

Polish adaptation of the Actively Open-Minded Thinking (AOT) questionnaire

Pomiar otwartego i elastycznego myślenia (POEM)

Anna Błaszczak*

Institute of Psychology

Maria-Curie Skłodowska University in Lublin

0000–0003–0597–5117

Marta Klocek

Institute of Psychology

Maria-Curie Skłodowska University in Lublin

0009–0007–4406–6844

ABSTRACT

Aims

The aim of this research was to develop a Polish adaptation of the Actively Open-Minded Thinking (AOT) questionnaire. This questionnaire allows to assess the disposition to flexibly consider different problems from various perspectives, regardless of one's initially favoured options, as well as the critical attitude in analysing information and forming opinions.

Methods

The research was conducted among two groups of adult participants. Validity of the AOT scale was assessed using confirmatory factor analysis, and correlation coefficient analyses between the scale's results and the measures of need for cognition, need for closure or cognitive reflection. Internal consistency of the Polish adaptation of the AOT scale was evaluated using Cronbach's alpha measure.

Results

Polish adaptation of the AOT questionnaire (Pomiar otwartego i elastycznego myślenia; POEM) satisfied the established reliability and validity criteria. Based on the conducted analyses, the instrument was significantly shortened relative to the original version of

* Correspondence address: Anna Błaszczak, Institute of Psychology UMCS, Głęboka Street 45, 20–612, Lublin. E-mail: anna.blaszczak2@mail.umcs.pl.

the AOT scale, with a similar approach seen in past research exploring the optimal item set with highest attainable psychometric properties. Internal consistency indices were sufficiently high and, in line with the researchers' predictions, the instrument correlated with other measures of corresponding cognitive constructs.

Conclusions

AOT is a dynamic variable, dependent upon the influence of training or experience. Polish adaptation of the scale measuring the phenomenon is an instrument with promising psychometric properties, worth further investigation. The POEM scale can be valuable in assessing open-minded thinking in educational settings, in research investigating the conditions facilitating social categorisation or radicalisation.

Keywords: actively open-minded thinking, flexible thinking, AOT, Polish adaptation

Introduction

Actively open-minded thinking (AOT), devoid of cognitive biases, seems to be presently of immense value. The need for critical verification of diverse information, opinions and perspectives incoming from a variety of sources simultaneously, requires an adequate capacity to cope with this diversity, as well as the ability to move beyond one's own preconceived notions of oneself and others (Jarymowicz, 2002). Open-mindedness, flexibility and reflectivity of thought, devoid of simplifications and rigid categorisations, appears to be a crucial competency of a modern-day person.

Baron (2008) and Stanovich and West (1997) characterise the nature of AOT, emphasising four core dimensions of such thinking: the tendency to consider alternative opinions and evidence; willingness to move beyond the readily available information and context (i.e., perspective-switch); willingness to shift one's perspective and consider a topic from a new point of view (i.e., decontextualise); lack of epistemological absolutism. The first component relies on actively seeking out opinions which differ from one's initially favoured conclusions, by weighting various arguments for and against said conclusions. Dimensions two and three, the willingness to perspective-switch and decontextualise, are two co-occurring and inter-related concepts. Moving beyond readily available information and context requires acknowledging data outside one's pre-existing perspectives. Key to this particular component of AOT is the ability to understand another person's standpoint and the underlying logic, regardless of whether we agree with that logic or not (Metz, Baelen, & Yu, 2020). The final dimension emphasised by Stanovich and West (1997) crucial to AOT, is lack of epistemological absolutism. It refers to an approach of acquiring knowledge and seeking information which contradicts dogmatism and unconditional acceptance of the "only one right way". It requires cognitive finesse and refinement, encouraging a person to make an effort to actively seek information in the name of acquiring new knowledge and broadening one's horizons (Metz et al., 2020).

Based on works by Baron (1993), in which the measure of AOT was the performance in tasks requiring participants to generate diverse arguments and

declarations with respect to different aspects of thinking, Stanovich and West (1997) developed a standardised questionnaire measuring this cognitive disposition. Their instrument includes statements about one's styles of thinking, problem-solving and decision-making, and the strategies they employ in these processes (e.g., "A person should always consider new possibilities"). A person's convictions regarding the importance of changing their opinions and verifying their views based on incoming novel information are an indicator of their AOT capabilities.

The aim of the present study was to develop a Polish version of the AOT scale and to ascertain its psychometric properties: factor structure, internal consistency, construct validity, and external validity. The original questionnaire was translated into Polish, applied in a population sample and subjected to statistical analyses. Factor structure was verified using confirmatory factor analysis and the goodness-of-fit statistics. The researchers tested the original AOT model and other models present in the literature based on the original model. These analyses served to identify a model which represented the best fit for the overall data. The instrument was modified according to factor loading analysis and interpretation (i.e., items with low loading factor were successively deleted). Such a modified model yielded high goodness-of-fit indices. The modified model, which only included the selected items which best assessed the nature of the construct of interest, was then applied in a population sample one more time. Internal consistency was assessed using Cronbach's alpha. External validity was evaluated by analysing correlation coefficients between responses from the Polish version of the scale, and the levels of need for cognition, need for closure, as well as reflective, bias-free thinking.

Experiment 1. Developing and Validating the Structure of the Poem Scale

Methods

Participants. The participants were 432 adults (304 female, 123 male, 5 non-binary). The sample consisted mostly of higher education students ($n = 313$), with the remaining participants being professionals who completed secondary education ($n = 29$), vocational training ($n = 3$), and higher education ($n = 87$). The mean age of the participants was approximately 25 years old ($SD = 9.09$), ranging from 18 to 64 years old.

