PRIMENJENA
PSIHOLOGIJA

No1, 2023

Contents

Primenjena psihologija
Vol. 16, No. 1, pp. 3-28, 2023

Research Article

A Rasch analysis of the International Personality Item Pool Big Five Markers Questionnaire: Is longer better?

Hanif Akhtar [1,2] ✉ and Bambang Sumintono[3]

[1] *Faculty of Psychology, Universitas Muhammadiyah Malang*
[2] *Doctoral School of Psychology, ELTE Eötvös Loránd University*
[3] *Faculty of Education, Universitas Islam Internasional Indonesia*

ABSTRACT

The 50-item International Personality Item Pool version of the Big Five Markers (IPIP-BFM) is an open-source and widely used measure of the big five personality traits. A short version of this measure (IPIP-BFM-25) has been developed using the classical test theory approach. No study was performed to examine the psychometric properties of a longer and shorter version of IPIP-BFM Indonesia using modern test theory. This study aimed to evaluate the psychometric properties of the Indonesian version of IPIP-BFM as well as IPIP-BFM-25 using Rasch analysis. The analysis was conducted in order to test their dimensionality, rating scale functioning, item properties, person responses, targeting, reliability, and item bias on 1003 Indonesian samples. The findings showed that both IPIP-BFM and IPIP-BFM-25 Indonesia have some adequate psychometric properties, especially regarding category function, item properties, reliability, and item bias. However, the emotional stability and intellect scales did not meet the assumption of unidimensionality, and all items on the scales were too easy to endorse by participants. In general, longer measures outperformed shorter measures in terms of person separation and reliability. Further testing and refinement must be conducted.

Keywords: Five-Factor model, IPIP Big Five Markers, Rasch analysis, DIF

UDC: 159.923.072
DOI: 10.19090/pp.v16i1.2401
Received: 11.04.2022.
Revised: 16.06.2022.
Accepted: 18.07.2022.

✉ Corresponding author email: hanifakhtar@umm.ac.id

Introduction

The Big Five personality trait is the most widely recognized personality model in psychology to date. This model explains that there are five main factors in an individual's personality; namely extraversion, agreeableness, conscientiousness, emotional stability, and intellect (Goldberg, 1992). Since many researchers use the Big Five model as a predictor of several outcomes in their studies, many instruments have been developed, such as the Big Five Inventory (BFI; John & Srivastava, 1999), the NEO PI-R (Costa & McCrae, 1995), the Ten-Item Personality Inventory (TIPI; Gosling et al., 2003), and the Trait Descriptive Adjective (TDA; Goldberg, 1992).

One of the most popular instruments to measure The Big Five personality traits is the 50-item International Personality Item Pool representation of the Goldberg (1992) markers for the Big-Five factor structure, hereinafter referred to as IPIP-BFM. This instrument is available on the International Personality Item Pool (IPIP) website and is free to use by anyone (Goldberg et al., 2006). The main advantage of IPIP-BFM is that this scale is cost-free and widely used by researchers, making the research findings comparable to existing studies. This measure has been adapted and validated in numerous different countries, such as Croatia (Mlačić & Goldberg, 2007), Poland (Strus et al., 2017), Scotland (Gow, Whiteman, Pattie, & Deary, 2005), China (Zheng et al., 2008), New Zealand (Guenole & Chernyshenko, 2005), Portugal (Oliveira, 2017), and Indonesia (Akhtar & Azwar, 2018).

Akhtar and Azwar (2018) have adapted IPIP-BFM in Indonesian samples using forward-backward-translation methods. Moreover, they developed a short version of the measure, IPIP-BFM-25, using the classical test theory (CCT) approach. The items for IPIP-BFM-25 were selected from the parent measure to maximize loading on the primary factor and minimize the cross-loading factor from the exploratory factor analysis (EFA). The study indicated that the Indonesian version of both IPIP-BFM and IPIP-BFM-25 has adequate Cronbach's alpha (ranging from .70 to .86), satisfactory factorial validity, and has a high correlation with BFI. However, due to its intrinsic limits,

the CTT cannot protect the IPIP-BFM from psychometric criticism. The CTT, for example, cannot determine the response category functioning of IPIP-BFM (see Kean et al., 2017; Petrillo et al., 2015). In addition, selecting items based on the loading factor only may narrow item content, restricting the breadth of the item content on the full scale (Smith et al., 2000). Kline (2000) also noted that the main drawback of using EFA is the tendency to select items that are essentially paraphrases of each other in order to form a factor based on correlational analysis. Although redundant in the meaning, those items will have a high correlation and thus will have high loading on a factor.

Given the extensive usage of IPIP-BFM, it is critical to evaluate its structural validity using modern test theory, which avoids many of the CTT's flaws (see Bond & Fox, 2015). Rasch rating scale model can be beneficial for several reasons. First, Rasch models can look at how people of various abilities respond to a set of IPIP-BFM items (Raykov & Marcoulides, 2015). Rasch analysis can reveal the relative endorsability of items using item-person maps, which display both items and persons on the same logit scale according to item difficulty estimates (Bond & Fox, 2015). Second, Rasch analysis can convert the ordinal scale data obtained using the IPIP-BFM into linear, interval scale data using the raw score-to-logit transformation (see Andrich & Marais, 2019). Third, the Rasch model enables the examination of the functioning of the ordered response categories. In this respect, it is assumed that the category thresholds will be arranged in value in the same order as the response categories (Adams et al., 2012). Fourth, Rasch model measurement analysis provides reliability figures for items in the measurement instrument and persons. Rasch model analysis uses the separation to measure not only the person reliability but also item reliability (Fisher, 1992). In order to achieve this aim, the current study employed Rasch analysis to evaluate the psychometric properties of the IPIP-BFM.

Previous studies aimed to validate the IPIP-BFM using the Rasch model were conducted by Apple and Neff (2012) in a Japanese sample. The results of their study indicated the possible existence of additional factors within the Intellect and Agreeableness factors, as well as additional item fit problems within each hypothesized construct. Moreover, emotional stability

items had a moderate floor effect, indicating that some items for this construct may have been too difficult for participants to endorse. On the other hand, the Agreeableness item-person map suggests a strong ceiling effect, indicating that the items for this particular construct were too easily endorsable by the participants. These findings could be a reference point to examine the appropriateness of the American-developed five-factor model personality trait instrument for measuring an Indonesian population.

