objects is criticized in Chapter 4 (it was also criticized earlier by several authors). However, I think that the basic objection of circularity (we presuppose that there are two a priori distinct objects) is rather misplaced, whether in this case or other similar ones, since it can be applied to Leibniz’s Principle itself, making its discussion essentially pointless. The important point, however, is that weak discernibility does not ensure the possibility of referring to a particular single object and no other one.

The core of the argument is the topic of Chapters 5 and 6. Let’s enumerate once more the goals. So, firstly, from the point of view of experimental physics, we would like to have the possibility of

relating to individual particles, secondly, the possibility of numerical individualization of particles (so, among other things, the possibility that the number of them in a given system is, for example, equal to 2). Within the approach called by the author the “orthodox” one, we assign to each particle “its” Hilbert space (“factorism”), and then we would like to measure a given observable (e.g., spin component) in one of these spaces, which would allow statements “the spin component of the k -th particle in a given direction has this particular value”. As it is shown in the previous chapters of the book, and shortly discussed above, such a program is not viable, due to symmetrization postulate. Tomasz Bigaj calls the proposes and discusses in detail the “heterodox” approach, where individualization, i.e. possibility of assigning different properties to particular particles, is carried out using the only admissible observables, i.e. observables symmetric with respect to permutations. The detailed construction of such observables is presented in the book. The goal is achieved by abandoning factorism in favor of such formulations for a system of particles of the same type as, e.g. “one of the particles has a spin component of some value and the other of them of a different value,” with the terms ‘one’ and ‘the other’ not referring to specific Hilbert space numbers of individual particles but meaning “one of two particles,” etc.

In Chapter 7. the author compares the “orthodox” and “unorthodox” approaches in detail, and above all, it discusses the problems that the “heterodox” approach poses (we already know the problems of the “orthodox” approach). Two of them seem particularly important. First, the constructed symmetric observable, in addition to measuring specific values of the physical quantity associated with it, individuates particles; another observable, associated with another physical quantity, leads to another individuation (“different particles”). Second, to the well-known problems of diachronic identity in the macroscopic

world, it adds specifically quantum problems, discussed in detail in the book. There is no doubt for me that these problems have not yet found a definitive solution, as the author also states in several cases.

For the philosopher, the last chapter of the book devoted to the metaphysics of quantum objects should be particularly attractive. That the ontology of quantum mechanics can be fundamentally different from that of classical objects has been, more or less, clear since the dawn of quantum mechanics. To some extent, Bohr's description of measurement in quantum mechanics, emphasizing that we are condemned to use the language of classical theory to describe phenomena that are not subject to this theory, can be seen as an anticipation of problems arising when we try to force the ontology of the quantum world to fit into the known classical world (if only by using such concepts as "particle"). A departure from this treatment of quantum objects is offered by quantum field theory, which uses "quanta" instead of "particles", i.e. in this context, excitations of the quantum field. Such a formalism is also discussed in Chapter 7, but its discussion left me somewhat unsatisfied. It was, however, largely reduced by reading the last chapter, where different ontological concepts of quantum objects are discussed and, above all, the metaphysical consequences of the "orthodox" and "heterodox" approaches discussed above are compared.

From my point of view, it will be interesting to include a short analysis of the problem of a part and the whole in the case of particles of the same type, e.g., in the mereological setting. I know only one position in the literature (Caulton, 2015) dealing with these issues, and definitely, more can be done here.

The potential reader should be warned that most of the reasonings carried out in the book require some knowledge of quantum mechanics and the formalism used therein. Fortunately, the author restricts

himself mostly to quantum mechanics in finite-dimensional Hilbert spaces (which allows, for example, the analysis of spin systems), where the mathematical formalism of quantum mechanics uses only elementary linear algebra. The basic concepts of quantum mechanics are described in the Appendix that closes the book. Whether this is enough to make the work accessible to non-physicists is hard for me (as a physicist) to say. I think that readers not familiar with the basics of quantum mechanics will find it quite difficult since the conclusions drawn by the author are based on quantum-mechanical calculations. Good familiarity with the philosophical foundations of quantum mechanics is also expedient. So both for physicists interested in the basics of quantum mechanics and philosophers of science, reading the book can be quite a challenge. However, the efforts of going through it will be highly rewarded. The book is the best and most comprehensive work on the physics and philosophy of quantum particle systems of the same type in world literature.

Abstract

Leibniz's *principium identitatis indiscernibilium* excludes the existence of two different objects possessing all properties identical. Although perfectly acceptable for macroscopic systems, it becomes questionable in quantum mechanics, where the concept of identical particles is quite natural and has measurable consequences. On the other hand, Leibniz's principle seems to be indispensable when we want to individuate an item and ascribe to it particular property (e.g. value of the projection of spin on a chosen axis). We may thus abandon the principle on the quantum level, claiming its falsity here, or (better) try to find other ways of individuation of objects, possibly by

adopting appropriately the very concept of it. All these problems, and many other connected with identity and indiscernibility of quantum objects, are thoroughly discussed in the book of Tomasz Bigaj, unique in the world literature due to its comprehensiveness.

Keywords

quantum identical particle, indiscernibility, Leibniz Principle.

Bibliography

- Bigaj, T., 2022. *Identity and Indiscernibility in Quantum Mechanics*. Cham: Palgrave Macmillan.
- Black, M., 1952. The Identity of Indiscernibles. *Mind* [Online], 61(242), pp.153–164. Available at: <<https://www.jstor.org/stable/2252291>> [visited on 12 September 2022].
- Castellani, E. and Mittelstaedt, P., 2000. Leibniz's Principle, Physics, and the Language of Physics. *Foundations of Physics* [Online], 30(10), pp.1587–1604. <https://doi.org/10.1023/A:1026489931449>.
- Caulton, A., 2014. *Qualitative Individuation in Permutation-Invariant Quantum Mechanics* [Online]. arXiv. <https://doi.org/10.48550/arXiv.1409.0247>.
- Caulton, A., 2015. Is Mereology Empirical? Composition for Fermions. In: T. Bigaj and C. Wüthrich, eds. *Metaphysics in Contemporary Physics* [Online]. Brill, pp.293–321. https://doi.org/10.1163/9789004310827_013.
- Cortes, A., 1976. Leibniz's Principle of the Identity of Indiscernibles: A False Principle. *Philosophy of Science* [Online], 43(4), pp.491–505. Available at: <<https://www.jstor.org/stable/187115>> [visited on 12 September 2022].
- French, S. and Redhead, M., 1988. Quantum Physics and the Identity of Indiscernibles. *The British Journal for the Philosophy of Science* [Online], 39(2), pp.233–246. Available at: <<https://www.jstor.org/stable/687269>> [visited on 12 September 2022].

- Ginsberg, A., 1981. Quantum Theory and the Identity of Indiscernibles Revisited. *Philosophy of Science* [Online], 48(3), pp.487–491. Available at: <<https://www.jstor.org/stable/186992>> [visited on 12 September 2022].
- Guay, A. and Pradeu, T., eds., 2016. *Individuals Across the Sciences*. Oxford; New York: Oxford University Press.
- Redhead, M. and Teller, P., 1992. Particle Labels and the Theory of Indistinguishable Particles in Quantum Mechanics. *The British Journal for the Philosophy of Science* [Online], 43(2), pp.201–218. Available at: <<https://www.jstor.org/stable/687872>> [visited on 12 September 2022].
- Saunders, S., 2003. Physics and Leibniz's Principles. In: K. Brading and E. Castellani, eds. *Symmetries in Physics: Philosophical Reflections*. Cambridge, U.K.; New York, NY: Cambridge University Press, pp.289–307.