Materials. Due to the social distancing restrictions implemented amidst the COVID-19 pandemic, the study was conducted online between February to March 2022, and again in November 2022 in the same format. The questionnaire was distributed via social media and the researchers' personal networks. The study used a Polish translation of the AOT scale (Stanovich & West, 2007).

The instrument consisted of 41 statements assessing the respondents' tendency to consider problems from a range of diverse perspectives, regardless of initially favoured options, as well as their critical attitude towards established rules or analysing data and forming opinions. The participants were asked to consider each statement and rate the extent to which the statement applies to them, using a 6-point scale (1 = *strongly disagree*; 6 = *strongly agree*). A sumscore of the responses to the 41 items reflected the respondents' disposition towards AOT. Such an approach to result collation, suggests that AOT should be regarded as a unidimensional trait (Janssen et al., 2020), in spite of clearly distinguished subscales. Within the AOT scale, the researchers have distinguished six subscales (see Sá, West, & Stanovich, 1999; Stanovich & West, 2007), which will be briefly discussed below.

- Flexible thinking is a core dimension of the AOT scale, assessing the ability to consider alternative options and perspectives, reflectivity of thinking, and the willingness to change one's conviction in the face of reliable arguments and justifications (see Baron, 1993). It includes items such as "People should always take into consideration evidence that goes against their beliefs". This subscale consists of ten items.
- Openness values is a subscale which includes eight statements drawn from the NEO-FFI PI(R) inventory. These items measure the extent of one's appreciation of openness to new experiences as an essential trait in the modern-day world. It includes statements such as "I believe that laws and social policies should change to reflect the needs of a changing world".
- Dogmatism subscale includes nine items, three of which were taken from the short-form Rokeach's Dogmatism Scale (Troidahl & Powell, 1965), four from the survey's full version (Robinson, Shaver, & Wrightsman, 1991), and three from its adaptation by Paulhus and Reid (1991). An example of a statement which indicates high levels of dogmatism is "No one can talk me out of something I know is right". Statements indicative of high levels of dogmatism are simultaneously symptomatic of low AOT levels.
- Categorical thinking is a subscale allowing to capture one's tendency for thinking in terms of simplistic categories. It consists of three reverse scored items, e.g., "I tend to classify people as either for me or against me".
- Belief identification is a subscale introduced by Sá and colleagues (1999), and refers to the extent to which people associate their beliefs with their concept of self. It includes nine statements, such as "It makes me happy and proud when someone famous holds the same beliefs that I do".
- Counterfactual thinking is the final subscale included in the AOT scale, which consists of two items. These items assess beliefs indicating one's willingness to decontextualise and accept perspectives different to the ones initially favoured, e.g., "My beliefs would not have been very different if I had been raised by a different set of parents". The statements are reverse scored, which suggests that the higher the subscale score, the greater the respondents' likelihood to ignore evidence-based facts, which in turn hinders AOT.

Validating the original AOT scale. Developing an AOT scale, first attempted a quarter of century ago, is an endeavour which continues till present day (Baron, 2008; Haran, Ritov, & Mellers, 2013; Janssen et al., 2020; Sá et al., 1999; Sá, Kelley, Ho & Stanovich, 2005; Svedholm-Häkkinen & Lindeman, 2017). The scale's first version (Stanovich & West, 1997) included 56 items distributed across eight subscales: Flexible thinking, Openness values, Dogmatism, Categorical thinking, Openness – ideas, Absolutism, Superstitious thinking, and Counterfactual thinking. The psychometric values of this instrument were questionable. Only three of the original version's subscales yielded sufficiently high reliability indicators (Flexible thinking – .50, Openness values – .71, Dogmatism – .60, Openness – ideas – .77, Absolutism – .64, Superstitious thinking – .73; see Stanovich & West, 1997¹). Moreover, principal component analysis showed that the first six subscales loaded heavily on one component, accounting for the greatest percentage of variance (40.8%). Consequently, items included in the Superstitious thinking and Counterfactual thinking subscales were retracted, and the remaining subscales were used to compute a single composite score where AOT is treated as a unidimensional trait. This sumscore differentiates between respondents who exhibit openness to belief-change and cognitive flexibility, and those who are better characterised by a resistance to belief change and cognitive rigidity.

Ten years after publishing the original version of the instrument (Stanovich & West, 1997), the authors developed a 41-item AOT scale (Stanovich & West, 2007), which is presently the most commonly used version (see Janssen et al., 2020). The scale includes six subscales: four adapted from the original version (Flexible thinking, Openness values, Dogmatism, Categorical thinking), and two new ones, i.e., Belief identification and Counterfactual thinking. A sumscore of the 41 items (with 30 of them being reverse scored) allows to determine the respondents' disposition towards AOT. This version of the instrument yielded good split half reliability and internal consistency measures (.75 and .83 accordingly; see Stanovich & West, 2007). Although the reliabilities for the individual subscales of the AOT41 were not reported, it is possible that, just as in case of the prototype, these values were not high (see Janssen et al., 2020).

Furthermore, considering the significant inter-item correlations between different statements on the AOT41, the authors recommended summarising the score of each item to obtain a single value (Janssen et al., 2020). In doing so, AOT is considered as a unitary construct, although past literature failed to report any evidence for this assumption. Various analyses undertaken in the past few years showed that this hypothesis is difficult to confirm, leading to multiple attempts at modifying the original scale including 41 items (AOT41; see Janssen et al., 2020). The authors of the original instrument sometimes used a shortened version of the scale (AOT30), which only included items from the Flexible thinking, Belief identification, Dogmatism and Categorical thinking subscales (Kokis,

¹ The original study did not report reliability values for Categorical thinking and Counterfactual thinking (see Stanovich & West, 1997).