Currently, there is no study to analyze the IPIP-BFM using the Rasch model in Indonesia. Moreover, although a short version of IPIP-BFM has been performed well in EFA and Cronbach's alpha analysis, there is no evidence that the chosen items are not redundant. Considering the benefits of Rasch analysis, this study aims to fill the gap in analyzing the psychometric properties of the Indonesian IPIP-BFM using Rasch analysis. Thus, this study aims to test IPIP-BFM dimensionality, rating scale functioning, item properties, person responses, targeting, reliability, and differential item functioning (DIF) across gender using the Rasch model. Second, this study aims to compare the targeting and reliability of the 50-item and 25-item versions of IPIP-BFM Indonesia.

Methods

Participants

A total of 1019 participants participated in this study. An online survey was used for the data collection of this study. Participants were recruited during 2021 using various strategies, including advertisements on social media (Facebook, Instagram, and WhatsApp groups) and encouraging our colleagues to share the advertisements. After data screening to filter participants with careless responses (*e.g.*, responding '*Strongly Agree*' to all items despite reverse wording), the final dataset contains the data from 1003 participants. The total sample consisted of 409 (40.8%) males and 594 (59.2%) females aged 15 to 50 (M = 19.9, SD = 4.8). The education of the sample consisted of 334 (33.3%) junior high school, 526 (52.4%) senior high school, 97

(9.6%) bachelor, 37 (3.6%) master, and 9 (0.8%) doctoral. On the first page of the questionnaire, it was stated that it was strictly anonymous and voluntary to address ethical concerns. Thus, by completing the questionnaire, the respondents have given their consent. All procedures have been approved by the ethical committee of the Faculty of Psychology, Universitas Muhammadiyah Malang.

Instrument

The Indonesian version of the IPIP-BFM was administered to all participants. The IPIP-BFM Indonesia contains 50 items that are used to measure five personality traits: extraversion, emotional stability, agreeableness, conscientiousness, and intellect. Each of the five personality traits is evaluated using a 10-item response, with each item rated using a 5-point Likert rating scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). IPIP-BFM-25 was a short version of IPIP-BFM that contains 25 items. The items have been adapted and validated into Indonesian by Akhtar and Azwar (2018) using the forward-backward-translation method.

Analysis procedures

The Rasch model analysis was performed using the Rasch Rating Scale Model (RSM), an extension of the Rasch model for polytomous items (Andrich, 1978; Andrich & Marais, 2019) in Winsteps 3.73 computer software (Linacre, 2012). The analysis was conducted for each dimension separately. Following Lim and colleagues' (2009) recommendations, this study conducted seven key Rasch evaluations for validity evidence: dimensionality, rating scale functioning, item properties, person responses, targeting, reliability, and item bias. The data analysis processes and a summary of the objectives of each process are shown in Figure 1. The scale level properties (dimensionality, targeting, and reliability) of IPIP-BFM-25 were compared. The analysis for IPIP-BFM-25 was conducted using the same dataset but only involving items from IPIP-BFM-25.

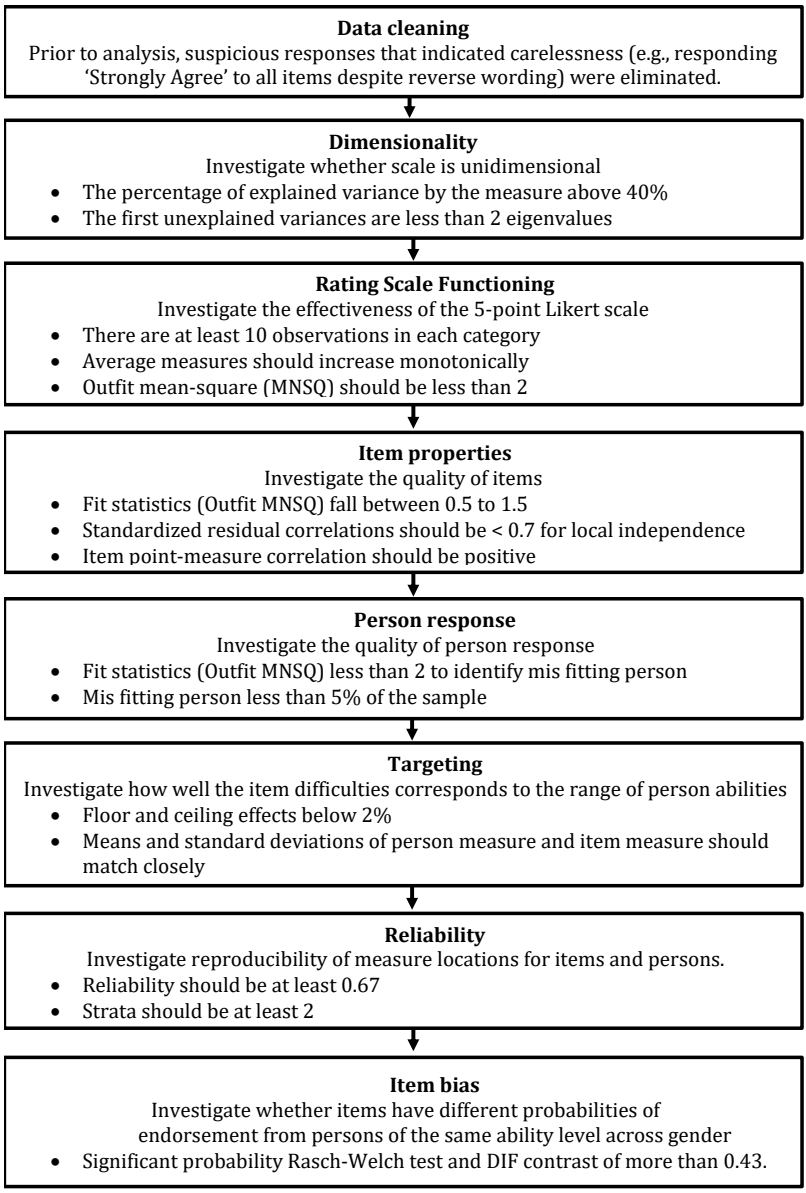

Figure 1. Data analyses flow

Results

Dimensionality

Dimensionality analysis investigated whether all scales are unidimensional. The unidimensionality was investigated using the Principal Components Analysis of Rasch measures and residuals. The scale is fundamentally unidimensional if the percentage of explained variance by the measure is greater than 40%, and the first unexplained variances are less than 2 eigenvalues (Linacre, 2012).

