

STUDIA PSYCHOLOGICA

T H E O R I A E T P R A X I S



WYDAWNICTWO NAUKOWE
UNIwersytetu Kardynała Stefana Wyszyńskiego

Warszawa 2021

**UNIWERSYTET KARDYNAŁA STEFANA WYSZYŃSKIEGO
W WARSZAWIE**

INSTYTUT PSYCHOLOGII

**TOM 21
NUMER 1
2021**

REDAKTOR NACZELNY:

JAN CIECIUCH, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska

RADA NAUKOWA:

W. KEITH CAMPBELL, University of Georgia, Stany Zjednoczone Ameryki Północnej

DAVID ENTWISTLE, Malone University, Stany Zjednoczone Ameryki Północnej

HENRYK GASIUL, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska

HUBERT HERMANS, International Institute for the Dialogical Self, Holandia

PETER JONASON, Western Sydney University, Australia

HELMUT LÜCK, Fern Universität in Hagen, Niemcy

JOHN MALTBY, University of Leicester, Wielka Brytania

PETER SCHMIDT, University of Giessen, Niemcy

ANDRZEJ STRZAŁECKI, Uniwersytet Humanistycznospołeczny SWPS, Polska

JAN F. TERELAK, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska

MICHELE VECCHIONE, Sapienza University of Rome, Włochy

MAGDALENA ŻEMOJTEL-PIOTROWSKA, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska

REDAKTORZY JĘZYKOWI:

ANNA POSTAWSKA, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska

LESZEK LECHOWICZ, Northeastern University, Stany Zjednoczone Ameryki Północnej

ADMINISTRATOR STRONY WWW:

DOMINIK GOŁUCH, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska

REDAKCJA STUDIA PSYCHOLOGICA: THEORIA ET PRAXIS

Instytut Psychologii UKSW

ul. Wóycickiego 1/3 bud. 14

01-938 Warszawa

e-mail: studia_psychologica@uksw.edu.pl

SKŁAD I ŁAMANIE:

Studio DTP Academicon | Patrycja Waleszczak

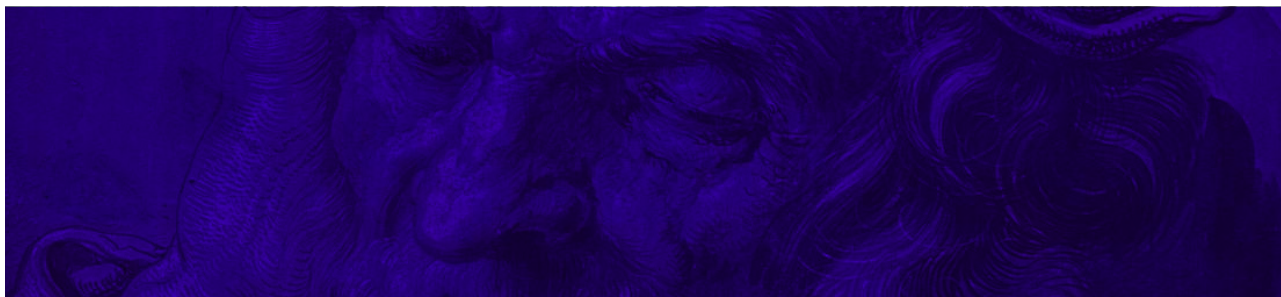
dtp@academicon.pl, dtp.academicon.pl

KOREKTA JĘZYKOWA:

Anna Postawska

WWW: <http://czasopisma.uksw.edu.pl/index.php/sp/>
ISSN (PRINT): 1642-2473 ISSN (ONLINE): 2449-5360

SPIS TREŚCI



5

Larry C. Bernard
Jan Ciecuch
Andrew Lac
Barbara Žuro
Dino Krupić
Michael Richter
Nicolas Silvestrini
Bettina von Helversen

A CROSS-CULTURAL STUDY
OF PURPOSIVE "TRAITS OF
ACTION": MEASUREMENT
INVARIANCE OF SCALES BASED
ON THE ACTION–TRAIT THEORY
OF HUMAN MOTIVATION
USING EXPLORATORY
STRUCTURAL EQUATION
MODELING

AIM-Q

motivation

motive

evolutionary psychology
personality

31

Włodzisław Zeidler

KURT LEWIN (1890–1947) –
W OCZACH JEGO BIOGRAFÓW

biografistyka

przestrzeń życiowa

źródła informacji

49

Jolanta Rytel

KONTROWERSJE
WOKÓŁ POJĘCIA TRAFNOŚCI

trafność

pomiar psychologiczny

standardy stosowania
testów psychologicznych

A CROSS-CULTURAL STUDY OF PURPOSIVE “TRAITS OF ACTION”: MEASUREMENT INVARIANCE OF SCALES BASED ON THE ACTION–TRAIT THEORY OF HUMAN MOTIVATION USING EXPLORATORY STRUCTURAL EQUATION MODELING

LARRY C. BERNARD* 

Loyola Marymount University

JAN CIECIUCH 

Cardinal Stefan Wyszyński University
and University of Zurich

ANDREW LAC 

University of Colorado – Colorado Springs

BARBARA ŽURO 

Peter McVerry Trust

DINO KRUPIC 

University of Osijek

MICHAEL RICHTER 

Liverpool John Moores University

NICOLAS SILVESTRINI 

University of Geneva

BETTINA VON HELVERSEN 

University of Basel

ABSTRACT

The Action–Trait theory of human motivation posits that individual differences in predispositional traits of action may account for variance in contemporary purposeful human behavior. Prior research has supported the theory, psychometric properties of scales designed to assess the motive dimensions of the theory, and the utility of these scales to predict an array of behaviors, but this is the first study to evaluate the cross-linguistical invariance of the 15-factor theoretical model. This study evaluated translations of the English language 60-item Quick AIM in 5 samples – Croatian ($N = 614$), French ($N = 246$), German ($N = 154$), Polish ($M = 314$), and U.S. English ($N = 490$) – recruited from 4 countries (Croatia, Poland, Switzerland, and the U.S.). Exploratory structural equation modeling (ESEM) supported the theoretical model on which the traits of action are based and scrutinized the measurement invariance (configural, metric, scalar invariance) of the scale across the languages.

AIM-Q
MOTIVATION
MOTIVE
EVOLUTIONARY PSYCHOLOGY
PERSONALITY

KEYWORDS

6	THEORETICAL BACKGROUND
7	Development and Validation of AIM-Q Scales
9	Measurement Validity of the Motive Scales
10	CURRENT STUDY
10	MATERIALS AND METHODS
10	Translation
10	Data Collection
11	Instrument
11	Samples
11	Statistical Analysis
13	RESULTS
13	Descriptive Statistics
16	Initial Models
16	Evaluating Measurement Invariance
17	DISCUSSION
19	REFERENCES
21	APPENDIX (LANGUAGE VERSIONS OF THE QUICK AIM 60-ITEM SCALE)



THEORETICAL BACKGROUND

This is a study of factor invariance of a 15-dimensional individual differences model of human motivation. Such multidimensional theories of human motivation are rare, although their utility has been acknowledged (Baumeister, 2016). Bernard, Mills, Swenson, and Walsh (2005) proposed a multidimensional theory of individual differences in human motivation. According to this theory, motivated behavior is purposeful and solves *individual fitness* challenges – involving individual survival and reproduction – and *inclusive fitness* challenges – involving interpersonal and social relationships and progeny (see Dawkins, 1982; Hamilton, 1964; Tooby & Cosmides, 1992). Bugental (2000) proposed that separate human neurohormonal systems evolved in order to overcome these challenges within different sized social domains. This led to consideration of the social domains in which more complex and inclusive fitness behaviors may operate, which include: Individual, Dyadic, Small Group, Large Group, and Very Large Group (see Bernard, 2012).

Evolution does not produce perfect universal adaptations; some individuals are better adapted to their environment. Therefore, individuals should vary in the strength of motive adaptations and this variation should allow motives to be measured as individual differences constructs. If, as Hogan (1996) contended, personality dispositions reflect reputation that influenced past survival and fitness, then perhaps motivational dispositions also reflect reputation that influenced them as well.

Bernard et al.'s (2005; Bernard, Mills, Swenson, & Walsh, 2008; see also Bernard, 2012) action–trait (A-T) theory of human motivation conceptualized motivational dispositions as “traits of action” that operate within narrow domains of human behavior. A-T theory drew on Holt’s (1931) and Tolman’s (1932) “purposive psychology”¹ and evolutionary principles suspected to drive the development of these predispositions. A-T theory offered specific criteria for the identification of a small set of independent, narrow domain motives which incorporated roles for emotion, culture, learning, conscious, and nonconscious processes, and tentatively proposed neuropsychological circuits presumed to underlie the traits of action (see Bernard, Mills, Swenson, & Walsh, 2005). The multidimensionality of A-T theory contrasts with uni- or bi-dimensional approaches that are primarily cognitive, e.g., self-regulation (Oettingen, Bulgarella, Henderson, & Gollwitzer, 2004) achievement motivation (McClelland, Atkinson, Clark, & Lowell, 1976), and intrinsic-extrinsic motivation (Deci & Ryan, 1987). The field could benefit from multidimensional theories that incorporate biological, behavioral, and environmental influences. Moreover, evolutionary theory offers a solid ground for motivational theory as it tackles the why of contemporary behavior and sets logical constraints on theory (Bernard et al., 2005). The theory posits 15 motives based on behaviors that solved specific fitness problems in specific environments. The theory also suggests specific brain regions that have evolved to mediate specific motives (Bernard et al., 2005). The utility of considering evolutionary motives lies in the strength to predict future behavior – normal and abnormal (Bernard, 2012, 2014b). A longitudinal study showed that motive strength at Time 1 can predict a statistically significant proportion of the variance

¹ Purposive psychology was a teleological refinement of James’ (1890) instinct theory and is defined as “[a] systematic exposition of psychological facts and principles which emphasizes the ends to which all actions are directed and the needs or purposes which instincts serve for the organism” (Harriman, 1975, p. 157).

in 21 of 28 health-related, self-indulgent, socially valued, and socially discordant behaviors at Time 2 (Bernard & Lac, 2014a).

A-T theory is unique in its use of evolutionary principles to posit specific criteria for determining motive traits of action and in postulating an assessment strategy for motive dimensions. This assessment strategy is called the Assessment of Individual Motives – Questionnaire (AIM-Q; Bernard, 2013a).

DEVELOPMENT AND VALIDATION OF AIM-Q SCALES

After reviewing challenges humans probably encountered in different sized social groups in the past (e.g., Bugental, 2000; Kenrick, Li, & Butner, 2003), Bernard and colleagues (2005) derived a short list of such challenges (see Table 1). To avoid lengthy lists of instincts, drives, and needs that have plagued prior multidimensional motivational theories (see Dunlap, 1919), specific criteria were used to identify each motive: (1) it must address a recurring challenge in humans' past; (2) it had the potential to affect survival and fitness; and (3) it can be demonstrated to operate as a relatively independent individual differences dimension of purposive behavior today. The result was the identification of 15 motives that emerged based on putative evolutionary processes. The labels for the motives and their conceptual definitions (in terms of behavior) are listed in Table 1.

Table 1. Recurring Challenges, Motive Labels, Conceptual Definitions, and Example Items

Challenge	Motive	Conceptual Definition	Example Item
Integrity Motives			
Identification of environmental resources and hazards	Environmental Inquisitiveness	Use resources to explore the physical environment; evaluate resource availability/hazards in new and different things, places, and situations	I explore unknown places to understand them better.
	Illness Avoidance	Use resources to maintain bodily integrity and health	I am very careful not to touch anything after I wash my hands in public place.
	Threat Avoidance	Use resources to maintain personal safety; avoid challenges to one's person and resources	I spend a lot of effort to avoid physical threats of any kind.
	Interpersonal Inquisitiveness	Use resources to explore the social environment; test limits and how others act, react, and interact; compare oneself to others	I try to understand others by playfully teasing them to see how they react.
Counteracting despair over the awareness of mortality	Meaning	Use resources to identify with, construct, and maintain a philosophy, purpose, or rationalization for existence (and non-existence) and attempts to arrive at an understanding and peace with the presumed purpose of life	I learn about various religions or philosophies to understand the purpose of life better.
Competitive Motives			
Competing for control of resources	Aggression	Use resources challenge and intimidate others for control of resources; approach challenges to one's person and resources with combative and intimidating displays and actions	I use physical threats so that others know not to get in my way.

Challenge	Motive	Conceptual Definition	Example Item
Competitive Motives			
Competing to increase social status and mating desirability	Appearance	Use resources to compete on the basis of physical appearance	I spend money and time on stylish haircuts to improve my appearance.
	Mental	Use resources to compete on the basis of intellectual abilities and knowledge; develop and display other skills, abilities, and talents	I put a lot of effort into developing my own special skills and abilities to impress others.
	Physical	Use resources to compete on the basis of physical strength, endurance, size, shape, and stature	I train hard to have the reputation of being in peak physical condition.
	Wealth	Use resources to compete by acquiring material resources	I buy (or would buy) expensive or rare luxury goods so that others will know I am successful.
	Sex	Use resources to seek sexual activity	I spend (or have spent) a lot of time and energy to have sex.
Cooperative Motives			
Formation of cooperative relationships with mates and offspring	Commitment	Transfer resources to mates and offspring; develop tender, intimate, supportive attachments with mate and close kin	I give what is necessary of my time and resources in romantic relationships.
Formation of coalitions of kin through development of individual-to-individual reciprocity	Altruism	Transfer resources to kin without expectation of immediate self-benefit (but at a cost to oneself that is generally lower than the benefits to others multiplied by their degree of genetic relatedness)	I sacrifice (or would sacrifice) to help distant relatives.
Formation of coalitions of non-kin through development of individual-to-individual reciprocity	Social Exchange	Enter into reciprocal, mutually beneficial exchanges of resources with non-kin; share resources fairly and without cheating; do what is legally and socially prescribed and avoids what is proscribed	I do what is required to keep a contract or bargain I have made.
Formation of coalitions of non-kin through development of individual-to-group-to-individual reciprocity ("institutionalized reciprocity")	Legacy	Transfer resources to institutions that benefit non-kin as much as, or more than, kin without the expectation of direct reciprocity to oneself, but perhaps to offspring and kin	I contribute time, effort, and/or money to make the world better for future generations.

The conceptual definitions of the motive dimensions were used to generate an initial item pool of 180 items to assess the 15 motive factors (12 items per factor). The items in the factors were initially validated using Principal Components Analysis (Bernard et al., 2008). Furthermore, three preliminary versions of AIM-Q scales were constructed to permit a multitrait-multimethod (MTMM; Campbell & Fiske, 1959) analysis of the convergent and discriminant validities of the motive factors across different assessment methods (Bernard et al., 2008). The MTMM matrix was scrutinized using Pearson correlations involving

all 15 motive composites serving as the traits as applied to three methods that were completed by all participants. These three method types consisted of self-reports that varied based on instructional and response option modifications: (1) indicate the best use of your personal resources (responses ranging from *strongly disagree* to *strongly agree*); (2) endorse your own interests, desires, and likes (responses ranging from *completely untrue* to *completely true*); and (3) indicate your behavioral priorities (responses ranging from *no importance at all* to *great importance*). Results provided empirical support for the overall strategy, across the various measurement methods, of assessing individual differences in motivation with the AIM-Q motives.

The number of items was subsequently reduced to 120 in a revision that exhibited strong internal consistency reliabilities for each subscale. These subscales were demonstrated to be highly correlated with subscales from the original longer version (Bernard, 2013a). The next objective was to evaluate the factor structure of the 15 factors using confirmatory factor analysis and to reduce the inventory to an even more manageable length for participants (Bernard & Lac, 2014a). Applying confirmatory factor analysis to the 120-item variant of the scale, the model fit indices for the original 120 item scale were judged to be mediocre (CFI = .79, IFI = .79, RMSEA = .047). The inspection of post hoc modification indices revealed that poor model fit was attributed to items with substandard factor loadings. Accordingly, the four items with the highest standardized factor loadings for each factor were retained. This process yielded a 60-item 15-dimension version (named the “Quick AIM”) that produced satisfactory fit indices (CFI = .91, IFI = .92, and RMSEA = .041). Finally, using a different sample in the same study (Bernard & Lac, 2014a), confirmatory factor analysis successfully cross-validated the factor structure and yielded desirable fit indices. Additional confirmatory factor analyses using this sample examined competing factor structures, including multidimensional variants of orthogonal (non-correlated factors) and oblique (correlated factors) models and a single-factor model (Bernard & Lac, 2014a). Fit indices and factor loadings best supported the original, oblique 15-dimensional model. This version of the Quick AIM was used in the present study.

MEASUREMENT VALIDITY OF THE MOTIVE SCALES

Other reliability and validity studies have evaluated additional statistical properties of the AIM-Q motive scales (Bernard, 2013a) and supported the construct validity of AIM-Q and Quick AIM versions. First, AIM-Q scales have demonstrated convergent validity with reliable and valid external measures for aggression, need for cognition, playfulness, and sexuality (Bernard, 2007). Specifically, certain hypothesized motives correlated logically with measures of aggression (anger, hostility, physical aggression, verbal aggression) ranging from $r = .60$ to $.67$; measures of the need for cognition ranging from $r = -.54$ to $.72$; measures of playfulness (spontaneous, expressive, fun, creative, silly) ranging from $r = -.47$ to $.62$; and measures of sexuality ranging from $r = -.54$ to $.47$.

Second, self-ratings on AIM-Q motives have been correlated with observer ratings by friends ($r = .29$ to $.64$) and relatives ($r = .32$ to $.56$; Bernard, 2009). This suggests that self and other judgments of motives converge in terms of the interpretation and meaning of the motive dimensions as descriptors of behavior. Third, experimental research of the motives' validity demonstrated that differences in some motives can affect judgments of facial attractiveness (Bernard & Hardy, 2014).

Fourth, A-T theory predicted that motives should explain variance in both normal and abnormal personality traits. Indeed, studies have suggested that certain motives are correlated with normal personality traits as assessed by the HEXACO (Lee & Ashton, 2004) six factor model (Bernard, Lac, & Lukaszewski, 2013) and the Five Factor Model

of personality (Bernard, 2010). And, as predicted, certain motives are also correlated with abnormal personality traits such as borderline personality, psychopathy, and maladjustment (Bernard, 2013b).

Fifth, AIM-Q motives predicted a variety of self-reported behaviors confirmed by reports of observers (Bernard, 2009). These behaviors included smoking ($R^2_{Adj.} = .15$), drinking alcohol ($R^2_{Adj.} = .21$), using illegal substances ($R^2_{Adj.} = .17$), holding positions of leadership ($R^2_{Adj.} = .29$), and exercising ($R^2_{Adj.} = .42$). Finally, a longitudinal study to address the issue of predictive validity demonstrated that AIM-Q motives were incrementally valid predictors of *future* behaviors (Bernard & Lac, 2014b).

CURRENT STUDY

This programmatic series of studies have supported the measurement validity of the AIM-Q and Quick AIM scales up to this point. However, until now, all previous studies have used only U.S. English-speaking samples. It is currently unknown whether results would be similar if the research had been conducted with other cultures or languages. Thus, this is the first study to attempt to corroborate the 15-factor structure of motive dimensions across linguistic and cultural groups. For the theoretical framework and scale to be linguistically invariant, the individual differences in motive dimensions should not be unique to the U.S. English speaking samples. Thus, the present study was meant to examine the factor structure and measurement invariance (configural, metric, and scalar) of the 15 Quick AIM scales cross-linguistically for the first time and is a crucial step in theory development and hypothesis testing.

MATERIALS AND METHODS

TRANSLATION

The U.S. English Quick AIM was translated into four different languages: Croatian, Polish, German and French. The International Testing Commission (2008) guidelines were followed for the translations of the original U.S. English Quick AIM items and instructions. Bilingual doctoral level psychologists and graduate students participated in multiple forward and backward translations. The back translations into English were reviewed by Quick AIM author (L. Bernard). Edits were provided as feedback to translators who modified items as necessary. Modified items were once again back translated into English for a final check. All the language versions used in the current study are in the Appendix.

DATA COLLECTION

Participants were treated in accordance with the ethical standards of institutional and governmental organizations for the treatment of human participants. All administrations were performed anonymously, with minimal demographic data collected. The Quick AIM was administered online through Internet survey programs and took approximately 15 minutes to complete.

INSTRUMENT

Quick AIM. The Quick AIM is a U.S. English language 60-item (Bernard, 2013a; Bernard & Lac, 2014a) empirically validated version of the AIM-Q scales. Each factor consists of 4 items. Respondents endorsed each item as true or untrue using a 6-point scale: 1 (*completely untrue*), 2 (*mostly untrue*), 3 (*slightly untrue*), 4 (*slightly true*), 5 (*mostly true*), to 6 (*completely true*). The Quick AIM scales' internal consistency reliabilities obtained in the current study are listed in Table 2.

SAMPLES

Samples from five languages were recruited – Croatian, French, German, Polish, and English:

Croatian. The sample consisted of $N = 614$ participants ranging in age from 16 to 73 years ($M = 28.0$, $SD = 9.7$). Fifty-seven percent of the sample was female. Half of the sample were university students in Croatia studying psychology and half were non-students recruited by the students.

French. The sample consisted of $N = 246$ participants studying psychology in Switzerland. Due to local privacy concerns, data about age and sex were not obtained.

German. The sample consisted of $N = 154$ participants ranging in age from 18 to 57 years ($M = 26.0$, $SD = 7.8$). Sixty-seven percent of the sample was female. The sample was comprised of university students studying psychology in Switzerland.

Polish. The sample consisted of $N = 314$ participants ranging in age from 19 to 71 years ($M = 31.5$, $SD = 12.7$). Sixty-five percent of the sample was female. The majority of the sample was university students studying psychology in Poland.

U.S. English. The sample consisted of $N = 490$ participants ranging in age from 18 to 78 years ($M = 27.7$, $SD = 13.50$). One-third of the participants were psychology university students and two-thirds were non-student adults living in diverse geographic regions, mostly within California. Fifty-six percent of the sample was female.

STATISTICAL ANALYSIS

The skewness of the 60 items ranged from -0.83 to 2.87 for the Croatian sample, -2.03 to 2.06 for the French sample, -1.65 to 2.57 for the German sample, -1.24 to 2.23 for the Polish sample, and -1.55 to 1.42 for the U.S. sample. Thus, the items were found to be within reasonable boundaries of normality and therefore the data were appropriate for factor-analytic techniques (Tabachnick & Fidell, 2013; Warner, 2013). However, robust estimation will be applied to the statistical techniques to automatically correct for any departures from normality (Hoyle, 2012).

Considering that 15 factors will be estimated in the study, model testing with exploratory structural equation modeling (ESEM) was warranted. This approach represents the integration of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA is not conducive for testing measurement and structural invariance across groups, while CFA requires that each item is specified to load exclusively on its dominant factor and constrained to exactly zero loadings on all other factors. Thus, CFA has been criticized as overly restrictive and not reflecting the reality of the data. Previous investigations scrutinizing multi-factor structures, as the Big Five personality inventory, have suggested that ESEM yields better approximation of the underlying data (Marsh et al., 2010; Vazsonyi, Ksinan, Mikuška, & Jiskrova, 2015). Marsh, Morin, Parker, and Kaur (2014) stated it is almost impossible to generate an acceptable model fit in CFA when multiple factors (5 to 10) are

measured by a reasonable number of items (5 to 10) per subscale in an inventory of at least 50 items. The number of factors and total number of items in this investigation exceeds this recommendation and therefore indicates that ESEM would be the ideal technique.

ESEM was conducted using *Mplus* 7.13 (Muthén & Muthén, 2012). Loadings of items intended to measure a given factor were freely estimated while all cross-loadings were allowed and targeted to be close to zero using “target rotation” (Marsh, Morin, Parker, & Kaur, 2014; Morin, Boudrias, Marsh, Madore, & Desrumaux, 2016). Marsh and colleagues (2014) state that “target rotation is particularly appropriate when there is a clearly defined a priori factor structure and a reasonable approximation to simple structure” and that this rotation is “based on partial knowledge of the factor structure, and is consistent with the view that ESEM is more typically used for confirmatory rather than exploratory purposes” (Marsh et al., 2014). Thus, ESEM using target rotation was pursued given that we possess partial knowledge of the factor structure from previous studies using the English version (but not the other translated versions) and that the research purpose was to confirm the factor structure across the other languages. Consistent with ESEM, the 15 factors were permitted to be correlated. The estimator was maximum likelihood estimation with robust estimation to account for variable departures from normality. A model was estimated for each linguistic sample.

After obtaining acceptable factor structures for each linguistic group, multi-group analyses were conducted using recommended guidelines (Byrne, 2012; Davidov, Meuleman, Cieciuch, Schmidt, & Billiet, 2014; Lance & Vandenberg, 2002) to test for measurement invariance of the questionnaire. Three levels of invariance were tested: Configural, metric and scalar. Configural invariance requires that latent factors are loaded by the same items in all groups, with each group representing a different language. Metric invariance requires that factor loadings of the items on the latent factors are the same across groups. Scalar invariance requires that intercepts of all items are equal across groups, in addition to the equality constraint of loadings at the metric level. Metric invariance is a precondition for meaningful comparisons of factor correlations across groups and scalar invariance is a precondition of latent means comparison across groups (Chen, 2007; Lance & Vandenberg, 2002). Byrne, Shavelson, and Muthén (1989) introduced the concept of partial measurement invariance. It implies that at least two parameters (loadings at the metric level or loadings and intercepts at the scalar level) are found to be equal across groups, while other parameters are freely estimated. Partial measurement invariance is assumed to allow the same types of comparisons as full measurement invariance (Byrne, Shavelson, & Muthén, 1989).

Several fit indices were used to scrutinize the overall fit of models. The model chi-square test was reported. Goodness-of-fit values of the Comparative Fit Index (CFI) ranged from 0.00 to 1.00, with higher values signifying a better fitting model (Tabachnick & Fidell, 2007). A value above .90 indicates good fit and above .95 very good fit (Hu & Bentler, 1999). Two badness-of-fit indices were also evaluated. The Root Mean-Square Error of Approximation (RMSEA) is adequately sensitive in detecting model misspecification and offers appropriate information about model quality (MacCallum & Austin, 2000). A value below .05 indicates close fit, between .05 and .08 fair fit, between .08 and .10 mediocre fit, and above .10 poor fit (MacCallum, Browne, & Sugawara, 1996). According to Hu and Bentler (1999) RMSEA below .08 indicates a good model fit. Standardized Root Mean-Square Residual (SRMR) has shown to be useful for detecting model misspecification, with a recommended value below .08 diagnostic of good fit (Hu & Bentler, 1998).

