# BEYOND BIAS
# INDIVIDUAL AND SOCIAL ASPECTS OF PERFORMANCE

PhD thesis


**Roland Kasek**

Semmelweis University Doctoral School
Mental Health Sciences Division


| | |
|---|---|
| Supervisor: | Imre Lázár, PhD |
| Official reviewers: | Zsolt Datki, PhD |
| | Olivér Nagybányai-Nagy, PhD |

Head of the Complex Examination Committee:
László Rosivall, DSc

Members of the Complex Examination Committee:
Csaba Csíkos, D.Sc
György Bagdy, D.Sc
Fruzsina Albert, Ph.D
Szabolcs Kiss, Ph.D

Budapest, 2025

# 1. Introduction

The quest to comprehend ourselves and the world around us has been on the minds of scholars since the dawn of humanity: from the ancient Greek philosophers through the Empiricists who disputed Descartes' *fundamentum absolutum inconcossum* to Freud's suspicion-infused mythology, to become an *aporia* of a sort, and a cornerstone of the birth of psychology.

The Dunning-Kruger effect is often criticized for its statistical flaws; however, it has significant implications for human behaviour and metacognitive abilities. Associations between self-assessment, basic cognition and metacognitive abilities – metamemory and decisional awareness – avoiding the usual statistical artefacts that arise from including top and bottom performers may support our understanding of the results of our main study. The pandemic and its control measures affected the mental health of the general population regardless of confirmed viral exposure at an unprecedented scale. Studies have often reported a severe increase in mental illnesses and behavioural disorders during the COVID-19 crisis, especially in pathologies related to metacognitive performance, attributed to organic and functional deterioration. Metacognition and facial emotional expressions both play a major role in human social interactions [1,2] as inner narrative and primary communicational display, and both are limited by self-monitoring, control and their interaction with personal and social reference frames.

# 2. Objectives

The aim of our preliminary study was to investigate associations of basic cognition, self-assessment and metacognitive abilities –metamemory and decisional awareness – without the usual statistical artefacts that stem from the bottom and top extremities, leaving less room for mathematical misinterpretation. [3] Adverse experiences influence generations and sexes to different degrees [7], and as resilience [5], coping strategies [6], self-regulation [7], self-talk [8], and self-assessment are closely linked to metacognitive abilities, given the opportunity, we have decided to examine the periodical deviations in cognitive performance, metacognitive effectiveness and self-assessment in terms of generational and sex differences. [9] Our main goal was to develop a setting that allows more space for the observation of the life-like approach to facial emotional expression using artistic and artificial experience as tools to understand the differences in emotional expressions and experiences between man-made and machine generated stimuli.” [10]

**H1** Self-assessment bias is associated with metamemory and decisional awareness

**H2** COVID-19 affected sexes and generations to different extents regarding cognitive performance, metacognitive effectiveness and self-assessment in terms of generational and sex differences.

**H3** Self-confidence (measured in self-assessment bias), personality traits and facial emotional expressions are interrelated.

## 3. Methods

### 3.1. Sample characteristics

The study investigates a sample of self-reported mentally healthy, adult subjects (N=1394) based on availability and willingness who were tested online on their own devices. We then removed the top and bottom performers task by task after transforming their score to a scale ranging from 1 to 6 to match the scaling of the self-assessment, resulting in a final sample of 356 participants – 136 males (age=18-56; mean= 30.4; SD=9.16) and 220 females (age=18-55; mean= 29.9; SD=9.13) – who could therefore be at least 1 point optimistic or pessimistic in their self-assessment on a Likert scale of 6. [3]

The data of working-age subjects (N=1385, age=18-65) – Generation Z (GZ; born between 1997-2012), Millennials (GM; born between 1981-96), Generation X (GX; born between 1965-80), Boomers (GB; born between 1946-64) [11] – were collected in three major periods: before the outbreak (BL), after the first lockdown (W1) and at the end of the crisis (PC) and trimmed to investigate the effects of the pandemic of four generations. [9]

The main "study investigates a sample of 35 mentally healthy native Hungarian adult subjects – 14 men (age=28-51 years; M= 35.1 years) and 21 women (age=19-48; M = 30.8 years) who were attended to a screening of artistic and artificial stimuli, where their facial emotional expressions were recorded and analyzed by artificial intelligence." [10]

### 3.2. Apparatus

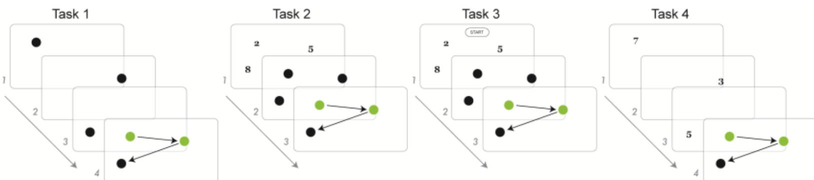


**Figure 1.** Test screens of the four computerized short term memory tasks. Task 1 and Task 4 require the recall of spatiotemporal sequence of dots (T1) or random single digit numerals (T4). Task 2 and 3 simultaneously presented single digit numerals in numeric order in automated (T2) and self-paced (T3) limited-hold settings [10].