For the IPIP-BFM, the analysis showed that the extraversion, agreeableness, and conscientiousness scales met the assumptions of unidimensionality. However, the emotional stability and intellect scales had eigenvalues of unexplained variance larger than 2. The raw variance explained by measures for extraversion, conscientiousness, agreeableness, emotional stability, and intellect was 48.7%, 41.6%, 45.3%, 49.3%, and 41.0%, respectively. The unexplained variances in the first contrast for extraversion, conscientiousness, agreeableness, emotional stability, and intellect were 1.9, 1.7, 1.9, 2.1, and 2.1, respectively.

For the IPIP-BFM-25, the analysis showed better unidimensionality with the extraversion, agreeableness, conscientiousness, and intellect scales met the assumptions of unidimensionality, but not for the emotional stability scales. The raw variance explained by measures for extraversion, conscientiousness, agreeableness, emotional stability, and intellect was 53.6%, 47.0%, 55.4%, 54.4%, and 49.0%, respectively. The unexplained variances in the first contrast for extraversion, conscientiousness, agreeableness, emotional stability, and intellect were 1.6, 1.6, 1.6, 2.1, and 1.8, respectively.

Table 1

Rating Scale Model category statistics for the total sample (N=1,003)

IPIP-BFM dimension	Category	Frequency	Percentage	Average measure	Outfit MNSQ	Andrich threshold
Extraversion	1 (SD)	571	6%	-1.87	1.13	NONE
	2 (D)	1916	19%	-.094	0.97	-2.60
	3 (N)	3824	38%	-0.02	0.87	-1.15
	4 (A)	3044	30%	0.97	1.01	0.68
	5 (SA)	675	7%	2.13	1.09	3.07
Agreeableness	1 (SD)	96	1%	-0.57	1.85	NONE
	2 (D)	512	5%	-0.16	1.20	-2.37
	3 (N)	2536	25%	0.41	0.84	-1.47
	4 (A)	5295	53%	1.61	0.89	0.29
	5 (SA)	1591	16%	3.13	1.02	3.55
Conscientiousness	1 (SD)	215	2%	-1.08	1.54	NONE
	2 (D)	1198	12%	-0.45	1.11	-2.64
	3 (N)	3348	33%	0.32	0.84	-1.06
	4 (A)	3974	40%	1.36	0.89	0.67
	5 (SA)	1295	13%	2.45	1.02	3.02
Emotional stability	1 (SD)	861	9%	-2.45	1.05	NONE
	2 (D)	3118	31%	-1.14	0.97	-3.09
	3 (N)	3359	33%	-0.12	0.84	-0.70
	4 (A)	2349	23%	0.77	1.04	0.66
	5 (SA)	343	3%	1.51	1.27	3.12
Intellect	1 (SD)	133	1%	-0.73	1.26	NONE
	2 (D)	1149	12%	-0.27	1.11	-2.83
	3 (N)	3894	39%	0.31	0.87	-1.20
	4 (A)	3742	37%	1.36	0.91	0.86
	5 (SA)	1112	11%	2.51	1.04	3.16

Note. SD = Strongly Disagree, D = Disagree, N = Neutral, A = Agree, SA = Strongly Agree

Rating scale functioning

Rating scale functioning investigated the effectiveness of the 5-point Likert scale (Van Zile-Tamsen, 2017) for each dimension. Several essential criteria suggested by Linacre (2002b, 2004) to diagnose the effectiveness of

the rating scale used are a) there are at least ten observations in each category; b) average measures should increase monotonically; c) outfit mean-square (MNSQ) should be less than 2.

Results illustrated that there were no disordered thresholds. At least 10 individuals chose each category in the five scales of IPIP-BFM Indonesia. The average measure by category moved up monotonically with the rating scale. All categories had outfit mean squares of less than 2.0 in each step. Hence, all the Likert scale categories were well functioning and fully understood by respondents. The rating scale model category is presented in Table 1.

Item properties

Item properties analysis examines the quality of items. Item fit is considered good if the fit statistics for outfit MNSQ fall between 0.5 and 1.5 (Linacre, 2002a). Standardized residual correlations represent item local dependence, with correlations greater than 0.7 indicating that two items share more than half of their random variance and only one has to be retained (Linacre, 2012). Meanwhile, the item point-measure correlation index should be positive, indicating no polarity. Negative correlations indicate reverse-coded item miskeying, whereas near-zero correlations indicate items that are very easy or difficult to endorse or that measure a different construct. A correlation of less than 0.4 can be used to identify items for wording investigation (Wolfe & Smith, 2007).

Out of the 50 items of IPIP-BFM, 49 had outfit mean squares that fell between 0.5 and 1.5, and one had outfit mean squares above 1.5. Furthermore, all of the items had point-measure correlations greater than 0.4. Item I8 was the most difficult item to endorse, and item A3 was the easiest item to endorse among all items. The complete item fit information for the five scales of IPIP-BFM is shown in Table 3. The largest standardized residual correlations ranged from -.30 to .22 for extraversion, from -.32 to .11 for agreeableness, from -.33 to .19 for conscientiousness, from -.30 to .48 for emotional stability, and

from -.36 to .24 for intellect, which indicated that the items could be viewed as locally independent.

Table 2

Item properties for the five scales of IPIP-BFM

Item	Measure	SE	Outfit MNSQ	PTME
Extraversion				
E1 Am the life of the party	-0.53	0.05	0.80	0.67
E2 Don't talk a lot	-0.11	0.04	0.95	0.70
E3 Feel comfortable around people	-0.52	0.05	0.96	0.64
E4 Keep in the background	0.63	0.04	1.27	0.62
E5 Start conversations	-0.66	0.05	0.91	0.65
E6 Have little to say	-0.05	0.04	0.96	0.73
E7 Talk to a lot of different people at parties	-0.34	0.05	0.75	0.76
E8 Don't like to draw attention to myself	0.69	0.04	0.98	0.66
E9 Don't mind being the centre of attention	0.45	0.04	1.29	0.59
E10 Am quiet around strangers	0.44	0.04	1.15	0.68
Agreeableness				
A1 Feel little concern for others	0.49	0.05	1.37	0.62
A2 Am interested in people	-0.64	0.06	0.74	0.68
A3 Insult people	-0.92	0.06	1.19	0.54
A4 Sympathise with others' feelings	-0.60	0.06	0.78	0.64
A5 Am not interested in other people's problems	0.56	0.05	1.10	0.62
A6 Have a soft heart	0.64	0.05	1.35	0.59
A7 Am not really interested in others	0.32	0.05	0.98	0.63
A8 Take time out for others	0.29	0.05	0.98	0.55
A9 Feel others' emotions	-0.18	0.05	0.78	0.63
A10 Make people feel at ease	0.04	0.05	0.81	0.66
Conscientiousness				
C1 Am always prepared	-0.50	0.05	0.87	0.67
C2 Leave my belongings around	-0.03	0.05	1.38	0.62