Science and its social grounding

Roman Krzanowski

Pontifical University of John Paul II in Kraków

Stuart Ritchie, *Science Fictions: Exposing Frauds, Bias, Negligence and Hype in Science*. Metropolitan Books Henry Holt and Company, New York 2020. pp.356.

We are constantly puzzled by the question of “What is science?” This comes despite the fact that most of our greatest achievements and failures are directly or indirectly related to science. The chaos of the COVID-19 pandemic stirred up various anti-scientific sentiments, intensifying the confusion and deep apprehension about scientific enterprise (see e.g. McGilchrist, 2021; Blaylock, 2022). Nevertheless, whatever notions laypeople, journalists, bloggers, politicians, actors, and other influencers have been entertaining about science, they have not stopped or hindered scientific work, or at least so it seems. Science mostly just continued without paying too much attention to what people thought about it, like it was somewhat oblivious to social moods; the Aristotelian claim *All men by nature desire to know* still holds firm. Now Stuart Ritchie comes along with a claim: “Science is a social construct.” This is his opening statement in the book *Science Fictions: Exposing Frauds, Bias, Negligence and Hype in Science* (2020), but who is Stuart Ritchie, and what is he actually saying?

Stuart Ritchie lectures in Social Genetic & Developmental Psychiatry at King’s College London, and he has authored several publications for experimental studies in cognitive research, brain studies,

intelligence, publication bias, meta studies, and secondary data analysis, among other things.¹ Thus, he is an insider and therefore seems qualified to comment on the nature of scientific enterprise.

The usual claim that science is a social phenomenon (see e.g. Bloor, 1976; 1991; MacKenzie, 1981; Longino, 2002) is that there is no direct link between science and nature. For Richie, science remains a quest search for objective truth. Its social dimension means that science as a profession is a social construct made by society and embedded into our social structures, so it is subject to social ills, just like any human activity is. Its social aspect comes into play not in the scientific methodology *per se* but rather in everything around science; science—as a complex of cultural systems comprising academia, papers, grants, tenures, hierarchies, competitions, and so on—is a social enterprise.

Richie begins his story by asking how does science make itself known? This generally happens through publications, with top scientific journals like *Science*, *Nature*, *The Lancet*, and such like carrying particular weight. These journals are supposed to present significant, top-quality research that is thoroughly vetted for coherence, verifiable results (which usually take the form of breakthrough discoveries or illuminating insights), and a strict adherence to accepted methodologies and the explicit and implicit ethical principles of science. It therefore comes as a surprise that many of the studies published in top journals, when put under close scrutiny, turn out to be fraudulent or at the minimum, poorly designed with misleading conclusions. This was the fate of Zimbardo's famous, or more accurately infamous, Stanford Prison Experiment, and the fate of the study of Stanley Milgram that

¹ See Stuart Richie on Google Scholar: <https://scholar.google.com/citations?user=9TsCy3IAAAAJ>

was supposed to demonstrate that people will blindly follow orders that go against their moral principles. The problem is that these are not isolated cases.

A study from 2019 reported that out of 3,000 papers published in the top medical journals, 396 studies that advised a significant change in medical practices were completely unfounded. The findings of the studies were never replicated (i.e., confirmed), even though the golden rule is that scientific results should be replicable to be deemed reliable. Furthermore, the flawed claims these studies made were not trivial matters. For example, they suggested changes in practice for childbirth, allergy treatment, and the treatment of heart attacks and strokes, so these changes were at worst harmful to patients or ineffective at best. A review published by the Cochrane Collaboration,² an organization dedicated to the quality of medical research, reassessed the published work and concluded that 45% of medical studies lacked sufficient evidence to support their conclusions. In other words, the findings of these studies could not have been replicated (i.e., verified).

The evil ways in which science gets distorted do not register as capital vices, but they still comprise an impressive collection. Among them, Richie lists fraud, bias, negligence, and hype. Fraud in science comes in many forms. It may be as simple as falsifying results for breakthrough surgical procedures, as was the case with Paolo Macchiaroni of The Karolinska Institute, whose results were published in *The Lancet* as being based on solid evidence. As one may guess, following investigation into his work, his supposedly life-saving procedures turned out to be ineffective if not outright harmful. Fraud also occurred in the case of Woo-Suk Hwang, who published a paper in

² Cochrane is an independent, global organization that collaborates to produce trusted synthesized evidence, making it accessible to all: <https://www.cochrane.org>.

Science about successfully cloning human stem cells. Unfortunately for him, further investigation revealed that the images used in his papers to prove his claim were doctored.

Next on the list of evils is bias. In publications, bias refers to papers focusing on reporting positive, significant results, because these are generally the ones that get published. Papers reporting null or negative results are usually desk-rejected, although things are slowly changing here, as Richie points out. After all, who is interested in a seemingly unsuccessful study? Successful studies, meanwhile, report significant experimental results, and they attract citations and readership, so ratings go up. Who wants to read about null results from an unsuccessful study? Actually, some may. Studies with negative or null results can still provide valuable information, such as by challenging the results of a successful study. Anyway, what is a “significant result”? It’s generally regarded as one that confirms an experimental hypothesis with a p-value of less than 0.05.³ This convention for evaluating the significance of experimental results was conceived by Ronald Fisher in the 1920s, but it has no deeper meaning. Indeed, it is arbitrary. The required p-value is arbitrary, yet it is a deciding factor in whether a paper gets accepted for publication or not. (Note that in high-energy physics a significant result must pass the five-sigma test, which is also a kind of p-test but much more stringent in that the required p-value is 3×10^{-7} , which would be a killer outside physics.)

Thus, an arbitrary number has become an arbiter of an experimental science. Like any statistic, however, a p-value can be manipulated in many ways (i.e., p-hacking) by changing data sets, altering experimental hypotheses, and manipulating experimental data. A hacked p-value means hacked research, meaning that science has been hacked. If you need successful results from a study, and the p-value is not

³ A statistical measurement used to validate a hypothesis against observed data.

cooperating with you, simply change the statistics you use. You need to be expert at massaging data, which is what p-hacking requires, so only an expert can unravel these practices. Where does this leave us, the public, and science, though? How can the public trust experts when the experts cannot trust each other?

Everything has its causes, so fraud, bias, negligence, and hype in science must as well. The main cause is us, and the main victims are us, the public, scientists, trust in science, and science itself, or at least our idea of it. But why does this happen? In the scientific world, like everywhere else, everyone wants to get ahead. There is nothing wrong with this, and it is what makes us human, at least for many of us in some way. It is how we go about doing it that may cause problems. Richie says that the system of incentives in the scientific world “incentivises scientists not to practice science but to meet its own perverse demand.” Simply put, they are perverse incentives. The scientific establishment imposes a simple measure of progress, namely the number of publications, which translates into the now infamous prime directive: publish or perish. It matters little what you publish as long as it is published in prestigious journals and in great quantity. Of course, there is a tacit assumption that these publications are the best stories produced by science, but Richie’s book showed that this is often not the case, maybe too often. But what does “too often” mean in this case, you may well wonder?

Incentives are originally provided to stimulate and support good research, and there is nothing wrong with that, but in reality, these incentives often work contrary to expectations. There are many ways in which incentives can lead to perverse outcomes. Grant-dispensing institutions, profit-oriented private companies, big and small pharmaceutical companies, and the chemicals industry tend to support research that benefits their bottom line. Studies showing anything

contrary to this—such as harmful side effects, placebo effects, or poor efficacy—are rejected or filed away in a basement somewhere. Studies that align with founding agencies’ objectives receive support (i.e., incentives),⁴ while other studies get nothing. It is therefore unsurprising when studies show the expected results (see also Strevens, 2022). If you publish more, you get awarded more funding, so you do whatever it takes to publish more. The tenure-track system also favors those who publish frequently in high-flying journals, rightly so in principle but without considering the essence of what they publish. Thus, you publish as much as you can, whatever and wherever you can. To aid this effort, predatory journals will publish anything for those who can pay the publication fee. This in turn encourages the so-called salami slicing of research. The term “salami slicing” relates to a practice where a large study whose findings should ideally be published as a single paper is instead divided into a series of shorter papers (i.e., salami slicing) covering different areas of the main study. When the number of publications counts, this is the way to increase it, but if what counts are reporting significant findings, this is a counterproductive practice. Measuring of academic productivity based solely on a range of publication metrics (see e.g. Carpenter, Cone and Sarli, 2014) loses the sight of what science is about and how science is actually done (Polak, 2011).