The four online short term memory tests (Figure 1) – 2 learning levels (span from 3 to 4 stimuli) and 5 test levels (span from 5 to 9 stimuli) each – consisted of the computerised version of Corsi Block-Tapping Test (CBTT) [12] as a baseline task (Task 1) for its clinical and developmental relevance [13,14], the Inoue-Matsuzawa Masked Memory Task [15] in its original, limited hold (Task 2) and self-paced (Task 3) settings redesigned to measure metamemory faculty of metacognitive ability, while the last task was created based on CBTT to measure how subjects inhibit numerospatiotemporal interference (Task 4) [16,17]. After each task, we have asked the participants to rate their performance compared to their peers on a Likert scale of 6.

### 3.3. Laboratory equipment and setting

"Close-up studio cameras were set up in front of the subjects while they listened to the three literary excerpts in Hungarian with a total duration of 16 minutes interpreted by a professional actor in the following order: first with actor-performed audio (sound only), then with actor-performed experience (sound and image), and finally with artificially generated audio (sound only). After the first screening the order of stimuli was shuffled to avoid artefacts that may arise from order effect confound, despite the fact that we were only going to investigate modality dependent changes and individually specific emotion expressions." [10]

### 3.4. Variables

The subjects reported their age and sex as biological variables. We measured cognitive performance at the learning and test levels separately task by task and accumulated scores in all learning ($P_{LA}$) and all test phases ($P_{TA}$). We logged the reaction times of hits (RT) and false alarms (FA) with millisecond accuracy levels by level and task by task and averaged them into new variables ($RT_{LA}$=learning level average reaction time; $RT_{TA}$=test level average reaction time; $FA_{LA}$= learning level average false alarm reaction time; $FA_{TA}$= test level average false alarm reaction time). Memory performance $P_{T1}$-$P_{T4}$ shows the number of hits across all the levels in the test phases of each task, $P_{TA}$ is the average of them. $RT_{T1}$-$RT_{T4}$ shows the average reaction time of hits across levels in each task, $RT_{TA}$ is the average of all of them. $FA_{T1}$-$FA_{T4}$ are the reaction times of false alarms, $FA_{TA}$ is the average reaction time before mistaken recalls across all tasks. $NP_{T1}$-$NP_{T4}$ are the 1-6 transformed versions of $P_{T1}$-$P_{T4}$ that enabled us to filter out low (1-16%) and high (84-100%) performers by percentage ranges.

The main index of metacognition often referred as metamemory, MM, shows the total score improvement between the self-paced ($IMMMT_{SP}$) and limited hold ($IMMMT_{LH}$) settings of the Inoue-Matsuzawa masked memory tasks. $DA_{T1}$-$DA_{T4}$ are another type of our metacognitive indices that show decisional awareness (DA), the ratio of average reaction times over average false alarm reaction times (DA=FA/RT) task by task, $DA_{TA}$ is the average reaction time ratio across all tasks and shows if a subject was hesitant – the sign of knowing that they do not remember correctly – before false recall.

"Facial emotional expressions were recorded and analysed throughout the entire screening time with a chosen analysis frame rate of 30 frames per second resolution [18]– a unique ability of the artificial intelligence – into cumulated percentage values by basic (happy, surprised, sad, scared, disgusted, angry), other and neutral emotion category and each modality ('HV Angry 28.36' for the total percentage of anger expressed during each Human Video setting), then averaged into a modality-independent facial expression index ('$A_E$ Angry 12.31').

Based on our preliminary exploration of the raw recordings and the collected data, we have identified the need to develop emotional expression ratios as new variables to be able to compare the characteristics across different modalities. Emotional Saturation ($S_E$) shows the ratio of 'dominant' to 'all other' emotions, to explore the subject's emotional span. Emotional Transparency ($T_E$) represents the ratio of 'basic six' to 'other' emotions and reveals how difficult might be to decode the social target's emotional state just by registering their facial expressions.

Since emotions and personality traits are interrelated [e.g. 19,20], based on therapeutic considerations [21] we have grouped basic emotions into three personality categories – Extroversion (happy and surprised), Neuroticism (sad and scared) and Hostility (disgusted and angry) by Izard's theoretical framework and results [22] – to be able to analyse the associations of self-confidence and personality dimensions as defining factors of how subjects empathize, experience and express emotions." [10]

### 3.5. Statistical analyses

Spearman correlations Kruskal-Wallis tests and Brunner-Munzel tests – the most reliable nonparametric procedure for relatively small sample comparisons [23] – were performed with jamovi version 2.5 [24,25]. The threshold for statistical significance was set at $p < .05$.

# 4. Results

## *4.1.* Performance, metacognitive abilities and self-assessment

### *4.1.1 Differences between the original and the filtered sample*

First, we wanted to understand whether there were any differences between our original and filtered samples. Removing the best and worst performers did not affect our original results much, subjects were correctly assessing their performance compared to their peers on average despite filtering extreme performers (Table 1).