C3 Pay attention to details	0.00	0.05	0.91	0.67
C4 Make a mess of things	-0.37	0.05	0.98	0.66
C5 Get chores done right away	0.80	0.04	0.97	0.65
C6 Often forget to put things back in their proper place	0.52	0.04	1.36	0.63
C7 Like order	-0.17	0.05	0.91	0.69
C8 Shirk my duties	-0.43	0.05	1.16	0.60
C9 Follow a schedule	-0.04	0.05	0.72	0.71
C10 Am exacting in my work	0.21	0.05	0.82	0.65
Emotional stability				
ES1 Get stressed out easily	-0.11	0.04	1.16	0.68
ES2 Am relaxed most of the time	-0.02	0.04	1.00	0.66
ES3 Worry about things	0.47	0.05	1.13	0.65
ES4 Seldom feel blue	0.05	0.04	0.90	0.71
ES5 Am easily disturbed	0.08	0.04	1.01	0.69
ES6 Get upset easily	-0.15	0.04	0.87	0.73
ES7 Change my mood a lot	0.47	0.05	0.89	0.71
ES8 Have frequent mood swings	0.03	0.04	0.89	0.75
ES9 Get irritated easily	-0.27	0.04	0.90	0.72
ES10 Often feel blue	-0.54	0.04	1.34	0.63
Intellect				
I1 Have a rich vocabulary	0.25	0.05	0.87	0.64
I2 Have difficulty understanding abstract ideas	0.73	0.05	1.02	0.64
I3 Have a vivid imagination	-0.56	0.05	1.09	0.64
I4 Am not interested in abstract ideas	0.20	0.05	1.03	0.61
I5 Have excellent ideas	-0.31	0.05	0.78	0.65
I6 Do not have a good imagination	-0.70	0.05	0.85	0.63
I7 Am quick to understand things	-0.08	0.05	0.96	0.60
I8 Use difficult words	*0.95*	*0.05*	*1.62*	*0.51*
I9 Spend time reflecting on things	-0.30	0.05	1.03	0.57
I10 Am full of ideas	-0.18	0.05	0.78	0.63

Note: italic item (I8) misfit for outfit MNSQ criteria. PTME = Point-measure correlation.

Table 3

Summary Statistics for Five Scales of IPIP-BFM

Construct	Item measure		Person measure		Person MNSQ	Outfit	Item MNSQ	Outfit	Item reliability	Person reliability	Item separation	Person Separation	Cronbach's alpha
	Mean	SD	Mean	SD	Mean	SD	Mean	SD					
IPIP-BFM													
Extraversion	0	0.49	0.16	1.33	1.00	0.78	1.00	0.17	0.99	0.85	10.55	2.34	0.87
Agreeableness	0	0.53	1.47	1.46	1.01	0.87	1.01	0.23	0.99	0.8	9.68	2.02	0.83
Conscientiousness	0	0.40	0.91	1.34	1.01	0.87	1.01	0.21	0.99	0.83	8.21	2.19	0.86
Emotional stability	0	0.29	-0.40	1.46	1.01	0.85	1.01	0.15	0.97	0.86	6.21	2.48	0.89
Intellect	0	0.51	0.89	1.26	1.00	0.8	1.00	0.23	0.99	0.8	10.36	2.00	0.82
IPIP-BFM-25													
Extraversion	0	0.47	0.43	1.61	1.00	0.92	1.00	0.19	0.99	0.76	9.15	1.76	0.77
Agreeableness	0	0.47	1.79	1.80	1.00	1.09	1.00	0.24	0.98	0.69	7.69	1.48	0.72
Conscientiousness	0	0.60	1.19	1.85	1.00	0.95	1.00	0.14	0.99	0.77	10.80	1.81	0.79
Emotional stability	0	0.34	-0.56	0.34	1.00	0.96	1.00	0.12	0.98	0.78	6.87	1.90	0.80
Intellect	0	0.62	1.20	1.61	1.00	0.93	1.00	0.12	0.99	0.71	11.65	1.56	0.72

Person response

Person response analysis examines the quality of a person's response. Person response was indicated by person fit statistics. Smith and Wolfe (2007) proposed using an outfit MNSQ value of 2 to identify misfits, which is expected to be less than 5% of the sample. Person misfits may indicate low-quality self-report data, for example, due to careless responses or misunderstanding of items (Kottorp, et al., 2003).

Out of the 1,003 participants, 911 (90.82%) reported an acceptable fit on the extraversion scale, 887 (88.43%) on the agreeableness scale, 898 (89.53%) on the conscientiousness scale, 899 (89.63%) on the emotional stability scale, and 896 (89.33%) on the intellect scale. At all scales, there were approximately 10% of misfitting persons. Although the percentage of misfitting persons was noticeably higher than expected, the average mean person's outfit MNSQ was at the ideal of approximately 1 (Table 3). The nature of the anonymous online survey might explain the occurrence of careless responses.

Targeting

Item-person targeting examined how well the distribution of item difficulties corresponds to the range of persons' abilities (Linacre, 2012). Floor and ceiling effects below 1% are very good, and effects between 1% and 2% are good (Fisher, 2007). In addition, the means and standard deviations of the person and item measures should match closely (Bond & Fox, 2015).

For IPIP-BFM, two participants (0.19%) scored maximum scores on the extraversion scale, ten participants (0.99%) on the agreeableness scale, six participants (0.19%) on the conscientiousness scale, six participants (0.59%) on the emotional stability scale, and seven participants (0.69%) on the intellect scale. Less than 1% of participants scored the maximum score of the five scales, and none of the participants scored minimum scores on the five scales of IPIP-BFM, indicating very good ceiling and floor effects. The means and SDs of the person and item measure are presented in Table 3.

For IPIP-BFM-25, six participants (0.59%) scored maximum scores on the extraversion scale, seventeen participants (1.69%) on the agreeableness scale, nine participants (0.89%) on the conscientiousness scale, fifteen participants (1.49%) on the emotional stability scale, and fourteen participants (1,39%) on the intellect scale. Less than 2% of participants scored the maximum score of the five scales, and none of the participants scored minimum scores on the five scales of IPIP-BFM, indicating good ceiling and floor effects. The means and SDs of the person and item measure are presented in Table 3.