In evaluating the extent that measurement invariance was established, we applied two criteria: the global fit proposed by Chen (2007) and the examination of local misspecification as proposed by Saris, Satorra, and van der Veld (2009) and applied to the test of measurement invariance by Cieciuch, Davidov, Oberski, and Algesheimer (2015). We used the first approach to test for the full measurement invariance and the second for

partial measurement invariance to identify noninvariant loadings and intercepts. For the first approach, metric noninvariance is indicated by a change larger than .010 in CFI, supplemented by a change larger than .015 in RMSEA or a change larger than .030 in SRMR compared with the configural invariance model. Regarding scalar invariance, noninvariance is indicated by a change larger than .010 in CFI, supplemented by a change larger than .015 in RMSEA or a change larger than .010 in SRMR compared with the metric invariance model. For the second approach, power of the test was taken into account and local misspecification was examined (Cieciuch, Davidov, Oberski, & Algesheimer, 2015). This way, we could identify parameters with noninvariant loadings and intercepts. We followed the suggestion of Saris et al. (2009) and treated deviation in loadings greater than .40 and in intercepts greater than .10 as a sign of serious misspecification.

RESULTS

DESCRIPTIVE STATISTICS

Descriptive statistics are presented in Tables 2–7. Table 2 presents means, standard deviations and Cronbach's alpha coefficients for the scales in each group and Table 3–7 present intercorrelations between scales in each country.

Table 2. Cronbach's Alpha for Quick AIM

	English (U.S.)		Croatian		German		Polish		French	
	α	<i>M</i> (SD)	α	<i>M</i> (SD)	α	<i>M</i> (SD)	α	<i>M</i> (SD)	α	<i>M</i> (SD)
Aggression	.86	2.27 (1.19)	.76	1.61 (0.77)	.86	1.67 (0.88)	.80	1.60 (0.73)	.72	1.88 (0.82)
Threat Avoidance	.69	3.98 (1.01)	.68	3.91 (1.01)	.53	3.43 (0.91)	.59	3.35 (0.89)	.63	3.67 (0.92)
Illness Avoidance	.69	3.45 (1.09)	.59	3.12 (0.96)	.67	2.82 (1.09)	.52	3.17 (0.90)	.60	2.90 (0.97)
Environment Inquisitiveness	.80	4.40 (0.93)	.87	4.47 (1.02)	.83	4.81 (0.87)	.69	4.16 (0.86)	.83	4.73 (0.82)
Interpersonal Inquisitiveness	.87	3.27 (1.30)	.91	2.43 (1.24)	.87	3.07 (1.25)	.88	2.38 (1.21)	.88	3.73 (1.19)
Sex	.86	2.33 (1.30)	.86	2.03 (1.05)	.84	2.47 (1.17)	.79	2.24 (1.06)	.91	2.30 (1.28)
Appearance	.84	3.28 (1.34)	.79	1.69 (0.87)	.88	2.57 (1.19)	.87	2.46 (1.26)	.81	2.69 (1.18)
Wealth	.87	2.93 (1.28)	.83	1.62 (0.84)	.86	1.83 (0.97)	.90	2.19 (1.18)	.88	2.09 (1.07)
Mental	.81	3.40 (1.12)	.84	2.56 (1.14)	.77	3.45 (1.06)	.84	2.94 (1.17)	.85	3.07 (1.16)
Physical	.91	3.48 (1.40)	.88	2.59 (1.25)	.92	2.95 (1.35)	.92	2.54 (1.24)	.74	3.14 (1.11)
Commitment	.88	4.95 (1.01)	.89	4.56 (1.16)	.80	4.74 (0.98)	.77	4.68 (0.97)	.72	4.96 (0.77)
Altruism	.82	4.45 (1.00)	.74	3.98 (0.98)	.84	4.32 (1.03)	.81	4.25 (0.90)	.81	4.49 (0.94)
Social Exchange	.77	5.15 (0.71)	.72	5.11 (0.63)	.73	5.38 (0.52)	.71	4.80 (0.66)	.74	5.13 (0.61)
Legacy	.84	3.88 (1.15)	.87	3.20 (1.24)	.85	3.35 (1.22)	.85	3.17 (1.11)	.91	3.14 (1.27)
Meaning	.89	3.49 (1.25)	.92	2.99 (1.42)	.90	3.75 (1.28)	.92	2.74 (1.26)	.91	3.46 (1.32)

Table 3. Intercorrelations between Quick AIM scales in Croatia

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Aggression														
(2) Threat Avoidance	-.22**													
(3) Illness Avoidance	.10*	.29**												
(4) Environment Inquisitiveness	-.03	.05	-.01											
(5) Interpersonal Inquisitiveness	.42**	-.13**	.03	.11**										
(6) Sex	.34**	-.14**	.00	-.04	.37**									
(7) Appearance	.17**	-.06	.09*	.09*	.11**	.15**								
(8) Wealth	.36**	-.11**	.09*	-.03	.27**	.32**	.65**							
(9) Mental	.35**	.01	.06	.11**	.41**	.28**	.28**	.49**						
(10) Physical	.30**	-.10*	.12**	.09*	.16**	.28**	.25**	.33**	.39**					
(11) Commitment	-.12**	.16**	.03	.16**	-.04	-.06	.02	-.12**	-.02	.02				
(12) Altruism	-.04	.07	.00	.19**	.02	.01	-.04	-.02	.05	.08	.25**			
(13) Social Exchange	-.21**	.24**	.12**	.34**	-.15**	-.18**	-.13**	-.24**	-.04	.01	.33**	.34**		
(14) Legacy	-.06	.09*	.06	.26**	-.13**	-.15**	.04	-.06	.00	.00	.05	.31**	.30**	
(15) Meaning	.14**	-.06	-.02	.35**	.13**	.08*	.03	-.03	.25**	.07	.07	.11**	.11**	.29**

Note. * $p < .05$; ** $p < .01$

Table 4. Intercorrelations between Quick AIM scales in France

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Aggression														
(2) Threat Avoidance	-.09													
(3) Illness Avoidance	.06	.30**												
(4) Environment Inquisitiveness	.03	-.07	.00											
(5) Interpersonal Inquisitiveness	.40**	-.08	-.01	.07										
(6) Sex	.39**	.01	.02	-.04	.32**									
(7) Appearance	.26**	.14*	.16*	.05	.15*	.23**								
(8) Wealth	.39**	.18**	.18**	-.04	.27**	.37**	.57**							
(9) Mental	.41**	.10	.03	.21**	.24**	.42**	.34**	.43**						
(10) Physical	.25**	.00	-.03	.17**	.14*	.28**	.31**	.25**	.45**					
(11) Commitment	.03	.04	-.04	.13*	.04	.01	.20**	.08	.18**	.21**				
(12) Altruism	-.06	-.03	.01	.16*	.00	-.07	.07	-.03	-.01	.10	.15*			
(13) Social Exchange	-.32**	.05	-.05	.27**	-.20**	-.33**	.05	-.24**	-.09	-.04	.27**	.32**		
(14) Legacy	.09	.01	.12	.33**	-.02	-.02	.01	-.11	.09	.16*	.01	.25**	.22**	
(15) Meaning	.22**	-.04	.08	.36**	.08	.13*	.00	.00	.31**	.31**	.08	.15*	.07	.36**

Note. * $p < .05$; ** $p < .01$

Table 5. Intercorrelations between Quick AIM scales in Germany

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Aggression														
(2) Threat Avoidance	.01													
(3) Illness Avoidance	.22**	.18*												
(4) Environment Inquisitiveness	.01	-.13	.16											
(5) Interpersonal Inquisitiveness	.46**	-.04	.20*	.20*										
(6) Sex	.44**	.07	.17*	.13	.43**									
(7) Appearance	.31**	-.02	.27**	.09	.20*	.27**								
(8) Wealth	.48**	.10	.21**	.04	.21*	.32**	.72**							
(9) Mental	.33**	.08	.08	.16*	.31**	.34**	.36**	.44**						
(10) Physical	.43**	.04	.23**	.17*	.36**	.45**	.32**	.39**	.47**					
(11) Commitment	.02	.07	.16*	.11	-.02	.09	-.04	-.02	.19*	.12				
(12) Altruism	.19*	-.09	-.05	.24**	.23**	.08	.08	.13	.13	.13	.13			
(13) Social Exchange	-.19*	-.08	.03	.27**	-.14	-.24**	.06	-.01	.08	-.01	.29**	.25**		
(14) Legacy	-.03	.08	.05	.25**	-.01	-.09	-.06	-.08	.06	-.01	.08	.08	.20*	
(15) Meaning	.13	.09	.17*	.48**	.13	.18*	.04	.06	.19*	.09	.02	.13	.17*	.34**

Note. * $p < .05$; ** $p < .01$

Table 6. Intercorrelations between Quick AIM scales in Poland

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Aggression														
(2) Threat Avoidance	.03													
(3) Illness Avoidance	.12*	.36**												
(4) Environment Inquisitiveness	-.03	.17**	.25**											
(5) Interpersonal Inquisitiveness	.44**	-.08	-.04	.23**										
(6) Sex	.39**	.03	.06	.18**	.41**									
(7) Appearance	.23**	.08	.25**	.18**	.18**	.40**								
(8) Wealth	.41**	.03	.15**	.14*	.36**	.48**	.74**							
(9) Mental	.37**	.19**	.13*	.39**	.40**	.43**	.47**	.61**						
(10) Physical	.33**	.03	.18**	.20**	.25**	.42**	.47**	.45**	.50**					
(11) Commitment	-.10	.13*	.05	.22**	.04	.12*	.13*	.10	.11	.05				
(12) Altruism	-.14*	.06	.05	.16**	-.08	-.10	.02	-.08	.00	.11*	.36**			
(13) Social Exchange	-.34**	.15**	.05	.18**	-.22**	-.21**	-.11	-.22**	-.07	-.13*	.31**	.47**		
(14) Legacy	.02	.10	.29**	.07	-.15**	-.07	.11	-.01	-.05	.08	.16**	.41**	.30**	
(15) Meaning	.19**	.04	.05	.429**	.23**	.23**	.14*	.10	.24**	.21**	.11	.14*	.03	.31**

Note. * $p < .05$; ** $p < .01$

Table 7. Intercorrelations between Quick AIM scales in U.S.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Aggression														
(2) Threat Avoidance	-.24**													
(3) Illness Avoidance	.03	.34**												
(4) Environment Inquisitiveness	-.05	.02	-.01											
(5) Interpersonal Inquisitiveness	.49**	-.13**	-.08	.11*										
(6) Sex	.51**	-.19**	-.16**	.05	.36**									
(7) Appearance	.19**	.19**	.12**	.05	.03	.18**								
(8) Wealth	.47**	.05	.05	.02	.30**	.35**	.63**							
(9) Mental	.45**	-0.04	-.01	.18**	.41**	.36**	.32**	.55**						
(10) Physical	.40**	-.05	.08	.08	.26**	.28**	.27**	.42**	.50**					
(11) Commitment	-.24**	.13**	.04	.32**	-.03	-.06	.14**	.01	.05	.06				
(12) Altruism	-.10*	.04	.10*	.30**	.01	.00	-.01	-.06	.13**	.17**	.30**			
(13) Social Exchange	-.45**	.18**	.08	.32**	-.24**	-.32**	-.10*	-.23**	-.09*	-.07	.46**	.37**		
(14) Legacy	-.08	.18**	.21**	.28**	-.15**	-.14**	.13**	-.03	.04	.04	.18**	.23**	.24**	
(15) Meaning	.11*	.02	.07	.42**	.11*	.12**	.05	.04	.22**	.22**	.18**	.22**	.13**	.36**

Note. * $p < .05$; ** $p < .01$

INITIAL MODELS

The 15-factor structure representing the 60-item Quick AIM was separately estimated for each of the five samples using ESEM. All the items were retained in all single-group models. Fit indices for all the linguistic groups were desirable: CFI ranged from .924 to .965, RMSEA from .033 to .051, and SRMR from .017 to .026 (see Table 8). Compared to the models for the other linguistic groups, the U. S. sample generated the best fit, not surprisingly, as the seminal version of the scale was developed in English. The factor structures for Croatian, German, and Polish also rendered acceptable fit indices. Although the software reported fit indices for French, convergence problems were indicated, so this structure should be interpreted with caution.

EVALUATING MEASUREMENT INVARIANCE

Due to estimation problems with the factor structure for French, this linguistic group was omitted from all subsequent analyses and multigroup ESEM was conducted using the U.S., Croatian, German, and Polish samples. No other problems were encountered in subsequent analyses. The multiple-group invariance tests for the four remaining languages indicated that item #52 (see Appendix for item wording) should be dropped, resulting in a 59-item model. Furthermore, the pair of error terms involving items #21 and #51; and items #41 and #56 were indicated to be strongly correlated, but these items were retained.

Table 8. Exploratory Structural Equation Modeling of the Quick AIM 60-Item Scale

Language	Model χ^2	CFI	RMSEA [90% CI]	SRMR
Croatian	1876.55	.952	.039 [.036–.041]	.019
French	1419.84	.944	.043 [.038–.048]	.023
German	1369.63	.924	.051 [.045–.057]	.026
Polish	1617.03	.936	.046 [.042–.050]	.022
English (U.S.)	1505.94	.965	.033 [.030–.037]	.017

Note. $df = 975$, all model chi-square tests were significant, $p < .01$.

Table 9. Global Fit Measures for the Multigroup Exploratory Structural Equation Modeling

Level of invariance	Model χ^2	df	RMSEA	SRMR	CFI
Configural	5871.57	3716	.038 [.037–.040]	.019	.955
Metric	9538.45	5696	.041 [.040–.043]	.045	.920
Scalar	10573.71	5828	.046 [.044–.047]	.046	.901

Note. Quick AIM with item #52 deleted. Two pairs of errors were correlated: Items #21 and #51, and items #41 and #56, but retained in the analysis. The French sample was not included.

The nested multi-group models involving the evaluation configural, metric, and scalar invariance are presented in Table 9. The multi-group configural model yielded satisfactory fit indices. Metric invariance was evaluated in the multi-group models next. It turned out that the deterioration of CFI and SRMR exceeds the recommended thresholds (Chen, 2007) that suggest the lack of metric measurement invariance. However, no local misspecifications were identified in applying the Saris et al. (2009) approach that suggests that the general result of metric noninvariance was caused by many small misspecifications rather than several large ones. Taking into account the large number of items (60) such results are not surprising.

Next, scalar invariance was evaluated in the multi-group model. Between the metric and scalar level, a deterioration of the indices was not exhibited (Chen, 2007). However, applying the local misspecification approach identified fewer than two invariant intercepts in some of the factors, which was not entirely supportive of scalar invariance.

Taken together, results suggest that the samples representing the different linguistic groups exhibited very good configural invariance of the Quick AIM scale after deleting 1 item. However metric and scalar measurement invariance is potentially problematic.

DISCUSSION

Obtaining empirical evidence to support a theory of the complexity of A-T theory has required a carefully planned, lengthy, iterative, and incremental development and testing process involving a variety of cross-sectional, longitudinal, and experimental designs (Bernard, 2013a). However, all previously published studies involved U.S. English-speaking samples, so the generalizability of Quick AIM scores and A-T theory remained an open question. The next logical step was to provide evidence of the measurement invariance of the model and the Quick AIM scales across samples of different languages and countries.

There was no reason to presume that *all* of the 60 Quick AIM items tapping the 15 motive dimensions in U.S. English samples should function well psychometrically in the different languages/cultures. Applying ESEM, the CFI suggested that the models were acceptable and the residual-based indices were more favorable for each of the five samples.

However, there were mixed results in the multi-group invariance analyses, with support for configural but not metric and scalar invariance. Nevertheless, it was encouraging that 59 of the original 60 Quick AIM items performed well enough to test the 15-dimensional model proposed by A-T theory and the configural level. The fact that the 15-dimensional model exhibited configural invariance across samples from four different languages is evidence of its utility for describing and assessing basic individual differences dimensions of human motivation. Due to the relatively small sample sizes for some of the groups, future research should cross-validate the results of the current study using a greater number of participants for each linguistic group in invariance analyses. Sample sizes of the groups might affect the interpretation of invariance analyses, as the RMSEA fit index is sensitive to sample sizes (Putnick & Bornstein, 2016).

A new consideration arising from the present study is that the Illness Avoidance and Threat Avoidance scales possessed considerably lower internal consistency reliabilities across the samples than any other scales (Table 2). Cronbach's alpha is known to be sensitive to scale length (increasing as additional items with similar inter-item correlations are added). However, all of the Quick AIM scales, at four items, are brief and good values for Cronbach's alpha were obtained for the other 13 subscales. Although Threat Avoidance and Illness Avoidance have had adequate internal consistency reliabilities in lengthier versions of the scales in past studies, they have been consistently among the lowest (Bernard, 2013a). It may be that further investigation could yield better items to represent these two dimensions. However, it should be recognized that these two dimensions are characterized as "avoidance" and may be fundamentally different from the other dimensions that are primarily characterized as "approach" dimensions in A-T theory.

Important limitation of this study is that there were no non-Indo-European language families or non-Western countries represented in these samples. Five samples in this study share a similar historical and cultural frame that may differ from the other language families. Therefore, while they offer a test of the cross-linguistical applicability of the Quick AIM motives, samples may not be dissimilar enough to fully test the cross-linguistical applicability and factorial invariance of the A-T model. Still, these results are promising and they do suggest that this study should be replicated with more diverse linguistic and cultural samples. A difficulty that affects this type of research is that it is very time consuming and costly to conduct, requiring cross-national cooperation. Future research of A-T theory may focus on the importance of basic demographic variables, such as age, gender and cultural background. Some motive-traits (e.g., appearance, illness avoidance) may vary depending on such variables. For example, men place a premium on physical attractiveness in their choice of romantic partners. Accordingly, women tend to attract or retain mates by enhancing their physical attractiveness (e.g., Buss, 1988). This is reflected in the ways that women spend more time making themselves appear physically attractive compared with men (Aune & Aune, 1994). Moreover, elderly people are considered to be more vulnerable to disease. Accordingly, illness avoidance behaviors are more frequent in elderly adults (Prohaska, Leventhal, Leventhal, & Keller, 1985).

Given the rarity of an individual differences approach to the study of human motivation, this study suggests that there is promise in this method. If additional cross-cultural studies also produce evidence of these dimensions' relevance, AIM-Q scales may be useful adjuncts in the study of motivation. Brief, reliable, and valid scales based on A-T theory can be included in a variety of research studies and may add an additional dimension to clinical assessments.

REFERENCES

- Aune, R. K., & Aune, K. S. (1994). The influence of culture, gender, and relational status on appearance management. *Journal of Cross-Cultural Psychology, 25*, 258–272. <https://doi.org/10.1177/0022022194252006>
- Baumeister, R. F. (2016). Toward a general theory of motivation: Problems, challenges, opportunities, and the big picture. *Motivation and Emotion, 40*, 1–10. <https://doi.org/10.1007/s11031-015-9521-y>
- Bernard, L. C. (2007). Assessing individual differences in motivation: Convergent validity of the Assessment of Individual Motives – Questionnaire [AIM-Q] and measures of aggression, cognition, playfulness, and sexuality. *Individual Differences Research, 5*, 158–174.
- Bernard, L. C. (2009). Consensual and behavioral validity of a measure of adaptive individual differences dimensions in human motivation. *Motivation and Emotion, 33*, 303–319. <https://doi.org/10.1007/s11031-009-9131-7>
- Bernard, L. C. (2010). Motivation and personality: Relationships between putative motive dimensions and the five factor model of personality. *Psychological Reports, 106*, 1–19. <https://doi.org/10.2466/PRO.106.2.613-631>
- Bernard, L. C. (2012). Evolved individual differences in human motivation. In R. M. Ryan (Ed.), *The Oxford Handbook of Human Motivation* (pp. 381–407). Oxford, UK: Oxford University Press.
- Bernard, L. C. (2013a). *Manual for the Assessment of Individual Motives – Questionnaire*. Department of Psychology, Loyola Marymount University, Los Angeles, CA.
- Bernard, L. C. (2013b). Relationships between individual differences in motivation and borderline personality disorder, psychopathy, and maladjustment. *Psychological Reports, 113*, 129–161. <https://doi.org/10.2466/14.02.09.PRO.113x14z9>
- Bernard, L. C., & Hardy, D. J. (2014). Motivated behavioral outcomes affect ratings of attractiveness. *Psychological Reports, 115*, 849–871. <https://doi.org/10.2466/21.02.PRO.115c30z0>
- Bernard, L. C., & Lac, A. (2014a). Testing a multidimensional model of putative evolved human motives. *Motivation and Emotion, 38*, 47–64. <https://doi.org/10.1007/s11031-013-9360-7>
- Bernard, L. C., & Lac, A. (2014b). The incremental validity of motive “traits of action:” Predicting behavior longitudinally. *Individual Differences Research, 12*, 79–100.
- Bernard, L. C., Lac, A., & Lukaszewski, A. (2013, May). *Motivation and personality: Relationships between QuickAIM motives and the HEXACO model of personality*. Poster presented at the meeting of the Society for the Study of Motivation, Washington, DC.
- Bernard, L. C., Mills, M., Swenson, L., & Walsh, R. P. (2005). An evolutionary theory of human motivation. *Genetic, Social, and General Psychology Monographs, 131*, 129–184. <https://doi.org/10.3200/MONO.131.2.129-184>
- Bernard, L. C., Mills, M., Swenson, L., & Walsh, R. P. (2008). Measuring Motivation Multidimensionally: Development of the Assessment of Individual Motives–Questionnaire [AIM-Q]. *Assessment, 15*, 16–35. <https://doi.org/10.1177/1073191107306131>
- Bugental, D. B. (2000). Acquisition of the algorithms of social life: A domain-based approach. *Psychological Bulletin, 126*, 187–219. <https://doi.org/10.1037/0033-2909.126.2.187>
- Buss, D. M. (1988). The evolution of human intrasexual competition: Tactics of mate attraction. *Journal of Personality and Social Psychology, 54*, 616–628. <https://doi.org/10.1037/0022-3514.54.4.616>
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge/Taylor & Francis. <https://doi.org/10.4324/9780203807644>
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466. <https://doi.org/10.1037/0033-2909.105.3.456>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. <https://doi.org/10.1037/h0046016>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 464–504. <https://doi.org/10.1080/10705510701301834>
- Cieciuch, J., Davidov, E., Oberski, D. L., & Algesheimer, R. (2015). Testing for measurement invariance by detecting local misspecification and an illustration across online and paper-and-pencil samples. *European Political Science, 14*, 521–538. <https://doi.org/10.1057/eps.2015.64>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology, 40*, 55–75. <https://doi.org/10.1146/annurev-soc-071913-043137>
- Dawkins, R. (1982). *The extended phenotype*. Oxford, UK: W. H. Freeman.
- Deci, E., & Ryan, R. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*, 1024–1037. <https://doi.org/10.1037/0022-3514.53.6.1024>
- Dunlap, K. (1919). Are there any instincts? *Journal of Abnormal Psychology, 14*, 35–50.
- Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology, 7*, 1–52. [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4)
- Harriman, P. L. (1975). *Handbook of psychological terms*. Totowa, NJ: Littlefield, Adams.
- Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 163–179). New York, NY: Guilford.

Bernard, L. C., Cieciuch, J., Lac, A., Žuro, B., Krupić, D., Richter, M., Silvestrini, N., von Helversen, B. (2021). A cross-cultural study of purposive “traits of action”: measurement invariance of scales based on the action–trait theory of human motivation using exploratory structural equation modeling. *Studia Psychologica: Theoria et Praxis, 21*(1), 5–29. <https://doi.org/10.21697/sp.2021.21.1.01>

- Holt, E. B. (1931). *Animal drive and the learning process: An essay toward radical empiricism*. New York, NY: Holt.
- Hoyle, R. H. (Ed.). (2012). *Handbook of structural equation modeling*. New York, NY: Guilford Press.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453. <https://doi.org/10.1037//1082-989x.3.4.424>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. <https://doi.org/10.1080/10705519909540118>
- International Test Commission. (2008). *International Test Commission Guidelines for Translating and Adapting Tests*. London, UK: The British Psychological Society. Retrieved from: www.InTestCom.org
- James, W. (1890). Instinct. In *Principles of Psychology* (Vol. II, pp. 383–441). New York, NY: Holt.
- Kenrick, D. T., Li, N. P., & Butner, J. (2003). Dynamical evolutionary psychology: Individual decision rules and emergent social norms. *Psychological Review, 110*, 3–28. <https://doi.org/10.1037/0033-295x.110.1.3>
- Lance, C. E., & Vandenberg, R. J. (2002). Confirmatory factor analysis. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behaviour in organizations: Advances in measurement and data analysis*. San Francisco, SF: Jossey-Bass.
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research, 39*, 329–358. https://doi.org/10.1207/s15327906mbr3902_8
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201–226. <https://doi.org/10.1146/annurev.psych.51.1.201>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149. <https://doi.org/10.1037/1082-989X.1.2.130>
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*, 471–491. <https://doi.org/10.1037/a0019227>
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110. <https://doi.org/10.1146/annurev-clinpsy-032813-153700>
- McClelland, D. C., Atkinson, J. W., Clark, R. W., & Lowell, E. L. (1976). *The achievement motive* (2nd ed.). New York, NY: Irvington.
- Morin, A. J., Boudrias, J. S., Marsh, H. W., Madore, I., & Desrumaux, P. (2016). Further reflections on disentangling shape and level effects in person-centered analyses: An illustration exploring the dimensionality of psychological health. *Structural Equation Modeling: A Multidisciplinary Journal, 23*, 438–454. <https://doi.org/10.1080/10705511.2015.1116077>
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Authors.
- Oettingen, G., Bulgarella, C., Henderson, M., & Gollwitzer, P. M. (2004). Self-awareness, self-motives, and self-motivation. In R. A. Wright, J. Greenberg, & S. S. Brehm (Eds.), *Motivational analyses of social behavior: Building on Jack Brehm's contributions to psychology* (pp. 225–244). Mahwah, NJ: Erlbaum.
- Prohaska, T. R., Leventhal, E. A., Leventhal, H., & Keller, M. L. (1985). Health practices and illness cognition in young, middle aged, and elderly adults. *Journal of Gerontology, 40*, 569–578. <https://doi.org/10.1093/geronj/40.5.569>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling, 16*, 561–582. <https://doi.org/10.1080/10705510903203433>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Upper Saddle River, NJ: Pearson Education.
- Tolman, E. C. (1932). *Purposive behaviour in animals and men*. London, UK: Century/Random House.
- Tooby, J., & Cosmides, L. (1992). The psychological foundation of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19–136). New York, NY: Oxford University Press.
- Vazsonyi, A. T., Ksinan, A., Mikuška, J., & Jiskrova, G. (2015). The big five and adolescent adjustment: An empirical test across six cultures. *Personality and Individual Differences, 83*, 234–244. <https://doi.org/10.1016/j.paid.2015.03.049>
- Warner, R. M. (2013). *Applied statistics: From bivariate through multivariate techniques* (2nd ed.). Thousand Oaks, CA: Sage.