**Table 1.** Based on self-assessment ($SA_{T1}$-$SA_{T4}$), Kruskal-Wallis tests show significant results with task performances ($P_{T1}$-$P_{T4}$) in both original and filtered samples. [3]

| | Original sample (N=1394) | | | | Filtered sample (N=356) | | | |
|---|---|---|---|---|---|---|---|---|
| | $\chi^2$ | df | p | $\varepsilon^2$ | $\chi^2$ | df | p | $\varepsilon^2$ |
| $P_{T1}$ x $SA_{T1}$ | 173*** | 5 | <.001 | .1240 | 34.0*** | 5 | <.001 | .0959 |
| $P_{T2}$ x $SA_{T2}$ | 363*** | 5 | <.001 | .2600 | 31.6*** | 5 | <.001 | .0890 |
| $P_{T3}$ x $SA_{T3}$ | 485*** | 5 | <.001 | .3480 | 64.4*** | 5 | <.001 | .1820 |
| $P_{T4}$ x $SA_{T4}$ | 485*** | 5 | <.001 | .3480 | 41.8*** | 5 | <.001 | .1180 |

*Note.* * $p < .05$, ** $p < .01$, *** $p < .001$

### *4.1.2 General associations*

Spearman correlations revealed that the age of the subjects had a positive tendency with reaction time and a positive effect on false alarm reaction time. Memory performance is positively associated with reaction time and decisional awareness, while negatively associated with false alarm reaction time, metamemory and self-assessment bias. Reaction time is positively associated with false alarm reaction time and decisional awareness but negatively affects metamemory.

**Table 2.** Spearman correlations and Bayes Factors ($BF_{10}$) show how performance indices (performance=P; reaction time= RT; false alarm reaction time= FA), self-assessment (=SAB) and metacognitive abilities – metamemory (=MM) and decisional avareness (=DA) are associated with each other. [3]

| | | AGE | $P_{TA}$ | $RT_{TA}$ | $FA_{TA}$ | $DA_{TA}$ | MM | $SAB_{PA}$ |
|---|---|---|---|---|---|---|---|---|
| **$P_{TA}$** | Spearman's rho | -.049 | — | | | | | |
| | p-value | .352 | — | | | | | |
| | $BF_{10}$ | .2892 | | | | | | |
| **$RT_{TA}$** | Spearman's rho | **.126*** | **.488***** | — | | | | |
| | p-value | .017 | <.001 | — | | | | |
| | $BF_{10}$ | 4.6222 | 1.44e +7 | | | | | |
| **$FA_{TA}$** | Spearman's rho | **.170**** | **-.416***** | **.225***** | — | | | |
| | p-value | .001 | <.001 | <.001 | — | | | |
| | $BF_{10}$ | 893.4652 | 67598.4169 | 742.581 | | | | |
| **$DA_{TA}$** | Spearman's rho | -.070 | **.520***** | **.362***** | **-.546***** | — | | |
| | p-value | .185 | <.001 | <.001 | <.001 | — | | |
| | $BF_{10}$ | 1.0544 | 1.48e +22 | 2102.005 | 1.11e +15 | | | |
| **MM** | Spearman's rho | -.073 | **-.171**** | **-.292***** | .060 | **-.187***** | — | |
| | p-value | .171 | .001 | <.001 | .263 | <.001 | — | |
| | $BF_{10}$ | .1690 | 4.1144 | 48330.620 | .0700 | 34.5539 | | |
| **$SAB_{PA}$** | Spearman's rho | .069 | **-.168**** | -.104 | **.117*** | **-.140***** | **.154**** | — |
| | p-value | .196 | .001 | .051 | .027 | .008 | .004 | — |
| | $BF_{10}$ | .3601 | 125.1165 | .169 | .1011 | 6.7650 | 5.044 | |
| **ABS $SAB_{PA}$** | Spearman's rho | .029 | .031 | .062 | .018 | .017 | -.089 | **-.520***** |
| | p-value | .591 | .563 | .242 | .734 | .752 | .093 | <.001 |
| | $BF_{10}$ | .0749 | .0824 | .117 | .0746 | .0907 | .1830 | 6.3e +8 |

***Note.*** *N=356; df=354; * p < .05, ** p < .01, *** p < .001*

False alarm reaction time is in negative association with decisional awareness, and positive tendency with self-assessment bias. Decisional awareness negatively associated with metamemory and self-assessment bias. Metamemory positively correlates with self-assessment bias. Self-assessment bias is also negatively associated with the absolute value of self-assessment bias (Table 2).

### 4.2. Performance and metacognitive abilities during the pandemic

#### *4.2.1 Baseline associations*

Older subjects before COVID-19 scored lower in both the learning ($P_{LA}$) and test phases ($P_{TA}$), were overconfident ($SAB_{PA}$) and tended to be more hesitant before mistakes ($FA_{TA}$ and $DA_{TA}$). The better learners ($P_{LA}$) were better performers ($P_{TA}$), scored slower ($RT_{TA}$) but failed quicker ($FA_{TA}$), improved less in the metamemory task (MM), depreciated their performance ($SAB_{PA}$) and were less aware of their wrong decisions ($DA_{TA}$; Table 3).