Item-person maps based on item difficulty measures were generated for each dimension to further examine the item-participant match. Figures 2 and 3 show the distribution of the person and item measures. The construct is laid out vertically, with the most difficult items at the bottom and the highest-abled person at the top. As shown in the figure, the difficulty of the items was lower than the levels of conscientiousness, agreeableness, and intellect of the participants, while the difficulty of the items was equal to the middle level of extraversion and emotional stability of the participants. Special attention should be directed to the agreeableness, conscientiousness, and intellect scales, as a majority of participants were above the range of the items, indicating that the items were too easy to endorse.

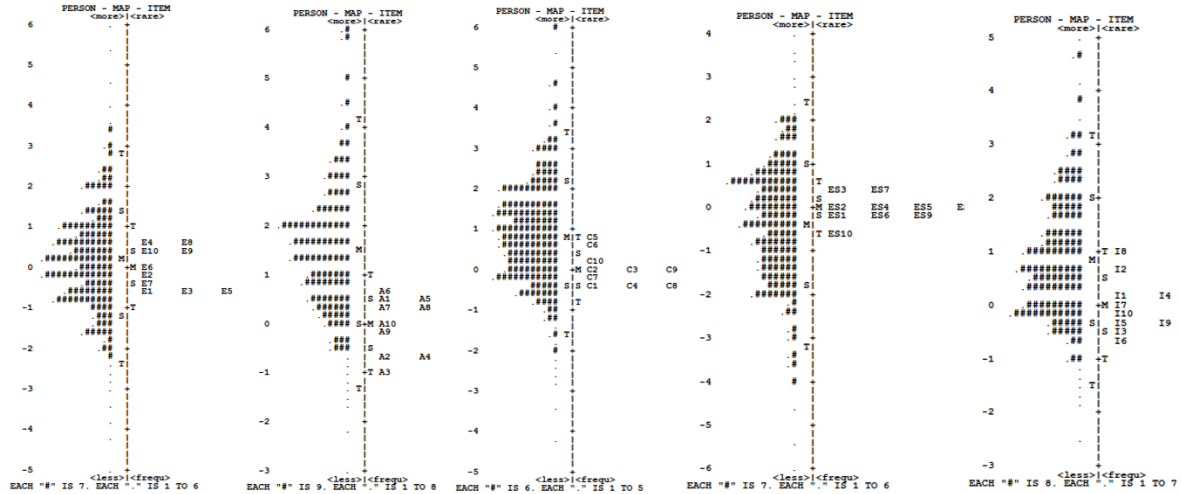

Figure 2. Wright map for all scale of IPIP-BFM

Figure 3. Wright map for all scale of IPIP-BFM-25

Reliability

Reliability analysis examined the reproducibility of measure locations for items and persons. We examined the reliability of the IPIP-BFM concerning the Rasch person separation reliability (Wright & Masters, 1982) and Cronbach's alpha. Fisher (2007) suggested that for fair quality, reliability and strata should be at least 0.67 and 2, and for good quality, they should be at least 0.81 and 3. The strata index is calculated as (4 S + 1)/3, where S refers to the person or item separation index (Wright & Masters, 1982).

All five scales in IPIP-BFM had satisfactory items, persons, and alpha reliability indices above 0.8 (Table 2). The item separation for all scales above six, and person separation for all scales above two, results in the item and person strata above three. It indicates that all scales have good quality (Fisher, 2007). The sample is large enough to confirm the item's hierarchical difficulty in the instrument, and all scales are sensitive enough to differentiate various levels of a person's abilities.

For the short scale, the downgrade of the measured quality was indicated by the lower person reliability and person separation. The scales are not sensitive enough to differentiate various levels of a person's abilities. A separation index implies that the scales can consistently identify only two levels of person's abilities. However, all five scales have a fair quality of reliability.

Item bias

Lastly, an item bias was tested using differential item functioning (DIF) by gender and education to test the generalized validity. According to ETS guidelines, a slight to moderate DIF was regarded to be present if the difficulty parameters had a significant probability and a DIF contrast of more than 0.43 (Zwick et al., 1999).

Overall, the differences in the item difficulty of the males and females were small. The highest contrast was -0.42 logits for item E6. However, none of the items with DIF contrast more than 0.43, indicating that females and males attached similar meanings to the items of IPIP-BFM. The differences in the item difficulty of people with secondary school and higher education

were slightly higher. The highest contrast was -1.01 logits for item A3. Eight items had a DIF contrast of more than 0.43 (Table 4), indicating that people with secondary school and higher education attached different meanings to some items of IPIP-BFM.

Table 4

Item bias for the five scales of IPIP-BFM

Item	Gender		Education	
	DIF Contrast[a]	Rasch-Welch t[a]	DIF Contrast[b]	Rasch-Welch t[b]
E4 Keep in the background	0.00	0.00	0.77	6.07*
A3 Insult people	-0.07	-0.63	-1.01	-5.84*
A5 Am not interested in other people's problems	0.02	0.23	0.68	5.06*
A6 Have a soft heart	0.08	0.84	0.61	4.58*
I1 Have a rich vocabulary	-0.11	-1.20	-0.57	-4.16*
I7 Am quick to understand things	0.00	0.00	-0.59	-4.26*
I8 Use difficult words	0.15	1.62	0.91	7.04*
I9 Spend time reflecting on things	-0.18	-1.87	-0.46	-3.27*

Notes: DIF contrast[a] = the difference in the difficulty of the item between males and females. A negative DIF contrast[a] indicates that the item is more difficult for females. DIF contrast[b] = the difference in the difficulty of the item between people with secondary school and higher education. A negative DIF contrast[b] indicates that the item is more difficult for people in secondary school.

* $p < .05$

Discussion

In general, our findings show that the IPIP-BFM Indonesia has some adequate psychometric properties, especially in terms of category function, item properties, and item bias. The use of five-point Likert scale categories was well functioning and fully understood by participants. No item was identified as DIF across gender. This indicates that females and males attached similar meanings to the items of IPIP-BFM. Therefore, comparing the big five personalities between genders can be conducted fairly using the

items in this IPIP-BFM. However, some items were detected to have DIF by education levels. If personality traits are compared across education levels for some reason, the conclusion should be made cautious since several items might be biased.