Can we do anything about it? It seems, according to Richie, that the chances of healing are, let’s say, slim. We are what we are, and our science practices reflect this. (We need to be careful to distinguish scientific practices from the methodology of science.) The whole system of science and its social construct—which is dependent on, and embedded in, the economic and political structures—does not

⁴ One can also read about the famous Eddington experiment for “confirming” Einstein’s general relativity theory to see how the human element enters science (Strevens, 2022).

bode well for the future. Science in this respect is an expression of our modern society with its drive for success and recognition, as well as its competitive culture of winner takes all. It does not mean that science will inevitably fail, though, because we will still have amazing discoveries and make progress in developing mind-boggling inventions. A sea of dirt will also be there, however, with tsunamis of hype swaying the public and feeding into all kinds of anti-whatever camps of less sensible but highly vocal groups. Is it therefore a surprise that “people do not believe in science,” as we have often heard during the pandemic? If people dismiss the Apollo moon landings as fake, why would they not question the COVID-19 vaccines? Are we surprised? However, do not blame science or the mythical “them,” because the fault lies with us, the ones doing the science and the ones writing about it.

Of course, there is much more detail to every story that Richie narrates. There are more examples of nuanced transgressions committed in scientific publications and research, fraud, and personal bias, both out of stupidity and arrogance—the list is long. Indeed, it reads like a crime novel, and it makes little difference that it is a white-collar crime. Again, you have to read about these in the book, but they certainly make for interesting and rewarding study. Now, does Richie’s book have any import for science outside of experimental research? After all, Richie focuses mostly on clinical, social, psychological studies? It seems that lessons in Richie’s book are for everyone “doing” some research, whatever it may be. The prime directive is real, and it works in every corner of academia, so the game of publications and grants is what today’s science is all about. Studies that do not rely on experiments unfortunately lack the truth serum of reproducibility, which is why it is difficult to spot the bogus research and ideas within them. Only time will ultimately reveal them.

So, what lessons does Richie's story have for all of us? First, science as practiced profession, is a survival game, and it must be played once you are in it. Some are good at science, some are good at games, and some are good at gaming science, but who wins? It's not necessarily those who are good at science. But in the long run, science seems to possess some self-cleansing properties. Second, do not read newspaper reports about miraculous new cures, foods, or whatever. These 'news' are certainly not science-based facts. So do not blame science for them. Thirds, in the long run, maybe too long for some of us, the truth, despite many doubting its value or existence, always eventually prevails in science, philosophy, and humanity, just maybe not during your lifetime.

So, who should read Stuart Richie's book? Anyone thinking about a career in science should certainly consider it. Indeed, young, idealistic scholars should realize they are not entering an idealistic field purely dedicated to the pursuit of truth and knowledge but rather a fierce neoliberal game of survival where anything is permitted as long as you do not get caught. In this world, you better learn the tricks of the trade or you will soon be gone unless you are independently wealthy. This is a perverse point of view, of course—it is infinitely preferable to invest in honest studies because they age much better, even if they are less spectacular.

Should anyone else read Stuart Richie's book? The universities and their governing bodies certainly should before they turn the screw to tighten the publishing requirements of their staff. Indeed, they should think twice about it and maybe start promoting research quality rather than quantity. Unfortunately, there is little hope for this. Quality research requires time, money, patience and a tolerance of failure, something that managers have very little time for. The managers responsible for the rules usually have nothing to do with science

and generally only care about the stats, because this is what they understand, if anything, and this is how they win the game that they are playing.

And what about Richie's original claim that science is a social construct? We believe that Richie proves his case. Science is embedded in societal structures, so it inherits all of their characteristics, both good and bad. This only applies to the operational aspects of science, however, and science in itself as the search for truth is unaffected in principle by this social dependency. The social dependencies hinder the reception of science in society, but the idea of science does not. This is why we still believe in it, at least some of us.

The optimistic message to take away is that science is still science, and it represents the greatest intellectual achievement of the human race so far, at least when done right. Science in its essence is not a playground of changing moods and social trends, as some would like to have us believe. In the absence of political and economic pressures, if such an ideal world could exist, or with sufficient time, science will refocus itself on its original objectives. This reflects what science was meant to be, and always will be: the search for truth (e.g. Grayling, 2022). The problem remains how to reconcile the objectives of science (as a search for truth) with external irrational influences (e.g., social, political, and cultural); the topic discussed in details by Liana (2019; 2020).

Abstract

Stuart Richie's book discusses social, political, and cultural influences on science. In a series of well documented cases Richie shows how many of top scientific journals publish poorly executed studies with

dubious conclusions. Such publications distort a public image of science as an unbiased search for truth. The roots of such practices, Richie traces to the way science enterprise is done in academia and in private research centers, where only positive and “expected” results are valued. While according to Richie there is a small chance to cure scientific practices from these ills, science itself is and remains the search for truth, even if our social moors make it so much harder.

Keywords

science as a social construct, practice of science, publication policies, politics of science.

Bibliography

- Blaylock, R.L., 2022. COVID UPDATE: What is the Truth? *Surgical Neurology International* [Online], 13, p.167. https://doi.org/10.25259/SNI_150_2022.
- Bloor, D., 1976. *Knowledge and Social Imagery*, Routledge Direct Editions. London: Routledge & Kegan Paul.
- Bloor, D., 1991. *Knowledge and Social Imagery*. 2nd ed. Chicago: University of Chicago Press.
- Carpenter, C.R., Cone, D.C. and Sarli, C.C., 2014. Using Publication Metrics to Highlight Academic Productivity and Research Impact. *Academic Emergency Medicine* [Online], 21(10). Ed. by G.M. Gaddis, pp.1160–1172. <https://doi.org/10.1111/acem.12482>.
- Grayling, A.C., 2022. *The Frontiers of Knowledge: What We Know About Science, History and the Mind - and How We Know It*. London: Penguin Books.
- Liana, Z., 2019. Nauka jako racjonalna doxa. Józefa Życińskiego koncepcja nauki i filozofii nauki – poza internalizmem i eksternalizmem. *Philosophy*

- ical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), pp.147–199. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/466>> [visited on 23 September 2019].
- Liana, Z., 2020. Józefa Życińskiego koncepcja racjonalizmu umiarkowanego: epistemologiczna i doxalogiczna funkcja podmiotowego commitment. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (68), pp.117–184. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/516>> [visited on 25 January 2021].
- Longino, H., 2002. The Social Dimensions of Scientific Knowledge. In: E.N. Zalta, ed. *Stanford Encyclopedia of Philosophy (Summer 2019 Edition)* [Online]. Stanford, Calif.: Metaphysics Research Lab, Stanford University. Available at: <<https://plato.stanford.edu/archives/sum2019/entries/scientific-knowledge-social/>> [visited on 10 September 2022].
- MacKenzie, D., 1981. Notes on the Science and Social Relations Debate. *Capital & Class* [Online], 5(2), pp.47–60. <https://doi.org/10.1177/030981688101400103>.
- McGilchrist, I., 2021. *The Matter with Things*. London: Perspectiva Press.
- Polak, P., 2011. Naukometryczne uwarunkowania pseudonauki. In: K. Korab, ed. *Pseudonauka. Choroba, magia czy biznes?* [Online]. Warszawa: Wydawnictwo Naukowe Scholar, pp.82–90. <https://doi.org/10.13140/2.1.3727.3924>.
- Ritchie, S., 2020. *Science Fictions: How Fraud, Bias, Negligence, and Hype Undermine the Search for Truth*. First edition. New York: Metropolitan Books; Henry Holt and Company.
- Strevens, M., 2022. *The Knowledge Machine: How an Unreasonable Idea Created Modern Science*. London: Penguin Books.