**Table 3.** Spearman correlations of the baseline sample (N=392) revealed associations with age, learning and test performance ($P_{LA}$;

$P_{LT}$), reaction time ($RT_{TA}$; $FA_{TA}$) and metacognitive ability (MM; $SAB_{PA}$; $DA_{TA}$). [9]

| BASELINE | | AGE | | $P_{LA}$ | | $P_{TA}$ | | $RT_{TA}$ | | $FA_{TA}$ | | MM | | $SAB_{PA}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P_{LA}$ | Spearman's rho | -.169 | *** | — | | | | | | | | | | | |
| | p value | <.001 | | — | | | | | | | | | | | |
| $P_{TA}$ | Spearman's rho | -.202 | *** | .431 | *** | — | | | | | | | | | |
| | p value | <.001 | | <.001 | | — | | | | | | | | | |
| $RT_{TA}$ | Spearman's rho | -.065 | | .185 | *** | .575 | *** | — | | | | | | | |
| | p value | .201 | | <.001 | | <.001 | | — | | | | | | | |
| $FA_{TA}$ | Spearman's rho | .113 | * | -.154 | ** | -.482 | *** | -.036 | | — | | | | | |
| | p value | .025 | | .002 | | <.001 | | .473 | | — | | | | | |
| MM | Spearman's rho | -.013 | | -.162 | ** | -.133 | ** | -.273 | *** | .198 | *** | — | | | |
| | p value | .792 | | .001 | | .008 | | <.001 | | <.001 | | — | | | |
| $SAB_{PA}$ | Spearman's rho | .153 | ** | -.195 | *** | -.281 | *** | -.142 | ** | .140 | ** | .028 | | — | |
| | p value | .002 | | <.001 | | <.001 | | .005 | | .005 | | .578 | | — | |
| $DA_{TA}$ | Spearman's rho | .109 | * | -.194 | *** | -.714 | *** | -.545 | *** | .704 | *** | .198 | *** | .156 | ** |
| | p value | .031 | | <.001 | | <.001 | | <.001 | | <.001 | | <.001 | | .002 | |

*Note. N= 392; * p < .05, ** p < .01, *** p < .001*

### *4.2.2 Before, during and after COVID-19*

In general, COVID-19 negatively affected short-term memory during the learning phase ($P_{LA}$) more strongly during the first lockdown (W1), whereas scores in the self-paced setting ($IMMMT_{SPT}$) of the tests and metamemory performance (MM) constantly decreased over time in the analysed sample.

Male subjects learned ($RT_{LA}$) and failed ($FA_{TA}$) more slowly but scored quicker ($RT_{TA}$) and more quickly in $CBTT_T$ and less so in $IMMMT_{SPT}$ and MM than did women, who were initially (BL) more confident, suddenly (W1) became more uncertain and then quicker again (PC) in their bad decisions ($DA_{TA}$).

The gender gap further deepened during the pandemic with respect to self-assessment ($SAB_{PA}$), as both sexes were negatively biased at the beginning of the sampling period, but while men improved their original accuracy, women remained almost as self-depreciating as they were (Figure 2).

Generation Z (GZ) could take the lead for a short while during the first wave (W1), whereas Boomers (BM) were the quickest (RT) and least successful (P). Metamemory performance (MM) decreased over time, with a spike in each generation at W1, similar to decisional awareness ($DA_{TA}$); however, the latter increased by the end of restrictions (PCs) for GX and GB.

The original order of self-assessment bias ($SAB_{PA}$) changed the most during the observed period for GB, whose overconfidence at the beginning further increased by W1 but became the least optimistic at the end. The baseline bottom Generation Z improved self-confidence ($SAB_{PA}$) almost to accuracy, whereas Millennials became the most pessimistic at W1, but all returned to the original ranks of bias with a slight overall increase in optimism. In contrast to other generations, GX improved in their self-assessment constantly and became the most accurate group at the end of the sampling period.