Macpherson, Toplak, West, & Stanovich, 2002), as best indicators of AOT. Other studies utilised only selected subscales (Baron, 2008; Haran et al., 2013; Sá et al., 1999; Sá et al., 2005). There have also been attempts at developing the questionnaire's shortened or national adaptations of substantially different length depending on the authors and the country.

The Swedish adaptation by Svedholm-Häkkinen and Lindeman (2017) consists of 17 statements (AOT17) distributed across four subscales. These subscales are: Dogmatism (6 items) including questions such as "I think there are many wrong ways, but only one right way, to almost anything"; Fact resistance (5 items) consisting of statements like "Certain beliefs are just too important to abandon no matter how good a case can be made against them"; Liberalism (3 items) referring to one's openness to new thoughts and ideas, e.g., "A person should always consider new possibilities"; as well as Belief personification (3 items) with items such as "My blood boils over whenever a person stubbornly refuses to admit he's wrong". The British version by Baron (2008) includes 10 to 11 items (Baron, 2018), however, these items are diverse and significantly deviating from the original items. The American version, used by Haran and colleagues (2013), consists of 7 statements which are not divided into any subscales.

Translating the AOT scale. The AOT41 scale was faithfully translated into Polish. The process was carried out by psychologists specialising in research into rational thinking who are fluent in English. The statements included in the Openness values subscale adapted from the NEO-PI(R) scale were taken from the pre-existing Polish version. Each statement was then translated again into English by a native speaker professional in psychology translations. In the case of seven statements, translations of which poorly reflected their original meaning, peer feedback (from a group of psychology researchers specialising in diagnosis and fluent in English) ruled that these should be replaced by more suitable Polish equivalents. In the adaptation process, the original instrument format was retained, whereby respondents indicate their answers on a 6-point scale (1 = *strongly disagree*; 6 = *strongly agree*). The translation process also suggested a Polish name of the scale – Pomiar otwartego i elastycznego myślenia (POEM).

Results

Reliability analysis and factor structure of the POEM scale prototype. First, the prototype's internal consistency was assessed based on Cronbach's alpha values, using questionnaire responses from all participants. The following reliability indices were obtained: for the entire scale – .83, for each subscale: Flexible thinking – .52; Openness values – .67; Dogmatism – .59; Categorical thinking – .49; Belief identification – .57; Counterfactual thinking – .68. These values suggest a good overall reliability of the instrument but – similarly to past

literature – insufficient or low reliability of specific subscales (Janssen et al., 2020).

Factor structure of the responses obtained via the POEM scale was verified using confirmatory factor analysis and the goodness-of-fit statistics. The proposed unidimensional model (adopted as per past literature's recommendation to consider only the instruments' sumscore) did not adequately capture the collected data ($\text{Chi}^2 = 2267.35$; $\text{df} = 779$; $\text{Chi}^2/\text{df} = 2.91$; $p < .001$; $\text{CFI} = .577$; $\text{RMSEA} = .067$ [.063; .070]).

The poor fit between the original model and the collected data encouraged further analyses to test alternative approaches. First, the scale's full version (AOT41) was tested again using a bifactor model, which is a more suitable approach when an instrument evaluates both the total sumscore and separate scores for each subscale. Then, the researchers verified how well the obtained results fit alternative models used in past research which also incorporated the original items from the scale's full version: AOT30 (Kokis et al., 2002), AOT17 (Svedholm-Häkkinen & Lindeman, 2017), as well as AOT7 (Haran et al., 2013). The results of these analyses are summarised in Table 1.

Table 1

Goodness-of-fit values for the models tested in Experiment 1

Scale version	Analysis method	Chi ²	df	p	Chi ² /df	CFI	RMSEA, 90% CA	SRMR
AOT41	Single-factor modelling	2267.35	779	.000	2.91	.577	.067 [.063; .070]	.075
AOT41	Bifactor modelling	1578.93	724	.000	2.11	.757	.052 [.049; .056]	.057
AOT30	Single-factor modelling	1230.90	434	.000	2.84	.650	.065 [.061; .070]	.070
AOT30	Bifactor modelling	975.97	403	.000	2.42	.748	.057 [.053; .062]	.072
AOT17	Single-factor modelling	496.04	119	.000	4.17	.706	.086 [.078; .094]	.075
AOT17	Bifactor modelling	288.93	102	.000	2.83	.854	.065 [.056; .074]	.075
AOT7	1 latent factor	71.07	14	.000	5.14	.859	.098 [.076; .121]	.060

These comparisons revealed that the AOT17 version, based on a bifactor model, was best suited to capture the collected data ($\text{Chi}^2 = 288.93$; $\text{df} = 102$; $\text{Chi}^2/\text{df} = 2.83$; $p < .001$; $\text{CFI} = .854$; $\text{RMSEA} = .065$ [.056; .074]). The observed CFI

value (.85) was nearly acceptable, and the RMSEA value (.065) indicated a satisfactory fit. Table 2 includes the set of items used in the AOT17 version alongside their respective factor loadings.