A companion to Kazimierz Twardowski

Ewelina Grądzka

Pontifical University of John Paul II in Kraków

J. Jadacki, ed., *Rozum i wola. Kazimierz Twardowski i jego wpływ na kształt kultury polskiej XX wieku*, Wydawnictwo Academicon, Lublin 2021, pp.573.

Kazimierz Twardowski is considered one of the influential Polish philosophers, if not the most influential, having founded a famous Central European analytical school, namely the Lvov–Warsaw School (LWS), which Alfred Tarski came from. However, anyone interested in learning more about Twardowski and the current state of the research will encounter some difficulty in finding an adequate source,¹ something like “A companion to...” Fortunately, Professor Jacek Jadacki has provided us with a book that, thanks to chapters authored mostly by established experts in the field, aims to guide the reader through the works and ideas of Twardowski and their influence on Polish and international philosophy. The book can therefore be recommended to anyone who has never encountered Twardowski’s *œuvre*, as well as those interested in the trends and shifts in the research of this.

The book under review is *Rozum i wola. Kazimierz Twardowski i jego wpływ na kształt kultury polskiej XX wieku* [Reason and will.

¹ It is worth noting that there is a wide range of publications about Twardowski and his school (e.g., Woleński, 1985; 1989; Poli, Coniglione and Woleński, 1993; Brandl and Woleński, 1999; Chybińska et al., 2016; Brożek, Stadler and Woleński, 2017; Drabarek, Woleński and Radzki, 2019), but each focuses on some particular aspects.

Kazimierz Twardowski and his influence on the shape of Polish culture in the 20th century], and it resulted from an interesting project called “Kazimierz Twardowski’s place in Polish culture and European philosophy,” which was financed by the National Science Center of Poland. It is also the fourth volume in a series of publications (cf. Brożek, 2020; Brożek, Będkowski et al., 2020; Jadacki and Cullen, 2020) by the Lvov–Warsaw School Research Center, which was established in 2020 at the Faculty of Philosophy, University of Warsaw² with the aim of promoting and developing research related to the LWS. The book can also be considered a great supplement to compilations of Twardowski’s texts’ (Twardowski, 1927; 1992; 2013; 2014; Brandl and Woleński, 1999; Brożek, Będkowski et al., 2020). It is also a book that—next to those of (Jadczak, 1993; Kleszcz, 2013; and Brożek, 2015)—aims to present Twardowski’s ideas in a comprehensive and critical way, with it going further than just focusing on one area of his activity.

The editor specializes in ontology and epistemology, logical semiotics, and the history of Polish philosophy, especially the Lvov–Warsaw School. Jadacki claims to be a follower of Jan Łukasiewicz’s program of philosophizing, so therefore considers himself a continuator of the LWS. The problem of LWS members and followers is an engaging and still debatable issue that leads to another discussion about what it means to belong to a school, both in general and for this particular example. More on this can be found in the work of (Woleński, 1985).

We first question whether the book can be referred to as a companion to Twardowski. Taking a look at the offerings of publishers like Cambridge, Oxford, Routledge, and Springer, one can find books where the editors attempt to refer step by step to particular, significant

² For more on (Brożek, 2020) and the center, see (Grądzka, 2021a,b).

aspects of a philosopher's ideas and publications. They are then later put into the context of their epoch and how they were received by philosophy/science. In this way, the reader is generally guided through the difficulties in understanding the philosopher's *œuvre* and helped to appreciate the questions or solutions that were presented by the philosopher. Giving such an extensive background enables the reader to comprehend the significance of the thoughts and critique them. So, does the book under review meet these aims? Let's take a closer look.

The book is divided into seventeen chapters grouped into six sections, each entitled "Twardowski as. . ." It seems the editor aims to present the Lvov philosopher as a personality who played multiple roles, which is indeed true.

The first section is titled "Twardowski jako klasyk" [Twardowski as a classic], and it refers to his role in Polish and global philosophy. It was written by some of the most significant Polish specialists on Twardowski, namely Jan Woleński, Jacek Jadacki, and Ryszard Kleszcz, who have great experience in this subject and the LWS in general (cf. Woleński, 1985; Jadacki, 1989; 2003; 2009; 2015; Brandl and Woleński, 1999; Jadacki and Paśniczek, 2006; Kleszcz, 2013; Chybińska et al., 2016; Brożek, Stadler and Woleński, 2017) Their contribution seeks to present Twardowski as an international personality, one that was important for his contemporaries (i.e., students, continuators, and opponents), and a person who applied his method of philosophizing to issues in daily life and all professional engagements.

They manage to convince the reader that Twardowski is "a classic" worth knowing. On the one hand, he had an impact on global philosophy (as described by Woleński on pp.17–47) thanks to his *Habilitationsschrift* entitled *On the Content and Object of Presentations* (1894), in which he distinguishes between act, content, and object of representation. It inspired Alexius Meinong's theory of

objects and Alois Höfler's and Edmund Husserl's concepts of intentionality. On the other hand, Twardowski's thought was integrated into global philosophy by his followers, particularly Jan Łukasiewicz and Alfred Tarski. It is continuously emphasized that many problems that the LWS members developed had originally taken root in Twardowski's ideas. (He is said to have not been a publish-or-perish sort of person, instead preferring to openly share his ideas with students and be pleased at seeing them advanced or refined.)

Woleński refers here to Twardowski's style of teaching logic, which was introduced by Tarski to the University of California, Berkeley, where it helped develop the Californian school of logic. The concept of absolute truth is also claimed to have provided the philosophical background for the many-valued logic of Jan Łukasiewicz, who later worked at the University of Dublin, and Tarski's logical semiotics. Woleński concludes that Twardowski's pedagogical success, with him having educated over thirty professors, would be enough alone to count as a great contribution to global philosophy.

How this achievement was made possible is explained in the next chapter by Kleszcz (pp.49–96). Twardowski, unlike other analytic philosophers, was able to raise “an army of intellectuals” thanks to his unprecedented engagement in working with students through proseminars, seminars, a library, the *Polish Philosophical Society*, and the *Polish Philosophical Congress*. He also established the journal *Ruch Filozoficzny* [Philosophical Movement], made himself accessible every day at particular hours, introduced a consistent philosophical curriculum, and demanded high methodological standards (i.e., clarity of expression, criticism, and application of logic and reasoning).

This was evidently so effective that the most fruitful and reasonable criticisms of Twardowski's works came from his students, which, according to Jadacki, is proof of his didactical success (Jadacki,

pp.97–199). It is an interesting contribution of Jadacki to present such criticism, because it gives the reader a feeling of how discussions took place within the LWS. The editor also touches upon another important issue, namely the state of Polish philosophy in the post-war period, when a new political and ideological system was imposed by the Soviet Union. He describes how Twardowski and his students were slandered and defamed, with them being accused of having bourgeois, anti-proletarian, imperialistic, reactionary, obscurantist, fideistic, and speculative ideas by prominent personalities like Adam Schaff, Henryk Holland, and Leszek Kołakowski.