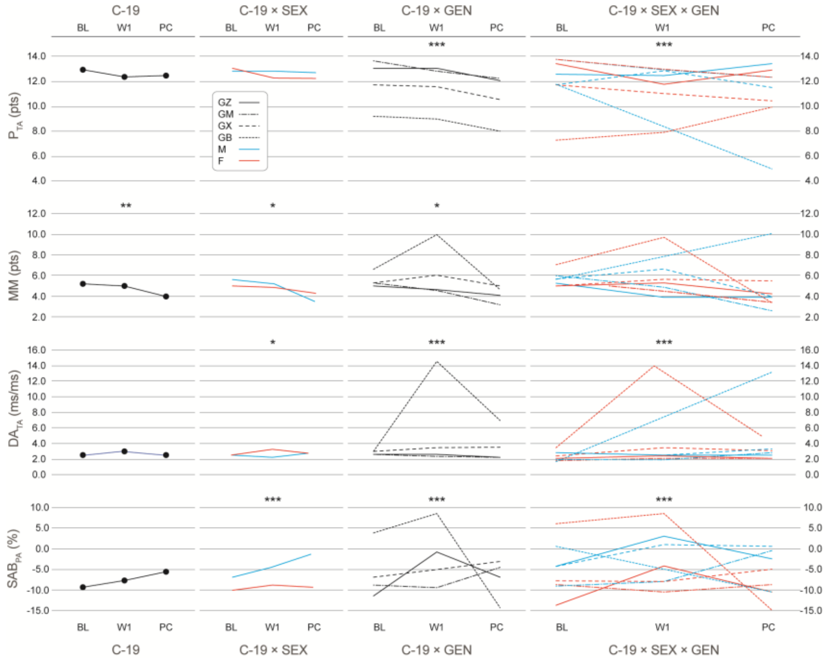


**Figure 2.** Kruskall–Wallis tests revealed differences between subjects by pandemic period in general (C-19), periodical sex (C-19×SEX; Male: blue line; Female: red line) and generational (C-19×GEN) differences and the interaction effects of sex and generations over time (C-19×SEX×GEN) on test performance ($P_{TA}$), metamemory (MM), decisional awareness ($DA_{TA}$) and self-assessment bias ($SAB_{PA}$). [9]

The within- and between-generational sex differences varied the most at the end of the first lockdown (W1), rearranging the rankings of performance ($P_{TA}$), decisional awareness ($DA_{TA}$) and self-assessment bias ($SAB_{PA}$) indices by the end of the pandemic, as expected, on the basis of our previous findings; however, there were no significant disparities in metamemory (MM) performance between the groups.

### 4.3. Performance, metacognitive abilities and emotion

"Brunner-Munzel testing confirmed no evidence of sex differences (Table 4) in our sample regarding age (Age), memory performance (P), self-assessment bias (SAB), self-assessment accuracy ($A_{SAB}$) and polarity ($P_{SAB}$)." [10]

**Table 4.** Brunner-Munzel tests confirms no evidence of differences between basic variables (Age, Memory Performance, Self-Assessment Bias, Accuracy and Polarity) and sex. [10]

| | Statistic | df | p | Relative effect |
|---|---|---|---|---|
| AGE | -1.559 | 31.8 | .129 | .350 |
| P | -0.182 | 29.9 | .857 | .481 |
| SAB | -0.441 | 28.9 | .662 | .459 |
| $A_{SAB}$ | 0.406 | 24.3 | .688 | .524 |
| $P_{SAB}$ | -0.538 | 27.9 | .595 | .452 |

*Note*. $H_a$ $\hat{P}(1 < 2) + ½\hat{P}(1 = 2) \neq ½$

"Compared to audio presentation (sound only), the full artistic experience (sound and video) seemingly rather dimmed than enhanced emotional expressions. Contrary to our expectations, Kruskal-Wallis testing of emotional expressions by modalities – Human Audio, Human Video, Artificial Audio and their modality-independent Average – confirmed no evidence of such phenomena (Table 5)." [10]

**Table 5.** Kruskal-Wallis tests confirm no evidence of differences between each modality (HA, HV, AA) and their average ($A_E$). [10]

| Emotion | $\chi^2$ | df | p | $\varepsilon^2$ |
|---|---|---|---|---|
| Neutral | 0.740 | 3 | .864 | .00532 |
| Happy | 0.741 | 3 | .863 | .00533 |
| Surprised | 2.249 | 3 | .522 | .01618 |
| Sad | 0.608 | 3 | .895 | .00437 |
| Scared | 2.015 | 3 | .569 | .01449 |
| Disgusted | 0.540 | 3 | .910 | .00388 |
| Angry | 3.074 | 3 | .380 | .02212 |
| Other | 5.107 | 3 | .164 | .03674 |
| $S_E$ | 1.801 | 3 | .615 | .01296 |
| $T_E$ | 0.343 | 3 | .952 | .00247 |
| Extroversion | 1.473 | 3 | .689 | .01060 |
| Neuroticism | 0.530 | 3 | .912 | .00382 |
| Hostility | 2.511 | 3 | .473 | .01807 |

"General analysis of the sample revealed that self-assessment bias is moderately associated with the expression of anger ($\chi^2$= 7.3736, df=2, p=0.25, $\varepsilon^2$=0.21687) and hostile tendency ($\chi^2$= 7.4567, df=2, p=0.24, $\varepsilon^2$=0.21931). Surprisingly and quite contrary to our expectations based on a previous phenomenological approach [**Hiba! A hivatkozási forrás nem található.**] the difference between the expressions of other basic emotions were not significant (Table 6)." [10]

"Categorical grouping and Brunner-Munzel testing (Table 7) of our sample by accurate or biased self-assessment ($A_{SAB}$) confirmed evidences that Neutrality (BM=2.97, df=5.58, p=0.014, RE=0.798) as lack the of detectable expressions and

Transparency (BM=2.33, df=26.01, p=0.014, RE=0.702) as the incidence of basic emotional expressions was strongly higher, while Saturation (BM=-2.62, df=9.08, p=0.005, RE=0.210) as the variety of emotional experiences was moderately lower in the biased group." [10]