Table 2

Factor loadings of the items included in the AOT17 version of the original AOT scale

Item nr in AOT41	Path				
	← AOT17	← Dogmatism	← Fact resistance	← Liberalism	← Belief personification
23	.69*	.12			
32	.30*	.31			
24	.60*	.19			
8	.46*	.24			
39	.24*	.25			
7	.36*	.14			
33 _R	.61*		.15		
15 _R	.68*		.37		
19 _R	.47*		.43		
37	.48*		.31		
41	.53*		.28		
18	.28*			.11	
4	.27*			.08	
27	.27*			.17	
3 _R	.01				.11
11 _R	.19*				.12
31 _R	.01				.10

R – reverse scored item; * $p < .001$

Next, based on factor loading values for each item, statements with lowest factor loading were deleted. First, the researchers deleted items 31 and 3, then 11, and then 27, 4 and 18 respectively as their factor loadings were still low. Finally, they deleted positions 32, 39 and 7, until all factor loadings on the AOT → item path were higher than .40 and the overall fit of the model was improved. In the process, the factor loading values for each subscale have also improved. The final set of items yielded a model with satisfactory fit parameters

($\chi^2 = 27.80$; $df = 12$; $\chi^2/df = 2.31$; $p < .001$; CFI = .980; RMSEA = .055 [.028; .082]; SRMR = .030). The remaining items with their respective factor loadings are presented in Table 3.

Table 3

Selected AOT items and their respective factor loadings

Item nr in AOT41	Original English items (Stanovich & West, 2007)	Polish translated items	Path	
			← AOT17	← subscale
AOT17 Factor 1: Dogmatism				
23	I believe that loyalty to one's ideals and principles is more important than open-mindedness.	Uważam, że lojalność w stosunku do własnych przekonań jest bardziej istotna niż „otwartość umysłu”.	.65*	.28
24	Of all the different philosophies which exist in the world there is probably only one which is correct.	Ze wszystkich filozofii, które istnieją na świecie, prawdopodobnie tylko jedna jest właściwa.	.52*	.48
8	I think there are many wrong ways, but only one right way, to almost anything.	Myślę, że prawie w każdej sytuacji jest wiele złych wyjść, ale tylko jedno właściwe.	.40*	.40
AOT17 Factor 2: Fact resistance				
33 _R	One should disregard evidence that conflicts with your established beliefs.	Należy lekceważyć dowody, które są sprzeczne z naszymi przekonaniami.	.59*	.19
15 _R	It is important to persevere in your beliefs even when evidence is brought to bear against them.	Ważne jest, aby wytrwać w swoich przekonaniach, nawet jeśli przedstawiane są dowody przeciwko nim.	.78*	.24
19 _R	Certain beliefs are just too important to abandon no matter how good a case can be made against them.	Pewne przekonania są zbyt ważne, by je porzucić bez względu na to, jak dobre argumenty przeciwko nim można przedstawić.	.55*	.27
37	Beliefs should always be revised in response to new information or evidence.	Poglądy zawsze powinny być weryfikowane w odpowiedzi na nowe informacje i dowody.	.43*	.44
41	People should always take into consideration evidence that goes against their beliefs.	Ludzie powinni zawsze brać pod uwagę dowody sprzeczne z ich przekonaniami.	.49*	.48

R – reverse scored item; * $p < .001$

These selected items were included in the POEM8 scale, which, as in the original model, distinguishes two subscales. The first three items were drawn from Svedholm-Häkkinen and Lindeman's (2017) Dogmatism subscale. The original subscale measures one's attachment to the "only one right way" of thinking and problem-solving, which is why the same subscale label was retained in the POEM8 version. The remaining five statements were taken from Svedholm-Häkkinen and Lindeman's (2017) Fact resistance subscale. Some of these items pertain to a tendency to ignore new evidence or arguments which disconfirm one's pre-existing convictions. They are reverse scored, meaning that the score reflects a behavioural disposition which opposes rigidly standing by one's convictions. Other items within the subscale refer to the importance of verifying one's pre-existing convictions against new evidence. Therefore, it seems that Openness to facts is a subscale label which better corresponds to the statements' content and the nature of the variable of interest. As such, this new proposed subscale label was used in the POEM8 scale.

The means and standard deviations were established for the POEM8 instrument. The mean sumscore was 4.79 ($SD = .70$), and for the subscales it was: Dogmatism – 5.02 ($SD = .84$), Openness to facts – 4.65 ($SD = .77$). Next, the internal consistency was assessed, with Cronbach's alpha value of .77, which is lower than the reliability of the scale's preliminary prototype, but still satisfactory (especially considering the significantly reduced number of statements included in this version compared to the prototype). The internal consistency of separate subscales within the POEM8 instrument was slightly lower than expected, with values of .67 for Dogmatism and .66 for Openness to facts.

Correlation between the POEM8 subscore and the original full version of the instrument (AOT41) was also evaluated. The correlation coefficient was .82 as was statistically significant ($p < .001$). The next experiment focused on further verifying the POEM8 scale, focusing on the reliability and external validity of the developed instrument.

Experiment 2. Analysing the Reliability and Validity of the Poem Scale

Methods

Participants. The participants were 144 adults (109 female, 34 male, 1 non-binary). The sample consisted mostly of higher education students ($n = 117$), with the remaining participants being professionals who completed higher education ($n = 27$). The mean age of the participants was approximately 24 years old ($SD = 7.87$), ranging from 18 to 54 years old.