APPENDIX

LANGUAGE VERSIONS OF THE QUICK AIM 60-ITEM SCALE

KEY

Aggression: 1, 16, 31, 46
 Threat Avoidance: 2, 17, 32, 47
 Illness Avoidance: 3, 18, 33, 48
 Environmental Inquisitiveness: 4, 19, 34, 49
 Interpersonal Inquisitiveness: 5, 20, 35, 50
 Sex: 6, 21, 36, 51
 Appearance: 7, 22, 37, 52
 Wealth: 8, 23, 38, 53
 Mental: 9, 24, 39, 54
 Physical: 10, 25, 40, 55
 Commitment: 11, 26, 41, 56
 Altruism: 12, 27, 42, 57
 Social Exchange: 13, 28, 43, 58
 Legacy: 14, 29, 44, 59
 Meaning: 15, 30, 45, 60

ENGLISH ITEMS

1. I try to intimidate others to keep them from challenging me.
2. I spend a lot of effort to avoid physical threats of any kind.
3. I wash my hands before I eat anything, even a quick snack.
4. I make the effort to explore something I have not seen before.
5. I play with, jest, or mock people to discover their reactions.
6. I spend time thinking about how to have more sex.
7. I spend money and time on stylish hair cuts to improve my appearance.
8. I buy (or would buy) expensive or rare luxury goods so that others will know I am successful.
9. I show off my understanding of abstract or complex ideas so people will respect me.
10. I train hard to have as strong a body as I can.
11. I take time to offer emotional support and affection to the person I am involved with romantically.
12. I sacrifice (or would sacrifice) to help my distant relatives.
13. I take care of all my responsibilities to others.
14. I contribute to charitable organizations that do good work in the world.
15. I put a lot of effort into investigating different philosophies and approaches to life.
16. I use physical threats so that others know not to get in my way.
17. I stay clear of people who are engaged in a conflict situation for fear of harm.
18. I am very careful not to touch any surfaces after I wash my hands in a public place.
19. I am curious about new and different places because they may have something to offer.
20. I tease friends and acquaintances because it helps me understand them better.
21. I spend (or have spent) a lot of time and energy to have sex.
22. I spend money or time to improve my appearance such as getting a tan, spa treatments, or other similar things.
23. I like to have valuable things like cars, watches, or jewelry that other people notice.

24. I engage in discussions or debates to develop a reputation as a smart and influential person.
25. I make the effort to stay in peak physical condition compared to others.
26. I make the effort to do the things that show I care for my spouse or partner.
27. I would let (or have let) distant relatives live with me temporarily if they needed shelter.
28. I do what is required to keep a contract or bargain I have made.
29. I donate my time, effort, or money to causes that promote health, reduce famine, or improve education in other countries.
30. I take time to investigate different religions and philosophies.
31. I use verbal threats so that others know not to get in my way.
32. I leave a situation when people become hostile toward each other.
33. I try to avoid any contact with people when there is an illness going around.
34. I explore unknown places to understand them better.
35. I like to play around with people to see how they react, even if they get angry at times.
36. I invest (or have invested) time seeking a lot of different sexual partners.
37. I buy expensive shoes or other accessories to look attractive.
38. I put a lot of time and energy into having expensive possessions.
39. I put a lot of effort into developing my own special skills and abilities to impress others.
40. I work on my body to be as physically competitive as I can.
41. I give of myself and my time in a romantic relationship.
42. I would take slight personal risks to my safety to help distant relatives.
43. I strive to be honest and fair.
44. I donate my time, effort, or money to non-profit community organizations that can make a better future for others.
45. I learn about various religions or philosophies to understand the purpose of life better.
46. I take physical risks so others will fear me.
47. I leave when a fight breaks out between people in order to stay safe.
48. I avoid traveling when infectious diseases are circulating.
49. I am interested in going to different places in order to learn something that could help me.
50. I try to understand others by playfully teasing them to see how they react.
51. I spend (or have spent) a lot of time pursuing sex.
52. I spend money on grooming (hair cuts, coloring, tanning) to look my best.
53. I spend a lot of money to buy things that make others envious.
54. I put effort into developing my talents and abilities in order to impress others.
55. I train hard to be to earn the reputation of being in peak physical condition.
56. I give what is necessary of my time and resources in romantic relationships.
57. I would invest time and effort to help a distant relative even if there were no immediate advantage to me.
58. I make an effort to treat others the way I want to be treated.
59. I contribute time, effort, and/or money to make the world better for future generations.
60. I study various philosophies of life and its meaning in order to find what is common in all.

CROATIAN ITEMS

1. Pokušavam zastrašiti druge kako bi ih spriječio da me izazivaju.
2. Dosta se trudim izbjeći fizičke opasnosti svih vrsta.
3. Perem ruke prije nego išta pojedem, makar i nešto na brzinu.
4. Trudim se istražiti mjesta koja do sada nisam vidio.
5. Poigravam se s ljudima, zadirkujem ih i podsmijavam im se kako bi vidio njihovu reakciju.
6. Provodim puno vremena razmišljajući kako se više seksati.
7. Trošim novac i vrijeme na stilizirane frizure kako bi poboljšao izgled.

8. Kupujem (ili bih kupovao) skupe i rijetke luksuzne stvari kako bi drugi znali da sam uspješan.
9. Razmećem se svojim razumijevanjem apstraktnih ili složenih ideja kako bi me ljudi poštovali.
10. Naporno treniram kako bih imao što snažnije tijelo.
11. Odvajam puno vremena kako bih pružio emocionalnu podršku i privrženost osobi s kojom se nalazim u romantičnoj vezi.
12. Žrtvujem se (ili bih se žrtvovao) kako bih pomogao svojoj daljoj rodbini.
13. Izvršavam sve svoje dužnosti prema drugima.
14. Doniram dobrotvornim organizacijama koje čine dobro u svijetu.
15. Ulažem puno truda u istraživanje različitih filozofija i pristupa životu.
16. Koristim se fizičkim prijetnjama kako bi drugi znali da mi ne trebaju stajati na putu.
17. Držim se podalje od ljudi uključenih u konfliktne situacije iz straha da ne bi nastradao.
18. Jako pazim da ne diram nikakve površine nakon što operem ruke na javnom mjestu.
19. Znatiželjan sam oko novih i drugačijih mjesta, jer bi mogla imati nešto za ponuditi.
20. Zadirkujem prijatelje i poznanike jer mi to pomaže da ih bolje upoznam.
21. Provodim (ili sam provodio/la) puno vremena i energije kako bih došao do seksa.
22. Trošim novac ili vrijeme kako bi poboljšao svoj izgled sunčanjem ili odlaskom na kvarcanje, spa tretmane i druge slične stvari.
23. Volim posjedovati vrijedne stvari poput automobila, satova ili nakita koje druge osobe mogu primjetiti.
24. Uključujem se u rasprave ili debate kako bih razvio reputaciju pametne i utjecajne osobe.
25. Trudim se da u odnosu na druge ostanem u naponu snage.
26. Trudim se raditi stvari koje mojem supružniku ili partneru pokazuju da mi je stalo do njega/nje.
27. Dopustio bih (ili sam dopustio) daljoj rodbini da privremeno žive kod mene ako im to zatreba.
28. Napraviti ću sve što je potrebno kako bih ispoštovao dogovor ili ugovor koji sam sklopio.
29. Doniram svoje vrijeme, trud ili novac događajima koji promoviraju zdravlje, smanjuju glad ili poboljšavaju obrazovanje u drugim zemljama.
30. Izdvajam vremena u istraživanje različitih religija i filozofija.
31. Koristim se verbalnim prijetnjama kako bi drugima dao/la do znanja da mi ne stoje na putu.
32. Napuštam situaciju u kojoj ljudi postanu agresivni (neprijateljski nastrojeni) jedni prema drugima.
33. Pokušavam izbjeći bilo koji kontakt s ljudima kada su bolesni.
34. Istražujem nepoznata mjesta kako bi ih bolje razumio.
35. Volim se poigravati s ljudima kako bih vidio kako će reagirati, čak iako se zbog toga znaju ponekad naljutiti.
36. Provodim (ili sam provodio) dosta vremena u traženju većeg broja različitih seksualnih partnera.
37. Kupujem skupe cipele i druge odjevne dodatke kako bih izgledao atraktivno.
38. Ulažem mnogo vremena i energije kako bih posjedovao skupocjene predmete.
39. Ulažem puno napora u razvijanje vlastitih posebnih vještina i sposobnosti kako bi impresionirao druge.
40. Radim na svom tijelu kako bi bio fizički spremniji u odnosu na ostale.
41. Dajem sebe i svoje vrijeme u romantičnoj vezi.
42. Malo bi riskirao svoju sigurnost kako bih pomogao svojoj daljoj rodbini.
43. Nastojim biti iskren i pošten.
44. Doniram svoje vrijeme, trud ili novac neprofitabilnim društvenim organizacijama koje mogu stvoriti bolju budućnost za druge.
45. Učim o različitim religijama ili filozofijama kako bih bolje razumio svrhu života.
46. Riskiram s fizičkim obračunima kako bi me se druge osobe bojale.
47. Otići ću kada izbije tuča između ljudi da bi ostao siguran/na.
48. Izbjegavam putovanja u vrijeme kada kruže zarazne bolesti.
49. Zainteresiran sam za odlazak na različita mjesta kako bi naučio nešto novo što bi mi moglo koristiti.
50. Pokušavam razumjeti druge na način da ih zadirkujem kako bi vidio kako će reagirati.

51. Provodim puno vremena kako bi došao do seksa.
52. Trošim novac za dotjerivanje (frizura, bojanje kose, kvarcanje) kako bi izgledao najbolje što mogu.
53. Trošim mnogo novaca na kupnju stvari na kojima mi drugi zavide.
54. Ulažem trud u razvijanje svojih talenata i sposobnosti kako bih impresionirao druge.
55. Naporno treniram kako bih ostvario reputaciju nekoga tko je u stalnoj fizičkoj kondiciji.
56. Dajem, koliko god da je potrebno, svoje vrijeme i resurse u romantičnoj vezi.
57. Uložio bih vrijeme i trud pomagajući daljoj rodbini, čak iako od toga ne bih imao izravne koristi.
58. Trudim se ponašati prema drugima onako kako bih želio da se oni ponašaju prema meni.
59. Doprinosim svojim vremenom, trudom i/ili novcem kako bih učinio svijet boljim za nadolazeće generacije.
60. Proučavam različite filozofije života i njihovo značenje u svrhu traženja onoga što im je zajedničko.

FRENCH ITEMS

1. J'essaie d'intimider les autres pour les empêcher de me défier.
2. Je fais beaucoup d'effort pour éviter les menaces physiques en tous genres.
3. Je lave mes mains avant de manger quoi que ce soit, même pour un snack rapide.
4. Je fais l'effort d'explorer quelque chose que je n'ai pas vu avant.
5. Je joue avec, plaisante, me moque des gens pour découvrir leurs réactions.
6. Je passe du temps à penser à comment avoir plus de rapports sexuels.
7. Je dépense de l'argent et du temps dans des coupes de cheveux élégantes pour améliorer mon apparence.
8. J'achète (ou j'achèterais) des articles de luxe chers ou rares comme ça les autres sauront que je réussis.
9. J'affiche fièrement ma compréhension d'idées complexes ou abstraites comme ça les gens me respecteront.
10. Je m'entraîne dur pour avoir un corps aussi fort que je peux.
11. Je prends du temps pour offrir du support émotionnel et de l'affection à la personne avec qui je suis engagé amoureusement.
12. Je me sacrifie (ou me sacrifierais) pour aider mes parents éloignés.
13. Je prends soin de toutes mes responsabilités envers les autres.
14. Je contribue à des organisations caritatives qui font du bon travail dans le monde.
15. Je mets beaucoup d'effort pour étudier différentes philosophies et approches de la vie.
16. J'utilise des menaces physiques de manière à ce que les autres sachent qu'ils ne doivent pas se mettre sur mon chemin.
17. J'évite les gens qui sont engagés dans une situation de conflit par peur d'avoir mal.
18. Je fais très attention de ne pas toucher de surfaces après que j'ai lavé mes mains dans un espace public.
19. Je suis curieux à propos d'endroits nouveaux et différents parce qu'ils peuvent avoir quelque chose à offrir.
20. Je taquine mes amis et connaissances parce que cela m'aide à mieux les comprendre.
21. Je dépense (ou ai dépensé) beaucoup de temps et d'énergie pour avoir des rapports sexuels.
22. Je dépense de l'argent et du temps pour améliorer mon apparence comme en allant bronzer, en recevant des soins esthétiques ou d'autres choses similaires.
23. J'aime avoir des choses qui ont de la valeur comme des voitures, des montres ou des bijoux que les autres gens remarquent.
24. Je m'engage dans des discussions ou débats pour développer une réputation de personne intelligente et influente.
25. Je fais l'effort de rester au meilleur de ma condition physique comparé aux autres.

26. Je fais l'effort de faire des choses qui montrent que je prends soin de mon épouse ou de ma partenaire.
27. Je laisserais (ou ai laissé) des proches vivre avec moi temporairement s'ils avaient besoin d'un abri.
28. Je fais ce qui est exigé pour honorer un contrat ou marché que j'ai fait.
29. Je donne mon temps, des efforts, ou de l'argent à des causes qui promeuvent la santé, réduisent la famine ou améliorent l'éducation dans d'autres pays.
30. Je prends du temps pour étudier différentes religions et philosophies.
31. J'utilise des menaces verbales de manière à ce que les autres sachent qu'ils ne doivent pas se mettre sur mon chemin.
32. Je quitte une situation lorsque les gens deviennent hostiles les uns envers les autres.
33. J'essaie d'éviter tout contact avec des gens lorsqu'il y a une maladie qui tourne.
34. J'explore des endroits inconnus pour mieux les comprendre.
35. J'aime m'amuser avec les gens pour voir comment ils réagissent, même s'ils se mettent parfois en colère.
36. J'investis (ou ai investi) du temps à chercher beaucoup de partenaires sexuels différents.
37. J'achète des chaussures chères ou d'autres accessoires pour avoir l'air attirant(e).
38. Je mets beaucoup de temps et d'énergie pour avoir des biens chers.
39. Je fais beaucoup d'effort pour développer mes propres compétences et capacités spéciales pour impressionner les autres.
40. Je travaille sur mon corps pour être le plus physiquement compétitif que je peux.
41. Je donne de moi-même et de mon temps dans une relation amoureuse.
42. Je prendrais de légers risques personnels pour ma sécurité pour aider des parents éloignés.
43. Je m'efforce d'être honnête et juste.
44. Je donne mon temps, des efforts, ou de l'argent à des organisations communautaires sans but lucratif qui peuvent contribuer à un meilleur futur pour les autres.
45. J'apprends sur diverses religions ou philosophies pour mieux comprendre le but de notre vie.
46. Je prends des risques physiques comme ça les autres me craignent.
47. Je pars lorsqu'une bagarre éclate entre des gens pour rester en sécurité.
48. J'évite de voyager lorsque des maladies infectieuses circulent.
49. Je suis intéressé d'aller dans des endroits différents pour apprendre quelque chose qui pourrait m'aider.
50. J'essaie de comprendre les autres en les taquinant malicieusement pour voir comment ils réagissent.
51. Je passe (ou ai passé) beaucoup de temps à chercher à avoir des rapports sexuels.
52. Je dépense de l'argent en tenues (coupes de cheveux, couleurs, bronzage) pour apparaître sous mon meilleur jour.
53. Je dépense beaucoup d'argent pour acheter des choses qui rendent les autres envieux.
54. Je fais des efforts pour développer mes talents et capacités pour impressionner les autres.
55. Je m'entraîne dur pour gagner la réputation d'être au meilleur de ma condition physique.
56. Je donne ce qui est nécessaire de mon temps et de mes ressources dans des relations amoureuses.
57. J'investirais du temps et de l'effort pour aider un parent éloigné même s'il n'y avait pas d'avantage immédiat pour moi.
58. Je fais un effort pour traiter les autres de la manière avec laquelle je veux être traité.
59. Je contribue avec mon temps, mes efforts et/ou mon argent pour rendre le monde meilleur pour les générations futures.
60. J'étudie diverses philosophies de vie et leur sens pour trouver ce qui est commun à toutes.

GERMAN ITEMS

1. Ich versuche Andere einzuschüchtern, damit sie mich nicht herausfordern.
2. Ich verwende viel Anstrengung darauf, körperliche Bedrohungen jeglicher Art zu vermeiden.

3. Ich wasche meine Hände bevor ich etwas esse, auch wenn es nur ein schneller Snack ist.
4. Ich bemühe mich etwas zu entdecken, was ich noch nie zuvor gesehen habe.
5. Ich spiele mit anderen Menschen, mache Scherze über sie oder veralbere sie, um ihre Reaktionen zu beobachten.
6. Ich denke häufig darüber nach, wie ich mehr Sex haben kann.
7. Ich investiere Zeit und Geld in stylische Frisuren, um mein Aussehen zu verbessern.
8. Ich kaufe teure und seltene Luxusgüter, damit andere glauben, dass ich erfolgreich bin / Ich würde teure und seltene Luxusgüter kaufen, damit andere glauben, dass ich erfolgreich bin.
9. Ich zeige, dass ich abstrakte oder komplexe Ideen verstehe, damit mich andere Menschen respektieren.
10. Ich trainiere hart, um einen möglichst starken Körper zu bekommen.
11. Ich nehme mir die Zeit, der Person mit der ich eine romantische Beziehung habe, emotionale Unterstützung und Zuneigung zu geben.
12. Ich gehe Opfer ein, um entfernten Verwandten zu helfen / Ich würde Opfer eingehen, um entfernten Verwandten zu helfen.
13. Ich nehme meine Verantwortungen gegenüber Anderen wahr.
14. Ich trage zu mildtätigen Organisationen bei, die in der Welt Gutes tun.
15. Ich verwende viel Anstrengung darauf, verschiedene Lebensphilosophien und Lebenseinstellungen kennenzulernen.
16. Ich setze körperliche Drohungen ein, damit Andere mir nicht in die Quere kommen.
17. Ich meide Personen, die sich in konfliktreichen Situationen befinden, da ich Angst habe verletzt zu werden.
18. Ich bemühe mich keine Oberflächen anzufassen, nachdem ich meine Hände in einem öffentlich zugänglichen Raum gewaschen habe.
19. Ich bin neugierig auf neue und unterschiedliche Orte, da sie möglicherweise etwas zubieten haben.
20. Ich necke Freunde und Bekannte, da es mir hilft, sie besser zu verstehen.
21. Ich verwende viel Zeit und Energie darauf Sex zu haben / Ich habe viel Zeit und Energie darauf verwandt Sex zu haben.
22. Ich verwende Geld und Zeit, um mein Aussehen zu verbessern (beispielsweise für Solariumbesuche, Schönheitsbehandlungen oder ähnlichen).
23. Ich besitze gerne wertvolle Dinge wie Autos, Uhren oder Schmuck, die andere Leute wahrnehmen.
24. Ich engagiere mich in Diskussionen oder Debatten, um einen Ruf als smarte und einflussreiche Person zu entwickeln.
25. Ich bemühe mich im Vergleich zu Anderen in bester körperlicher Verfassung zu bleiben.
26. Ich strengte mich an Dinge zu tun, die zeigen, dass ich mich um meinen Ehepartner oder Partner kümmere.
27. Ich habe entfernte Verwandte bei mir vorübergehend wohnen lassen, wenn sie eine Unterkunft brauchten / Ich würde entfernte Verwandte bei mir vorübergehend wohnen lassen, wenn sie eine Unterkunft brauchen.
28. Ich tue was nötig ist, um einen Vertrag oder ein Geschäft, das ich abgeschlossen habe, einzuhalten.
29. Ich verwende meine Zeit, Anstrengung oder Geld für Zwecke, die in anderen Ländern Gesundheit fördern, Hunger reduzieren oder die Bildung verbessern.
30. Ich nehme mir die Zeit verschiedene Religionen und Philosophien kennenzulernen.
31. Ich setze verbale Drohungen ein, damit Andere mir nicht in die Quere kommen.
32. Ich verlasse Situationen, in denen Menschen feindselig zu einander werden.
33. Ich vermeide jeglichen Kontakt mit Menschen, wenn eine Krankheit umgeht.
34. Ich erkunde unbekannte Orte genau, um sie besser zu verstehen.
35. Ich spiele gerne mit anderen Menschen, um zu sehen wie sie reagieren, auch wenn sie hin und wieder ärgerlich werden.

36. Ich verwende/verwendete viel Zeit darauf viele unterschiedliche Sexualpartner zu haben.
37. Ich kaufe teure Schuhe oder andere Accessoires um attraktiv auszusehen.
38. Ich investiere eine Menge Zeit und Energie darauf teure Güter zu besitzen.
39. Ich strenge mich an, meine eigenen speziellen Fähigkeiten und Fertigkeiten zu entwickeln, um damit Andere zu beeindrucken.
40. Ich arbeite an meinem Körper, um physisch so konkurrenzfähig wie möglich zu sein.
41. Ich bringe mich selbst und meine Zeit in eine romantische Beziehung ein.
42. Ich würde ein kleines persönliches Sicherheitsrisiko eingehen, um entfernten Verwandten zu helfen.
43. Ich bemühe mich ehrlich und fair zu sein.
44. Ich spende gemeinnützigen Organisationen, die eine bessere Zukunft für Andere schaffen können, meine Zeit, Anstrengung oder Geld.
45. Ich informiere mich über verschiedene Religionen oder Philosophien, um den Sinn des Lebens besser zu verstehen.
46. Ich gehe körperliche Risiken ein, damit Andere mich fürchten.
47. Ich verlasse Situationen, wenn es zwischen Personen zu Kämpfen kommt, damit mir nichts passiert.
48. Ich vermeide Reisen, wenn ansteckende Krankheiten zirkulieren.
49. Ich bin daran interessiert unterschiedliche Orte zu besuchen, um etwas zu lernen was mir helfen könnte.
50. Ich versuche andere zu verstehen, indem ich sie spasseshalber aufziehe, um zu sehen, wie sie reagieren.
51. Ich verwende/verwendete viel Zeit darauf, Sex zu haben.
52. Ich verwende Geld darauf mich zurechtzumachen (Haarschnitt, Haare färben, braun zu werden) um bestmöglich auszusehen.
53. Ich gebe viel Geld aus, um Dinge zu kaufen, die andere eifersüchtig machen.
54. Ich strenge mich an, meine Talente und Fähigkeiten zu entwickeln, um Andere zu beeindrucken.
55. Ich trainiere hart, um den Ruf zu erlangen, in einem physischen Topzustand zu sein.
56. Ich investiere in romantischen Beziehungen die Zeit und Ressourcen, die nötig sind.
57. Ich würde Zeit und Anstrengung investieren, um einem entfernten Verwandten zu helfen, auch wenn mir dabei kein unmittelbarer Vorteil entstehen würde.
58. Ich bemühe mich andere so zu behandeln, wie ich selbst gerne behandelt werden möchte.
59. Ich verwende Zeit, Anstrengung und/oder Geld darauf, die Welt für zukünftige Generationen zu verbessern.
60. Ich studiere verschiedene Lebensphilosophien und ihre Bedeutung um herauszufinden, was ihnen allen gemeinsam ist.

POLISH ITEMS

1. Staram się zastraszyć innych, aby nie rywalizowali ze mną.
2. Wkładam wiele wysiłku, aby uniknąć jakichkolwiek fizycznych zagrożeń.
3. Przed każdym jedzeniem myję ręce, nawet jeśli jest to szybka przekąska.
4. Staram się dokładnie badać rzeczy, które widzę po raz pierwszy.
5. Gram z ludźmi, kpię z nich lub drwię, aby zobaczyć, jaka będzie ich reakcja.
6. Dużo się zastanawiam, co zrobić, aby częściej uprawiać seks.
7. Poświęcam pieniądze i czas na dobrego fryzjera, żeby poprawić swój wygląd.
8. Kupuję (lub kupowałam(a)bym) drogie lub rzadkie przedmioty luksusowe, żeby inni widzieli, że jestem człowiekiem sukcesu.
9. Pokazuję innym, że rozumiem jakieś abstrakcyjne lub złożone problemy, aby mnie szanowali.
10. Ciężko trenuję, aby mieć jak najsilniejsze ciało.