**Table 6.** Kruskal-Wallis tests confirm evidence of differences between self-assessment bias (SAB) and the average of facial emotional expressions ($A_E$ Angry, $A_E$ Hostility). [10]

| Emotion | $\chi^2$ | df | p | $\varepsilon^2$ |
|---|---|---|---|---|
| Neutral | 3.7451 | 2 | .154 | .11015 |
| Happy | 2.3276 | 2 | .312 | .06846 |
| Surprised | 0.0597 | 2 | .971 | .00176 |
| Sad | 2.0917 | 2 | .351 | .06152 |
| Scared | 1.5000 | 2 | .472 | .04412 |
| Disgusted | 2.6545 | 2 | .265 | .07807 |
| Angry | 7.3736 | 2 | .025* | .21687 |
| Other | 5.3169 | 2 | .070 | .15638 |
| $S_E$ | 3.5787 | 2 | .167 | .10526 |
| $T_E$ | 2.7109 | 2 | .258 | .07973 |
| Extroversion | 1.3546 | 2 | .508 | .03984 |
| Neuroticism | 2.0917 | 2 | .351 | .06152 |
| Hostility | 7.4567 | 2 | .024* | .21931 |

**Table 7.** Brunner-Munzel tests confirm evidence of differences by self-assessment bias accuracy ($A_{SAB}$) and emotional neutrality, emotional saturation ($S_E$) and emotional transparency ($T_E$) across modalities and their average ($A_E$). [10]

| Emotion | Modality | Statistic | df | p | Relative effect |
|---|---|---|---|---|---|
| Neutral | HA | 1.88 | 8.96 | .047* | .694 |
| | HV | 2.87 | 4.20 | .021* | .815 |
| | AA | 4.03 | 5.50 | .004** | .847 |
| | $A_E$ | 2.97 | 5.58 | .014* | .798 |
| $S_E$ | HA | -2.72 | 14.51 | .008** | .258 |
| | HV | -1.87 | 4.27 | .065 | .258 |
| | AA | -2.62 | 7.56 | .016* | .242 |
| | $A_E$ | -3.26 | 9.08 | .005** | .210 |
| $T_E$ | HA | 1.33 | 3.96 | .128 | .694 |
| | HV | 1.14 | 4.18 | .158 | .665 |
| | AA | 4.11 | 12.28 | < .001*** | .819 |
| | $A_E$ | 2.33 | 26.01 | .014* | .702 |

*Note. n=35;* $H_a$ $\hat{P}(1 < 2) + ½\hat{P}(1 = 2) > ½$

"Exploring the differences between modalities, the artificial audio (AA) setting emerged both in emotional Neutrality (BM=4.03, df=5.50, p=0.004, RE=0.847) and Transparency (BM=4.11, df=12.28, p<0.001, RE=0.819) in the biased group with a strong relative effect, while the human audio (HA) setting elicited moderately Saturated expressions (BM=-2.72, df=14.51, p=0.008, RE=0.258) in the accurate group.

Brunner-Munzel tests revealed (Table 8) relatively strong associations of positive bias polarity ($P_{SAB}$) and Hostility across modalities and their Average (BM=2.91, df=33.0, p=0.003, RE=0.732), as a result of strong synergetic tendencies of Disgust (BM=1.74, df=27.0, p=0.047, RE=0.608) and Anger (BM=2.35, df=27.0, p=0.013, RE=0.670).

Altogether, the subjects of the observed sample reacted typically with neutral expressions to the presented stimuli, which can be explained by the reduced attentional capacity and the resulting lower receptivity and empathy due to the fear of the epidemic and the stress arising from the restrictions, as well as by the laboratory situation itself." [10]

**Table 8.** Brunner-Munzel tests confirm evidence of differences by self-assessment bias polarity ($P_{SAB}$) and Disgust, Anger and Hostility across modalities and their average ($A_E$). [10]

| Emotion | Modality | Statistic | df | p | Relative effect |
|---|---|---|---|---|---|
| **Disgusted** | **HA** | **1.74** | **27.0** | **.047*** | **.608** |
| | HV | 1.67 | 29.0 | .053 | .605 |
| | AA | 0.97 | 30.1 | .171 | .552 |
| | $A_E$ | 1.67 | 29.0 | .053 | .605 |
| **Angry** | **HA** | **2.35** | **27.0** | **.013*** | **.670** |
| | HV | 1.67 | 29.0 | .053 | .605 |
| | AA | 1.46 | 25.0 | .078 | .585 |
| | **$A_E$** | **2.40** | **31.9** | **.011*** | **.681** |
| **Hostility** | **HA** | **3.38** | **31.3** | **< .001***** | **.755** |
| | **HV** | **1.93** | **32.9** | **.031*** | **.647** |
| | **AA** | **1.82** | **30.7** | **.039*** | **.632** |
| | **$A_E$** | **2.91** | **33.0** | **.003**** | **.732** |

*Note.* $n=35$, $H_a$ $\hat{P}(1 < 2) + ½\hat{P}(1 = 2) > ½$

# 5. Conclusions

Our results indicate that the psychological explanations of the Dunning-Kruger paradigm regarding metacognition are correct and remind us that challenging, or even depreciating scientific results based on intentionally designed mathematical models are sometimes nothing but throwing the baby out with the water, although may improve research design and lead to stronger conclusions.