Materials. Since the POEM prototype in Experiment 1 was tested online, due to the social distancing restrictions implemented amidst the COVID-19 pandemic,

these testing conditions were replicated in Experiment 2. The participants completed the POEM8 questionnaire following a demographic survey which included instruments described below. To ensure randomisation, they were presented with the instruments in six random orders. The instruments other than POEM8 included in Experiment 2 included:

The Need for Closure Scale (Skala potrzeby poznawczego domknięcia) developed by Webster and Kruglanski (1994) and adapted by Kossowska (2003) allows to capture individual differences in the respondents' tendency to seek out and possess clear and certain knowledge reducing cognitive uncertainty. In order to measure the participants' need for closure (potrzeba poznania poznawczego; PPD), the present study used the full version of the instrument, which includes 32 statements referring to one's desire for order and predictability as opposed to ambiguity and uncertainty in various real-life scenarios. The respondents are to rate how closely each statement reflects their personal views on a 6-point scale (1 = *strongly disagree*; 6 = *strongly agree*). These statements are distributed across subscales, four of which: Need for order, Need for predictability, Avoidance of ambiguity, and Close-mindedness contribute to one factor (Kossowska, 2003; Kossowska et al., 2012). The second factor, independent of the first one, consists of the Decisiveness subscale. The instrument has a high reliability (Cronbach's alpha in the previously mentioned studies was .81 [see Kossowska, 2003; Kossowska, Hanusz, & Trejtowicz, 2012] and .82 in the present sample).

High PPD encourages a superficial analysis of the incoming information, and motivates searching for information consistent with one's pre-existing knowledge. Such a tendency is contradictory to AOT. Meanwhile, low PPD levels are indicative of a more discerning, diversified situation analysis, as well as an openness to new and potentially unfamiliar perspectives, solutions and options. Low PPD scale scores suggest one's capacity for more complex and unconventional information processing style, a tendency to consider alternative interpretations, and an ability to assimilate data even when they disconfirm one's pre-existing convictions. Therefore, in the present analyses, the researchers assumed a negative correlation between the results of the POEM scale and the PPD scale, especially in its Close-mindedness subscale.

The Need for Cognition Scale (Skala potrzeby poznania) developed by Cecioppo and Petty (1982) and adapted in Polish by Matusz and colleagues (2011) allows to measure one's need for cognition (potrzeba poznania; PP), which is the tendency to willingly engage in cognitively demanding activities and derive pleasure from them. The scale consists of 36 items pertaining to one's cognitive motivation and enjoyment of effortful cognitive tasks, e.g., "I really enjoy a task that involves coming up with new solutions to problems". The respondents are to indicate the extent to which they agree with each statement using a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*). The sumscore reflects their need for cognition. Thus, the scale explores one's cognitive motivation and captures individual differences in this disposition. The reliability of the Polish adaptation of the instrument is high (Cronbach's alpha of the Polish adaptation – .90; see Matusz, Traczyk, & Gašiorowska, 2011), just as the reliability observed in the present study (.90).

Individuals with high PP levels are characterised by a generally more positive attitude to tasks and situations which require more careful consideration in problem-solving (regardless of their attitude to other activities, e.g., cleaning). They will be more willing to process incoming novel information and actively seek them out. Past research utilising the PP scale suggests that people with a high need for cognition are more sensitive to the quality of the argument and less sensitive to peripheral cues when processing persuasive messages (Cacioppo, Petty, & Morris, 1983), which overall facilitates objective information processing. In the event of cognitive biases, individuals with high PP levels are more likely to correct their decisions and judgements (Kossowska, 2009). Such a disposition should promote AOT, allowing an assumption that the PP scale results will be positively correlated with the POEM scale indices.

Szydłowski's (2015) Polish adaptation of the Cognitive Reflection Test (CRT) by Fredrick (2005) consists of three mathematical tasks, which provoke quick, automatic solutions, which often leads to incorrect answers. An example of a CRT task is: "A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?". The automatic response, 10 cents, is incorrect. According to Fredrick (2005), this occurs because information is first processed by an intuitive system which does not require much attention. Solving the task requires the activation of a reflective system, since reaching the correct answer requires overriding the initially posited intuitive answers. This short test is characterised by well-established validity (Fredrick, 2005) and satisfactory reliability (Cronbach's $\alpha = .73$; Szydłowski, 2019). These findings were supported by analyses of the present sample (Cronbach's $\alpha = .72$). However, past studies showed only moderate correlations between CRT results and cognitive measures, suggesting that the construct measured by this instrument is distinct from intelligence, more indicative of one's cognitive tendencies in information processing (Szydłowski, 2019). Reflective thinking is one of the foundations of AOT, thus justifying the use of this test in the present analyses.

The set of tasks requiring an individual to consider probability taking into account base-rate information (Base-Rate Tasks, BRT) by Kahneman and Tversky (1973; Polish adaptation: Szydłowski, 2015), generate a dissonance between a stereotype presented in the task description and the objective probability. An example BRT scenario describes a hypothetical study and a participant with a gender-neutral name Jo: "One thousand people took part in the study. The participants included 4 men and 996 women. From this group, a participant was selected at random and assigned the pseudonym Jo. Jo is 23 and will soon complete their vocational training. On Friday evenings, Jo likes to go out with their friends to listen to loud music and drink beer". The respondents are asked to specify whether Jo is a woman or a man. This task was designed to trigger intuitive reasoning, based on the available contextual data and gender stereotypes. However, the correct, fallacy-free answer requires analytical reasoning, i.e., thinking beyond initially posited intuitive solutions and paying attention to the probability data in the hypothetical scenario, whilst ignoring simplistic, stereotype-based categorisations. Applying logical principles in reasoning and considering all available information and options, in spite of the intuitive solutions, seems to be a vital component of

AOT. Therefore, the ability to correctly solve this type of tasks should be positively correlated with the POEM8 scale, confirming its validity.