11. Poświęcam czas, aby dać emocjonalne wsparcie i czułość osobie, którą kocham.
12. Poświęcam się (lub poświęcił(a)bym) się, aby pomóc swoim dalszym krewnym.
13. Wypełniam wszystkie swoje zobowiązania wobec innych.
14. Daję datki na organizacje charytatywne, które czynią coś dobrego dla świata.
15. Wkładam wiele wysiłku w badanie różnych filozofii i podejść do życia.
16. Stosuję groźby fizyczne, aby inni nie wchodzili mi w drogę.
17. Z obawy przed niebezpieczeństwem trzymam się z dala od ludzi zaangażowanych w jakiegokolwiek konflikty.
18. Bardzo uważam, aby po umyciu rąk nie dotykać żadnych powierzchni w miejscu publicznym.
19. Ciekawią mnie nowe, inne miejsca, bo mogą oferować coś szczególnego.
20. Dokuczam przyjaciółom i znajomym, ponieważ pozwala mi to lepiej ich rozumieć.
21. Poświęcam wiele czasu i energii na uprawianie seksu.
22. Poświęcam czas i pieniądze, aby poprawić swój wygląd, np. chodzę do solarium, spa itd.
23. Lubię mieć cenne rzeczy, na które inni zwracają uwagę, takie jak: samochody, zegarki lub biżuterię.
24. Angażuję się w dyskusje lub debaty, aby wyrobić sobie reputację osoby mądrej i wpływowej.
25. Staram się, aby w porównaniu z innymi mieć bardzo dobrą kondycję fizyczną.
26. Staram się robić rzeczy świadczące o tym, że troszczę się o swego współmałżonka lub partnera/partnerkę.
27. Pozwalam (lub pozwolił(a)bym) swoim dalszym krewnym mieszkać ze mną przez jakiś określony czas, jeśli potrzebowaliby dachu nad głową.
28. Robię wszystko, co jest wymagane w umowie, na którą się godzę.
29. Przeznaczam swój czas, wysiłek lub pieniądze na organizacje, które promują zdrowie, zmniejszają klęskę głodu lub wspierają edukację w innych krajach.
30. Poświęcam swój czas na zgłębianie różnych religii i filozofii.
31. Stosuję groźby werbalne, aby inni nie wchodzili mi w drogę.
32. Kiedy w jakiejś sytuacji ludzie zaczynają być wobec siebie wrodozy, odchodzę stamtąd.
33. Kiedy panują jakieś choroby, staram się unikać kontaktu z innymi ludźmi.
34. Badam miejsca, których nie znam, bo chcę je zrozumieć.
35. Lubię igrzać z ludźmi, aby zobaczyć ich reakcje, nawet jeśli się wtedy czasem złościę.
36. Poświęcam czas na poszukiwanie wielu różnych partnerów seksualnych.
37. Kupuję drogie buty lub inne dodatki, aby wyglądać atrakcyjnie.
38. Inwestuję wiele czasu i energii, żeby posiadać kosztowne rzeczy.
39. Wkładam wiele wysiłku w rozwijanie swoich zdolności i umiejętności, aby zrobić wrażenie na innych.
40. Pracuję nad swoim ciałem, aby było fizycznie bezkonkurencyjne.
41. Daję siebie i swój czas osobie, którą kocham.
42. Zaryzykował(a)bym nieco własne bezpieczeństwo, aby pomóc dalekiemu krewnemu.
43. Staram się być uczciwy/a i sprawiedliwy/a.
44. Przeznaczam swój czas, wysiłek lub pieniądze dla społecznych organizacji non-profit, które dbają o lepszą przyszłość dla innych.
45. Zdobywam informacje o różnych religiach lub filozofiach, żeby lepiej rozumieć cel życia.
46. Podejmuję fizyczne ryzyko, aby inni bali się mnie.
47. Gdy między ludźmi rozpoczyna się bójka, oddalam się, aby pozostać bezpiecznym.
48. Unikam podróży, gdy jest dużo zachorowań na choroby zakaźne.
49. Z zainteresowaniem chodzę w różne miejsca, żeby nauczyć się czegoś, co może mi się przydać.
50. Staram się zrozumieć innych, drocząc się z nimi i obserwując ich reakcję.
51. Poświęcam wiele czasu na poszukiwanie seksu.
52. Wydaję pieniądze na różne zabiegi, żeby wyglądać jak najlepiej (nowe fryzury, farbowanie włosów, opalanie).
53. Wydaję dużo pieniędzy na rzeczy, których inni będą mi zazdrościć.
54. Staram się rozwijać moje talenty i zdolności, aby zrobić wrażenie na innych.

55. Ciężko ćwiczę, aby mieć opinię osoby w doskonałej formie fizycznej.
56. W moim związku z kochaną osobą poświęcam konieczny czas i środki.
57. Poświęcił(a)bym czas i wysiłek, aby pomóc dalekiemu krewnemu, nawet jeśli nie miał(a)bym z tego bezpośrednich korzyści.
58. Staram się traktować ludzi tak, jak sam/a chcę być traktowany/a.
59. Poświęcam swój czas, wysiłek i/lub pieniądze, aby sprawić, że świat będzie lepszy dla przyszłych pokoleń.
60. Zdobywam wiedzę na temat różnych filozofii życia i jego sensu, aby znaleźć to, co je łączy.

KURT LEWIN (1890–1947) – W OCZACH JEGO BIOGRAFÓW

WŁODZISŁAW ZEIDLER*

Hohenwestedt

STRESZCZENIE

Poglądy i twórczość Kurta Lewina wywierały i nadal wywierają fascynujący wpływ na wielu autorów, którzy piszą o jego drodze życiowej oraz twórczości, opierając się na własnej wiedzy oraz na stanie badań nad ową twórczością. Kurt Lewin ma trwałe miejsce w historii psychologii, ale pomimo to, w zależności od punktu widzenia autora, obrazy jego dorobku są różne. Idąc śladem Graumanna (1992/2007), w artykule przeanalizowano jego pogląd, że psychologia XX wieku nie była przygotowana na przyjęcie i zrozumienie dzieł Kurta Lewina. Stanowisko Graumanna uzasadnia to, że poszczególni biografowie ograniczali swoje zainteresowania pracami Lewina albo do „okresu berlińskiego”, albo tego w USA. Tym samym zacierali złożoność i wyrazistość obrazu jego przestrzeni życiowej, która rozpoczynała się jednak już w Mogilnie i w Poznaniu.

BIOGRAFISTYKA
PRZESTRZEŃ ŻYCIOWA
ŹRÓDŁA INFORMACJI

SŁOWA KLUCZOWE

- | | |
|----|--|
| 33 | KURT LEWIN W HISTORII PSYCHOLOGII |
| 35 | OPOWIEŚCI BIOGRAFICZNE I PUNKT WIDZENIA |
| 36 | POCHODZENIE INFORMACJI BIOGRAFICZNYCH |
| 37 | Tożsamość Uczzonego |
| 38 | RECEPCJA PRAC KURTA LEWINA – TAKŻE W POLSCE |
| 38 | Właściwy punkt widzenia? |
| 40 | MOŻLIWE PUNKTY WIDZENIA |
| 41 | LEWIN I FILOZOFIA |
| 45 | OSOBLIWOŚĆ PRZESTRZENI ŻYCIOWEJ KURTA LEWINA |
| 46 | BIBLIOGRAFIA |



KURT LEWIN (1890–1947) – IN THE EYES OF HIS BIOGRAPHERS

ABSTRACT

Kurt Lewin's views and work have had, and continue to have, a fascinating influence on many authors who write about his life's journey and work, based on their own knowledge and the state of research into that work. Kurt Lewin has a permanent place in the history of psychology, but despite this, depending on the author's point of view, images of his work vary. Following in the footsteps of Graumann (1992/2007), this paper examines his view that 20th century psychology was not prepared to receive and understand the works of Kurt Lewin. Graumann's position is justified by the fact that individual biographers limited their focus on Lewin's work to either the „Berlin period” or that in the United States. By doing so, they blurred the complexity and clarity of the picture of his life space, which, however, already began in Mogilno and Poznan.

KEYWORDS

biography, living space, sources of information

Ohne schöpferische, selbständig denkende und urteilende Persönlichkeiten ist eine Höherentwicklung der Gesellschaft ebenso wenig denkbar wie die Entwicklung der einzelnen Persönlichkeit ohne den Nährboden der Gemeinschaft¹.
(Einstein, 1934/1979, s. 12)

KURT LEWIN W HISTORII PSYCHOLOGII

Dzieje każdej nauki powstają jako wynik interakcji poglądów oraz życiorysu uczonego z tym, co na dany temat już przed nim zostało powiedziane. Niektórzy spośród psychologów, których dzieła zaliczane są do historii psychologii, napisali własne biografie Lewina (Wundt, 1920; Heider, 1984; Obuchowski, 2009). Inni nie mieli na to czasu? Ponieważ dla zrozumienia wielkości dzieła uczonego, który zajął miejsce w historii nauki, potrzebna jest znajomość jego poglądów, życiorysu oraz klimatu czasu, w którym tworzył, napisania biografii wielkich twórców podejmowali się także inni. Wtedy, gdy widoczne było *zainteresowanie* właśnie tą biografią, która czekała na napisanie. Należy dodać, że owo zainteresowanie miało także rozmaity wpływ na prezentację danej biografii. Niekiedy prowadziło do ograniczenia się do jej fragmentów, innym razem – do jej deformacji. Bo biografię pisze się zawsze nie tylko *o kims*, lecz także *dla kogoś*. Dokładniej mówiąc, w porządku czasowym właśnie zainteresowanie daną osobą lub jej twórczością wyprzedza i uzasadnia pisanie biografii.

Nazwisko Kurta Lewina, poczynając od 2. poł. XX wieku, występuje niemal we wszystkich podręcznikach historii psychologii. W tych wcześniejszych zaledwie wzmiankowane (Brett, 1969) albo odnotowane na marginesie innych osiągnięć (Lück, 1997), w innych – na tle rozwoju ówczesnej psychologii (Schönpflug, 2000; Stachowski, 2000), ale także zamknięte w systemie interpretacyjnym „amerykańskiej psychologii” (Schultz, Schultz, 2008), wreszcie jako wielostronna i wyczerpująca nota informacyjna (Lück, 2008; Lück, Miller, Sewz-Vosshenrich, 2008). Dodam, że część spośród tych autorów (Lück, Schönpflug, Stachowski) opublikowała także osobne monografie na temat twórczości albo biografii Kurta Lewina. Nie dziwi zatem, że podawane tam przez tych samych autorów informacje dotyczące jego życiorysu i twórczości są obszerniejsze i bardziej interesujące.

W roku 1997 ukazała się obszerna publikacja przygotowana przez zespół wielu autorów z Europy i USA, pt. *A pictorial history of psychology*. Na łamach tego tomu wzmianki na temat Lewina występują 10 razy, ale poświęcono mu tylko jeden artykuł – sześć stron – pt. *Kurt Lewin – Filmmaker!* W tym kontekście widzimy problem recepcji badań i poglądów Kurta Lewina na tle ówczesnej psychologii akademickiej. A był to już rok 1997! Czy było zatem tak, jak to sobie wyobrażał Graumann (1992/2007), że dopiero wiek XXI będzie wiekiem, w którym język i dzieła Lewina staną się zrozumiałe? Ten renesans zapowiadały w Europie prace Lück (1996, 2001). Niewątpliwie właśnie on przyczynił się do zainteresowania twórczością Lewina, zarówno tą wczesną, w Berlinie, jak i późniejszą, w USA. Ze zrozumiałych

¹ Wyższy rozwój społeczeństwa bez osobowości, które potrafią w sposób twórczy oraz samodzielny myśleć i sądzić, jest równie trudny do pomyślenia, jak rozwój pojedynczej osobowości bez wspólnoty jako odżywczej gleby – tłum. W. Z.

względów zainteresowanie twórczością Lewina w Europie rozpoczęło się późno. Najpierw w Niemczech, w Republice Federalnej, później, po załamaniu się ideologii komunistycznej i zmianie sytuacji politycznej – także w Polsce. Poznawczą wartość miały materiały zebrane podczas konferencji w Bydgoszczy (2004) oraz w Warszawie (2009 i 2011), o których powiem niebawem więcej.

W roku 2015 Fahrenberg (ur. 1937, emerytowany profesor psychologii Uniwersytetu w Fryburgu) w pracy nt. teorii psychologicznych omówił teorie Lewina jako jednego spośród szczególnie utalentowanych psychologów XX wieku – tych, którzy chociaż nie stworzyli odrębnej szkoły², to jednak wnieśli do psychologii idee oryginalne i nowatorskie. Fahrenberg zna prace Lewina z okresu pracy w USA, jednak interesują go głównie te, które powstały w Berlinie³. Chociaż inni dostrzegali ich powinowactwa z filozofią, Fahrenberg ogranicza swoje spojrzenie do aspektów, które mają ewidentne znaczenie dla psychologii oraz jej metodologii. Dla niego Lewin był twórcą teorii psychologicznych. A teorie psychologiczne w jego ujęciu były już niezależne od filozofii. Tak myślał Fahrenberg, znakomity biograf Wilhelma Wundta!

Kilka lat przed publikacją Fahrenberga Trempała, Pepitone i Raven (2006) zebrali w jednym tomie materiały z konferencji „Kurt Lewin: contribution to contemporary psychology” (konferencja, zorganizowana przez Instytut Psychologii bydgoskiego UKW, odbyła się w dniach 10–12 września 2004 r. w Bydgoszczy i Mogilnie). Z wyjątkiem wystąpień prof. Stachowskiego z UAM w Poznaniu i prof. Schönpluga z Berlina były to referaty uczniów i współpracowników Kurta Lewina z USA. Wystąpienie na konferencji bratanka Kurta Lewina, Thomasa, bardzo ją wzbogaciło. Thomas Lewin scharakteryzował interesująco oraz plastycznie życie rodziny Lewinów, poczynając od poziomu generacji, która żyła w Mogilnie. Znane mu były także pewne szczegóły dotyczące środowisk akademickich, z którymi jego stryj miał kontakty w Berlinie i w USA (T. Lewin, 2006). W tomie tym znalazło się także bibliograficzne zestawienie 57 prac poświęconych Lewinowi opublikowanych w latach 1949–2005 w USA. Ów tom i wspomniana konferencja są świadectwem zainteresowania także polskich psychologów dziejami tego, co znalazło wyraz w tytule publikacji: *Lewinian psychology...* A całość cieszy, bo do Mogilna z Bydgoszczy było o wiele bliżej niż np. z... Berlina.

Zebrane we wskazanym tomie referaty stanowią dobre źródło informacji o działalności Lewina w USA, pokazują również najbardziej wartościowe elementy jego teorii sformułowanych i opublikowanych tam podczas jego pobytu, przemilczano jednak proces skomplikowanego i wielostronnego rozwoju Kurta Lewina w Europie! Proces ustosunkowania się do odmienności politycznych i społecznych, wybór dalszej drogi rozwoju naukowego, ale też opozycję wobec polityki nazistowskiej. Lewin w USA trafnie rozpoznawał różnice wariantów ideologicznych i politycznych, ponieważ już wcześniej zrezygnował z punktu widzenia *artylerzysty polowego*. Teraz oglądał otaczające go wydarzenia *okiem własnym*, już krytycznym, a otrzymane wnioski prezentował po transformacji, niemal aforystycznej, jako wyniki psychologicznych badań naukowych! Czy rzeczywiście niezależnych od wszelkich hierarchii wartości? Jego celem było bowiem znalezienie takiej teorii, która byłaby zarówno *dobra*, jak i *praktyczna*. Czy była to alternatywa, która miała zastąpić rozpamiętywanie tragedii oraz skutków nazistowskiego terroru? Nie będę się upierał, rzeczą istotną było to, że pomimo zmiany realiów Kurt Lewin nadal był w stanie uprawiać naukę w sposób twórczy i krytyczny!

² Nazwiska te przypomnę, gdyż porównanie z nimi ukazuje znaczenie Lewina: Georg E. Müller (1850–1934), Theodor Ziehen (1862–1950), Hugo Münsterberg (1863–1916), Oswald Külpe (1862–1915), Carl Stumpf (1848–1936), Hermann Ebbinghaus (1850–1909), Theodor Elsenhans (1862–1918), August Messer (1867–1937), Wolfgang Köhler (1887–1967), Kurt Lewin (1890–1947), William Stern (1871–1938).

³ Należy zauważyć, że kilka prac napisanych w Berlinie nie mogło być tam opublikowanych ze względów politycznych, dlatego ukazały się dopiero w USA, w języku angielskim.

Autorzy prac zebranych w tomie *Lewinian psychology...* pokazali, jak powstawały teorie Lewina. Pokazywali sytuacje, w których intuicja pomagała w programowaniu pionierskich badań psychologicznych. Sprawą otwartą było to, w jakim kontekście należało pojmować prezentowane wyniki. Czy rezultaty badań z lat 30. XX wieku były prezentowane w oryginalnym ujęciu Lewina czy już po adaptacji do języka psychologii USA w XX wieku? Bez wątplenia *Lewinian psychology...* ułatwiła dostęp do lewinowskich koncepcji także w Polsce, a jej wartością szczególną było to, że organizatorzy bydgoskiej konferencji mieli odwagę część obrad zlokalizować w Mogilnie. Bo przecież właśnie tutaj zaczęła się *droga życiowa* twórcy psychologii społecznej!

OPOWIEŚCI BIOGRAFICZNE I PUNKT WIDZENIA

Karl Popper (1928/2006) napisał, że ten sam obiekt można oglądać z różnych punktów widzenia. Była to prawda o szerszym zakresie obowiązywania niż te już dobrze znane prawa optyki. Myślał on bowiem o *widzeniu*, o którym mówimy zarówno w epistemologii, jak i w psychologii. On tę różnorodność akceptował i akcentował. Podobnie w jednym z wczesnych artykułów Lewin (1917) powiada, że jego opisy (sytuacji na froncie wojennym) są dokonywane z punktu widzenia fenomenologii, po czym za chwilę wyjaśnia: „Da den Infanteristen in manchen Fällen andere Landschaftsgebilde begegnen mögen, sei erwähnt, daß ich Feldartillerist bin” [Te krajobrazy, które spotyka piechota, mogą być inne, dlatego muszę dodać, że jestem artylerzystą polowym] (Lewin, 1917, t. IV, s. 315). To prawda. Kiedy Kurt Lewin pisał ten artykuł, oglądał ów krajobraz wojenny z punktu widzenia artylerzysty polowego. Ale to jakby wyznaczenie miało już wyraźny charakter metodologiczny. Było ono zgodne także z regułami, które wcześniej akcentował Immanuel Kant: „Das Erkenntnis besteht aus zwei Stücken der Anschauung und dem Denken” (2017, s. 295). Zatem na spostrzeżenie składają się dane zmysłowe oraz wiedza potrzebna do ich interpretacji lub zrozumienia. Zatem artylerzysta polowy, oglądając ten sam krajobraz wojenny, *postrzegał* go inaczej niż piechur. A tak być musiało, ponieważ mieli inne punkty widzenia oraz inne zadania militarne. W tym stwierdzeniu można już rozpoznać metodologiczną deklarację dotyczącą znaczenia sytuacji.

Biografię Kurta Lewina oraz całą jego twórczość można *ogłądać* z różnych punktów widzenia, na tle konkretnych, coraz to innych sytuacji! A te różne punkty widzenia były wyznaczane także przez kolejne miejsca jego zamieszkania: Mogilno, Poznań, Berlin, USA. Spośród biografów Lewina Tomasz Kardaś (2021) bardzo trafnie potrafił pokazać specyfikę tych miejsc, interpretując je jako odmienności *przestrzeni życiowej* badacza. Dokładniej: związek tych różnych przestrzeni życiowych z dziejami rozwoju uczonego, który celowo, świadomie i odważnie wchodził z nimi w interakcje. Tylko po to, aby je lepiej poznać?

Skoro jest mowa o biografach, należy zauważyć, że są one *opowiadaniem*! A ono jest zawsze dla kogoś i w określonym celu. To jest oczywiste. Warto jeszcze powiedzieć, czym jest *ogłądanie*. Mam na myśli źródło informacji, z którego korzysta *biograf*. Mogą to być informacje bezpośrednie, np. czerpane z obserwacji lub rozmów z uczonym. Częściej biograf korzysta ze źródeł pośrednich: z analizy spuścizny po uczonym, z analizy akt urzędowych, wreszcie z informacji przekazanych przez osoby trzecie. Biografie są to opracowania wykonane z określonego punktu widzenia, jak również oparte na różnym materiale. Te dwie zasady podziału należy traktować jako niezależne, ale równie ważne!

POCHODZENIE INFORMACJI BIOGRAFICZNYCH

Kurt Lewin miał wielu współpracowników oraz uczniów, którzy zachowali wspomnienia o nim, wielorakie – zarówno te przyjacielskie, jak i raczej dramatyczne. Najbardziej interesujące zostały wykorzystane przez Marrowa w jego znanej biografii: *The practical theorist. The life and work of Kurt Lewin* (1969; *Kurt Lewin. Leben und Werk* [1977/2002]). Miał on jako cel pokazać okoliczności oraz sam proces kształtowania się indywidualności Kurta Lewina i powstawania jego teorii. Marrow był współpracownikiem Lewina w okresie, gdy ten pracował w USA. Biografia jego autorstwa ma ambicję być biografią całego życia, zatem okres, który przebiegał w Europie, trzeba było uzupełnić, korzystając ze źródeł pośrednich. Jednym z ważniejszych były informacje uzyskane od Fritza Heidera (1896–1988), który od roku 1930 mieszkał w USA⁴, a obie rodziny (Lewinowie i Heiderowie) utrzymywały kontakty towarzyskie, przyjacielskie i naukowe⁵. Lück (2021) podaje, że przyjaźń Lewina z Heiderem przetrwała 25 trudnych lat, od spotkania w Berlinie w roku 1922 aż do śmierci Lewina w USA w roku 1947. Przy wyraźnej różnicy temperamentów znajdowali ciągle na nowo punkt, w którym spotykały się ich zainteresowania, wyraźnie zmierzające w kierunku rozległej problematyki społecznej⁶.

Tę linię rozwoju ujawniają i dobrze ilustrują opisy zawarte w autobiografii Fritza Heidera (1984), gdzie widzimy jego drogę przez Europę aż do USA jako ciąg spotkań z różnymi konkretnymi osobami i środowiskami psychologów w Wiedniu, Berlinie i Hamburgu – na początku XX wieku. A w tym korowodzie, w coraz to innym oddaleniu i odmiennej perspektywie, ukazuje się sylwetka Kurta Lewina. Zawsze w akcji i w interakcji z otoczeniem. Bo to wielorakie i zróżnicowane środowisko psychologów łączyło *abstrakcyjne* pojęcie psychologii, ale różnicowały je wzajemnie wrogie ideologie w służbie polityki. W okresie wzrostu w Europie tendencji nacjonalistycznych! Nie mogło to dziwić, ale powodowało występowanie wśród psychologów niepewności, napięć, niechęci oraz tendencji do ucieczki. A ta, kiedy poziom zagrożenia zapowiadał fizyczne unicestwienie, stawała się często jedynym efektywnym ratunkiem. Heider (1984) akcentował elementy, które psychologów mogły łączyć, a jednym z nich była metodologia. Ale teraz już ta własna, nowatorska, także ta Kurta Lewina. Dlatego też w swojej autobiografii wyraził zadowolenie z tego, że podczas pobytu w Berlinie mógł w kontakcie bezpośrednim z Kurtem Lewinem tak wiele się od niego nauczyć (por. Heider, 1984, s. 34: „Wieviel Glück hatte ich, daß ich ihm ein paar Jahre später in Berlin begegnete und von ihm lernen konnte”). Warto zauważyć: już w Berlinie!

⁴ Fritz Heider doktoryzował się w roku 1920 u Alexiusa von Meinonga w Grazu. Przez wiele lat interesował się psychologią postaci, od tej z Grazu poczynając. Od roku 1921 pracował w Monachium i Berlinie, gdzie miał kontakty z psychologami postaci i przyjaźnił się z Kurtem Lewinem (Lück, 2021). W roku 1926 opublikował *Ding und Medium* i został asystentem Sterna w Hamburgu, a w roku 1930 wyemigrował do USA.

⁵ W roku 1932 Lewin oddał w wydawnictwie G. Fischera w Berlinie maszynopis pracy *Grundlagen der topologischen Psychologie*. Radykalizacja sytuacji politycznej uniemożliwiła jej publikację w Niemczech. Po emigracji, w roku 1934 w USA Fritz i Grace Heider przetłumaczyli to dzieło na język angielski. *Principles of topological psychology* ukazało się w roku 1936 w Nowym Jorku (por. Heider, 1983, s. 114). Podczas wojny niemiecki oryginał zaginął. Tekst przetłumaczony zwrótnie na język niemiecki ukazał się ponownie w roku 1969 w wydawnictwie Hubera w Szwajcarii. Nakład został wyczerpany! (Lück, 1996).

⁶ W roku 1926, zachęcany przez Kurta Lewina i przy jego pomocy, Heider opublikował w Berlinie zauważoną niedawno na nowo pracę *Ding und Medium*, dotyczącą psychologii poznawania. Później, po śmierci Lewina, zajął się bardziej intensywnie *psychologią społeczną* i w roku 1958 opublikował szeroko znaną *The psychology of interpersonal relations*.

TOŻSAMOŚĆ UCZONEGO

Na kartach biografii Lewina napisanej przez Marrowa (1977/2002) rozpoznajemy grupę wybitnych dawnych berlińskich studentów Lewina, którym udało się ucieczka z nazistowskich Niemiec do USA, a wśród nich Tamarę Dembo i Marię Owsiankinę. Marrow podkreślił, że w tej grupie – w jej nastroju i oczekiwaniach – odżyła atmosfera twórczej pracy, przypominającej tę z lepszych lat... w Berlinie. Dlatego cytuje jednego z nich: „Wahre Wissenschaft duldet keine Stagnation. Immerwährende Veränderung – das ist die Essenz der Wissenschaft” [Prawdziwa nauka nie znosi stagnacji. Istotą nauki stanowią ciągle przemiany] (Marrow, 1977, s. 55). W twórczości naukowej Kurta Lewina, także w nowych warunkach, w USA, zaraz na początku pojawiły się elementy, które charakteryzowały ją już w Berlinie. Ten obraz uczonego, który opisał Wolfgang Köhler (1968), utrwalił się w pamięci jego berlińskich studentów. Można go rozpoznać jeszcze w wywiadzie, który z Blumą Zeigarnik w roku 1984 w Moskwie przeprowadził prof. Andrzej Gołąb. W roku 1984!

Autobiografia Kurta Lewina jest nieznaną. Dlatego ważne są źródła informacji zawarte w dokumentach pochodzących od osób trzecich, bliskich, przyjaciół, współpracowników. Ale równie wartościowym źródłem informacji są same wyniki naukowej twórczości, które jasno dokumentują kolejne etapy naukowego rozwoju autora. Wiadomo, że Kurt Lewin swój dorobek naukowy spisywał w dwóch językach. W okresie berlińskim był to język niemiecki, a podczas pobytu w USA – angielski. Wyniki badań trzeba bowiem było zaprezentować w języku zrozumiałym dla czytelnika i słuchacza. Na tym tle i w tym kontekście należy pojmować i oceniać opublikowanie w roku 1936 w USA *Principles of topological psychology*, która zawierała prace powstałe już w Berlinie.

Wydaje się, że za tym, co uwagi te wyrażają *prima facie*, ukrywa się problem trudniejszy. Wówczas w USA, ze względu na zainteresowanie twórczością Lewina, pojawiła się potrzeba udostępnienia jego studentom i czytelnikom prac, które powstały w Europie – szczególnie ważnych, tych, dla których podstawę stanowiła specyfika kultury europejskiej, specyfika europejskich stosunków społecznych, a także poziom i specyfika rozwoju europejskiej psychologii – głównie tych, które prezentowały już zarys jego własnej metodologii (zob. też: Schönplflug, 2006). Powiem wprost: prace Kurta Lewina wymagały od jego studentów w USA wyjścia poza ograniczenia w myśleniu, które wyznaczał dominujący tam behawioryzm! Bo specyfika myślenia Kurta Lewina była ważna i budziła zainteresowanie; w myśleniu tym do głosu dochodziły intuicje powstałe pod wpływem filozofii, głównie tej Cassirera – w USA w tym czasie mało znanej.