We have found strong links between baseline cognitive performance, metacognitive effectiveness—measured in metamemory performance and decisional awareness—and self-assessment, which deviated the most from default after the initial shock in men and women of four generations during the COVID-19 crisis, however further investigation may be required to define whether the associations are direct or under another, yet hidden influence.

We have confirmed that performance and self-assessment ratings were 1) task dependent and 2) in line with the results of previous investigations (e.g.: the dual burden effect), now explored in basic cognition. Subjects in general were able to assess their performance compared to each other correctly despite their biases, and since we have also confirmed that both empirical metacognitive indices – metamemory and decisional awareness – correlate with self-assessment, we may safely say that earlier arguments of previous psychological explanations also stand. Separating causes from effects and understanding what, when, how and why certain changes affected individuals during the pandemic are extremely difficult. Although we were able to confirm the decrease in metacognitive effectiveness that we expected on the basis of the increased prevalence of related mental disorders, the exploration of whether and how and to what extent the neural correlates of metamemory are affected by direct exposure to the virus, immunization, or psychosocial discomfort as a result of the control measures requires further investigation.

Sample characteristics allow us to generalize our results to mentally healthy working-age populations with the usual limitations of academic research [26]; hence, the prevention of an increase in psychosocial discomfort and a decline in mental health arising from such synergy of multiple negative events requires more thoughtful communication and improved action preparedness from future crisis managers, both globally and locally.

# 6. Bibliography of the candidate's publications

### 6.1. Publications related to the thesis:

Kasek, R., Sepsi, E. & Lázár, I. (2025). Overconfident, but angry at least. AI-Based investigation of facial emotional expressions and self-assessment bias in human adults. *BMC psychology, 13*(1), 1-9.; https://doi.org/10.1186/s40359-025-02590-7; PMid:40065465 PMCid:PMC11895137 IF: 3.0

Kasek, R., Sepsi, E. & Lázár, I. (2024). Teljesítmény, magabiztosság és mimika: az arcon megjelenő érzelem-kifejezés és a metakognitív képességek vizsgálata *In: Lázár Imre (szerk.) Sorsunk hálóiban: Szocioszomatikus hálózatok nyomában* Budapest, Magyarország: L'Harmattan Kiadó, Károli Gáspár Református Egyetem, pp. 257-269.

Bozsányi, S., Czurkó, N., Becske, M., Kasek, R., Lázár, B. K., Boostani, M., ... & Lázár, I. (2023). Assessment of Frontal Hemispherical Lateralization in Plaque Psoriasis and Atopic Dermatitis. *Journal of Clinical Medicine*, *12*(13), 4194. IF:3.0

### 6.2. Publications not related to the thesis:

Sepsi, E., Kasek, R. & Lázár, I. (2021). Művészeti befogadás pszichofiziológiai vizsgálata Noldus FaceReader8 segítségével. *Károli Gáspár Református Egyetem L'Harmattan Kiadó,* Budapest, pp 212-230.

Kasek, R. (2019). Rewarding self-liberation: The use of NLPt in lifestyle improvement and weight-control. *Journal of Experiential Psychotherapy*. 22 (1(85)): 11-20.

Kasek, R. (2014) A good therapist amuses the client - so that he heals himself. *Psychologia Hungarica Caroliensis*, *2.* / 1.

Surányi, Z., Rózsa, S., Kasek, R., & Aluja, A. (2014). Person-oriented analyses in the Catalan–Hungarian comparison of the Zuckerman–Kuhlman–Aluja Personality Questionnaire (ZKA-PQ). *Personality and Individual Differences*, *60*, S39-S40.

**∑IF: 6.0**

# 7. References


1. Bruner, J. S. (1964). The course of cognitive growth. American psychologist, 19(1), 1.; https://doi.org/10.1037/h0044160

2. Flavell, J. (1979). Metacognition and Cognitive Monitoring, American Psychologist, Vol. 34, No. 10, 906-911 https://doi.org/10.1037//0003-066X.34.10.906

3. Kasek, R., Fejes E., Bogdányi, R., Gergely, B., Takács, S. & Lázár, I. (2025) Metacognitive abilities and the Dunning-Kruger effect in basic cognition. BMC Psychology (in review)

4. Harari, T.T.E., Sela, Y. and Bareket-Bojmel, L. (2023). Gen Z during the COVID-19 crisis: A comparative analysis of the differences between Gen Z and Gen X in resilience, values and attitudes. Current Psychology, 42(28), pp.24223-24232. https://doi.org/10.1007/s12144-022-03501-4 PMid:35967492 PMCid:PMC9362676

5. Hatami, M. & Hobi, M. B. (2022). The role of metacognition, tolerance of ambiguity and psychological flexibility in predicting resilience. Journal of psychological science, 21(117), 1799-1814. https://doi.org/10.52547/JPS.21.117.1799

6. Zhang, H., Lin, Z., Li, T., Kong, M., Li, J., Wu, W. & He, Z. (2023). The association between metacognition, coping style and fear of cancer recurrence among postmastectomy patients: a cross sectional study. European Journal of Oncology Nursing, 67, 102460. https://doi.org/10.1016/j.ejon.2023.102460 PMid:37951072

7. Zimmerman, B. J. & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In Handbook of metacognition in education (pp. 299-315). Routledge.