In terms of scale-level psychometric properties, the longer version of the test outperforms the shorter one. As predicted, IPIP-BFM has higher reliability, separation, and better item targeting than IPIP-BFM-25. However, for research purposes, IPIP-BFM-25 provides adequate reliability as all scales have reliability above 0.70. Therefore, when the resource is limited, and a shorter measure should be used, then IPIP-BFM-25 provides an adequate measure of the big five personality traits. For example, researchers who wish to place personality, not as the primary variable of interest are suggested to use IPIP-BFM-25.

In terms of unidimensionality, the emotional stability and intellect scales did not meet the assumption of unidimensionality. This finding is similar to Apple and Neff (2012) findings. The shorter scale has better unidimensionality in this case. This is not surprising since personality is a complex construct with many facets (Cooper et al, 2010). Selecting items with high loading and low cross-loading factor would ensure unidimensionality, but it could narrow the content and range of endorsability. However, these reports are better than the Japanese version of IPIP-BFM, especially in terms of reliability and item properties (Apple & Neff, 2012). The explanation can be addressed in the adaptation process since Apple and Neff (2012) changed the original version of five-point Likert-type categories into four response categories by omitting the middle category response.

In general, all items in extraversion, agreeableness, conscientiousness, and intellect scales were too easy to endorse by participants. The inability of the scale items to cover participants with maximum scores of the constructs in our findings is in line with Apple and Neff (2012) findings. This is not surprising because the instrument was developed based on a CTT approach, which has two major problems: having sets of redundant items and having skewed response categories for most items (Petrillo, et al., 2015). This problem seems more severe on a shorter scale.

Since the items for the short scale were selected based on the loading factor, some items are redundant for the content and the item difficulty. The present research suggests that adding more difficult items to the scales could enhance their measurement precision. Specifically for short scale, items selected should have a broader range of endorsability, which may improve future psychological research involving Indonesian participants.

The mean for a person measure on the agreeableness scale was the highest among the five scales in IPIP-BFM and IPIP-BFM-25, indicating that the item on the agreeableness scale is too easy to be endorsed by Indonesian samples. This result could be explained by previous literature in cultural psychology, where it is popularly termed the "interdependent self" (Markus & Kitayama, 1998). Eastern societies (such as Indonesia) are considered more group-oriented than Western societies, which are more individualistic (Hofstede & McCrae, 2004; Kitayama, Markus, & Kurokawa, 2000). It may be like Indonesians to consider others' feelings in their daily life, which affects the participants' responses, as such in the current instrument.

One limitation of the study is that participants were recruited using an online survey, which means they may not be representative of the Indonesian population. The internet has not yet reached several regions of Indonesia. It is possible that Indonesians from rural areas would interpret the IPIP-BFM content differently than expected. This must be empirically examined. A second possible limitation is that our study solely focused on the internal psychometric characteristics of the IPIP-BFM. While these findings are promising, the external validity measure could address to what degree this measure corresponds to any theoretically related constructs. A third limitation is that the investigation of IPIP-BFM-25 used the same dataset as the IPIP-BFM. Smith and colleagues (2000) suggested administering the short form on an independent sample to avoid overestimating the correlation. However, the procedure used in this study is considered enough to describe the reliability and separation loss of using the short form.

Conclusion

IPIP-BFM Indonesia has some adequate psychometric properties, especially in terms of category function, person and item reliability, and item properties. In general, longer measures outperformed shorter measures regarding person separation and reliability. However, further testing and refinement must be developed. Based on the Rasch model analysis in this study, the following suggestions can be made to improve the test's precision. First, more difficult items should be added to represent a wider variety of endorsement abilities. Second, the item detected as a misfit (I8 Use difficult words) should either be revised or deleted. Third, if a short version is needed, then the item selection for the short scale should not solely be chosen based on the loading factor on EFA but also considering the broad of the content and item difficulty.

Authors note

HA performed study design, data collection, statistical analysis, data interpretation, manuscript preparation, literature search, and funds collection. BS performed statistical analysis, data interpretation, manuscript preparation, and literature search.

Conflict of Interests

The authors declare that there is no conflict of interest.

Data availability statement

The datasets presented in this study can be found in online repositories at https://osf.io/uca3p/.

Funding

H.A. received funding from the Faculty of Psychology, University of Muhammadiyah Malang, Indonesia: Grant no. E.2.b/749.a/F.Psi-UMM/VIII/2021

References

Adams, R. J., Wu, M. W., & Wilson, M. (2012). The Rasch rating model and the disordered threshold controversy. *Educational and Psychological Measurement, 72*(4), 547–573. https://doi.org/10.1177/0013164411432166

Akhtar, H., & Azwar, S. (2018). Development and validation of a short scale for measuring big five personality traits: The IPIP-BFM-25 Indonesia. *Journal of Innovation in Psychology, Education and Didactics, 22*(2), 167–174.

Andrich, D. (1978). Scaling attitude items constructed and scored in the Likert tradition. *Educational and Psychological Measurement, 38*(3), 665–680. https://doi.org/10.1177/001316447803800308

Andrich, D. & Marais, I. (2019*). A Course in Rasch Measurement Theory, measuring in the educational, social and health sciences*. Singapore: Springer.

Apple, M. T., & Neff, P. (2012). Using Rasch Measurement to Validate the Big Five Factor Marker Questionnaire for a Japanese University Population. *Journal of Applied Measurement*, *13*(3), 1–17.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (Third edition). New York ; London: Routledge, Taylor and Francis Group.

Costa Jr, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*(1), 21–50. https://doi.org/10.1207/s15327752jpa6401_2

Fisher, W. P., Jr. (1992). Reliability statistics. *Rasch Measurement Transactions*, 6, 238

Fisher, W.P. Jr. (2007). Rating scale instrument quality criteria. *Rasch measurement transactions*, 21(1), 1095. (Downloaded from http://www.rasch.org/rmt/rmt211m.htm)

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26–42. https://doi.org/10.1037/1040-3590.4.1.26

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Gow, A. J., Whiteman, M. C., Pattie, A., & Deary, I. J. (2005). Goldberg's 'IPIP' Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences, 39*(2), 317–329. https://doi.org/10.1016/j.paid.2005.01.011

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Guenole, N., & Chernyshenko, O. S. (2005). The Suitability of Goldberg's Big Five IPIP Personality Markers in New Zealand: A Dimensionality, Bias, and Criterion Validity Evaluation. *New Zealand Journal of Psychology, 34*(2), 86–96.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research, 2*(1999), 102–138.