Jak pisał Graumann (1992/2007), zainteresowanie twórczością Lewina w Europie nastąpiło dopiero po roku 1968. Umożliwiały je zarówno publikacje jego dzieł oryginalnych, jak i ich opracowania, np. Lüscke (1996, 2001). Dodam, że zainteresowanie pracami Lewina inspirowały i utrudniały jednocześnie te same czynniki: ich bogactwo i różnorodność! (por. dobre i dokładne zestawienie publikacji Kurta Lewina: Kardaś, 2021, s. 388–405). A te określenia odnoszą się zarówno do tematyki prac, ich teoretycznych podstaw, języka, interpretacji oraz dostępności. Bywało też tak, że wcześniejsze poglądy Lewina powracały po latach w ujęciu innych autorów. Przykładem niech będzie tematyka, której dotyczył mój referat o krytyce taylorizmu przedstawionej przez Lewina w roku 1920 (W. Zeidler, 2011a). Należy przyjąć, że wczesne publikacje Lewina dotyczą tematyki, która go interesowała w sposób spontaniczny, jako poszukiwanie celu i sensu własnej aktywności. Punkt wyjścia stanowiła przy tym sytuacja rzeczywista. W przypadku pierwszego artykułu, *Krajobraz wojenny* (por. W. Zeidler, 2009), punkt widzenia należy do artylerzysty polowego. Kiedy wojna się skończyła, w roku 1919 na łamach *Zeitschrift für angewandte Psychologie* Lewin (1919) opublikował artykuł na temat racjonalizacji gospodarstwa rolniczego za pomocą środków psychologii stosowanej. Treści i praktyczne wskazówki w tym artykule stanowią udane transpozycje zasad ówczesnej *psychologii pracy*, głównie wg wskazań Hugona Münsterberga

(1912b). Uwagi Lewina koncentrują się na produkcji, ale ich podstawę stanowią trafne psychologiczne *analizy czynności*! Sam Kurt Lewin w życiorysie dołączonym do pracy doktorskiej napisał, że urodził się w Mogilnie jako syn kupca i właściciela posiadłości ziemskiej. Być może to pochodzenie obligowało go do racjonalizacji gospodarstwa rolniczego? Kardaś (2021) podaje, że rodzina Lewinów poczyniła inwestycje, które mogły wskazywać na chęć umocnienia swoich pozycji w Mogilnie, ale od roku 1905 (10 kwietnia) zamieszkiwała już w Poznaniu. Tak czy inaczej, trudno sobie wyobrazić, czym mógłby zajmować się wówczas w Mogilnie doktor psychologii.

RECEPCJA PRAC KURTA LEWINA – TAKŻE W POLSCE

Wiadomo, że w Europie przez wiele lat po zakończeniu II wojny światowej zainteresowanie twórczością Kurta Lewina tłumili agresywne frazeologie ideologiczne. Kiedy zwiędły, był czas na restytucję „niewygodnej wiedzy”. Zamiast klasowych wrogości dało się teraz odczuć bardziej subtelne różnice między grupami społecznymi, często dalekie od wzajemnej niechęci. Możliwości te wynikały z teorii i wyników badań Kurta Lewina. Teraz na przypomnienie zasługują dwie inicjatywy. Pierwszą była inicjatywa Graumanna z lat 1981–1982: *Kurt-Lewin-Werkausgabe*. Niestety, spośród planowanych siedmiu udało się opublikować tylko cztery tomy. Resztę zablokowały spory dotyczące praw autorskich. Kolejną inicjatywą stał się tom *Kurt Lewin: contribution to contemporary psychology...* (Trempała, Pepitone, Raven, 2006), w którym zebrano referaty wygłoszone na wspomnianej już bydgosko-mogileńskiej konferencji. W roku 2009 Lück wydał zbiór prac Lewina: *Schriften zur angewandten Psychologie*. W latach 2009–2011 w Wyższej Szkole Finansów i Zarządzania w Warszawie odbyły się dwie konferencje międzynarodowe pt. *Psychologia europejska w okresie międzywojennym*. Na obydwu konferencjach wygłoszono 36 referatów, których teksty wydano w publikacjach zwartych (W. Zeidler, 2009; W. Zeidler, H. E. Lück, 2011). Spośród nich jedna czwarta była poświęcona Kurtowi Lewinowi, tzn. różnym aspektom jego twórczości. Zainteresowanie Lewinem było znaczne i wyraźne. Dlaczego? Również w roku 2009 ukazała się w Warszawie pod redakcją Kozińskiego praca *Nowe idee w psychologii*. Nazwisko Lewina występuje w tej książce *dwa razy*. Raz w tekście (str. 164) i drugi raz w indeksie osobowym. Dwa słowa – to jednak niewiele? Okoliczność tę wyjaśniają uwagi Kozińskiego zawarte we wstępie do tej pracy. Ta pierwsza wzmianka słusznie odnosi tematykę prac zamieszczonych w tomie do popularnej w Polsce koncepcji regulacji, druga zaś stanowi konstatację, że prace na te tematy są „obiektywnie nowe i wartościowe” (Koziński, 2009, s. 17). Tak było, w tym czasie bowiem dostęp do oryginalnych prac Lewina i opracowań źródłowych był w Polsce nadal ograniczony.

WŁAŚCIWY PUNKT WIDZENIA?

W sensie idealnym właściwym punktem widzenia będzie ten, który pomoże zrekonstruować biografię danego autora oraz teorie twórcy lub uczonego. Gdy poznajemy życiorys, a potem interesują nas dzieła lub teorie danego twórcy, oglądamy za każdym razem tę samą postać, ale na innym tle lub w innej perspektywie. Istniejące biografie Kurta Lewina są różnorodne. Chociaż różne, mogą być *prawdziwe*, ponieważ zostały napisane z różnych punktów widzenia. Biografowie Uczzonego bowiem w różnym miejscu i różnym czasie kierowali swoją uwagę ku różnym aspektom jego życia i twórczości i starali się je opisać i zrozumieć. Punkt widzenia warunkuje i determinuje zarówno poznanie, jak i zrozumienie zjawiska, życiorysu lub dzieła? A punkt widzenia wybieramy nie wtedy, gdy chcemy

rzeczywistość zobaczyć, ale wtedy, gdy chcemy ją lepiej poznać, a jeśli ktoś woli: zrozumieć? Włączyć do systemu naszej wiedzy, tej własnej, żywej i „codziennej”.

Oddzielanie życiorysu od twórczości jest uproszczeniem. Ale bez niego poznanie jako takie trudno byłoby sobie wyobrazić. Poznanie ma swoje prawa. Ale również rzeczywistość funkcjonuje według własnych praw. Równoczesne stosowanie jednych i drugich jest trudne, ale interesujące, bo obiecuje poznanie całości, a nie tylko jej części. I chyba właśnie stąd bierze się zapotrzebowanie na biografie, a także ich urok. Zrozumienie i uznanie budzi to, że jeden z uczniów Lewina, Marrow (2002), opisał jego „życie i dzieło”, zaraz gdy ich ślady trwały jeszcze w pamięci przyjaciół i uczniów Uczonego. Rekonstrukcja biografii opierała się na dwóch rodzajach źródeł: były to własne wspomnienia i wypowiedzi osób trzecich oraz opis i analiza dzieł Lewina. Fragmenty wspomnień stanowiły źródło dla ważnych, chociaż wycinkowych opisów biograficznych, takich jak w opisach Fritza Heidera (1984) albo w wypowiedziach Blumy Zeigarnik zanotowanych przez Gołąba (1984). Myślę, że Tomasz Kardaś (2021) oparł swoją pracę na innych założeniach. On nie miał do dyspozycji świadków, tylko dokumenty, do których dotarcie nie było wcale łatwe. Z tym większym respektem akceptuję jego starania o to, aby swoje stwierdzenia opierać na starannej analizie dokumentów, tych leżących niemal pod ręką (w Mogilnie i Poznaniu), ale także zdeponowanych na innym kontynencie.

Myślę, że należy mieć odwagę i powiedzieć, że zainteresowanie osobą oraz dorobkiem Kurta Lewina w Europie zaczęło się pojawiać po roku 1968, czyli po tej studenckiej rewolucji, która ożywiła cały kontynent. W Niemczech spowodowała odnowienie pamięci ofiar nazistowskiej tyranii. Również psychologów – kobiet i mężczyzn! – a wśród nich także Kurta Lewina. Na początku lat 80. miało go przypomnieć zbiorowe wydanie prac pod redakcją Carla Friedricha Graumanna⁷: *Kurt-Lewin-Werkausgabe* (1981). Ocena i recepcja dzieł Lewina były utrudnione przez to, że w Europie jego prace opublikowane w USA nie były znane. W roku 1992 w wydawnictwie Peter Lang we Frankfurcie nad Menem ukazały się dwie publikacje zbiorowe. *Zur Geschichte der Psychologie in Berlin*, pod redakcją Sprunga i Schönpfluga, prezentowała wydarzenia ważne dla psychologii berlińskiej, od oświecenia poczynając. Wśród tych „wydarzeń” brakowało jednak Kurta Lewina. A ta druga praca, również pod redakcją Schönpfluga (1992/2007), pt. *Kurt Lewin – Person, Werk, Umfeld*, składała się z 15 opracowań różnych autorów, w sposób wybiórczy skoncentrowanych albo na fragmentach jego życiorysu, albo na aspektach jego działalności naukowej, zarówno w Berlinie, jak i USA. Ten tom zamyka opracowanie Graumanna (1992/2007) zatytułowane *Lewin zu Beginn des 21. Jahrhunderts* (Lewin na progu XXI wieku). Tekst opracowania powstał na podstawie referatu, który Graumann wygłosił na Kongresie Niemieckiego Towarzystwa Psychologicznego (Kongreß der Deutschen Gesellschaft für Psychologie) w roku 1990 w Kilonii.

Wydawca dzieł zebranych Kurta Lewina miał bardzo dobrą orientację w jego twórczości, zarówno tej wcześniejszej, w Niemczech, jak i późniejszej, w USA. Miał wiele do powiedzenia także na temat recepcji jego prac i poglądów. Zapowiedź w tytule artykułu Graumanna (1992/2007) zasługuje na uwagę, ponieważ jest w nim mowa nie tylko o tym, co się wydarzyło, lecz także o znaczeniu twórczości Lewina dla dziejów psychologii, dla jej przyszłości! Zaletą artykułu jest to, że autor potraktował recepcję jako proces otwarty. Odbiór prac Lewina był teraz wydarzeniem i procesem, który się dokonywał, ale jeszcze nie zakończył, gdyż psychologia XX wieku nie była przygotowana na przyjęcie wielu idei Lewina, a powierzchowne interpretacje ukrywały ich sens wynikający z założeń teoretycznych autora. Przedstawiciele poszczególnych dyscyplin, którzy w jego koncepcjach znajdowali coś, co ich interesowało, akceptowali to wybiórczo, na swój sposób, i dopasowywali do swoich teorii i swojego języka: psychologii rozwojowej, psychologii motywacji, psychologii społecznej itd.

⁷ Carl-Friedrich Graumann (1923–2007) w latach 1963–1991 był profesorem psychologii w Uniwersytecie w Heidelbergu.

W tych adaptacjach gubiło się zarazem to, co w twórczości Lewina było oryginalne, a dla dziejów psychologii jeszcze obce, bo wręcz prekursorskie!

Wyraźne kłopoty stwarzało przełożenie na język psychologii tego, co Lewin (1931/1981b) wyraził w artykule *Der Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie* (Graumann, 1992/2007). Tak musiało być, ponieważ psychologowie w USA nie mieli dobrej znajomości filozofii europejskiej, szczególnie tej z kręgu neokantyzmu. Przy tym Graumann (1992/2007) w tytułowym *przejściu* (*Übergang*) nie widzi procesu, który się zakończył, tylko jego zapowiedź. Bo Lewin owego procesu nie diagnozował, tylko wskazał jego możliwość⁸. Sens tej uwagi widzę w tym, że „wczesna” psychologia oparła swój system na filozofii Arystotelesa, tymczasem filozofia XX wieku wędrowała już własnymi drogami. Czy „dopasowanie” ówczesnej psychologii do filozofii Galileusza było zadaniem sensownym? Kiedy ówczesna psychologia chciała jednak być nauką od filozofii niezależną... Lewin natomiast, znając Cassirera, miał własne pomysły na temat tego, jak mógłby wyglądać dialog psychologii z filozofią, a myślał przy tym już konkretnie o filozofii neokantyzmu. Graumann był zdania, że w wieku XXI nadejdzie właściwy czas dla zintegrowania dorobku Lewina oraz wcześniejszej wiedzy psychologicznej, także tej, którą właśnie próbowano tworzyć! To było życzenie słuszne, ale trudne do realizacji, ponieważ część idei Lewina została już zakotwiczona w metodologiach podporządkowanych życzeniom praktyki.

MOŻLIWE PUNKTY WIDZENIA

Dla Tomasza Kardasia (2021) Lewin nie był *przypadkiem*, tylko *procesem* – wielostronnym oraz wieloetapowym, który wypełnił cały czas jego życia. A on to bogactwo intensyfikował. Zatem nie chciał swoich zainteresowań zredukować ani do filozofii, ani do psychologii. Szeroko zakrojone badania w różnych grupach społecznych prowadziły do odkrycia mechanizmu psychologicznego nazywanego dzisiaj rezyliencją. Następcy Lewina nie chcieli interpretować wyników jego badań w terminach lewinowskich kategorii wyjściowych, tylko zredukowali je do kategorii aktualnych kierunków psychologii w USA.

Ten model interpretacji występuje w wielu opracowaniach biograficznych. Jest to model, który nobilituje sam siebie przez postulat, aby psychologia zachowała całkowitą niezależność w stosunku do filozofii. Ale czy ten behawiorystyczny dogmatyzm jest psychologii w XXI wieku potrzebny? Rezygnacja z niego nie udała się także Fahrenbergowi (2015). Autorzy, których teksty zawiera *Levinian psychology*, konsekwentnie idą jego śladem i traktują cały dorobek Kurta Lewina jako produkcję jednostki twórczej. Zaslugą Kardasia (2021) jest to, że w przeciwieństwie do innych, wcześniejszych biografii Kurta Lewina nie tylko wyszczególnia dwa różne okresy: ten w Berlinie i USA, lecz także opisuje środowiska wcześniejsze: Mogilno i Poznań. Nie tylko lokalizuje je na mapie Europy, ale też pokazuje ich miejsce w Europie tamtego czasu, w jej historii i kulturze. Jest to wielka zaleta, ale i trudność. Bo lokalizacja w czasie europejskim musi się łączyć z przypomnieniem tragedii rozbiorów Polski. W odróżnieniu od innych biografów Stachowski (2006) miał odwagę i możliwość pokazania fragmentu dziejów Mogilna jako skrawka poznańskiej prowincji będącej częścią Prus. Niestety, Norman Davies, którego Stachowski cytuje, przedstawił sytuację Mogilna oraz rozbiorowej Polski w sposób uproszczony. Procesy narodotwórcze w przypadku państwa polskiego były bardziej zaawansowane, niż oceniał to brytyjsko-polski historyk. Prusy natomiast przez cały wiek XIX dążyły do restytucji swoich zasobów ludnościowych

⁸ Miriam Lewin (ur. 1931) wskazała na *Übergang* jako wybór nowego paradygmatu dla naukowych badań psychologicznych (1986, s. 13–18). W ten sposób różnica, lokowana w płaszczyźnie filozofii, otrzymała teraz miejsce na poziomie metodologii.

zmniejszonych podczas wojen napoleońskich! Przyjmowały jako swoich obywateli tych, którzy z rozmaitych względów chcieli zmienić miejsce zamieszkania i wyzwolić się z panującego poddaństwa albo uniknąć prześladowań religijnych. Do czasu rozbiorów Polski w Mogilnie brak jest wyraźnych śladów osadnictwa Żydów. Brak śladów antysemityzmu. W roku 1812 na terenie zaboru pruskiego Żydzi otrzymali prawa obywatelskie. Od tego momentu zamieszkiwały na terenie Mogilna trzy nacje: polska, żydowska i „pruska”, mające *swoją* własną religię. Sytuację komplikował dodatkowo fakt, że w Prusach językiem oficjalnym był język niemiecki, co reglamentowało możliwości społecznego awansu.

W biografii Kurta Lewina napisanej przez Marrowa (1969) dwie trzecie tekstu zajmują opisy prac oraz twórczości Lewina w USA. Myślę, że stało się tak, ponieważ dla każdego z biografów droga życiowa Kurta Lewina miała specyficzne akcenty, różnie „uzasadnione”. Chociaż interesowała ich twórczość Lewina, nie mogli ignorować miejsca jego zamieszkania, zwłaszcza gdy chodziło o różne kontynenty oraz różne języki i tym bardziej że różnice tego rodzaju zostały już dostrzeżone przez znanych i wysoko cenionych uczonych tamtego czasu, takich jak Münsterberg (1912a, 1915), Simmel (1903/2014) czy Einstein (1934/1979).

Marrow (1969) mógł swój opis oprzeć na relacjach z drugiej ręki. Kardaś (2021) musiał szukać potrzebnych mu informacji w archiwach, które szczęśliwym zbiegiem okoliczności ocalały z wojennych kataklizmów. Rzeczą oczywistą jest to, że część z nich znajdowała się w Mogilnie, Inowrocławiu, Poznaniu i Berlinie, a nie tylko w Stanach Zjednoczonych. Myślę, że ta uwaga jest tutaj potrzebna, bo trudno wyobrazić sobie pisanie biografii w oderwaniu od *rzeczywistości*, bez pewności, z jakiego punktu widzenia dany przedmiot oglądamy.

Różnorodność otaczających Lewina struktur społecznych w kolejnych miejscach jego zamieszkania tworzyła osobliwy raster, który wyznaczał dla społecznych jakości (wyznania, narodowości itp.) swoistą walencję. Ten fakt musiał determinować już to, co w ówczesnej psychologii nazywano *przeżyciem*. Kurt Lewin gromadził je i na swój sposób porządkował. Już w Mogilnie, potem w Poznaniu i Berlinie. Gdy znalazł się w USA, wiedział już nie tylko, że te przeżycia istnieją, ale i to, że aby z nimi móc współżyć, należy je przedtem co najmniej uporządkować. W tym kontekście pojmuję uwagę Kardasia (2021) przy powołaniu się na świadectwo Helgi Sprung na temat percepcji społecznej Kurta Lewina, gdy powiada, że studenci uwielbiali Lewina za jego demokratyczny styl pracy i tolerancję. Także w USA potrafił Lewin przeżycia nie tylko dostrzegać, lecz także operacjonalizować, na tyle dokładnie, że mogły stać się problemem dla badań naukowych. To była jego zasługa. Nie będę twierdził, że to było jego odkrycie, dokonał go bowiem nie tam, gdzie panował behawioryzm, ale już wcześniej – w tej starej, pocziwej Europie, gdzie było już wiadomo, że reakcja to jednak coś innego niż przeżycie, a przeżycie to więcej niż reakcja!

LEWIN I FILOZOFIA

Szczegółowa analiza związków Lewina z filozofią nie mieści się w ramach tego artykułu, zatem ograniczę się do wskazania tych aspektów, które mają duże znaczenie: (a) związki z filozofią na początku rozwoju naukowego, (b) zainteresowanie Lewina tylko wybranymi koncepcjami filozofii, (c) czas trwania zainteresowania Lewina filozofią – od początku twórczości do końca życia.

(a) Heider w swojej *Autobiografii* przypomina wypowiedź Lewina, w której wyraził on wdzięczność wobec Ernsta Cassirera za to, że zarówno w teoriach, jak i badaniach mógł się opierać na jego poglądach („...die ich in seinen Ansichten über Wissenschaft und Forschung fand...” – Heider, 1984, s. 78). Ten trwały wpływ poglądów Cassirera na myślenie

Kurta Lewina potwierdził także Kurt W. Zeidler⁹ (2011), wskazując zarazem, że chodziło tutaj o inspiracje wynikające z teorii poznania i teorii nauki Cassirera (K. W. Zeidler, 2011, s. 292). Stwierdzenie to zawęża problem do tylko dwóch dyscyplin filozofii. Przypomnijmy, że w roku 1910 Cassirer (2000) opublikował pracę *Substanzbegriff und Funktionsbegriff*, w której przedstawił konsekwencje wynikające z matematyki dla teorii poznania. Jego analizy dały podwaliny wyraźnego rozróżnienia pojęć substancji i funkcji. Lewin wykorzystał to różnicowanie jako podstawę rewizji pojęć tradycyjnej psychologii, co w roku 1931 pozwoliło mu zakreślić punkt wyjścia dla wyróżnienia dwóch poglądów na naukę: Arystotelesa i Galileusza. Ale już wcześniej, w roku 1914, Lewin pokazał różnice dzielące świadomość jednostki od poznania naukowego. Pominę wyliczanie kolejnych jego prac na temat epistemologii i teorii nauki, wskażę jeszcze na tę, którą napisał pod koniec życia, a która ukazała się po jego śmierci, w roku 1949: *Cassirers Wissenschaftsphilosophie und die Sozialwissenschaften* (Lewin, 1981a). Jest to kompendium metodologii badań nowoczesnej psychologii. W Polsce praca ta miała wpływ na praktykę badań psychologicznych tylko okazjonalnie lub symbolicznie. Dlaczego?

(b) Nie ma tutaj miejsca na dokładniejszą analizę zainteresowań Kurta Lewina filozofią, ale wskażę na jego zainteresowania możliwością wykorzystania matematyki w psychologii. Wtedy gdy inni psychologowie ograniczali zastosowanie matematyki do statystyki, Lewin miał już odwagę korzystania z niej dla tworzenia teorii psychologicznych: topologii i psychologii pola. Być może ze świadomością przestrogi Einsteina: „Insofern sich die Sätze der Mathematik auf die Wirklichkeit beziehen, sind sie nicht sicher, und insofern sie sicher sind, beziehen sie sich nicht auf die Wirklichkeit” [Twierdzenia matematyki dotyczące rzeczywistości nie są pewne; kiedy są pewne, nie dotyczą rzeczywistości] (1979, s. 141). Zasługą Lewina jest to, że dostrzegł różnicę między myśleniem potocznym i poznaniem naukowym.

(c) Oryginalne i nowatorskie elementy poglądów Lewina były inspirowane przez filozofię Cassirera. Były to inspiracje trwałe, zauważono tylko ich elementy. Kardaś (2021) przedstawił staranny i zwarty opis studiów Lewina, który po krótkim pobycie w Monachium oraz Fryburgu Bryzgowijskim, w dniu 16 kwietnia 1910 roku został wpisany na listę studentów Uniwersytetu Fryderyka Wilhelma w Berlinie. Do tego punktu w czasie należy odnieść młodzieńcze prace Lewina, które znalazły się w tomie 1 *Werkausgabe* (1981), a dotyczą zasady zachowania (*Erhaltungsprinzip*) w psychologii, oraz następną: *Die Verwechslung von Wissenschaftssubjekt und psychischen Bewusstsein in ihrer Bedeutung für die Psychologie* (1914/1981c; Konsekwencje dla psychologii pomylenia myślenia poznającego podmiotu ze świadomością psychiczną). Métraux opatrzył ją adnotacją, że została opublikowana w roku 1914, w zeszycie jubileuszowym oferowanym przez studentów prof. Aloisowi Riehlowi¹⁰. Praca ta cieszy się nadal uznaniem, prawdopodobnie dlatego że prezentuje krytyczną teorię poznania. Jest to dzieło, które zamiast rozstrzygnąć mnoży raczej pytania. Jaką wartość ma poznanie jednostkowe? – pytał Kurt Lewin, mając 24 lata! Tylko czy filozof Alois Riehl mógł udzielić zadowalającej odpowiedzi psychologowi, który jeszcze nie chciał zredukować swojego punktu widzenia do horyzontów psychologii eksperymentalnej? Będzie dobrze, jeżeli w tym miejscu zauważymy, że Kurt Lewin będzie potem oglądał interesujące go problemy zwykle w kontekście filozoficznych propozycji neokantystów, najczęściej Cassirera. W przywołanej już pracy z roku 1949, *Cassirers Wissenschaftsphilosophie und die Sozialwissenschaften*, Lewin (1949/1981a) stwierdził, że w latach 1910–1946 korzystał z jego idei, aby zrozumieć, jakie zadania ma nauka i badania naukowe (s. 347: „die Hilfe anzuerkennen,

⁹ Kurt W. Zeidler (ur. 1953) jest profesorem filozofii w Uniwersytecie Wiedeńskim, specjalizuje się m.in. w zakresie teorii poznania. Szeroko znana jest jego praca z roku 2000, *Prolegomena zur Wissenschaftstheorie*.

¹⁰ Alois Riehl (1844–1924), wykształcony w Grazu, w Uniwersytecie w Berlinie był profesorem na stanowisku zajmowanym poprzednio przez Wilhelma Diltheya. Należał do neokantystów.

die Cassirers Auffassung über die Natur von Wissenschaft und Forschung anbietet”). Stąd też jego uwaga krytyczna, iż psychologowie nie doceniają zjawisk kulturowych i są skłonni mniemać, że istnieją wyłącznie ludzkie indywidua (Lewin, 1949/1981, s. 355: „Manche Psychologen blicken mit Mißtrauen auf jene Kulturtatsachen, mit denen sich der Anthropologe beschäftigt. Sie neigen dazu, nur Individuen als real zu betrachten...”). Znaczenie tej uwagi zrozumiemy, gdy uświadomimy sobie specyfikę panujących poglądów filozoficznych¹¹ oraz tak silnie odczuwaną przez psychologów potrzebę niezależności od filozofii!