Additionally, Władysław Tatarkiewicz, Roman Ingarden (the attacks were also directed against phenomenology and neotomism), and Izydora Dąmbska were banned from teaching students. A short Appendix II with citations from *Krótki słownik filozoficzny* [A short philosophical dictionary] helps understand the meanings of some typical Marxism–Leninism concepts, and this is a great idea, especially for the younger generation.

The communist persecution of Polish philosophers is still relatively fresh, however, and it is difficult to remain moderate and present the issue in a restrained way, which is understandable, especially for those with experience of this system. Jadacki's work is relevant, but he does not manage to be neutral, and this can unfortunately be discouraging, like it was in the case of the books by Kuliniak, Pandura, and Ratajczak (2018, 2019). Maybe it will be the task of the next generation to prepare material that will expose the brutal and shameful truth but in a reserved, detached manner.

Twardowski was a great inspirer, and Woleński's efforts (pp.203–235) to emphasize the importance of *Action and Products* (cf. Brandl and Woleński, 1999) are especially significant nowadays,

when we face a sort of resurgence of neopositivism in the form of “neuroscientism,” a belief that all our behavior and awareness can be explained by research into the functioning of the brain.

Meanwhile, Alicja Chybińska offers the intriguing hypothesis that Twardowski’s methodological requirements of clarity of expression influenced Tadeusz Kotarbiński’s reism (pp.237–252). As stated before, Twardowski’s students were strongly influenced by their master, but they also often resisted his solutions. It is claimed here that Kotarbiński was “inspired negatively” in the sense that he was motivated to contradict his teacher by rejecting “hypostasis.”

Nevertheless, the methodological postulate of clarity, so emphasized by Twardowski, is tricky and challenging to follow, even for the members of the LWS. Łukasiewicz, one of the most famous of these, is considered to have been a master of precision in terms of the theories he created, but he failed while considering the relations between logic and philosophy. He even claimed that during his investigations, he faced some kind of ideal construction that a philosopher-believer would call God’s thought. This is surprisingly close to the statements of people like Michael Heller (2019) and Albert Einstein (the mind of God). The problem of the relation of logic to philosophy or even theology is a question for debates that can be followed in publications like those of Dadaczyński (2014) and Heller and Awodey (2020).

The importance of logic and analysis, not just for philosophers but also those in other disciplines, can be gleaned in the texts of the third part of the book. Anna Brożek, clear and bright as ever (Brożek, 2020; cf. Grądzka, 2021a) introduces how Twardowski conceived his analysis of concepts (pp.271–297). This provides significant background for the next chapter, which exposes a rarely mentioned aspect of Twardowski’s engagement, namely his teaching of logic to students of medicine. It turns out that the Lvov philosopher also influenced the

Polish School of the Philosophy of Medicine and one of its main contributors, Władysław Szumowski (Aleksandra Horecka, pp.299–323). One again gets a feeling that some groundbreaking ideas were introduced long ago, ones that we are ignorant of now. Modern medicine is far removed from recognizing the importance of logic and philosophy for making progress, reflecting over its methods, and reconsidering the modern form of materialism as a paradigm, and no one seems to care if “illness” really exists or not. This chapter is therefore especially recommended to doctors and other healthcare professionals, especially in the wake of the recent pandemic, when so many paradigms were challenged.

In his introduction to the book, Jadacki admits that the publication has a problem in repeating some issues in different texts. He considers this not just a weakness but also a strength, because it allows readers to follow each article independently. This impression prevails on reaching some chapters, namely Jadacki’s “On Twardowski’s contributions to development of logic in Poland” (pp.325–343) and “Twardowski’s postulate of Clarity in the Eyes of an Anti-Irrationalist” (pp.385–406), as well as Dariusz Łukasiewicz’s “Analysis of the relationship between Philosophy, Science and the Worldview in Terms of Twardowski” (pp.347–383). However, it must be admitted that despite some basic information being repeated in the texts, the authors manage to present a fairly systematic and thorough analysis of the problems, and there are new aspects that are worth reading, such as the comparison of Twardowski’s, Salamucha’s, and Kleszcz’s ideas on the worldview and Twardowski’s theory of induction, which comes thanks to investigating some unpublished lectures about logic.

Finally, we reach the two final chapters, which may be somewhat controversial because they attempt to interpret Twardowski’s activities as an expression of his political (Ryszard Mordarski pp.409–428,

Jacek J. Jadacki pp.429–459) and pedagogical (Tadeusz Czeżowski pp.463–468, Wojciech Rechlewicz pp.469–530, Jacek J. Jadacki pp.531–562) ideas and theories. The Lvov philosopher lived and worked in an interesting time when political science and pedagogy were gaining independence from philosophy. He himself was motivated not just by his way of philosophizing but also his upbringing, and he engaged in many projects related to building a well-educated nation of rational patriots.

Nowadays, we would say he was a theoretician and a practitioner, an interdisciplinary one like in so many other cases. However, methodological questions remain as to whether he should be considered a philosopher of politics or education, or possibly a political scientist or pedagogist in modern terms. It is fair to say, however, that we should judge the thinker by the standards of his times. Indeed, “the case of Twardowski” is captivating because it makes the reader or investigator deliberate with much precision, but also meditation, on what the fundamentals or differentiations are between disciplines or fields. This might be challenging for beginners and even some advanced researchers, but it is a good exercise for anyone interested in the philosophy of science.

In conclusion, the answer to the question of whether the book meets the requirements of being a companion to Kazimierz Twardowski’s intellectual, pedagogical, and practical heritage is a positive one. The editor brought together a valuable group of researchers who then managed to inspire their readers to further investigate the works and accomplishments of the Lvov philosopher. It will be a challenge for a future publisher to go further in preparing a subsequent

companion to Twardowski, because such work is never completed. Meanwhile, it would be useful to translate the book into English, so non-Polish speakers can also better appreciate Tarski's background.³

Abstract

There have been many significant publications on Kazimierz Twardowski. Jacek Jadacki intends to add to this list another book *Rozum i wola. Kazimierz Twardowski i jego wpływ na kształt kultury polskiej XX wieku*. In the review it is appraised whether it can be called “a companion to...”. It provides introductory information that can help readers better understand the role of Twardowski in Polish philosophy and culture. Updated findings by contemporary scholars are also included. The quality of the articles is guaranteed by such authors as J. Woleński, R. Kleszcz, A. Brożek and J. Jadacki. However, new authors are also present as well as less common topics like Twardowski's influence on the Polish School of Philosophy of Medicine and his roles as political scientist, educational theorist, and historian of Ancient philosophy. The authors manage to convince the reader that Twardowski is “a classic” worth knowing, in consequence the book can be treated as a “companion to Twardowski”. It also inspires readers to further investigate the works and accomplishments of the Lvov philosopher.

Keywords

Kazimierz Twardowski, Lvov-Warsaw School, analytic philosophy, Alfred Tarski.

³ For now, there is a valuable introduction to Twardowski in the Stanford Encyclopedia (Betti, 2019).