Results

First, mean scores were established for the POEM8 instrument. The mean sum-score of the scale was 4.78 ($SD = .75$). The mean scores for each subscale were as follows: Dogmatism – 5.04 ($SD = .83$), Openness to facts – 4.63 ($SD = .82$). Then, the researchers evaluated the instrument's internal consistency and validity.

Analysing reliability of the POEM scale. Reliability of the POEM8 scale used in Experiment 2 was evaluated based on internal consistency values. Cronbach's alpha for the whole scale was .81 and, respectively, .74 for the Dogmatism subscale and .71 for the Openness to facts subscale.

Verifying the structure of the POEM scale. Factor structure of POEM8 results collected in Experiment 2 was verified using confirmatory factor analysis and the goodness-of-fit statistics. The obtained fit values were good ($\text{Chi}^2 = 21.44$; $df = 12$; $\text{Chi}^2/df = 1.787$; $p = .006$; $CFI = .973$; $RMSEA = .074$ [.012; .124], $SRMR = .029$). Moreover, the fit indices of the model tested in Experiment 2 were similar to those yielded in Experiment 1 (the difference in CFI was lower than .010, the difference in RMSEA lower than .015, and the difference in SRMR lower than .010; see Lubiewska & Głogowska, 2018).

Figure 1 illustrates the factor structure of the POEM scale responses.

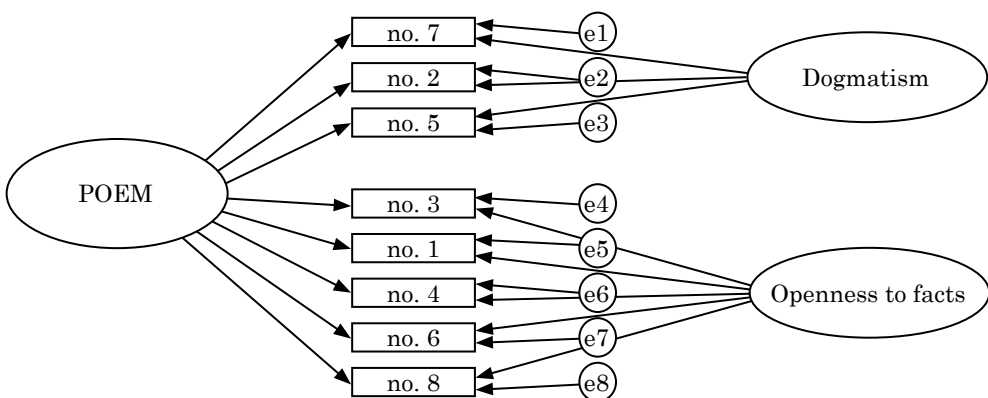


Figure 1. Factor structure of the POEM scale results.

Factor loading values obtained in Experiment 2 for each POEM8 item are summarised in Table 4 (p. 148).

Table 4

Factor loading of items included in the POEM8 scale

Item nr in POEM8	Original English items (Stanovich & West, 2007)	Polish translated items	Path	
			← POEM8	← subscale
	Dogmatism	Dogmatyzm		
7	I believe that loyalty to one's ideals and principles is more important than open-mindedness.	Uważam, że lojalność w stosunku do własnych przekonań jest bardziej istotna niż „otwartość umysłu”.	.65*	.42
2	Of all the different philosophies which exist in the world there is probably only one which is correct.	Ze wszystkich filozofii, które istnieją na świecie, prawdopodobnie tylko jedna jest właściwa.	.52*	.39
5	I think there are many wrong ways, but only one right way, to almost anything.	Myślę, że prawie w każdej sytuacji jest wiele złych wyjść, ale tylko jedno właściwe.	.40*	.38
	Openness to facts	Otwartość na fakty		
3 _R	One should disregard evidence that conflicts with your established beliefs.	Należy lekceważyć dowody, które są sprzeczne z naszymi przekonaniem.	.59*	.11
1 _R	It is important to persevere in your beliefs even when evidence is brought to bear against them.	Ważne jest, aby wytrwać w swoich przekonaniach, nawet jeśli przedstawiane są dowody przeciwko nim.	.78*	.26
4 _R	Certain beliefs are just too important to abandon no matter how good a case can be made against them.	Pewne przekonania są zbyt ważne, by je porzucić bez względu na to, jak dobre argumenty przeciwko nim można przedstawić.	.55*	.40
6	Beliefs should always be revised in response to new information or evidence.	Poglądy zawsze powinny być weryfikowane w odpowiedzi na nowe informacje i dowody.	.43*	.42
8	People should always take into consideration evidence that goes against their beliefs.	Ludzie powinni zawsze brać pod uwagę dowody sprzeczne z ich przekonaniem.	.49*	.42

R – reverse scored item; * $p < .001$

Analysing the validity of the POEM scale. Next, the researchers evaluated convergent validity by analysing correlation coefficients between the responses on the proposed POEM scale, and the need for cognition, need for closure (in

particular tolerance to ambiguity and fact resistance), as well as CRT and BRT results. The obtained results are presented in Table 5.