Kean, J., Bisson, E. F., Brodke, D. S., Biber, J., Gross, P. H. (2017). An introduction to item response theory and Rasch analysis: *Application using the Eating Assessment Tool (EAT-10). Brain Impairment, 19*(1), 91–102. https://doi.org/10.1017/BrImp.2017.31

Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, Emotion, and Well-being: Good Feelings in Japan and the United States. *Cognition & Emotion, 14*(1), 93–124. https://doi.org/10.1080/026999300379003

Kline, P. (2000). The future of personality measurement. In J. Mohan (Ed.), *Personality across cultures: Recent developments and debates* (pp. 336-351). New Delhi, India: Oxford University Press.

Kottorp, A., Bernspång, B., & Fisher, A.G. (2003). Validity of a performance assessment of activities of daily living for people with developmental disabilities. *Journal of Intellectual Disability Research, 47*(8), 597–605. https://doi.org/10.1046/j.1365-2788.2003.00475.x

Lim, S.M., Rodger, S., & Brown, T. (2009). Using Rasch analysis to establish the construct validity of rehabilitation assessment tools. *International Journal of Therapy and Rehabilitation, 16*(5), 251–260. https://doi.org/10.12968/ijtr.2009.16.5.42102

Linacre, J. M. (2002a). What do Infit and Outfit, Mean-square and Standardized mean? *Rasch Measurement Transactions, 16*(2), 878.

Linacre, J. M. (2002b). Optimizing Rating Scale Category Effectiveness. *Journal of Applied Measurement, 3*(1), 85–106.

Linacre, J. M. (2012). *A User's Guide to WINSTEP & MINISTEP: Rasch-Model Computer Programs*. Chicago: Winsteps.com.

Markus, H. R., & Kitayama, S. (1998). The cultural psychology of personality. *Journal of Cross-Cultural Psychology*, *29*(1), 63–87. https://doi.org/10.1177/0022022198291004

Mlačić, B., & Goldberg, L. R. (2007). An Analysis of a Cross-Cultural Personality Inventory: The IPIP Big-Five Factor Markers in Croatia. *Journal of Personality Assessment*, *88*(2), 168–177. https://doi.org/10.1080/00223890701267993

Oliveira, J. P. (2017). Psychometric Properties of the Portuguese Version of the Mini-IPIP five-Factor Model Personality Scale. *Current Psychology, 38,* 432–439. https://doi.org/10.1007/s12144-017-9625-5

Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2015). Using Classical Test Theory, Item Response Theory, and Rasch Measurement Theory to Evaluate Patient-Reported Outcome Measures: A Comparison of Worked Examples. *Value in Health*, *18*(1), 25–34. https://doi.org/10.1016/j.jval.2014.10.005

Raykov, T., Marcoulides, G. A. (2015). On the relationship between classical test theory and item response theory. *Educational and Psychological Measurement, 76*(2), 325–338. https://doi.org/10.1177/0013164415576958

Smith, E. V., Jr., & Wolfe, E. W. (2007). Understanding Rasch measurement: Instrument development tools and activities for measure validation using Rasch models: Part 2–Validation activities. *Journal of Applied Measurement, 8*(2), 204–234.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12(*1), 102–111, https://doi.org/10.1037//1040-3590.12.1.102.

Strus, W., Cieciuch, J., & Rowiński, T. (2017). The Polish adaptation of the IPIP-BFM-50 questionnaire for measuring five personality traits in the lexical approach. *Roczniki Psychologiczne/Annals of Psychology*, *17*(2), 347–366.

Van Zile-Tamsen, C. (2017). Using Rasch Analysis to Inform Rating Scale Development. *Research in Higher Education,* 58(8), 922–933. https://doi.org/10.1007/s11162-017-9448-0

Wolfe, E.W., & Smith, E.V. (2007). Instrument development tools and activities for measure validation using Rasch models: Validation activities. *Journal of Applied Measurement, 8*(2), 204–234.

Wright, B., Masters, G. (1982). *Rating scale analysis: Rasch measurement*. Chicago, IL: MESA Press.

Zheng, L., Goldberg, L. R., Zheng, Y., Zhao, Y., Tang, Y., & Liu, L. (2008). Reliability and concurrent validation of the IPIP Big-Five factor markers in China: Consistencies in factor structure between Internet-obtained heterosexual and homosexual samples. *Personality and Individual Differences*, *45*(7), 649–654. https://doi.org/10.1016/j.paid.2008.07.009

Zwick, R., Thayer, D.T., Lewis, C. (1999). An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis. *Journal of Educational Measurement, 36(*1), 1–28. https://doi.org/10.1111/j.1745-3984.1999.tb00543.x

Primenjena psihologija
Vol. 16, No. 1, pp. 29-58, 2023

Research Article

Does nature work? Effects of workplace greenery on employee well-being

Boris Popov[1]✉, Snežana Popov[2] and Mojca Nastran[3]

[1] *Department of Psychology, Faculty of Philosophy, University of Novi Sad, Serbia*

[2] *Department of Biology and Ecology, Faculty of Science, University of Novi Sad, Serbia*

[3] *Department of Forestry and Renewable Forest Resources, Biotechnical Faculty, University of Ljubljana, Slovenia*

ABSTRACT

A significant number of studies have been conducted in order to explore the effects of greenery on various aspects of human functioning. However, little is known about how natural elements affect indicators of well-being at work, such as work burnout or work engagement. Two studies (Study 1, winter – Ljubljana, Slovenia; Study 2, spring – Novi Sad, Serbia) were performed in order to: 1) assess the effect of natural elements on work burnout and work engagement in two different seasons (winter and spring); 2) explore if gender moderates the effects of workplace greenery on work burnout and work engagement. The results from Study 1 (winter) showed that, after the exclusion of outliers from the dataset, neither indoor nor outdoor greenery had a significant effect on work engagement and work burnout. Contrary to this, in Study 2 (spring) outdoor greenery exerted a significant effect on both burnout and engagement, while indoor greenery did not. The moderating effect of gender was not clear and further studies on this topic are needed. Thus, the current research extends the existing literature on workplace greenery and demonstrates that natural elements can, to some extent, affect indicators of employee well-being, such as burnout and work engagement.

Keywords: natural elements, workplace greenery, employee well-being, job demands

UDC: 331.101.37
DOI: 10.19090/pp.v16i1.2409
Received: 17.05.2022.
Revised: 22.10.2022.
Accepted: 27.12.2022.