Bibliography

- Awodey, S. and Heller, M., 2020. The homunculus brain and categorical logic. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (69), pp.253–280. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/510>> [visited on 24 February 2021].
- Betti, A., 2019. *Kazimierz Twardowski*. Available at: <<https://plato.stanford.edu/archives/sum2019/entrieswardowski/>> [visited on 14 January 2021].
- Brandl, J.L. and Woleński, J., eds., 1999. *On Actions, Products and Other Topics in Philosophy. Kazimierz Twardowski* (A. Szylewicz, Trans.), *Poznań Studies in the Philosophy of the Sciences and the Humanities*, 67. Amsterdam; Atlanta: Rodopi.
- Brożek, A., ed., 2015. *Kazimierz Twardowski we Lwowie: o wielkim myślicielu, nauczycielu i obywatelu*. Bydgoszcz: Oficyna Wydawnicza Epigram.
- Brożek, A., 2020. *Analiza i konstrukcja: o metodach badania pojęć w Szkole Lwowsko-Warszawskiej*. I. Vol. 3, *Centrum Badań nad Tradycją Szkoły Lwowsko-Warszawskiej*. Kraków: Copernicus Center Press.
- Brożek, A., Będkowski, M. et al., 2020. *Anti-irrationalism. Philosophical Methods in the Lvov-Warsaw School*. Warszawa: Semper – Wydawnictwo Naukowe.
- Brożek, A., Stadler, F. and Woleński, J., eds., 2017. *The Significance of the Lvov-Warsaw School in the European Culture* [Online], *Vienna Circle Institute Yearbook*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-52869-4>.
- Chybińska, A., Jadacki, J.J., Woleński, J. and Brożek, A., eds., 2016. *Tradition of the Lvov-Warsaw school: ideas and continuations, Polish Analytical Philosophy*, 106. Leiden; Boston: Brill-Rodopi.
- Dadaczyński, J., 2014. Some remarks on the logic of theology. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (57), pp.33–58. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/17>> [visited on 6 January 2022].

- Drabarek, A., Woleński, J. and Radzki, M.M., 2019. *Interdisciplinary Investigations into the Lvov-Warsaw School*. Cham, Switzerland: Palgrave Macmillan.
- Grądzka, E., 2021a. Putting analysis and construction of concepts into its legitimate position. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (71), pp.155–173. Available at: <<https://zfn.edu.pl/index.php/zfn/article/view/544/544>>.
- Grądzka, E., 2021b. Report from the philosophical workshop organized by The Lvov-Warsaw School Research Center and Kazimierz Twardowski Philosophical Society of Lviv. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)*, (71), pp.131–152.
- Heller, M., 2019. *God and geometry: when space was God* (P. Krasnowolski, Trans.). Kraków: Copernicus Center Press.
- Jadacki, J.J., 1989. *Semiotyka deskryptywna Kazimierza Twardowskiego: analiza krytyczna*. Toruń: Polskie Towarzystwo Filozoficzne.
- Jadacki, J.J., 2003. *From the viewpoint of the Lvov-Warsaw School, Poznań Studies in the Philosophy of the Sciences and the Humanities*, 78. Amsterdam; New York: Rodopi.
- Jadacki, J.J., 2009. *Polish analytical philosophy: studies on its heritage: with the appendix containing the bibliography of Polish logic from the second half of the 14th century to the first half of the 20th century*. Warszawa: Wydawnictwo Naukowe Semper.
- Jadacki, J.J., 2015. *Polish philosophy of the 19th and 20th centuries: heritage studies*. Warsaw: Semper, Wydawnictwo Naukowe.
- Jadacki, J.J. and Cullen, K., 2020. *Stanisław Leśniewski: Genius of Logic, The Lvov-Warsaw School Research Center Publishing Series*, 2. Bydgoszcz: Oficyna Wydawnicza Epigram.
- Jadacki, J.J. and Pańniczek, J., eds., 2006. *The Lvov-Warsaw School - the new generation, Polish Analytical Philosophy*, 6. Amsterdam; New York: Rodopi.
- Jadczak, R., 1993. *Człowiek szukający etyki: filozofia moralna Kazimierza Twardowskiego, Rozprawy / Uniwersytet Mikołaja Kopernika*. Toruń: Wydaw. UMK.

- Kleszcz, R., 2013. *Metoda i wartości: metafizologia Kazimierza Twardowskiego*. Warszawa: Wydawnictwo Naukowe Semper.
- Poli, R., Coniglione, F. and Woleński, J., eds., 1993. *Polish scientific philosophy: the Lvov-Warsaw School, Poznań Studies in the Philosophy of the Sciences and the Humanities*, 28. Amsterdam; Atlanta: Rodopi.
- Twardowski, K., 1894. *Zur Lehre vom Inhalt und Gegenstand der Vorstellungen*. Wien: Alfred Hölder, K. U. K. Hof- und Universitäts-Buchhändler.
- Twardowski, K., 1927. *Rozprawy i artykuły filozoficzne*. Lwów: Księgarnia „Książnica-Atlas”: nakładem Komitetu Uczniów.
- Twardowski, K., 1992. *Wybór pism psychologicznych i pedagogicznych*. Ed. by M. Decewicz. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
- Twardowski, K., 2013. *Mysł, mowa i czyn*. Ed. by A. Brożek and J.J. Jadacki, *Bibliothèque des Philosophes*, 4. Kraków: Copernicus Center Press.
- Twardowski, K., 2014. *Mysł, mowa i czyn. Cz. 2*. Ed. by A. Brożek and J.J. Jadacki, *Bibliothèque des Philosophes / Towarzystwo Naukowe Warszawskie*, 7. Warszawa: Semper - Wydawnictwo Naukowe.
- Woleński, J., 1985. *Filozoficzna szkoła Lwowsko-Warszawska*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Woleński, J., 1989. *Logic and Philosophy in the Lvov-Warsaw School, Synthese Library: studies in epistemology, logic, methodology and philosophy of science*, 198. Dordrecht: Kluwer Academic Publishers.

Idee i ideaty

Michał Heller

Centrum Kopernika Badań Interdyscyplinarnych

Janusz Mączka

Uniwersytet Papieski Jana Pawła II w Krakowie

Józef Życiński, *Struktura rewolucji metanaukowej. Studium rozwoju współczesnej nauki*, tłum. Michał Furman, Copernicus Center Press,

Kraków 2013.

Konferencja naukowa zwołana w 2021 roku w ramach obchodów dziesiątej rocznicy śmierci abpa Józefa Życińskiego nie tylko stworzyła okazję, lecz również zachęciła do ponownej lektury jego najważniejszych książek. Niewątpliwie należy do nich *Struktura rewolucji metanaukowej* (Życiński, 2013). Oryginalnie została ona opublikowana w języku angielskim (Życiński, 1988), ale w ostatnich latach znów stała się przedmiotem rozważań w związku z przetłumaczeniem jej na język polski (zob. np. Liana, 2019; 2020). Książka ta zawiera przemyślenia Józefa Życińskiego dotyczące filozofii nauki istotne dla całego jego stylu filozofowania. Zostało w niej także zapoczątkowanych szereg wątków, które i dziś warto rozwijać. Niektóre z nich przybliżymy w tym niewielkim opracowaniu.

Tytuł tej książki, choć nawiązuje do znanej książki Thomasa Kuhna *Struktura rewolucji naukowych* (Kuhn, 1968), nie jest jednak jego kopią. Różnica z tytułem książki Kuhna sprowadza się do dwóch drobnych zmian, które jednak wprowadzają istotną różnicę. Te dwie zmiany to: mały przedrostek „meta” i liczba pojedyncza w określeniu „metanaukowa rewolucja”. Przedrostek „meta” określa

perspektywę Życińskiego, która jest odmienna od perspektywy Kuhna. Kuhn poddawał analizie rewolucje, które dokonywały się w nauce i które powodowały istotne zmiany w jej pojmowaniu. Przykładami takich rewolucji są przejście od nauki średniowiecznej do nowożytnej oraz przejście od fizyki klasycznej do mechaniki kwantowej i relatywistycznej. Poddając analizie te i inne rewolucje naukowe, Kuhn podważał obowiązujące dotychczas w filozofii nauki poglądy, związane zwykle z nazwiskiem Karla Poppera i filozofią nauki Koła Wiedeńskiego. O ile Popper i neopozytywiści wiedeńscy dokonywali analiz naukowych przy użyciu narzędzi logicznych i metodologicznych, o tyle Kuhn posłużył się metodą analizy historycznej i socjologicznej. Propozycja Kuhna odbiła się szerokim echem i zapoczątkowała ciąg prac, których autorzy prześcigali się w wynajdywaniu coraz to nowych ujęć rozwoju nauki. Zagadnienia te są zbyt dobrze znane, by je tutaj referować.