Table 5

Correlation coefficients between POEM scale results and other instruments used in Experiment 2

POEM8 Scale	Total score	Dogmatism	Openness to facts
Need for closure			
Total score	-.25**	-.29**	-.20**
Need for order	-.16	-.29**	-.07
Need for predictability	-.06	-.14	-.01
Avoidance of ambiguity	-.06	-.08	-.04
Close-mindedness	-.38**	-.42**	-.31**
Decisiveness	-.14	-.15	-.12
Need for cognition			
Total score	.20**	.18**	.19*
Cognitive Reflection Test			
Total score	.30**	.32**	.23**
Base-Rate Tasks			
BRT_I	.38*	.32**	.35**
BRT_C	-.10	-.08	-.09

** $p < .005$; * $p < .05$

Overall, a negative correlation was found between POEM scale results and the need for cognitive closure, with an especially clear, but moderate relationship between the total POEM scale sumscore and its Fact resistance subscale. As predicted, the researchers also identified a (weak) positive correlation between the POEM scale sumscore and the need for cognition levels. A moderate positive correlation was observed between POEM scale results and reflective thinking tendencies, as well as performance in tasks requiring an individual to consider probability taking into account base-rate information. The stronger were the indicators of respondents' AOT disposition, the less likely they were to commit a base-rate fallacy under the influence of the suggested stereotype. They were also more likely to process information in a way which considers objective probability and base-rate information. In summary, the researchers observed the expected relationships between the construct measured by the POEM scale and other tests measuring similar constructs, suggesting that this adaptation of the AOT scale is a valid instrument.

Discussion

The present study described the process of Polish adaptation of the AOT scale and developing an instrument to measure this disposition, called POEM, which satisfied the established reliability and validity criteria. Crucially, this Polish version included only eight statements from the original AOT scale. Similar approach was undertaken in past research seeking an optimal item set and translation of the instrument in order to best capture the phenomenon of interest (see Janssen et al., 2020). The varying number of statements included in different versions of the scale can be attributed not only to issues in translating the nuanced expression of AOT disposition across languages, but also the strong contextual nature of said disposition. For instance, statements pertaining to limiting freedom of speech among different political factions will be more meaningful to respondents from the USA compared to Poland. Interestingly, even the American and British version of the AOT scale vary in content, the main difference being the importance of staying faithful to one's personal values (British version) and the importance of intuitive thinking in decision-making (American version). This suggests that, in some aspects of AOT-specific behavioural tendencies and experiences, it is difficult to determine a universal, objective and context-independent set of AOT indices. What constitutes a conservative approach to information processing can have diverse expressions depending on one's culture and upbringing. Possibly, these are the reasons why many statements included in the original AOT scale were not included in other versions of the instrument, particularly the POEM version proposed in this study. This also means that comparing the AOT disposition more broadly, requires for it to be measured in more general terms (just as was the case in the Polish adaptation of the AOT scale statements such as "Beliefs should always be revised in response to new information or evidence"). Unfortunately, this limitation of the present instrument and its other adaptations means that each national adaptation may actually measure a slightly different construct. Consequently, cross-cultural comparisons of the AOT disposition are, at present, virtually unfeasible.

Internal consistency indices calculated based on the sumscore of the proposed POEM scale were satisfactory. Cronbach's alpha values exceeded the established threshold of .70, justifying the use of the instrument in subsequent research. The internal consistency indices for each subscale are a more complicated matter. In Experiment 1, these values were slightly lower than expected, whereas in Experiment 2, where the dataset was collected from a smaller sample, were slightly higher than expected. Moreover, in both experiments, the factor loading values for the subscales were not high. Low reliability of the subscales is undoubtedly a limitation of the proposed AOT scale adaptation. To overcome this limitation, a similar approach should be taken as in the past AOT adaptation attempts which encountered this very issue – the AOT disposition should be measured by using the scale's sumscore.

An interesting aspect, although not explored in the present study, is the scale's response stability, and the potential effect of the current socio-political circumstances on the participants' reported openness to facts or rigidly standing by their convictions. AOT disposition is facilitated by diverse experiences, and

an active training in perspective-switching and decontextualisation. The present study was conducted in times of significant socio-political changes (war in Ukraine; influx of Ukrainian refugees to Poland; involvement of Polish citizens, particularly students, in humanitarian aid), which could have affected the participants' openness to diverse world-views. A longitudinal study would expand the existing knowledge on the dynamic nature of AOT and the diverse factors affecting this disposition over time. Moreover, this method would allow to verify the POEM scale's response stability.

Comparing the POEM scale results with those of other instruments measuring need for closure, need for cognition or reflective thinking resulted in predicted correlations, thus supporting the scale's validity. The clearest relationships were observed between the AOT disposition and the need for cognitive closure. The stronger were the respondents' convictions indicative of AOT, the lower were their need for closure indicators, especially close-mindedness. As for the relationship between AOT and the need for cognition, weak but positive correlations were observed. The researchers found that AOT is related to the tendency to willingly and actively seek out novel information, but not necessarily to deriving pleasure from engaging with said information to solve problems. The need for cognition scale measures one's cognitive motivation, whereas AOT is considered more of a cognitive strategy or style. Both characteristics are related to reflective thinking, considering how reflective thinking relies on one's capacity for considering different options and re-evaluating pre-existing convictions. Regardless, AOT and need for cognition are not unequivocal concepts, which is why only moderate positive correlations were observed between the present POEM scale adaptation and the Cognitive Reflection Test (CRT; Frederick, 2005). Further arguments supporting the validity of the POEM scale were found when comparing the scale's results with the measures of rational thinking via the Base-rate Tasks (BRT; Kahneman & Tversky, 1973). The participants' greater willingness to think in an open-minded, flexible manner, the more likely they were to actively consider objective premises, whilst ignoring intuitive stereotype-based suggestions.