8. Morin, A. (2005). Possible links between self-awareness and inner speech theoretical background, underlying mechanisms, and empirical evidence. Journal of Consciousness Studies, 12(4-5), 115-134.

9. Kasek, R., Fejes E., Bogdányi, R., Gergely, B., Takács, S. & Lázár, I. (2025) Intergenerational and sex differences in metacognitive effectiveness during COVID-19. Scientific Reports (in review)

10. Kasek, R., Sepsi, E. & Lázár, I. (2025). Overconfident, but angry at least. AI-Based investigation of facial emotional expressions and self-assessment bias in human adults. BMC psychology, 13(1), 1-9.; https://doi.org/10.1186/s40359-025-02590-7; PMid:40065465 PMCid:PMC11895137

11. Pew Research Center (2019). Defining generations: where millennials end and generation Z begins, available at: https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/

12. Corsi, P. M. (1972). Human memory and the medial temporal region of the brain.

13. Kessels, R.P., Van Zandvoort, M.J., Postma, A., Kappelle, L.J. & De Haan, E.H. (2000). The Corsi block-tapping task: standardization and normative data. Applied neuropsychology, 7(4), pp.252-258. https://doi.org/10.1207/S15324826AN0704_8 PMid:11296689

14. Farrell Pagulayan, K., Busch, R.M., Medina, K.L., Bartok, J.A. & Krikorian, R. (2006). Developmental normative data for the Corsi Block-tapping task. Journal of clinical and experimental neuropsychology, 28(6), pp.1043-1052. https://doi.org/10.1080/13803390500350977 PMid:16822742

15. Inoue, S. & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. Current Biology, 17(23), R1004-R1005. https://doi.org/10.1016/j.cub.2007.10.027 PMid:18054758

16. Kasek, R. & Albu, M., (2014). Are visual types basically visual? DuCog VI Language and conceptual development, Dubrovnik, Croatia, May 23 (in Program and abstracts, pp. 41.)

17. Kasek, R., Takacs, S. & Konya, A., (2015). Visual expertise in immediate spatial localisation tasks. Cognitive Processing, 16, pp. S91.

18. Borges, N., Lindblom, L., Clarke, B., Gander, A. & Lowe, R. (2019). Classifying confusion: autodetection of communicative misunderstandings using facial action units. In *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)* (pp. 401-406). IEEE. https://doi.org/10.1109/ACIIW.2019.8925037; PMCid:PMC6898424

19. Reisenzein, R. & Weber, H. (2009). Personality and emotion. The Cambridge handbook of personality psychology, 1, 54-71. https://doi.org/10.1017/CBO9780511596544.007

20. Martínez-Tejada, L. A., Maruyama, Y., Yoshimura, N. & Koike, Y. (2020). Analysis of personality and EEG features in emotion recognition using machine learning techniques to classify arousal and valence labels. Machine Learning and Knowledge Extraction, 2(2), 7. https://doi.org/10.3390/make2020007

21. Harkness, A. R., Reynolds, S. M. & Lilienfeld, S. O. (2014). A review of systems for psychology and psychiatry: Adaptive systems, personality psychopathology

five (PSY–5), and the DSM–5. Journal of personality assessment, 96(2), 121-139. https://doi.org/10.1080/00223891.2013.823438, PMid:23941204

22. Izard, C. E., Libero, D. Z., Putnam, P. & Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. Journal of personality and social psychology, 64(5), 847. https://doi.org/10.1037/0022-3514.64.5.847, PMid:8505713

23. Karch, J. D. (2021). Psychologists should use Brunner-Munzel's instead of Mann-Whitney's U test as the default nonpara-metric procedure. Advances in Methods and Practices in Psychological Science, 4(2), 2515245921999602. https://doi.org/10.1177/2515245921999602

24. Core Team (2023). R: A Language and environment for statistical computing. (Version 4.3) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2024-01-09).

25. The jamovi project (2024). jamovi. (Version 2.5) [Computer Software]. Retrieved from https://www.jamovi.org.

26. Henrich, J., Heine, S.J. & Norenzayan, A. (2010). The weirdest people in the world?. Behavioral and brain sciences, 33(2-3), pp.61-83. https://doi.org/10.1017/S0140525X0999152X PMid:20550733

## 8. Acknowledgements

The author would like to thank first of all whose insights, encouragements and support were invaluable to reach the final form of the study – Dr Anikó Kónya, Réka Bogdányi, Emese Fejes, Dr Szabolcs Takács, Dr Enikő Sepsi, Dr Imre Lázár, Ádám Ritter, József Hornyik, Dr Szilveszter E Vizi, Bence Gergely, Dr Tamás Dömötör Szalai, Dr Róbert Bódizs, Csilla Raduch, Bence Döbrössy, Dr Edmund Girasek, Dr Mónika Albu, Dr Judit Páli, Dr Zsuzsanna Kövi, Dr Marianna Török and Márta Buczkó – the participants, the colleagues and the students who contributed in the research. The authors were provided with the screening materials, the recording equipment, and post-production services by the Art Sciences and Media Education Division of Benda Kálmán College for Advanced Studies at Károli University.