Oryginalne spojrzenie Życińskiego sprowadza się do spostrzeżenia – głównie dzięki książce Kuhna – iż niemal na naszych oczach dokonana się wielka rewolucja, ale nie w nauce, lecz w filozofii nauki, czyli rewolucja metanaukowa. Na niej to właśnie skupia się Życiński w swojej książce. Warto zauważyć, że w trakcie pisania książki rewolucja jeszcze się dokonywała. Autor dostrzega, że rewolucji metanaukowych w historii nauki było więcej i poświęca im nieco uwagi. Wspomina o zmianach w poglądach na rozumienie świata w starożytnej Grecji, w chrześcijańskim średniowieczu i w renesansie.

W trakcie dokonywania się rewolucji naukowych następuje stopniowa zmiana przedzałożeń, które obowiązywały w dotychczasowym okresie. Zbiór twierdzeń filozoficznych, będących częścią przedzałożeń, często określa się mianem „ideologii”. Życiński pisze: „Zbiór twierdzeń ideologicznych uzależnionych od konkretnych interpretacji w naukach przyrodniczych nie jest jednorodny epistemologicznie.

Zawiera zasady metodologiczne oraz założenia epistemologiczne, heurystyczne wskazówki i twierdzenia o charakterze ontologicznym. Te same terminy, które tworzą «ideologiczne» twierdzenia, mogą denotować rzeczywistość, która jest odmienna w innych programach badawczych” (Życiński, 2013, s. 33). Celem ukonkretnienia swoich analiz, Życiński proponuje „pozanaukowe składniki teorii naukowych” nazwać *ideatami*. Dalej czytamy: „Ideaty są zarówno pojedynczymi, podstawowymi terminami w konkretnych interpretacjach, na przykład atom, ewolucja czy stworzenie, jak i asercjami inspirującymi określony sposób wyrażania rzeczywistości oraz wykonywania podstawowych funkcji metodami charakterystycznymi dla danych naukowych programów badawczych” (Życiński, 2013, s. 34). Według Życińskiego, ideaty „inspirują poszczególne programy badawcze oraz współtworzą ich twardy rdzeń” (Życiński, 2013, s. 34). Jednakże według niego, nie wszystkie pozanaukowe elementy, związane z uprawianiem nauki, zasługują na miano ideatów. Niektóre z nich są przypadkowymi wtętami bez większego wpływu na całość programu badawczego.

Życiński wymienia ideaty o charakterze ontologicznym, epistemologicznym i metodologicznym. Tego rodzaju ideaty nie muszą wynikać z jakichś głębszych przemyśleń filozoficznych; często ich treść „zależy wyłącznie od socjologicznych i psychologicznych uwarunkowań, a nie od bardziej podstawowej ontologii” (Życiński, 2013, s. 51).

Pojęcie ideatu, obok samej idei rewolucji metanaukowej, należy uznać za najbardziej oryginalny pomysł w tej książce. Życiński przyznaje, że rola ideatów jest „taka sama” jak rola tzw. *themata* opisanych przez G. Holtona w jego *Thematic Origins of Scientific Thought* (1988). Ponieważ jednak Życiński podał jedynie dość ogólnikowy zarys tego co powinno się rozumieć pod pojęciem ideatu, spróbujmy

poszukać w jego książce przykładów ideatów, na które się powołuje. Poniższa lista nie jest listą kompletną, ale daje ona wyobrażenie o tym, co Życiński rozumiał pod pojęciem ideatu.

- I tak zaraz po wprowadzeniu pojęcia ideatu, Życiński pisze o ideatach bliskich filozofii Arystotelesa, które „kładą szczególny nacisk na kryteria zdrowego rozsądku i racjonalności” (Życiński, 2013, s. 35).
- Natomiast ideatami bliskimi filozofii Platona są, według niego, „piękno teorii, jej prostota, symetria i tym podobne” (s. 36).
- Życiński mówi również o „Euklidesowskich ideatach logicyzmu”, które „inspirowały nowe próby coraz bardziej wyrafinowanych poszukiwań głębszych podstaw” (s. 167).
- Wspomina także o „racjonalistycznych ideatach Euklidesa i Kartezjusza” (s. 183).
- Ideaty funkcjonują również w nauce nowożytnej. Na przykład: „W ramach dziewiętnastowiecznej nauki podobny racjonalistyczny redukcjonizm wyrażał się w próbach wyjaśniania wszystkich procesów w kategoriach mechaniki. Po upadku mechanicyzmu ideaty te przeniknęły w obszar badań meta naukowych” (s. 328).
- Wedle Życińskiego istnieją ideaty, wykazujące niezwykłą trwałość na poziomie meta „niezależnie od lokalnych nieciągłości spowodowanych rewolucjami naukowymi”. „Jednym z nich jest powszechnie rozumiany ideat racjonalności, którego różne warianty występują na różnych etapach rozwoju nauki” (s. 334).

- „Innym ważnym ideatem wyrażającym racjonalizm metanaukowy był ideał poznawalności, wyrażający przekonanie, że możliwe jest poznanie świata za pomocą środków dostępnych człowiekowi” (s. 334).
- Jeszcze innym ideatem jest „ideał pozytywizmu metodologicznego, który miał decydujący wpływ na sformułowanie zasad metodologicznych współczesnej nauki” (s. 335–336).
- Z zasadami metodologicznymi wiąże się także „ideał istnienia w nauce różnorodnych ograniczeń, które zostały dowiedzione na podstawie przesłanek dostarczonych przez nią samą” (s. 336).
- Zaprzeczenie ograniczeń poznania naukowego prowadzi do „ideatu totalizmu naukowego”, który wyraża slogan „Nie ma zbawienia poza nauką” (s. 337).
- Wreszcie Życiński zwięźle omawia „ideał unifikacji”, który przybierał różne formy: Od „poszukiwania jednej formuły stosowalnej do wszystkich zjawisk, typowej dla magii”, poprzez redukowanie wszystkiego do mechaniki, aż do „programu logicystycznego, w którym cała matematyka miała zostać zredukowana do małej liczby aksjomatów i reguł przekształcania” (s. 340–341).

Należy przestrzec Czytelnika, że lista ta nie jest w żadnym sensie systematyczna. Życiński wymienia te ideaty przy okazji omawiania innych tematów, zakładając, że pojęcie ideatu jest już znane. Jak widać z tej listy, Życiński pojęcie ideatu rozumie szeroko i w sposób rozmyty. Przypuszczalnie nie jest to efekt niezamierzony, gdyż, jak z tych przykładów wynika, ideaty działają w znacznej mierze dzięki swojemu niedookreśleniu, sterując niejako z ukrycia sposobami myślenia w danej epoce. Co więcej, ten sam ideał w różnych epokach

może przybierać inne treści zależnie od tego, w jaki sposób jest wykorzystywany. Ideaty nie tylko kształtują myślenie w danej epoce, lecz są również kształtowane przez to myślenie.

Życiński wspomina także „o niekrytycznych ideatach”. Pozytywizm i scjentyzm XIX w. stworzył przekonanie, „że wypłenienie ideologicznych wtrętów z nauki było tylko kwestią czasu” (Życiński, 2013, s. 49); wkrótce jednak okazało się, że „znacznie łatwiej było krytykować ideologiczne programy badawcze szkodzące nauce w przeszłości, niż unikać niekrytycznych ideatów w nowych programach badawczych” (Życiński, 2013, s. 49). Tego rodzaju niekrytyczne ideaty niekiedy z warstwy meta wpływają na ewolucję teorii naukowych. Próby ich wyeliminowania „prowadziły do wprowadzenia nowych, ukrytych ideatów, które bez uzasadnienia były podnoszone do rangi twierdzeń naukowych” (Życiński, 2013, s. 59–60).