The conducted adaptation of the AOT scale and the collected POEM scale results from an adult sample appear to be promising. This adaptation may be useful in further research exploring individual differences in cognitive biases or simplified social categorisations. It also seems that the instrument can be applied as a control measure in the development and evaluation of educational interventions, which aim to enhance AOT abilities, as well as to estimate cognitive dispositions which facilitate overcoming maladaptive beliefs in psychotherapy. In the age of radicalising worldviews, growing social polarisation and escalating geo-political conflicts, AOT seems to be crucial to welcome diverse perspectives and open a constructive dialogue. Thus, this cognitive disposition is worth developing and strengthening, and the presented POEM scale can be an easily-applicable tool to capture the disposition. To summarise, the described Polish adaptation of an AOT disposition measurement scale requires further analyses to corroborate its psychometric properties. Regardless, it allows new perspectives for future research into how this disposition is acquired, including factors facilitating its development and its consequences in everyday functioning.

References

- Baron, J. (1993). Why Teach Thinking? – An Essay. Target article with commentary. *Applied Psychology*, 42(3), 191–214.
- Baron, J. (2008). *Thinking and deciding* (4th ed.). New York: Cambridge University Press.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.
- Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, 45, 805–818.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42. DOI: 10.1257/089533005775196732.
- Haran, U., Ritov, I., & Mellers, B. A. (2013). The Role of Actively Open-Minded Thinking in Information Acquisition, Accuracy, and Calibration. *Judgment and Decision Making*, 8(3), 188–201.
- Janssen, E. M., Verkoeijen, P. P. J. L., Heijltjes, A. E. G., Mainhard, T., van Peppen, L. M., & van Gog, T. (2020). Psychometric Properties of the Actively Open-Minded Thinking Scale. *Thinking Skills and Creativity*, 100659.
- Jarymowicz, M. (red.) (2002). *Poza egocentryczną perspektywą widzenia siebie i świata* [Beyond the egocentric perspective of seeing yourself and the world]. Warszawa: Wydawnictwo Instytutu Psychologii PAN.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4), 237–251. DOI: 10.1037/h0034747.
- Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, 83, 26–52. DOI: 10.1016/S0022-0965(02)00121-2.
- Kossowska, M. (2003). Różnice indywidualne w potrzebie poznawczego domknięcia [Individual differences in the need for cognitive closure]. *Przegląd Psychologiczny* [The Review of Psychology], 46(4), 355–374.
- Kossowska, M. (2009). Nowe poznawcze wymiary osobowości a społeczne poznanie i działanie [New cognitive dimensions of personality versus social cognition and Action]. In M. Kofta & M. Kossowska (Eds.), *Psychologia poznania społecznego: Nowe tendencje* [The psychology of social cognition: New trends] (pp. 225–244). Warszawa: Wydawnictwo Naukowe PWN.
- Kossowska, M., Hanusz, K., & Trejtowicz, M. (2012). Różnice indywidualne w potrzebie poznawczego domknięcia [Individual differences in the need for cognitive closure]. *Psychologia Społeczna* [Social Psychology], 7(1), 89–99.
- Lubiewska, K., & Głogowska, K. (2018). Zastosowanie analizy równoważności pomiarowej w badaniach psychologicznych [Application of measurement equivalence analysis in psychological research]. *Polskie Forum Psychologiczne* [Polish Psychological Forum], 23, 330–356.
- Matusz, P. J., Traczyk, J., & Gąsiorowska, A. (2011). Kwestionariusz Potrzeby Poznania – konstrukcja i weryfikacja empiryczna narzędzia mierzącego motywację poznawczą

- [Need for Cognition Questionnaire – construction and empirical validation of a tool measuring cognitive motivation]. *Psychologia Społeczna [Social Psychology]*, 2(17), 113–128.
- Metz, S. E., Baelen, R. N., & Yu, A. (2020). Actively open-minded thinking in American adolescents. *Review of Education*, 8(3), 768–799.
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60(2), 307–317.
- Rachev, N. R., Geiger, S. J., Vintr, J., Kirilova, D., Nabutovsky, A., & Nelsson, J. (2022). Actively open-minded thinking, bullshit receptivity, and susceptibility to framing: Evaluating the dual-process account in North America and Bulgaria. *European Journal of Psychological Assessment*, 38(6), 440–451. DOI: 10.1027/1015-5759/a000685.
- Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.). (1991). *Measures of personality and social psychological attitudes*. Academic Press.
- Sá, W. C., Kelley, C. N., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: Individual differences in the coordination of theory and evidence. *Personality and Individual Differences*, 38, 1149–1161. DOI: 10.1016/j.paid.2004.07.012.
- Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 91(3), 497–510. DOI: 10.1037/0022-0663.91.3.497.
- Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, 89, 342–357.
- Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, 13(3), 225–247.
- Svedholm-Häkkinen, A.M., & Lindeman, M. (2017). Actively open-minded thinking: Development of a shortened scale and disentangling attitudes towards knowledge and people. *Thinking & Reasoning*, 24, 21–40. DOI: 10.1080/13546783.2017.1378723.
- Szydłowski, P. (2015). Pomiar stylu analitycznego przetwarzania informacji. Wstępna weryfikacja narzędzi: Testu Cognitive Reflection Test (CRT) i zadań Base-Rate Tasks (BRT) [Measuring analytical style of information processing. Pre-verification of tools: Cognitive Reflection Test (CRT) and Base-Rate Tasks (BRT)]. *Studia Psychologica*, 15(2), 57–70.
- Szydłowski, P. (2019). *Style poznawcze a religijność [Cognitive styles and religiosity]*. Warszawa: Wydawnictwo Liberi Libri.
- Troldahl, V., & Powell, F. (1965). A short-form dogmatism scale for use in field studies. *Social Forces*, 44, 211–215.
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67, 1049–1062.