Daniel Cervone (2006) oraz Ryszard Stachowski (2011) przedstawili analizy artykułu Kurta Lewina (1931/1981b), który właściwie sygnalizował osiągnięcie przez niego w rozwoju naukowym punktu, który można nazwać dojrzałością i samodzielnością: *Der Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie*. Owszem, *Übergang* znaczy ‘przejście’, ale tutaj ma sens zmiany w trakcie rozwoju. Bo właśnie rozwój myślenia w nauce był dla Lewina sprawą najważniejszą. Redukowanie jego poglądów do potrzeb psychologii zajętej porządkowaniem problemów adaptacyjnych byłoby błędem. Lewin pracował wówczas nad modelem myślenia, który zaprezentował w pracy *Grundlagen der topologischen Psychologie* (zob. przyp. 5). Lück (1996) przedstawił kolejne fazy i etapy powstawania psychologii pola oraz topologii. Także on widzi pewne podobieństwo tych poglądów do intuicji przedstawionych w roku 1920 przez Wolfganga Köhlera w jego pracy *Die physischen Gestalten in Ruhe und stationären Zustand* (por. W. Zeidler, 2012). Jest to praca przez psychologów rzadko czytana i cytowana; powstała podczas pobytu Köhlera na Teneryfie (1914–1920) i została przyjęta jako rozprawa habilitacyjna (Berlin). Jej tematykę modelowała obfita korespondencja autora z fizykiem Hansem Geitelem (1855–1923) na temat teorii względności, drugiej zasady dynamiki i pojęcia postaci (por. W. Zeidler, 2011b). Köhler napisał do niej dwie różne przedmowy. Jedną *dla filozofów i biologów*, a drugą *dla fizyków*. W rzeczywistości był to głos w sporze logicyzmu z psychologizmem, ale „na rzecz” psychologii jako nauki przyrodniczej. Dodam, że jako klucz do zrozumienia psychologii pola Graumann przypomniał zdanie Cassirera: „Die Frage kann hier niemals lauten, wie wir von den Teilen zum Ganzen, sondern wie wir von den Ganzen zu den Teilen gelangen” [Nie należy pytać, w jaki sposób tworzymy z części całość, tylko w jaki sposób od całości dochodzimy do poznania jej części] (Cassirer, 1910, za: Graumann, 1982, s. 21). A to jest jednoznacznie problem epistemologii. Dla zrozumienia specyfiki psychologii topologicznej teorii fizyki oraz matematyki są niezbędne (por. Einstein, 1979, s. 135–150), wszakże przy właściwym zrozumieniu używanych pojęć. W tym kontekście zaletą opisów Kardasia (2021) jest to, że w sposób interesujący pokazuje rozwój poglądów Lewina, które niebawem utworzą jego psychologię topologiczną. Na podkreślenie zasługuje, że ukazuje nam przy tej okazji, albo z tej racji, przestrzeń życiową Lewina. Tę szczególnie ważną – Uniwersytetu Berlińskiego, a w niej także tamtejszą psychologię postaci *in statu nascendi*, jeszcze „naturalną”, wolną od deformacji przez tzw. gestaltizm.

Gdy próbuję wyobrazić sobie, jak wyglądały filozoficzne poglądy Kurta Lewina, nie mogę pominąć informacji dotyczących psychologii, tej z pogranicza XIX i XX wieku. W roku 1928 w swoim doktoracie Karl Popper napisał: „Es ist wohl ein Verdienst Wundts,

¹¹ Po tej pracy z roku 1910, w latach 1923–1929 Cassirer napisał swoje dzieło: *Philosophie der symbolischen Formen*. Podczas pobytu w USA na życzenie nowych przyjaciół napisał w języku angielskim skrót tego dzieła. *Esej o człowieku* ukazał się w języku polskim w roku 1977. Sens poglądów Cassirera sprowadza się do tego, że człowiek jako istota ludzka żyje w świecie symbolicznym – świecie języka, religii, sztuki i nauki. Teraz, gdy to zostało już powiedziane, możemy temat zakończyć albo przypomnimy sobie, że te same tematy wymieniał już Wilhelm Wundt jako centralne zagadnienia *Völkerpsychologie*. Wspomniany *Esej o człowieku* zaopatrzył przedmową B. Suchodolski, który stwierdził, że zainteresowania Cassirera należały do utopijnych, ale na szczęście jego filozofia należy „do minionego już okresu mieszczańskiej świadomości” (Cassirer, 1977, s. 19). W tej sytuacji można dociekać, czy wiedza polskich psychologów nt. Cassirera pozwala na właściwe zrozumienie poglądów Kurta Lewina, zwłaszcza tych, które dotyczą metodologii badań psychologicznych.

die Bedeutung der *Völkerpsychologie* für die Psychologie des Denkens hervorgehoben zu haben” [Ważną zasługą Wundta było to, że pokazał znaczenie *Völkerpsychologie* dla psychologii myślenia] (1928/2006, s. 244). To uznanie dziwi, ale w rzeczywistości katalog problemów, które powstały poza repertuarem psychologii akademickiej, był szerszy. W roku 1913 Carlo Sganzi (1881–1948; zob. W. Zeidler, 2018, s. 145–150), profesor Uniwersytetu w Bernie, atestował znaczenie *Völkerpsychologie* w ujęciu Wilhelma Wundta, ale nie pominął swoistej inspiracji tej problematyki, którą przyniosła praca Carla Cattaneo¹². Podążając tym śladem dalej, możemy stwierdzić, że zarówno u Wundta, jak i u Cassirera podstawowe problemy *Völkerpsychologie* oraz *Filozofii form symbolicznych* można ująć w te same grupy: język, religia, sztuka, nauka¹³. Byłby to krok ryzykowny, gdybym mniemał, że Kurt Lewin wspomniane właśnie poglądy Wundta i Cassirera znał i akceptował. Równie lekkomyślne – właśnie: lekkomyślne – byłoby twierdzenie przeciwne: że z tymi poglądami się nie zetknął.

...życie bez pracy jest puste i niepełne...¹⁴

...napisał Lewin w roku 1920 w krytyce taylorizmu. Praca ta powstała w czasie, gdy utrzymywał kontakty z Karlem Korschem (1866–1961), marksistą. Ta informacja skłania do przypuszczenia, że owa krytyka taylorizmu Lewina odbywa się z punktu widzenia marksizmu. Tak nie jest. Widać to, gdy uświadomimy sobie, że w cytowanym zdaniu ukryte są te wartości, dla których w marksizmie nie ma miejsca. Nie było też dla nich miejsca w ówczesnej przyrodniczo zorientowanej psychologii. Stąd też były tutaj potrzebne uwagi na temat stosunku Lewina do filozofii Cassirera. Bo także Lewin wiedział, iż *rzeczywistość ludzka* jest czymś więcej niż *rzeczywistość biologiczna*! Obecność elementów aksjologicznych jest w niej niezbędna. Myślę, że oryginalność i wartość prac Lewina buduje i uzasadnia dorobek całej nauki europejskiej z przełomu wieków XIX i XX. Komentatorzy niemieccy i amerykańscy zauważyli w tym dorobku to, co chcieli albo co byli w stanie zauważyć. W polskiej biografii natomiast Kardaś (2021) wyszedł poza kredowy krąg rutynowych ocen. Zauważył nieco więcej! I to jest jego zasługą, ponieważ otworzył możliwość przywrócenia myśli Kurta Lewina europejskiej kulturze. Utorował jej drogę także do polskiej psychologii, przede wszystkim tej akademickiej!

Można powiedzieć, że w ten sposób wskazałem palcem pewną całość – kilka problemów, które wymagają identyfikacji, werbalizacji i operacjonalizacji. Tą całością jest nic innego jak twórczość Lewina. Owszem, obszerna, bo obejmuje kilka tysięcy stron, trudno dostępnych. Tym pierwszym problemem byłoby ich udostępnienie polskiemu czytelnikowi, nie tylko studentom. Tłumaczenie wybranych tekstów? Owszem, ale nie tylko. Dalej, idąc tropem Kardasia (2021), należy stwierdzić, że potrzebne są także opracowania krytyczne najważniejszych prac Kurta Lewina. Ta niezwykle staranna biografia sporządzona przez Kardasia domaga się uzupełnienia w postaci osobnej pracy poświęconej twórczości Lewina, co najmniej takiej jak znakomita praca Lücka z roku 1966, *Die Feldtheorie und Kurt Lewin*. Takiej, która ukaże specyfikę i odmienność tego, co w tej twórczości należy do psychologii i do epistemologii, do *teorii nauki*. Ukaże też stosunek myślenia w rozumieniu psychologii do poznania logicznego, i to na tyle dokładnie, że zrozumiemy absurdalność sporów toczonych między logicyzmem oraz psychologizmem i zbliżymy się do wizerunku psychologii jako nauki. A stosunek psychologii jako nauki do praktyki psychologicznej wygląda mniej więcej

¹² Carlo Cattaneo (1801–1869) – działacz włoskiego *risorgimento*, po upadku rewolucji 1848 roku nauczyciel filozofii w szwajcarskim Lugano, gdzie w latach 1853–1863 napisał *Psicologia delle menti associate*.

¹³ Chociaż Pankalla (2011) podąża inną drogą, można przyjąć, że jego wyniki nie są z tymi intuicjami sprzeczne.

¹⁴ „...das Leben ohne Arbeit hohl und halb ist...” (Lewin, 1920, s. 12).

tak jak ten, który opisał Einstein, mając na myśli wspomniany wcześniej stosunek twierdzeń matematyki do rzeczywistości.

OSOBLIWOŚĆ PRZESTRZENI ŻYCIOWEJ KURTA LEWINA

Książka Kardasia (2021) inspirowane pewne sugestie dotyczące tych aspektów, które pozwolą lepiej poznać biografię Lewina i wartość jego twórczości. Odpowiednie badania muszą być dziełem ośrodków akademickich. Nie można jednak przeoczyć tego, że Kardaś pokazał elementy przestrzeni życiowej Lewina, których inni... nie zauważyli albo nie chcieli ich opisać lub zinterpretować. Bo przecież już nie istnieją! Zaczę od tej szczególnej przestrzeni życiowej w Mogilnie, na obrzeżu prowincji pruskiej w Cesarstwie Niemieckim. Miała ona historyczne osobliwości wykraczające poza obszar problematyki rozbiorowej. Nie miała także nic wspólnego z problemami narodowej polityki państwa polskiego, które w tym czasie nie istniało! Wskażę jednak na tę państwowotwórczą politykę Oświeconego Monarchy i jego zabiegi o restytucję potencjału Prus zrujnowanego przez wojny napoleońskie. Jej narzędziem była tolerancja w stosunku do różnowierców, zbiegów szukających wyzwolenia z poddaństwa albo religijnej wolności!

Kardaś (2021) ujawnia ślad, którym biegła „tolerancja” Oświeconego Monarchy. Powołuje się przy tym na dokumenty. Swojsko, jako możliwy do identyfikacji, prezentuje się gmach dawnego Friedrich Wilhelm Gymnasium w Poznaniu. Dziennik gimnazjum, rozkład materiału, listy uczniów i stopnie Kurta Lewina w roku szkolnym 1902/1903 budzą nie tylko ciekawość, lecz także respekt dla „kwestora” poznańskiego archiwum! Respekt należy się za odwagę w ukazywaniu złożoności zjawisk. Jesteśmy w okresie zaborów, w rzeczywistości bardzo skomplikowanej. Obok zniewolonych istnieli w niej także zaborcy. Kardaś dodaje, że funkcjonował również podział ze względu na religię, oraz przekazuje, że w roku 1905, kiedy Kurt był uczniem klasy trzeciej i obchodzono bar micwę, do jego klasy uczęszczało: „20 ewangelików, 10 katolików i 17 Żydów” (Kardaś, 2021, s. 66). Dodał też, że interesuje go nie narodowość, ale *wyznanie*. A język? Ten był przecież tylko jeden – państwowy. I wyznaczał obszar tolerancji na usługach „oświeconej” polityki. A dalsze zmienne demograficzne w ówczesnej psychologii i tak nikogo nie interesowały.

Tymczasem te różne zasady podziału komplikowały wzajemne zrozumienie, a także funkcjonowanie członków należących do wyróżnionych zbiorowości. Chociaż tylko po części, to także na zasadzie przymusu. Także w relacjach interpersonalnych. Dzisiaj nie chcemy wiedzieć, ile narodowości i wyznań istniało w każdym z zaborów. Dobrze, że Kardaś (2021) te okoliczności przypomniał, dzięki czemu potwierdził wartość materiałów zgromadzonych w publikacji *Poznańscy Żydzi* (Karolczak, 2006; Witkowski, 2006), także tych, które dotyczą rodziny Lewinów. Raz obok, raz w tle tych informacji pierwszoplanowych: studia, publikacje, praca w uniwersytecie, Kardaś (2021) ma coś do powiedzenia także na temat rodziny Lewinów, dzięki czemu czytelnik może wytworzyć własny obraz tej rodziny – taki jak moje impresje, które powstały podczas spotkania z Thomasem Lewinem w roku 2004 w Mogilnie nad jeziorem. Miałem wrażenie, że był on dumny z tego, że droga życiowa jego stryja miała swój początek w tej pięknej okolicy. Zrozumiałem wówczas, jak wiele treści zawierała jego wypowiedź, którą scharakteryzował Kurta Lewina w swoim oficjalnym wystąpieniu: „He also loved music” (T. Lewin, 2006, s. 296). Czulem, że ta przestrzeń życiowa w Mogilnie była przyjazna, a on to jezioro... *rozpoznawał!*

Okres młodości Kurta Lewina – w Mogilnie, Poznaniu oraz w Charlottenburgu przedstawił Kardaś (2021) na 80 stronach, dokładniej niż w innych opracowaniach. Opisy te uświadamiają, że dla Kurta Lewina jego przestrzeń życiowa zawierała szereg sygnałów ważnych i potrzebnych w procesie tworzenia poczucia własnej tożsamości, dla porządkowania

i zrozumienia otaczającego świata. Był to swoisty raster, specyficzny, własny sposób widzenia świata! Trudno byłoby sądzić, że Kurt Lewin go nie zauważył. Od początku edukacji Kurt wynosił ze swoich kolejnych doświadczeń refleksje dotyczące istnienia grup społecznych, także wstępny pogląd na temat istnienia i różnorodności grup społecznych, które go otaczały. Później, podczas studiów w Berlinie, zauważył, że grupy społeczne jako takie istnieją w każdym społeczeństwie. Zauważył to już tutaj, w Europie! A potem te *przednaukowe intuicje* sprawdzał i potwierdzał w kolejnych programach badawczych w USA.

Hohenwestedt, 22.06.2021 r.

BIBLIOGRAFIA

- Brett, G. S. (1969). *Historia psychologii*. Warszawa: PWN.
- Bringmann, W. G., Lück, H. E., Miller, R., Early, Ch. E. (1997). *A pictorial history of psychology*. Chicago: Quintessence.
- Cassirer, E. (1977). *Esej o człowieku. Wstęp do filozofii kultury*. Warszawa: Czytelnik.
- Cassirer, E. (2000). Substanzbegriff und Funktionsbegriff. Untersuchungen über die Grundfragen der Erkenntniskritik W: B. Recki (red.), *Ernst Cassirer Gesammelte Werke: Hamburger Ausgabe*, t. 6. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Einstein, A. (1934/1979). *Mein Weltbild*. Zürich: Europa Verlag.
- Fahrenberg, J. (2015). *Theoretische Psychologie. Eine Systematik der Kontroverse*. Langerich: Pabst Science.
- Gołąb, A. (1984). Erinnerungen an Kurt Lewin. Ein Interview mit Bluma Zeigarnik. *Gruppendynamik. Zeitschrift für angewandte Sozialpsychologie*, 15(1), 103–110.
- Graumann, C.-F. (red.). (1981–1983). *Kurt-Lewin-Werkausgabe*, t. 1–4. Berno–Stuttgart: Hans Huber, Klett-Cotta.
- Graumann, C.-F. (1982). *Zur Einführung in diesem Band*. W: C.-F. Graumann (red.), *Kurt-Lewin-Werkausgabe*, t. 4 (s. 11–40). Berno–Stuttgart: Hans Huber, Klett-Cotta.
- Graumann, C.-F. (2007). *Lewin zu Beginn des 21. Jahrhunderts*. W: W. Schönplflug (red.), *Kurt Lewin – Person, Werk, Umfeld* (s. 283–296). Frankfurt nad Menem: Peter Lang.
- Heider, F. (1984). *Das Leben eines Psychologen*. Berno: Hans Huber.
- Kant, I. (2017). *Anthropologie in pragmatischer Hinsicht*. Ditzingen: Reclam.
- Kardaś, T. (2021). *Droga Kurta Lewina z Mogilna*. Poznań: FNCE.
- Karolczak, W. (2006). Żydzi stowarzyszeni. Życie organizacyjne gminy żydowskiej w Poznaniu na przełomie XIX i XX wieku. *Kronika Miasta Poznań. Poznańscy Żydzi*, 3, 223–242.
- Kozielecki, J. (red.). (2009). *Nowe idee w psychologii*. Gdańsk: GWP.
- Köhler, W. (1968). *Werte und Tatsachen* (M. Kofka, tłum.). Berlin: Springer.
- Lewin, K. (1917). *Kriegslandschaft*. *Zeitschrift für angewandte Psychologie*, 12, 440–447.
- Lewin, K. (1920). Die Sozialisierung des Taylorsystems. Eine grundsätzliche Untersuchung zur Arbeits- und Berufs-Psychologie. W: K. Korsch (red.), *Schriftenreihe Praktischer Sozialismus*, t. 4 (s. 3–36). Berlin-Fichtenau: Gesellschaft für Erziehung.
- Lewin, K. (1981a). *Cassirers Wissenschaftsphilosophie und die Sozialwissenschaften*. W: C.-F. Graumann (red.), *Kurt-Lewin-Werkausgabe*, t. 1 (s. 347–365). Berno–Stuttgart: Hans Huber, Klett-Cotta.
- Lewin, K. (1981b). Der Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie. W: C.-F. Graumann (red.), *Kurt-Lewin-Werkausgabe*, t. 1 (s. 233–277). Berno–Stuttgart: Hans Huber, Klett-Cotta.
- Lewin, K. (1981c). *Die Verwechslung von Wissenschaftssubjekt und psychischen Bewußtsein in ihrer Bedeutung für die Psychologie*. W: C.-F. Graumann (red.), *Kurt-Lewin-Werkausgabe*, t. 1 (s. 111–125). Berno–Stuttgart: Hans Huber, Klett-Cotta.
- Lewin, M. (1986). *Psychologische Forschung im Umriß*. Berlin: Springer.
- Lewin, T. (2006). The Kurt Lewin I remember. W: J. Trempala, A. Pepitone, B. H. Raven. (red.), *Lewinian psychology. Proceedings of the international conference „Kurt Lewin: contribution to contemporary psychology“ september10–12.2004, Bydgoszcz–Mogilno, Poland* (s. 291–297). Bydgoszcz: Wydawnictwo Uniwersytetu Kazimierza Wielkiego.
- Lück, H. E. (1996). *Die Feldtheorie und Kurt Lewin*. Weinheim: Psychologie Verlags Union.
- Lück, H. E. (1997). Kurt Lewin – Filmmaker. W: W. Bringmann, H. E. Lück, R. Miller, Ch. Early, C. E. (red.), *A pictorial history of psychology* (s. 282–287). Chicago: Quintessence.
- Lück, H. E. (2001). *Kurt Lewin. Eine Einführung in sein Werk*. Weinheim: Beltz.
- Lück, H. E. (2008). *Historia psychologii. Orientacje, szkoły, kierunki rozwoju*. Warszawa: Vizja Press&IT.

- Lück, H. E. (red.). (2009). *Kurt Lewin. Schriften zur angewandten Psychologie*. Wiedeń: Krammer.
- Lück, H. E. (2021). Kurt Lewin und Fritz Heider. Ihre Freundschaft, ihre Filme und ihre Theorien. W: J. McElvenny, A. Ploder (red.), *Holisms of communication. The early history of audio-visual sequence analysis* (s. 3–23). Berlin: Language Science Press.
- Lück, H. E., Miller, R., Sewz-Vosshenrich, G. (2008). *Klasycy psychologii*. Kraków: WAM.
- Marrow, A. J. (2002). *Kurt Lewin. Leben und Werk*. Weinheim: Beltz.
- Münsterberg, H. (1912a). *Die Amerikaner*. Berlin, E. S. Mittler & Sohn.
- Münsterberg, H. (1912b). *Psychologie und Wirtschaftsleben*. Lipsk: J. A. Barth.
- Münsterberg, H. (1915). *Amerika und der Weltkrieg*. Lipsk: J. A. Barth.
- Obuchowski, K. (2009). *Refleksje autobiograficzne psychologa*. Łódź: Wydawnictwo WSHE.
- Pankalla, A. (2011). *Mitocentryczna psychologia kulturowa Ernesta Boescha*. Poznań: Wydawnictwo Naukowe UAM.
- Popper K. (2006). *Zur Methodenfrage der Denkpsychologie*. W: T. E. Hansen (red.), *K. R. Popper. Gesammelte Werke in deutscher Sprache. T. 1. Frühe Schriften* (s. 187–260). Tübingen: Mohr Siebeck.
- Schönplflug, W. (2000). *Geschichte und Systematik der Psychologie*. Weinheim: Psychologie Verlags Union.
- Schönplflug, W. (2006). Science in context: Lewin in Germany, and Lewin in the United States of America. W: J. Trempała, A. Pepitone, B. H. Raven (red.), *Lewinian psychology. Proceedings of the international conference „Kurt Lewin: contribution to contemporary psychology” september 10–12.2004, Bydgoszcz–Mogilno, Poland* (s. 33–45). Bydgoszcz: Wydawnictwo Uniwersytetu Kazimierza Wielkiego.
- Schönplflug, W. (red.). (2007). *Kurt Lewin – Person, Werk, Umfeld*. Frankfurt nad Menem: Peter Lang.
- Schultz, D. S., Schultz, S. E. (2008). *Historia współczesnej psychologii*. Kraków: Wyd. UJ.
- Sprung, L., Schönplflug, W. (red.). (1992). *Zur Geschichte der Psychologie in Berlin*. Frankfurt nad Menem: Peter Lang.
- Stachowski, R. (2000). *Historia współczesnej myśli psychologicznej*. Warszawa: Scholar.
- Stachowski, R. (2006). Kurt Lewin's Prussian/Polish Mogilno episode. W: J. Trempała, A. Pepitone, B. H. Raven (red), *Lewinian psychology. Proceedings of the international conference „Kurt Lewin: contribution to contemporary psychology” september 10–12.2004, Bydgoszcz–Mogilno, Poland* (s. 23–32). Bydgoszcz: Wydawnictwo Uniwersytetu Kazimierza Wielkiego.
- Trempała, J., Pepitone, A., Raven, B. H. (red.). (2006). *Lewinian Psychology. Proceedings of the international conference „Kurt Lewin: contribution to contemporary psychology” september 10–12.2004, Bydgoszcz–Mogilno, Poland*. Bydgoszcz: Wydawnictwo Uniwersytetu Kazimierza Wielkiego.
- Witkowski, R. (2006). Lista wyborcza gminy poznańskiej z 1935 roku, do druku przygotował Rafał Witkowski. *Kronika Miasta Poznania. Poznańscy Żydzi*, 3, 275–289.
- Wundt, W. (1920). *Erlebtes und Erkanntes*. Stuttgart: A. Kröner.
- Zeidler, K. W. (2011). Neokantyci i psychologia. Przyczynek do metodologii Kurta Lewina. W: W. Zeidler, H. Lück (red.), *Psychologia europejska w okresie międzywojennym* (s. 291–300). Warszawa: Vizja Press&IT.
- Zeidler, W. (2009). „Krajobraz wojenny” Kurta Lewina: reportaż czy projekt nowej metodologii. W: W. Zeidler (red.), *Psychologia w okresie międzywojennym* (s. 159–178). Warszawa: Vizja Press&IT.
- Zeidler, W. (2011a). Krytyka taylorizmu jako kolejny krok w rozwoju metodologii Kurta Lewina. W: W. Zeidler, H. Lück (red.), *Psychologia europejska w okresie międzywojennym. Sylwetki, osiągnięcia, problemy* (s. 371–389). Warszawa: Vizja Press&IT.
- Zeidler, W. (2011b). Wolfgang Köhler–Hans Geitel: korespondencja z lat 1907–1920. *Studia Psychologica*, 11(2), 129–148.
- Zeidler, W. (2012). Wolfgang Köhler: die Verantwortung für eigene Identität. W: T. Herrmann, W. Zeidler (red.), *Psychologen in autoritären Systemen* (s. 248–269). Frankfurt nad Menem: Peter Lang.
- Zeidler, W. (2016). Czy polska psychologia ma swoją historię? Różne oblicza historii psychologii. *Studia Psychologica*, 16(1), 5–21. <https://doi.org/10.21697/sp.2016.16.1.01>
- Zeidler, W. (2018). *Psychologia i psychotechnika w twórczości Franciszki Baumgarten*. Warszawa: Liberi Libri.
- Zeidler, W., Lück, H. E. (2011). *Psychologia europejska w okresie międzywojennym. Sylwetki, osiągnięcia, problemy*. Warszawa: Vizja Press&IT.

KONTROWERSJE WOKÓŁ POJĘCIA TRAFNOŚCI

JOLANTA RYTEL* 

Instytut Psychologii, Uniwersytet Kardynała Stefana Wyszyńskiego

STRESZCZENIE

W artykule zrelacjonowano zmieniające się na przestrzeni ponad 100 lat koncepcje pojęcia trafności pomiaru testowego. Aktualnie pojęcie trafności odnosi się do stopnia, w jakim dane empiryczne oraz teoria uzasadniają interpretację wyników testowych w zakładanym kierunku (*American Educational Research Association, American Psychological Association, National Council on Measurement in Education [AERA, APA, NCME], 2007, s. 31*). Przedstawiono 5 podstawowych źródeł danych dotyczących trafności oraz problemy związane z integracją dowodów na rzecz trafności w spójną argumentację. Podkreślono użyteczność zaproponowanego przez Kane'a podejścia do walidacji opartego na argumentacji, odwołującego się do logiki nieformalnej i struktury argumentu wprowadzonej przez Toulmina. Omówiono także różnice stanowisk zajmowanych przez badaczy w odniesieniu do 2 podstawowych kwestii: czemu przysługuje trafność i jaki jest właściwy sposób jej ustalania?

TRAFNOŚĆ
POMIAR PSYCHOLOGICZNY
STANDARDY STOSOWANIA TESTÓW PSYCHOLOGICZNYCH

SŁOWA KLUCZOWE

- | | |
|----|---|
| 51 | MODEL TRAFNOŚCI OPARTY NA Kryterium |
| 52 | MODEL TRAFNOŚCI OPARTY NA TREŚCI |
| 52 | MODEL TRAFNOŚCI TEORETYCZNEJ |
| 53 | W KIERUNKU UJEDNOLICENIA POJĘCIA TRAFNOŚCI |
| 55 | USTALANIE TRAFNOŚCI JAKO PROCES INTEGRACJI DANYCH POCHODZĄCYCH Z RÓŻNYCH ŹRÓDEŁ |
| 57 | PROBLEMY ZE WSPÓŁCZESNĄ KONCEPCJĄ TRAFNOŚCI |
| 61 | PODSUMOWANIE |
| 61 | BIBLIOGRAFIA |



CONTROVERSIES OVER THE CONCEPT OF VALIDITY

ABSTRACT

In the article the key changes which have occurred in conceptualization of validity are considered. Validity is currently defined in term of the degree to which a proposed interpretation of test scores is justified by evidence and theory (American Educational Research Association, American Psychological Association, National Council on Measurement in Education [AERA, APA, NCME], 2007, p. 31). Five types of validity evidence are described, and problems with the integrations of various strands of evidence in sound validity argument are discussed. Usefulness of Kane's argument-based approach to validity referring to informal logic and the structure of the argument introduced by Toulmin is stressed in the article. There is also emphasized the lack of uncontroversial definition of validity.