W kolejnych rozdziałach Życiński piętnuje nadużycia ideologiczne mechanicyzmu (rozdz. 2) i fizykalizmu (rozdz. 3). W obydwu tych rozdziałach pojawiają się wątki historyczne, ale historia jest traktowana bądź jako ilustracja tez stawianych przez Życińskiego, bądź w charakterze argumentów przeciwko tezom zwalczanym przez niego. Należy pamiętać o czasie, w którym Życiński pisał tę książkę. Wówczas twierdzenia pozytywizmu logicznego były ciągle żywe, a jedną z cech pisarstwa Życińskiego jest to, że zawsze był wrażliwy na głos współczesności. Niejednokrotnie ta wrażliwość zamieniała się w styl dziennikarski i polemiczny i często przejawia się to w tej książce. Na przykład, polemizując ze zbyt pospiesznymi próbami „sformułowania uniwersalnej teorii rewolucji naukowych na podstawie wiedzy o rewolucjach Galileusza – Newtona oraz Einsteina – Plancka” (Życiński, 2013, s. 45), Życiński porównuje je do „opracowania teorii społeczeństwa na podstawie obserwacji dwóch rencistów” (Życiński, 2013, s. 45). Albo polemizując z dziewiętnastowiecznym materializmem, pi-

sze: „w powyższym kontekście sama idea ‘obiektów matematycznych’ wydaje się jedynie tytułem honorowym używanym grzecznościowo lub przez sentyment do tradycyjnej terminologii” (Życiński, 2013, s. 138). Czytając tę książkę, niekiedy trudno oprzeć się wrażeniu, że jeżeli tylko nadarzy się okazja użycia zgrabnego sformułowania lub ciętej riposty, Życiński natychmiast ulegnie pokusie.

Życińskiego koncepcja rewolucji metanaukowej wyrosła z jego fascynacji rewolucją, jaka dokonała się w podstawach matematyki w pierwszych dekadach dwudziestego wieku. Jego metodologiczna uwaga jest głównie skierowana na twierdzenia Gödla i inne twierdzenia limitacyjne. Twierdzenia te zburzyły dotychczasowe przekonanie o niewzruszalności podstaw matematyki i tym samym przyczyniły się do zmiany paradygmatu w matematyce z *episteme* na *doxa*. Pojęcia te, nawiązujące do poglądów Platona i Arystotelesa, odgrywają istotną rolę w analizach Życińskiego. Przejście od *episteme* do *doxa* uważa on za główną charakterystykę rewolucji metanaukowej, jaka dokonała się XX wieku. Naukę typu *episteme* charakteryzuje dążenie do sformułowania niepodważalnych podstaw dla wszelkich nauk oraz poszukiwanie uniwersalnej metody. Naukę typu *doxa* charakteryzuje świadomość ograniczeń metodologicznych i zadowalanie się argumentami probabilistycznymi, gdy inne są nieosiągalne. Przejście od *episteme* do *doxa* nie oznacza jednak rezygnacji z racjonalności, lecz jedynie dostosowanie racjonalności do realnych możliwości metody. Temu wątkowi poświęcony jest rozdział 5, w którym Życiński piętnuje poglądy radykalnego socjologizmu w filozofii nauki, głoszącego koniec dotychczasowej racjonalności i dominację pozaracjonalnych składników nauki.

Ważnym spostrzeżeniem książki jest to, że istotne przemiany dokonują się na poziomie meta, a więc w metarewolucji. Czy działają tam ideaty, czy *themata*, czy metaparyadygmaty, to już mniej istotne.

O tym, czy jakaś zmiana w nauce zasługuje na miano rewolucji, decyduje to, czy przemiany, którym uległa, sięgały również do poziomu meta. Dlatego też akcja książki Życińskiego rozgrywa się głównie na tym poziomie, jednakże poziom meta nieustannie oddziałuje z tym, co dzieje się w samej nauce.

Ponieważ Życiński szuka filozoficznych uzasadnień tego, co dzieje się w nauce, na poziomie meta, książka miejscami sprawia wrażenie, jakby była opracowaniem z historii nauki, ale historii nauki w specyficznym wydaniu, gdyż nie jest to systematyczny wykład dziejów nauki, lecz raczej „próbkiwanie” historii licznymi przykładami. Jednak zbliżając się w lekturze ku końcowi, czytelnik ma ogłód „całości” dziejów nauki (choć jest to ogłód dość ogólny) od czasów starożytnych do polemik Życińskiego ze współczesnymi autorami.

Do książki Życińskiego warto powracać. Zawiera ona wiele ważkich myśli, które do dziś nie utraciły swojej doniosłości. Wśród twórców ludzkiej myśli trudno wskazać obszar, który zmieniałby się w sposób bardziej dynamiczny niż właśnie nauka, ale spojrzenie na nią z perspektywy meta pozwala uchwycić te jej cechy, które działają w znacznie szerszej czasowej skali.

Ideals and ideats

Abstract

The original view of Joseph Życiński, presented in his book *The Structure of the Metascientific Revolution* (1988), boils down to the observation that almost before our eyes a great revolution took place, not in science, but in the philosophy of science, that is the meta-scientific revolution. His concept of the meta-scientific revolution grew out of his fascination with the revolution that took place in the foundations of mathematics in the first decades of the twentieth cen-

ture. Whether a change in science deserves to be called a revolution is determined by whether the transformations it underwent also reached the meta-level. The set of presuppositions underlying transformations on the meta-level Życiński calls ideata. One of the aims of this article is to critically reconstruct the meaning of this term.

The action of Życiński's book takes place mainly on meta-level, but the meta-level constantly interacts with what is happening in science itself. The book sometimes makes an impression as if it were a study of the history of science, but history of science in a specific sense – something like a “sampling” of history with numerous examples. Among the creations of human thought, it is difficult to point to an area that changes more dynamically than science itself, but looking at it from a meta-perspective allows us to grasp those of its features that operate on a much broader scale.

Keywords

scientific revolution, meta-scientific revolution, ideats, themata, *episteme*, *doxa*.

Bibliografia

- Holton, G., 1988. *Thematic origins of scientific thoughts. Kepler to Einstein*. Cambridge: Harvard University Press.
- Kuhn, T.S., 1968. *Struktura rewolucji naukowych*. Red. S. Amsterdamski (H. Ostromecka, tłum.). Warszawa: Państwowe Wydawnictwo Naukowe.
- Liana, Z., 2019. Nauka jako racjonalna doxa. Józefa Życińskiego koncepcja nauki i filozofii nauki – poza internalizmem i eksternalizmem. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), ss. 147–199. Dostępne na: <<https://zfn.edu.pl/index.php/zfn/article/view/466>> [ostatni dostęp: 23 września 2019].

- Liana, Z., 2020. Józefa Życińskiego koncepcja racjonalizmu umiarkowanego: epistemologiczna i doxalogiczna funkcja podmiotowego commitment. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (68), ss. 117–184. Dostępne na: <<https://zfn.edu.pl/index.php/zfn/article/view/516>> [ostatni dostęp: 25 stycznia 2021].
- Życiński, J., 1988. *The Structure of the Metascientific Revolution: An Essay on the Growth of Modern Science*, seria: *Philosophy in Science Library*, 3. Tucson: Pachart Publishing House.
- Życiński, J., 2013. *Struktura rewolucji metanaukowej: studium rozwoju współczesnej nauki* (M. Furman, tłum.). Kraków: Copernicus Center Press.