KEYWORDS

validity, psychological measurement, testing standards

Przez ostatnie ponad sto lat poglądy na temat trafności zmieniały się kilkakrotnie, jednakże stanowisko, że sama trafność wyróżnia się wśród innych pojęć psychometrycznych, pozostało niezmiennie. Trafność od zawsze uważano w psychometrii za własność najbardziej fundamentalną i najważniejszą (Angoff, 1988). Zgodnie z definicją podaną w *Standardach dla testów stosowanych w psychologii i pedagogice* „pojęcie trafności odnosi się do stopnia, w jakim dane empiryczne oraz teoria uzasadniają interpretację wyników testowych w zakładanym kierunku. Trafność jest zatem najbardziej podstawową kategorią w procesie tworzenia i oceny testu” (American Educational Research Association, American Psychological Association, National Council on Measurement in Education [AERA, APA, NCME], 2007, s. 31). Podstawowa rola pojęcia trafności jest niekwestionowana, jednakże określenie, czym jest trafność, nadal wzbudza kontrowersje.

MODEL TRAFNOŚCI OPARTY NA KRYTERIUM

Okolo 1915 roku, choć nie istniała jeszcze formalna definicja trafności, zaczął się kształtować model, który Kane (2001) określa jako *model trafności oparty na kryterium*. Z pewnością wpływ na to miało opublikowanie przez Pearsona w roku 1896 pracy dotyczącej współczynnika korelacji. Był to ekscytujący nowy wskaźnik statystyczny, który można było zastosować bezpośrednio, rozwiązując kwestię tego, jak dobrze wyniki poddanych ocenie osób wiążą się z innymi przejawami testowanej własności (Sireci, 2009). Pogląd, że test jest trafny dla wszystkiego, z czym koreluje (Guilford, 1946), utrzymywał się do połowy XX wieku. Choć większość badań nad trafnością prowadzonych w latach 30. i 40. XX wieku koncentrowała się na ustalaniu trafności prognostycznej, pojawił się też inny jej rodzaj – trafność diagnostyczna. Pojęcie trafności diagnostycznej związane jest ze stopniem, w jakim wyniki testowe umożliwiają ocenę kryterium (pewnej zmiennej pozatestowej, której dotyczy wnioskowanie na podstawie wyników testu), a oba pomiary: pomiar testowy i pomiar kryterium przeprowadzane są w zbliżonym bądź w tym samym czasie, a więc ocena dotyczy stanu aktualnego. W przypadku trafności prognostycznej pomiar kryterium dokonywany jest w czasie późniejszym niż pomiar testowy, zatem ocena odnosi się do możliwości przewidywania stanu przyszłego. Testy były wykorzystywane przede wszystkim do celów selekcji i klasyfikacji. Nie dziwi zatem, że te dwie formy trafności odzwierciedlały cel, w jakim testy były powszechnie w tym czasie przeprowadzane. Celem konstrukcji testu było przewidywanie pewnych zewnętrznych zachowań (Lissitz, Samuelson, 2007). Do ewolucji poglądów na temat trafności niewątpliwie przyczyniło się również opracowanie przez Spearmana metody analizy czynnikowej. Zwolennikiem wykorzystania tej metody do szacowania trafności był Guilford (1946), postulujący dwa rodzaje trafności: praktyczną – odnoszącą się do korelacji między wynikami testu i miarami kryterium – oraz czynnikową.

Trafność czynnikową testu charakteryzują jego ładunki na znaczeniowo zinterpretowanych wspólnych, istotnych czynnikach. To jest ten rodzaj trafności, o który rzeczywiście chodzi, gdy pada pytanie: Czy test mierzy to, do mierzenia czego został opracowany? Stosowniejszym pytaniem winno być: Co mierzy test? Przewiduję, że przyjdzie czas, gdy od każdego autora będzie się oczekiwać podania informacji na temat kompozycji czynnikowej jego testu (Guilford, 1946, s. 428 – tłum. J. R.).

MODEL TRAFNOŚCI OPARTY NA TREŚCI

Kłopoty z modelem trafności opartym na kryterium dotyczą konieczności odwołania się do dobrze zdefiniowanego i trafnego kryterium, które nie zawsze jest łatwo osiągalne. Pojawienie się testów osiągnięć, dla których nie było lepszego kryterium niż sam test, przyczyniło się do powstania kolejnego modelu trafności – modelu opartego na treści (Kane, 2008). Trafność treściowa jest wbudowana w test od samego początku przez wybór odpowiednich pozycji (Anastasi, Urbina, 1999). Pojęcie trafności treściowej dotyczy oceny stopnia, w jakim treść pozycji testowych reprezentuje dziedzinę będącą przedmiotem pomiaru. Spotkało się ono z krytyką (por. Kane, 2006a; Sireci, 1998), która wskazywała na subiektywizm ocen wydawanych na podstawie analizowania treści pozycji testowych, a także na możliwość obciążenia tych ocen silną tendencją do potwierdzania, jako że analiza treści pozycji zwykle przeprowadzana jest w trakcie tworzenia narzędzia bądź wkrótce po jego opracowaniu przez osoby zaangażowane w ten proces. Ponadto – co podkreślał Messick (1989) – ocena trafności treściowej, jako niezwiązana z wynikami testu, nie może stanowić uzasadnienia dla interpretacji wyników testu. Guion (1978) wręcz odrzucał pojęcie trafności treściowej, przestrzegając, że rozważności przy konstruowaniu testu czy wyborze próbek z uniwersum pozycji nie powinno się mylić z trafnością treściową. Warto zauważyć, że problem użyteczności szacowania trafności treściowej nadal budzi kontrowersje wśród badaczy, o czym świadczy dyskusja, jaka odbyła się na łamach pisma *Industrial and Organizational Psychology*, zapoczątkowana artykułem Murphy'ego (2009) zatytułowanym prowokacyjnie: *Content validation is useful for many things, but validity isn't one of them*.

MODEL TRAFNOŚCI TEORETYCZNEJ

Trafność teoretyczna badana jest zwykle, gdy testujący nie dysponuje żadną dokładnie określoną miarą kryterium cechy, którą jest zainteresowany, i musi wykorzystać miary pośrednie. W takim przypadku właśnie cecha czy własność leżąca u podstaw testu ma zasadnicze znaczenie, a nie zachowania, które angażuje wykonanie testu czy wyniki otrzymane w zakresie kryteriów. (Cronbach, Meehl, 1955, s. 282 – tłum. J. R.)

Cytat ten, zamieszczony w przełomowym artykule autorstwa Cronbacha i Meehla (1955), zatytułowanym *Construct validity in psychological test*, pochodzi z dokumentu *Technical recommendations for psychological tests and diagnostic techniques*. W dokumencie tym opublikowano wnioski uzgodnione przez Komitet Testów Psychologicznych Amerykańskiego Towarzystwa Psychologicznego, który w latach 1950–1954 starał się określić, jakie właściwości testu winny być przedmiotem analizy, zanim zostanie on opublikowany. Cronbach i Meehl (1955), wyjaśniając pojęcie trafności teoretycznej, stwierdzają: „Z określeniem trafności teoretycznej mamy do czynienia, ilekroć test ma być interpretowany jako miara pewnego atrybutu lub własności, która nie jest «zdefiniowana operacyjnie». Problem, przed którym stoi badacz, można sformułować następująco: Jakie konstrukty wyjaśniają wariację poziomu wykonania testu?» (s. 282 – tłum. J. R.).

W niedługim czasie Campbell i Fiske (1959) opracowali metodę badania trafności teoretycznej za pomocą macierzy wielu cech-wielu metod. Metoda ta dotyczyła przede wszystkim adekwatności testów jako miar konstruktów, a nie adekwatności konstruktów określanej przez potwierdzenie przewidywanych teoretycznie związków z miarami innych konstruktów. „Sądzymy, że przed przystąpieniem do testowania związków między daną cechą

a innymi cechami należy nabrać zaufania do miary tej cechy. To zaufanie mogą podtrzymać dowody trafności zbieżnej i różnicowej” (Campbell, Fiske, 1959, s. 100 – tłum. J. R.).

Osadzając znaczenie konstruktów w sieci nomologicznej, Cronbach i Meehl (1955) wskazują na konieczność sformułowania teorii danej cechy: „Dowiedzenie się więcej” o konstrukcie teoretycznym jest ich zdaniem kwestią opracowania sieci nomologicznej, w której konstrukt występuje, bądź zwiększenia precyzji definiowania jego składników. Wskazują tym samym na konieczność zdefiniowania konstruktów przed przystąpieniem do ustalania trafności testu. Stwierdzając natomiast, że badanie trafności teoretycznej testu nie różni się zasadniczo od ogólnych naukowych procedur tworzenia i potwierdzania teorii, autorzy akcentują, że metodologia ustalania trafności nie może sprowadzać się do podania wartości jednego współczynnika (Cronbach, Meehl, 1955).

Trafność ujmowana z perspektywy teorii przestaje być trafnością związaną z samym testem, a staje się trafnością dotyczącą inferencji na temat znaczenia wyników testu czy też trafnością dotyczącą nadawanej im przez badacza interpretacji. „Nie ustala się trafności testu, lecz trafność interpretacji danych uzyskanych za pomocą specyficznej procedury” (Cronbach, 1971, s. 447 – tłum. J. R.), a ustalanie trafności odnosi się do wszystkich interpretacji wyników testowych. Różne zastosowania danego testu wiążą się z innymi interpretacjami jego wyników, a każda z tych interpretacji wymaga uzasadnienia. „Ponieważ każdą interpretację charakteryzuje właściwy dla niej stopień trafności, nigdy nie da się wyprowadzić prostego wniosku, że dany test «jest trafny»” (Cronbach, 1971, s. 447 – tłum. J. R.).

Trafność teoretyczna stanowiła kolejny typ trafności, jednakże Cronbach i Meehl (1995) wyrażali przekonanie, że określenie, które konstrukty psychologiczne wyjaśniają wykonanie testu, jest pożądane dla niemal każdego testu. Podkreślali też, iż z trafnością teoretyczną wiąże się wiele typów dowodów, łącznie z: trafnością treściową, korelacjami między pozycjami, korelacjami między testami, korelacjami między testem i «kryterium», badaniami stabilności w czasie oraz stabilności w różnych warunkach eksperymentalnych (Cronbach, Meehl, 1995). Tym samym pojęcie trafności teoretycznej stało się najlepszym kandydatem do objęcia nim wszystkich dotychczasowych rodzajów trafności.

W KIERUNKU UJEDNOLICENIA POJĘCIA TRAFNOŚCI

To, że pojęcie trafności teoretycznej stanowi doskonałą podstawę do ujednoczenia wszystkich rodzajów trafności, zostało rozpoznane przez badaczy wcześniej. Już w 1957 roku Loevinger twierdziła, że trafność teoretyczna stanowi całościowe ujęcie trafności z naukowego punktu widzenia. Loevinger (1957) wyróżniła trzy aspekty trafności teoretycznej powiązane z trzema stadiami procesu konstrukcji testu. Aspekt esencjalny związany z tworzeniem puli pozycji testowych dotyczy zakresu, w jakim treść pozycji można wyjaśnić w terminach mierzonej cechy i w kontekście teorii psychologicznej. Aspekt strukturalny wiąże się z analizą struktury wewnętrznej puli pozycji i ich selekcją w celu opracowania procedury obliczania wyników, odnosząc się do odpowiedniości czy też – używając terminu Loevinger – wierności modelu strukturalnego w stosunku do charakterystyk strukturalnych pozatestowych manifestacji cechy i samej struktury utworzonej przez pozycje testowe. Aspekt zewnętrzny obejmuje analizę związków wyników testowych z miarami zachowań pozatestowych i wynikami innych testów.

Trzy dekady po rozważaniach Loevinger podobny pogląd wyraził sam Cronbach (1988): „licząca 30 lat idea trzech typów trafności, rozłącznych, choć może sobie równych, to idea, której czas minął. Większość teoretyków trafności uznała, że trafność treściowa i kryterialna nie są niczym więcej niż żyłami w kablu argumentacji na rzecz trafności” (s. 4 – tłum. J. R.). Postulując ujmowanie procesu walidacji jako konstruowania argumentacji

na rzecz trafności, Cronbach podkreślał, że argumentacja taka kierowana jest do różnorodnego i potencjalnie krytycznego audytorium i dlatego „musi łączyć pojęcia, dowody empiryczne, konsekwencje społeczne i indywidualne oraz wartości” (s. 4 – tłum. J. R.).

Wymiar konsekwencji społecznych związanych z interpretacją wyników testowych i ich wykorzystaniem został wprowadzony przez Messicka (1989; 1995) jako jeden z aspektów trafności teoretycznej stanowiącej podstawę ujednoczonego modelu ujmowania trafności. Na pozostałe aspekty trafności teoretycznej składają się wskazane już przez Loevinger (1957): esencjalny, strukturalny i zewnętrzny oraz aspekt treściowy, obejmujący dowody na rzecz treściowej ważności, reprezentatywności i jakości technicznej, a także aspekt ogólności, odnoszący się do stopnia generalizowalności wyników testowych i ich interpretacji. Konsekwencje mogą być pozytywne bądź negatywne, a zgromadzenie dowodów na rzecz obu tych rodzajów jest równie istotne. Szczególnie ważne – zwłaszcza dla konsekwencji potencjalnie niekorzystnych – jest ustalenie, czy nie wynikają one z niedoreprezentowania konstruktów lub wariancji niezwiązanej z konstruktem. W pierwszym przypadku test nie obejmuje ważnych wymiarów konstruktów, tym samym interpretacja wyników zostaje zawężona, w drugim zaś na wyniki testowe wpływają czynniki niezwiązane z konstruktem, a powiązane z innymi konstrukcjami, samym testem lub warunkami badania. W takim ujęciu pojęcie trafności dotyczy nie tylko interpretacji wyników testowych, lecz także pożądaných i niepożądanych konsekwencji testowania: „ustalenie trafności jest empiryczną oceną znaczenia i konsekwencji pomiaru. Termin *ocena empiryczna* został użyty dla podkreślenia, że proces walidacji ma zarówno naukowy, jak i retoryczny charakter, wymaga zatem i dowodów, i argumentów” (Messick, 1995, s. 747 – tłum. J. R.).

Jednolite ujęcie trafności przez Messicka (1989) stało się podstawą dla współczesnego jej pojmowania: „trafność to całościowy osąd stopnia, w jakim dowody empiryczne i racjonalne uzasadnienia natury teoretycznej potwierdzają adekwatność i poprawność interpretacji i działań podejmowanych na podstawie wyników testu lub innych sposobów oceny”. Messick (1995) dobitnie podkreśla, że:

(...) trafność nie jest własnością testu lub oceny jako takiej, a stanowi ona znaczenie wyników testowych. Wyniki te są funkcją nie tylko pozycji testowych lub warunków bodźcowych, ale również odpowiadających na nie osób badanych i kontekstu przeprowadzania oceny. Tym, co powinno być trafne, jest znaczenie lub interpretacja wyniku, podobnie jak każda implikacja dla działań, jaką to znaczenie za sobą pociąga (Cronbach, 1971). Problem zasięgu, w jakim znaczenie wyniku i implikacje działań obejmują osoby oraz układy warunków lub konteksty, stanowi zawsze aktualne i odwieczne pytanie empiryczne. Trafność jest własnością ewoluującą, a proces jej ustalania ma charakter ciągły przede wszystkim z tego właśnie powodu. (s. 741 – tłum. J. R.).

Główne zmiany, jakie zaszły w rozumieniu pojęcia trafności, można podsumować w następujący sposób:

1. Rozumienie trafności zmieniło się od praktycznego, empirycznego sposobu jej ujmowania, a tym samym wprowadzania różnych rodzajów trafności, do osadzenia pojęcia trafności w teorii, dzięki czemu stało się możliwe jednolite jej ujęcie. Podstaw do zintegrowania różnych rodzajów trafności dostarczyło pojęcie trafności teoretycznej. Współcześnie trafność traktowana jest jako pojęcie jednolite (trafność teoretyczna), mające kilka aspektów. „Dane pochodzące z poszczególnych źródeł mogą potwierdzać różne aspekty trafności, ale nie reprezentują różnych rodzajów trafności. Trafność jest pojęciem spójnym” (AERA, APA, NCME, 2007, s. 35).

2. Trafność nie jest własnością testu. Jest trafnością uzasadnienia proponowanej interpretacji i wykorzystania wyników testowych dostarczanego z punktu widzenia teorii leżącej u podstaw testu oraz zgromadzonych dowodów empirycznych. Tym samym twierdzenia

na temat trafności nie mają charakteru dychotomicznego (trafny vs nietrafny), ale ujmowane są jako kontinuum. Trafność to „stopień, w jakim wszystkie kumulujące się dane potwierdzają zamierzoną interpretację wyników testowych” (AERA, APA, NCME, 2007, s. 35). Zmianę tę doskonale ujmuje stwierdzenie Messicka (1980): „trafność jest oceną dowodów, osądem, a nie samoistnym bytem” (s. 1020).

3. Trafność powinna być ustalona dla każdego sposobu interpretacji i wykorzystania wyników testowych.

Twierdzenia dotyczące trafności powinny odnosić się do konkretnych sposobów interpretacji czy zastosowania wyników testowych. Posługiwanie się sformułowaniem: *trafność wyników testowych* nie jest poprawne. Żaden test nie jest trafny w wypadku wszystkich celów i w każdej sytuacji. Wszystkie rekomendowane zastosowania czy interpretacje testu wymagają potwierdzenia trafności (AERA, APA, NCME, 2007, s. 45).

4. Ustalanie trafności wymaga uwzględnienia danych pochodzących z różnych źródeł. „Obejmuje to dane wynikające z analizy procesu tworzenia testu, dane dotyczące adekwatnej rzetelności testu, właściwej procedury badania testem i procedury obliczania wyników testowych, procedury skalowania i wyrównywania wyników oraz analizy testu z punktu widzenia potencjalnej stronniczości testu (...)” (AERA, APA, NCME, 2007, s. 44). Dane te służą jako podstawa zbudowania wiarygodnej naukowo argumentacji na rzecz trafności przyjmowanej interpretacji wyników.

5. Ustalanie trafności nie jest jednorazowym przedsięwzięciem, stanowi ciągły proces. Wymaga ono starannego zaplanowania i zrealizowania programu badawczego, a nie przeprowadzenia pojedynczego badania. „Proces walidacji obejmuje ciągłe zbieranie danych w celu dostarczenia mocnych podstaw naukowych proponowanej interpretacji wyników testowych” (AERA, APA, NCME, 2007, s. 31).

USTALANIE TRAFNOŚCI JAKO PROCES INTEGRACJI DANYCH POCHODZĄCYCH Z RÓŻNYCH ŹRÓDEŁ

W *Standardach dla testów stosowanych w psychologii i pedagogice* (AERA, APA, NCME, 2007) wymieniono pięć podstawowych źródeł danych dotyczących trafności.

1. Zgromadzenie danych opartych na treści testu (włączając poruszane tematy, sformułowania słowne, format pozycji testowych, zadań lub pytań tworzących test oraz procedury badania i obliczania wyników) umożliwia ocenę stopnia, w jakim treść testu reprezentuje dziedzinę będącą przedmiotem pomiaru. Dane te mogą pochodzić z analiz logicznych czy empirycznych dotyczących stopnia reprezentatywności treści testu oraz stopnia, w jakim owa treść wiąże się z proponowaną interpretacją wyników. Ważnym źródłem tych danych może być przeprowadzona przez ekspertów ocena stopnia, w jakim pozycje czy części testu odpowiadają definicji mierzonego konstruktowi. Oceny ekspertów mogą być wykorzystane do określenia stopnia adekwatności, ważności i precyzji sformułowania pozycji testowych oraz zakresu, w jakim niedoreprezentowanie konstruktowi lub włączenie niepowiązanych z nim treści może dostarczać nieuczciwej przewagi jakiejś grupie badanych osób.

2. Informacji o stopniu zgodności między konstruktowi a zachowaniami czy odpowiedziami pojawiającymi się w procesie rozwiązywania testu mogą dostarczyć dane oparte na analizie procesu udzielania odpowiedzi. Dane te można uzyskać, śledząc proces konstruowania odpowiedzi czy przeprowadzając wywiad z osobami badanymi i analizując odpowiedzi indywidualne na pytania dotyczące wykorzystywanych strategii, zachowań czy zasad odpowiadania na konkretne pozycje testu. Może ich też dostarczyć analiza procesu rozwiązywania testu uwzględniająca podobieństwa i różnice w sposobach odpowiadania, przeprowadzona w różnych grupach. Analiza taka przyczynia się do określenia zakresu,

w jakim umiejętności mające niewielkie znaczenie lub nieistotne z punktu widzenia konstruktów wpływają na proces odpowiadania. Proces walidacji może obejmować także badanie sposobów gromadzenia, rejestrowania i interpretowania danych przez ekspertów, obserwatorów i osoby przeprowadzające wywiady. W takim przypadku istotna jest analiza poprawności tych procesów i ustalenie zgodności przyjmowanych kryteriów bądź wydawanych ocen z planowaną interpretacją wyników testowych.

3. Dane wynikające z analizy struktury wewnętrznej testu pozwalają na ustalenie stopnia, w jakim relacje między pozycjami czy składowymi testu odpowiadają zdefiniowanemu konstruktowi. Źródło takich danych może stanowić analiza związków między pozycjami testu oraz badanie jego struktury czynnikowej (w szczególności za pomocą konfirmacyjnej analizy czynnikowej). Analiza pozycji testowych może ujawnić te z nich, które funkcjonują odmiennie w różnych (np. ze względu na płeć lub warunki środowiskowe czy kulturowe) grupach badanych, nieróżniących się jednak ze względu na mierzoną wielkość. Zróżnicowane funkcjonowanie pozycji (*differential item functioning* – DIF) związane jest z wystąpieniem nieoczekiwanej różnicy między grupami osób badanych, które przypuszczalnie powinny być porównywalne co do atrybutu mierzonego przez pozycję i test, w którym ona występuje (Dorans, Holland, 1993). Taka pozycja może mierzyć coś innego niż pozostała część testu (np. składniki niezwiązane z konstruktem) lub też może mierzyć to samo, ale z różnym stopniem precyzji w różnych grupach. Doskonały przegląd zarówno podejść do problemu zróżnicowanego funkcjonowania pozycji, jak i metod służących identyfikowaniu tego zjawiska zawiera artykuł Zumbo (2007).

4. Kolejne źródło danych walidacyjnych stanowią dane oparte na analizie związków z innymi zmiennymi. Interpretacja takich danych umożliwia określenie rodzaju i zakresu zależności między wynikami testu a innymi zmiennymi zewnętrznymi oraz ustalenie stopnia, w jakim te zależności są spójne z konstruktem leżącym u podstaw proponowanej interpretacji wyników. W skład tych danych wchodzi dane walidacyjne zbieżne, dotyczące związków (silnych) między wynikami testu a innymi miarami odzwierciedlającymi podobne konstrukty oraz dane walidacyjne o charakterze rozbieżnym, odnoszące się do związków (słabych bądź ich braku) z miarami reprezentującymi inne konstrukty. Dane te mogą pochodzić z analizy związków między testem a kryterium, mierzonym w zbliżonym bądź tym samym czasie (dane diagnostyczne) lub też mierzonym w czasie późniejszym niż pomiar testowy (dane prognostyczne). Badania nad efektywnością decyzji selekcyjnych, klasyfikacyjnych czy lokacyjnych także stanowią źródło takich danych. Mogą ich dostarczyć badania różnic międzygrupowych, w których sprawdzana jest, sformułowana na podstawie teorii mierzonej wielkości, hipoteza dotycząca wystąpienia oczekiwanych różnic w wynikach testu między badanymi grupami lub badania eksperymentalne, których celem jest wykazanie wystąpienia zmiany w wynikach testu, przewidywanej w następstwie wprowadzonego oddziaływania eksperymentalnego.

Rezultaty badań wykorzystujących technikę metaanalizy mogą posłużyć do uogólnienia danych dotyczących trafności. Zanim pojawiła się przełomowa praca Schmidta i Huntera (1977) wprowadzająca metaanalizę do badań nad generalizacją trafności, sądzono, że bardzo duże zróżnicowanie współczynników trafności testów uzdolnień stosowanych do selekcji personelu, dla których kryterium stanowił poziom wykonania pracy, związane jest ze specyfiką sytuacji badania lub specyfiką danej pracy. Przekonanie o *sytuacyjnej specyficzności* (Schmidt, Hunter, 1998) powodowało konieczność każdorazowego ustalania trafności w nowej sytuacji wykorzystania testów, poważnie ograniczając ich użyteczność. Okazało się, że większość tych różnic jest artefaktem wynikającym z błędu próby, zazwyczaj niewielkich rozmiarów (od 40 do 70 osób). To, na ile dane walidacyjne uzyskane w konkretnej sytuacji badawczej mogą być uogólnione na nowe sytuacje, stanowi bardzo ważny problem, zwłaszcza w obszarze edukacji i pracy. Warto zaznaczyć, że nie zawsze jest możliwe zgromadzenie

danych z przeprowadzonych już badań walidacyjnych adekwatnych do sytuacji, na którą mają być uogólniane. Wtedy przeprowadzenie badań lokalnych jest nieodzwonne.

5. Ostatni rodzaj danych to dane oparte na konsekwencjach testowania. Jeżeli zamierzony kierunek interpretacji wyników testu wiąże się z odniesieniem określonych korzyści związanych z jego stosowaniem (np. ulepszenie programów nauczania, efektywności procesu rekrutacji pracowników, skuteczności programów terapeutycznych), to w badaniach walidacyjnych należy wykazać, że dane korzyści uda się uzyskać. Dane dotyczące konsekwencji testowania, zarówno pożądaných, jak i niepożądanych, winny być odróżniane od danych, które mogą wpłynąć na decyzje dotyczące polityki społecznej, jednakże z trafnością niezwiązane. Jest to szczególnie ważne, gdy badanie różnych grup osób ma określone konsekwencje społeczne. I tak np. gdy w przypadku zastosowania testu jako narzędzia selekcyjnego na gruncie zawodowym czy edukacyjnym otrzymana różnica wyników między grupami stanowi odzwierciedlenie zróżnicowanego rozkładu umiejętności, które test ma mierzyć, to nie świadczy ona o braku trafności wnioskowania. Może jednak wynikać z wrażliwości testu na niektóre cechy osób badanych nienależące do zakresu konstruktów bądź z jego niedoreprezentowania, wtedy bezpośrednio wiąże się z trafnością. Jak podkreślono w *Standardach dla testów...* (AERA, APA, NCME, 2007):

Chociaż informacje na temat konsekwencji testowania mogą wpłynąć na decyzje o zastosowaniu testu, to same konsekwencje jako takie nie mają wpływu na trafność zamierzonej interpretacji wyników. Sądy dotyczące trafności lub jej braku, formułowane w świetle konsekwencji testowania, powinny zależeć raczej od racjonalnego wyjaśnienia ich źródeł. (s. 42)

Żaden rodzaj danych nie jest preferowany na tle innych, a zadaniem badacza jest synteza informacji pochodzących z różnych źródeł i przedstawienie ich w postaci spójnej argumentacji na rzecz zamierzonej interpretacji wyników testu i sposobów jego wykorzystania. Problem polega na tym, w jaki sposób to zrobić.

PROBLEMY ZE WSPÓŁCZESNĄ KONCEPCJĄ TRAFNOŚCI

Najbardziej kontrowersyjnym z wymienionych wyżej źródeł danych walidacyjnych jest ostatnie z nich. Dyskusje nad zasadnością włączania danych opartych na konsekwencjach testowania w proces walidacyjny prowadzone są od dawna. Już w 1997 roku poświęcono tej problematyce cały numer czasopisma *Educational Measurement: Issues and Practice* (t. 16, nr 2), podobnie stało się w roku kolejnym (t. 17, nr 2). Od tego czasu, podobnie jak od czasu opublikowania w roku 1999 przełożonego na język polski (w roku 2007) wydania *Standardów dla testów...*, upłynęło ponad 20 lat, warto zatem przyjrzeć się roli, jaką w badaniach walidacyjnych odgrywa ten aspekt trafności.

Cizek, Rosenberg i Koons (2008) dokonali przeglądu informacji na temat aspektów trafności relacjonowanych w odniesieniu do 283 testów psychologicznych i pedagogicznych opublikowanych w *Mental Measurements Yearbook* (Spies, Plake, 2005). Okazało się, że informacje dotyczące trafności rzadko były podawane w terminach zgodnych z jej współczesnym ujęciem. Podczas gdy dane dotyczące trafności teoretycznej (58% wszystkich testów), diagnostycznej (50,9%) i treściowej (48,4%) były zamieszczane stosunkowo często, to informacje na temat aspektu trafności związanego z konsekwencjami testowania pojawiły się jedynie w przypadku dwóch testów (0,7%). W kolejnym badaniu autorzy (Cizek, Bowen, Church, 2010) objęli analizą opublikowane w latach 1999–2008 numery ośmiu wiodących czasopism (takich jak *Educational and Psychological Measurement* czy *Practical Assessment, Research & Evaluation*). Spośród 2408 artykułów problematyki trafności dotyczyło 1007

(41,8%), i aspekt trafności związane z konsekwencjami nie został wymieniony w żadnym z nich. Taki sam rezultat uzyskano, analizując treść wystąpień prezentowanych na corocznych, odbywających się w latach 2006–2008, spotkaniach trzech organizacji sponsorujących wydawanie *Standardów...*: American Educational Research Association, American Psychological Association i National Council on Measurement in Education. Autorzy konkludują (Cizek, Rosenberg, Koons, 2008), że aspekt trafności dotyczący konsekwencji testowania nie jest przez badaczy uwzględniany, ponieważ aspektu tego po prostu nie da się w logiczny sposób połączyć z pozostałymi. Tym samym stanowi on usterkę we współczesnej teorii trafności (Cizek i in., 2008, s. 410).

Ustalanie trafności proponowanej interpretacji wyników testu i uzasadnianie sposobów jego wykorzystania obejmujące związane z tym konsekwencje stanowią zadania, które należy od siebie oddzielić, choć niewątpliwie zgromadzenie przekonujących dowodów na rzecz trafności interpretacji wyników testu stanowi warunek konieczny, ale niewystarczający wykorzystania testu. Nie oznacza to umniejszania wagi konsekwencji testowania – wręcz przeciwnie. W związku z tym, że aspekt ten jest niemal całkowicie ignorowany w praktyce walidacyjnej, pojawia się potrzeba opracowania wskazań dotyczących gromadzenia i ewaluacji danych uzasadniających różne sposoby wykorzystania testów (Cizek, 2012, 2016, 2020). Ponadto autorzy podkreślają, że w dokonany przez nich przeglądzie nie udało im się znaleźć ani jednego przypadku, w którym dane pochodzące z różnych źródeł zostałyby zintegrowane w spójną argumentację na rzecz (lub przeciw) trafności.

Problem z integracją dowodów na rzecz trafności w spójną argumentację sprowadza się do odpowiedzi na pytanie, czym jest argument na rzecz trafności i w jaki sposób powinien być konstruowany. Z problemem tym badacz musi się zmierzyć niezależnie od tego, czy integruje wszystkie rodzaje dowodów czy też tylko cztery (pierwsze) z nich, zważywszy na odmienny status logiczny danych opartych na konsekwencjach testowania. Użyteczne może się okazać rozwijane przez Kane'a (1992, 2004, 2006b, 2013, 2016) podejście do walidacji oparte na argumentacji (*argument-based validation*). Kane w swoich pracach odwołuje się do logiki nieformalnej i struktury argumentu wprowadzonej przez Toulmina (1958). Na strukturę argumentu składają się trzy podstawowe elementy: wynikająca z jakiejś przesłanki (*datum*) teza (*claim*) i uzasadniająca łącząca je relację gwarancja (*warrant*). Teza nie musi być wyrażona z całkowitą pewnością, może być poprzedzona jakimś wyrażeniem modalnym, np. *prawdopodobnie*, pełniącym funkcję modyfikatora (*qualifier*). Gwarancja, do której się odwołujemy, może wymagać wskazania, w jakich przypadkach stanowiących odstępstwa od reguły tezy nie da się utrzymać, czyli wymagać de facto wprowadzenia ograniczenia lub obalenia (*rebuttal*). Sama gwarancja może też wymagać uzasadnienia, określanego jako wsparcie dla niej (*backing*). Argumentacją posługujemy się w celu poparcia stwierdzeń, co do których nie mamy pewności, i czynimy to, odwołując się do stwierdzeń, względem których pewność tę posiadamy. Oceniając poprawność argumentacji, rozważamy akceptowalność wszystkich wykorzystanych przesłanek oraz to, czy stanowią one dostateczne uzasadnienie dla proponowanej tezy. W przypadku oceny poprawności argumentacji budowanej zgodnie z modelem Toulmina kluczową rolę odgrywa ocena poprawności stwierdzeń występujących w funkcji gwarancji, a tym samym – także twierdzeń funkcjonujących jako jej poparcie. W tym sensie zgodnie z poglądem Toulmina (1958) poprawność argumentacji jest zrelatywizowana do dziedziny, której dotyczy.

Proponowane przez Kane'a podejście obejmuje dwa etapy i odnosi się do dwóch rodzajów argumentów. Najpierw konstruowany jest argument dotyczący interpretacji (*interpretive argument*), obejmujący każdą interpretację i wykorzystanie wyników testu przedstawione w formie wniosku prowadzącego od wyników testu do nadawanej im interpretacji / sposobu ich wykorzystania oraz uzasadniającej je gwarancji. I tak np. gdy za pomocą równania regresji na podstawie wyników testu przewidujemy średnią ocen, to wyniki testu są przesłanką dla tezy, którą jest przewidywana średnia ocen, gwarancję natomiast stanowi

równanie regresji. Na drugim etapie budowany jest argument na rzecz trafności (*validity argument*), dzięki któremu możliwe jest wykazanie trafności proponowanych interpretacji i sposobów wykorzystania wyników testu przez dostarczenie stosownego wsparcia dla gwarancji, na których opierają się wnioski. Odwołując się do powyższego przykładu: równanie regresji pełniące funkcję gwarancji wymaga odpowiedniego wsparcia. Takiego wsparcia dla korzystania z niego w celu przewidywania średniej ocen dostarczają przeprowadzone badania empiryczne, w których równanie zostało otrzymane. W przypadku wykazania, że któreś z wniosków jest nieuprawnione, argument na rzecz trafności należy odrzucić.

Jak podkreśla Kane (2009):

...konstruując argument dotyczący interpretacji, mamy do czynienia z istotnym 'coś za coś'. Wykazanie trafności prostej interpretacji wymaga stosunkowo skromnej inwestycji w zgromadzenie danych na rzecz trafności. Interpretacja ambitniejsza, związana z bardziej rozbudowaną strukturą wniosków i wspierających je założeń, wymaga więcej wsparcia, by osiągnąć porównywalny poziom trafności, ale ambitniejsza interpretacja może być bardziej użyteczna, jeśli zostanie uzasadniona. Jeżeli nie twierdzimy, że test będzie przewidywał przyszłe zachowanie, to nie jesteśmy zobowiązani do zgromadzenia danych odnośnie do trafności prognozy, jednakże powstały argument na rzecz trafności nie będzie uzasadniał takich przewidywań. (s. 48–49 – tłum. J. R.)

Chapelle, Enright i Jamieson (2010) przedstawiają strategię konstruowania argumentu na rzecz trafności zrewidowanej wersji TOEFL® (Test of English as a Foreign Language). Autorki jako jedne z pierwszych wykorzystują podejście Kane'a do syntezy danych walidacyjnych. Chapelle i in. (2010) oceniają różnice między podejściem przedstawionym w *Standardach dla testów...* (AERA, APA, NCME, 2007) i tym proponowanym przez Kane'a odnośnie do sposobu formułowania zamierzonej interpretacji wyników testu, prezentacji przeprowadzonych badań walidacyjnych, syntetyzowania wyników tych badań w argument na rzecz trafności i prób jego podważenia. W ich opinii opierająca się na argumentacji strategia walidacyjna stanowi cenne uzupełnienie rekomendacji przedstawionych w *Standardach dla testów...* (AERA, APA, NCME, 2007), dostarczając zarówno ogólną strukturę dla przeprowadzenia syntezy danych pochodzących z różnych źródeł, jak i precyzyjne wskazówki co do sposobu, w jaki należy to zrobić. Warto zaznaczyć, że wydane ostatnio *Standards for educational and psychological testing* (AERA, APA, NCME, 2014) zalecają podejście do walidacji oparte na argumentacji, obejmujące jasne sformułowanie zamierzonej interpretacji i wykorzystania wyników testów, w tym zidentyfikowanie podstawowych założeń i wniosków (argument na rzecz interpretacji / wykorzystania wyników testowych) i zgromadzenie dowodów na ich poparcie lub obalenie.

Badacze od dawna wyrażali obawy, że podnieśli pojęcie trafności teoretycznej na tak wysoki poziom, że wydaje się ono nieosiągalnym celem (Fremer, 2000). Coraz częściej pojawiają się badania, których rezultaty wskazują na rozdzźwięk między teorią i praktyką walidacyjną, co stanowi poważny problem zwłaszcza w przypadku testów mających wielu interesariuszy, a ich autorzy konkludują, że teorię trafności bardzo trudno przełożyć na działania praktyczne wobec braku wskazówek odnośnie do syntetyzowania danych dotyczących trafności i ich interpretacji (por. np. Wolming, Wikström, 2010). „Teoria trafności stopniowo doszła do traktowania każdej ważnej kwestii wiążącej się z testem jako istotnej dla pojęcia trafności i dąży do integracji wszystkich tych kwestii pod jednym nagłówkiem. Jednakże postępując w taki sposób, nie służy ani teoretycznie zorientowanym psychologom, ani stosującym testy praktykom” (Borsboom, Mellenbergh, van Heerden, 2004, s. 1061 – tłum. J. R.). Borsboom (2006) stwierdza, że trafność teoretyczna działa jak czarna dziura, z której nic nie może uciec: gdy dana kwestia zostanie określona jako problem trafności teoretycznej, jego trudność uważa się za nadludzką, a rozwiązanie – za przekraczające

możliwości zwykłego śmiertelnika. Trudność to być może nie nadludzka, ale jednak bardzo poważna, co potwierdza aktualna praktyka. Dowody na rzecz trafności zazwyczaj mają charakter fragmentaryczny, a badacze nadal bardzo często, prezentując dane na temat trafności, odwołują się do klasycznych terminów trafności teoretycznej, treściowej, diagnostycznej i prognostycznej.

Kolejną konsekwencją tego stanu rzeczy są propozycje odmiennych sposobów ujmowania trafności, odrzucające jednolitą teorię trafności; zgodnie z nimi trafność jest własnością testu, a nie własnością interpretacji jego wyników. Lissitz i Samuelsen (2007) postulują, że współczesna teoria trafności, skoncentrowana na trafności teoretycznej, ma bardzo niewiele do zaoferowania badaniom testowym prowadzonym w obszarze edukacji. W proponowanym przez nich ujęciu trafności centralną rolę odgrywa trafność treściowa, którą łącznie z rzetelnością testu określają mianem trafności wewnętrznej. Tym samym ustalenie definicji mierzonej przez test wielkości w procesie jego konstrukcji oraz stabilność wyników mają zasadnicze znaczenie dla trafności testu. Wewnętrzne charakterystyki testu nie zależą od czynników zewnętrznych, np. związków wyników testu z innymi miarami. Aspekty zewnętrzne w stosunku do testu, związane z oceną użyteczności wykorzystania testu w danym celu, wsparciem ze strony teorii oraz konsekwencjami testowania, pomimo że bardzo ważne, nie powinny być włączane do definicji trafności.

Te odrębne charakterystyki – które aktualnie określa się jako ustalanie trafności kryterialnej i teoretycznej – są bardzo ważne, jednakże użytkownik tych technik powinien zdawać sobie sprawę, że dostarczają one odpowiedzi na fundamentalnie różne pytania. Te pytania nie powinny odciągać uwagi badacza od skoncentrowania się na trafności wewnętrznej. (Lissitz, Samuelsen, 2007, s. 446 – tłum. J. R.)

Borsboom i współpracownicy (Borsboom i in., 2004; Borsboom, Cramer, Kievit, Scholten, Franic, 2009) krytykują pojęcie trafności teoretycznej, określając je mianem doktryny, i wskazują, że jednolita teoria trafności nie jest potrzebna psychologii, ponieważ nie ma czego ujednoclić. Integracja wszystkich związanych z testem działań (*validation*) nie jest trafnością (*validity*). Trafność jest własnością przysługującą testowi, a zadaniem badacza jest sprawdzenie, czy test taką własność posiada (czy test mierzy to, co powinien mierzyć). „Trafność jest pojęciem takim jak prawda: reprezentuje idealną lub pożądaną sytuację” (Borsboom i in., 2004, s. 1063). Zgodnie z proponowaną przez autorów koncepcją trafności test jest trafną miarą danego atrybutu, gdy ten atrybut istnieje i jego zróżnicowanie stanowi przyczynę zróżnicowania wyników testu, co oznacza, że relacja między wynikami testu i atrybutami ma charakter przyczynowy, a nie korelacyjny. Tym samym wnioskowanie o trafności wyłącznie na podstawie macierzy wielu cech-wielu metod nie jest możliwe, jako że pożądana konfiguracja współczynników korelacji nie ustanawia trafności. Tym, co stanowi o trafności, jest istnienie atrybutu i wykazanie jego przyczynowego wpływu na wyniki testu. „Problem trafności testu odnosi się do kwestii, czy atrybut psychologiczny (np. ‘inteligencja ogólna’) istnieje i odpowiada atrybutowi, który test rzeczywiście mierzy (przyjmując, że jakiś mierzy). W tym przypadku terminy ‘inteligencja ogólna’ i ‘atrybut, który test mierzy’ odnoszą się do tej samej struktury. Walidacja testu to po prostu przeprowadzenie badania w celu dowiedzenia się, czy jest to prawda, czy nie” (Borsboom i in., 2009, s. 153 – tłum. J. R.). Ujmowana w ten sposób trafność staje się własnością jakościową. Określanie „stopnia trafności” traci sens, a na pytanie o to, czy dany test jest trafny, można udzielić prostej odpowiedzi: tak lub nie.

PODSUMOWANIE

Prowadzona od ponad stu lat dyskusja nad konceptualizacją trafności jako pojęcia psychometrycznego trwa. Zainteresowany Czytelnik może zapoznać się z aktualnie podejmowanymi problemami, sięgając choćby do poświęconego pojęciu trafności wydania czasopisma *Assessment in Education: Principles, Policy & Practice* (2016, t. 23, nr 2) czy też do wydania *Educational Assessment* (2020, t. 25, nr 1), dedykowanego zagadnieniom praktyki walidacyjnej.

Czy tocząca się dyskusja doprowadziła do jakichś konkluzyjnych ustaleń? Odpowiedź na to pytanie musi być – niestety! – przecząca. Kontrowersje nadal dotyczą kluczowej kwestii, związanej z tym, do czego się trafność odnosi: czy jest ona własnością narzędzia, czy związana jest z interpretacją wyników testu czy też z konkretnymi jego zastosowaniami? Pod szyldem jednolitej teorii trafności każdy z ważnych aspektów dotyczących testu i procesu testowania został objęty terminem *trafność*, w efekcie czego sam termin stał się wyjątkowo złożony znaczeniowo i – rzecz by można – nieco uciążliwy. Nie dziwi zatem, że niektórzy badacze postulują, że nadszedł czas, by z niego zrezygnować: „Wziąwszy pod uwagę powszechność nieprecyzyjnego i niejednoznacznego użycia oraz fakt, że nawet specjaliści używają tego słowa w całkiem różny sposób, trudno zrozumieć, jak cokolwiek koncepcyjnie fundamentalnego mogłoby zostać utracone, gdyby słowo to przeszło na emeryturę” (Newton, Shaw, 2016, s. 190 – tłum. J. R.). Termin *trafność* można by zastąpić bardziej ogólnym terminem *jakość* (*quality*), który – obejmując znaczeniowo wszystkie odniesienia terminu *trafność* – wymagałby również od użytkowników sprecyzowania jego znaczenia w konkretnym kontekście użycia (Newton, Shaw, 2013). Z różnicami w ujmowaniu trafności związane są kontrowersje dotyczące właściwego sposobu jej ustalania.

Wyniki badania przeprowadzonego przez Camargo, Herrere i Traynor (2018) za pomocą metody delfickiej (którym to badaniem objęto siedmiu ekspertów zajmujących się problematyką trafności) wydają się potwierdzać różnice w odniesieniu do tych dwóch podstawowych kwestii: czemu przysługuje trafność i jaki jest właściwy sposób jej ustalania? Biorący udział w badaniu eksperci uznali definicję trafności zamieszczoną w ostatnim wydaniu *Standardów dla testów...* za niejasną i interpretacyjnie wieloznaczną, zgadzając się, że określenie, do czego odnosi się trafność, jest utrudnione przez brak jasnej definicji tego terminu. Zgodzili się także co do potrzeby sformułowania jasnych wskazań odnośnie do przeprowadzania procesu walidacji, nie osiągając jednak zgody co do roli, jaką konsekwencje testowania miałyby w nim odgrywać, ani też w odniesieniu do tego, ile danych i dane jakiego typu są konieczne lub wystarczające do ustalenia trafności. Eksperti nie osiągnęli również zgody w kwestii zastąpienia terminu *trafność* innym terminem, takim jak np. *jakość*, czy w sprawie definiowania walidacji w sposób obejmujący wszystko, co wiąże się z naukową oceną jakości pomiaru. Pozwala to sądzić, że termin *trafność* pozostanie w słowniku nauki, a dyskusja nad konceptualizacją trafności będzie toczyć się nadal.

BIBLIOGRAFIA

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2007). *Standardy dla testów stosowanych w psychologii i pedagogice*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). *Standards for educational and psychological testing* [wersja Adobe Digital Edition]. Pobrane z: https://www.testingstandards.net/uploads/7/16/6/4/76643089/standards_2014edition.pdf
- Anastasi, A., Urbina, S. (1999). *Testy psychologiczne*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Angoff, W. H. (1988). Validity: an evolving concept. W: H. Wainer, H. Braun (red.), *Test validity* (s. 19–32). Hillsdale, NJ: Lawrence Erlbaum.

- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*, 425–440. <https://doi.org/10.1007/s11336-006-1447-6>
- Borsboom, D., Mellenbergh, G. J., van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*, 1061–1071. <https://doi.org/10.1037/0033-295X.111.4.1061>
- Borsboom, D., Cramer, A. O. J., Kievit, R. A., Scholten, A. Z., Franic, S. (2009). The end of construct validity. W: R. Lissitz (red.), *The concept of validity. Revisions, new directions, and applications* (s. 135–170). Charlotte, NC: Information Age Publishing.
- Campbell, D. T., Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105. <https://doi.org/10.1037/h0046016>
- Camargo, S. L., Herrera, A. N., Traynor, A. (2018). Looking for a consensus in the discussion about the concept of validity: a Delphi study. *Methodology*, *14*, 146–155. <https://doi.org/10.1027/1614-2241/a000157>
- Chapelle, C. A., Enright, M. K., Jamieson, J. (2010). Does an argument-based approach to validity make a difference? *Educational Measurement: Issues and Practice*, *29*(1), 3–13. <https://doi.org/10.1111/j.1745-3992.2009.00165.x>
- Cizek, G. J. (2012). Defining and distinguishing validity: interpretations of score meaning and justifications of test use. *Psychological Methods*, *17*, 31–43. <https://doi.org/10.1037/a0026975>
- Cizek, G. J. (2016). Validating test score meaning and defending test score use: different aims, different methods. *Assessment in Education: Principles, Policy & Practice*, *23*, 212–225. <https://doi.org/10.1080/0969594X.2015.1063479>
- Cizek, G. J. (2020). *Validity: an integrated approach to test score meaning and use*. Nowy Jork, NY: Routledge.
- Cizek, G. J., Rosenberg, S., Koons, H. (2008). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement*, *68*, 397–412. <https://doi.org/10.1177/0013164410379323>
- Cizek, G. J., Bowen, D., Church, K. (2010). Sources of validity evidence for educational and psychological tests: a follow-up study. *Educational and Psychological Measurement*, *70*, 732–743. <https://doi.org/10.1177/0013164407310130>
- Cronbach, L. J. (1971). Test validation. W: R. L. Thorndike (red.), *Educational measurement* (wyd. 2, s. 443–507). Waszyngton, DC: American Council on Education.
- Cronbach, L. J. (1988). Five perspectives on validation argument. W: H. Wainer, H. Braun (red.), *Test validity* (s. 3–17). Hillsdale, NJ: Lawrence Erlbaum.
- Cronbach, L. J., Meehl, P. E. (1955). Construct validity in psychological test. *Psychological Bulletin*, *52*, 281–302. <https://doi.org/10.1037/h0040957>
- Dorans, N. J., Holland, P. W. (1993). DIF detection and description: Mantel-Haenzel and standardization. W: P. W. Holland, H. Wainer (red.), *Differential item functioning* (s. 35–66). Hillsdale, NJ: Lawrence Erlbaum.
- Fremer, J. (2000). Promoting high standards and the „problem” with construct validation. *NCME Newsletter*, *8*(3), 1.
- Guilford, J. P. (1946). New standards for test evaluation. *Educational and Psychological Measurement*, *6*, 427–438. <https://doi.org/10.1177/001316444600600401>
- Guion, R. M. (1978). „Content validity” in moderation. *Personnel Psychology*, *31*, 205–213. <https://doi.org/10.1111/j.1744-6570.1978.tb00440.x>
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, *112*, 527–535. <https://doi.org/10.1037/0033-2909.112.3.527>
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, *38*, 319–342. <https://doi.org/10.1111/j.1745-3984.2001.tb01130.x>
- Kane, M. T. (2004). Certification testing as a illustration of argument-based validation. *Measurement*, *2*, 135–170. https://doi.org/10.1207/s15366359mea0203_1
- Kane, M. T. (2006a). Content-related validity evidence in test development. W: S. M. Downing, T. M. Haladyna (red.), *Handbook of test development* (s. 115–154). Mahwah, NJ: Lawrence Erlbaum.
- Kane, M. T. (2006b). Validation. W: R. L. Brennan (red.), *Educational measurement* (wyd. 4, s. 17–64). Westport, CT: Greenwood Publishing.
- Kane, M. T. (2008). Terminology, emphasis, and utility in validation. *Journal of Educational Measurement*, *37*, 76–82. <https://doi.org/10.3102/0013189X08315390>
- Kane, M. T. (2009). Validating the interpretations and uses of test scores. W: R. Lissitz (red.), *The concept of validity. Revisions, new directions, and applications* (s. 39–64). Charlotte, NC: Information Age Publishing.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*, 1–73. <https://doi.org/10.1111/jedm.12000>
- Kane, M. T. (2016). Explicating validity. *Assessment in Education: Principles, Policy & Practice*, *23*, 198–211. <https://doi.org/10.1080/0969594X.2015.1060192>
- Lissitz, R. W., Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educational Researcher*, *36*, 437–448. <https://doi.org/10.3102/0013189X07311286>
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, Monograph Supplement*, *3*, 635–694. <https://doi.org/10.2466/pr0.1957.3.3.635>
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, *35*, 1012–1027. <https://doi.org/10.1037/0003-066X.35.11.1012>
- Messick, S. (1989). Validity. W: R. L. Linn (red.), *Educational measurement* (s. 13–103). Nowy Jork, NY: American Council on Education and Macmillan.

- Messick, S. (1995). Validity of psychological assessment. Validation of inferences of persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Murphy, K. R. (2009). Content validation is useful for many things, but validity isn't one of them. *Industrial and Organizational Psychology*, 2, 453–464. <https://doi.org/10.1111/j.1754-9434.2009.01173.x>
- Newton, P. E., Shaw, S. D. (2013). Standards for talking and thinking about validity. *Psychological Methods*, 18, 301–319. <https://doi.org/10.1037/a0032969>
- Newton, P. E., Shaw, S. D. (2016). Disagreement over the best way to use the word 'validity' and options for reaching consensus. *Assessment in Education: Principles, Policy & Practice*, 23, 178–197. <https://doi.org/10.1080/0969594X.2015.1037241>
- Schmidt, F. L., Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 529–540. <https://doi.org/10.1037/0021-9010.62.5.529>
- Schmidt, F. L., Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262–274. <https://doi.org/10.1037/0033-2909.124.2.262>
- Sireci, S. (1998). The construct of content validity. *Social Indicators Research*, 45, 83–117. <https://doi.org/10.1023/A:1006985528729>
- Sireci, S. (2009). Packing and unpacking sources of validity evidence: History repeats itself again. W: R. Lissitz (red.), *The concept of validity. Revisions, new directions, and applications* (s. 19–37). Charlotte, NC: Information Age Publishing.
- Spies, R. A., Plake, B. S. (red.). (2005). *The sixteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4, 223–233. <https://doi.org/10.1080/15434300701375832>
- Wolming, S., Wikström, C. (2010). The concept of validity in theory and practice. *Assessment in Education: Principles, Policy & Practice*, 17, 117–132. <https://doi.org/10.1080/09695941003693856>