✉ Corresponding author e-mail: boris.popov@ff.uns.ac.rs

Introduction

Rapid urbanization together with the growing pressures of workplace demands has isolated humans from exposure to nature, making it harder to maintain a balanced lifestyle. Part of the solution for bringing nature back and integrating it into urban infrastructures is the careful planning and management of urban green resources. A significant number of studies have been conducted in order to explore the effects of natural elements on various aspects of human functioning, including postoperative recovery (Ulrich, 1984), students' preferences and perceived restoration likelihood (Bogerd et al., 2018), stress levels (Repke et al., 2018), to name just a few. Greenery is a valuable asset to the urban workplace environment as well, where it has been shown that natural stimuli such as indoor plants, nature walks and/or window views of outdoor greenery, have beneficial effects on employees in terms of reducing stress (Lottrup et al., 2013), improving mood (Barton et al., 2009), stress recovery (Ulrich et al., 1991) and psychophysiological status (Chang & Chen, 2005), increasing job satisfaction (Dravigne et al., 2008), health and well-being (Bowler et al., 2010; Elings, 2006; Houlden et al., 2017; Sanchez et al., 2018; Smith & Pitt, 2009; Wang et al., 2019), and even diminishing intention to quit (Leather et al., 1998). Moreover, attention and memory at work are also affected by the presence of a natural environment (Raanaas et al. 2011).

There are at least two theoretical explanations of why natural elements have positive effects on human functioning. One is the Attention Restoration Theory (ART; Kaplan & Kaplan, 1989) and it focuses on the cognitive processes that are involved in information processing. According to ART, directed attention is limited and prone to fatigue, while certain environments have restorative qualities that can facilitate attentional restoration. In line with ART, restoration is more likely to happen when an individual becomes fascinated with, and the attention is effortlessly drawn to an interesting element in the environment. One such environment is nature. The second is Stress Recovery Theory (SRT; Ulrich et al., 1991) and it explains the physiological and affective changes observed in natural environments.

According to SRT, nature speeds up physical recovery through muscle tension and reduces blood pressure and heart rate. Also, the natural environment promotes positive change in affect and emotions - all leading to the conclusion that nature can facilitate stress recovery via changes in the autonomic nervous system that increase relaxation.

Burnout and work engagement

To date several definitions, models, and concepts of burnout have been developed. One of the most popular concepts (Maslach et al., 2001) recognizes burnout as a syndrome which consists of three dimensions: (1) emotional exhaustion: the feeling of being "drained" and the inability to summon sufficient energy for a new day, as well as a lack of enthusiasm; (2) depersonalization/cynicism: the feeling of detachment from work and from the people in the work environment, distancing and taking on a cynical attitude, and (3) reduced personal accomplishment: feelings of decline in one's competence and productivity and a lowered sense of self-efficacy. Other conceptualizations of burnout put more emphasis on fatigue, exhaustion, and emotional weariness. For example, Pines and Aronson (1988) argue that burnout is a "state of physical, emotional and mental exhaustion, caused by long-term involvement in emotionally demanding situations" (p. 9). Similar to Pines and Aronson (1988), as well as Shirom (1989), Schaufeli and Greenglass (2001) define burnout as "a state of physical, emotional, and mental exhaustion that results from long-term involvement in work situations that are emotionally demanding" (p. 501). Under the influence of the positive psychology movement, a group of researchers have defined a new concept that has been viewed as a counterpart to that of burnout, and they have developed a new construct – work engagement (Maslach et al., 2001; Schaufeli et al., 2002; González-Romá et al., 2006). Work engagement is defined as a positive, fulfilling, work-related state of mind that is characterised by vigour, dedication, and absorption. Vigour refers to the high level of energy and resilience while working, the willingness of the employee to invest effort in his/her work and persistence in dealing with difficulties.

Dedication refers to the sense of significance, enthusiasm, inspiration, and pride. Absorption is characterized by being fully concentrated on one's work, whereby time passes quickly and the employee finds it difficult to detach him/herself from work. This definition of absorption makes it very similar to what Csikszentmihalyi (1990) once called "flow".

Mediating/moderating variables in the relationship between natural environment and human functioning

There are various possible ways and mediating variables by which natural green elements promote human functioning. For example, Zhang et al. (2014) found that connectedness with nature is associated with greater psychological well-being and is dependent on the tendency to perceive natural beauty. Also, a natural environment can provide a setting for an activity or exercise programme and thus promote increased physical activity, which in turn has a positive impact on health (Bowler et al., 2010). As gender differences in responsiveness to natural elements are well established (Lottrup et al., 2013), they are described in more detail.

Generally speaking, the literature mentions several gender differences with respect to their reactivity to natural elements (see for example, Jiang et al., 2014), although the results are not consistent in all studies (see Shin, 2007). For example, Lottrup et al. (2013) found that both physical and visual access to workplace greenery were associated with lower stress for men, while such association was not significant for women. Furthermore, Astell-Burt et al. (2014) conducted a longitudinal study in which they found that men benefitted from green spaces in early adulthood. However, no association between green spaces and mental health was evident for women until later in life. Older women in the greenest and least green neighbourhoods reported similar levels of general health, but those with a moderate degree of exposure had the most favourable mental health scores. Shibata & Suzuki (2002) carried out an experimental study in which participants were asked to perform two different tasks (sorting cards and

word associations) in three different working environments (front plant, side plant, no plant). It has been shown that in the association task, male subjects working without plants performed worse than female subjects under the same conditions. Moreover, the task performances of the male subjects using the front plant arrangement were higher than that of the male subjects working without plants. Jiang et al. (2014) have reported different patterns of relationship between males and females regarding tree density following a stress induction test in a laboratory experiment. For women, no relationship has been found between varying densities of tree cover and stress recovery. For men, the dose-response curve proved to be an inverted-U shape: as tree cover density increased from 1.7% to 24%, stress recovery increased. Tree density between 24% and 34% resulted in no change in stress recovery. Tree densities above 34% were associated with a decrease in recovery.

Finally, there are two additional constructs of interest for this study, for they have been proved to exert a significant effect on personal well-being. One is the *Engagement with natural beauty* - i.e., the degree to which individuals perceive natural beauty and are emotionally aroused by nature's beauty (Zhang et al., 2014). There is an emerging line of research that suggests a positive relation between individual's tendency to perceive natural beauty and well-being across different cultures (Diessner et al., 2008; Capaldi et al., 2017). The second one is *Job demands*, defined as "physical, social, or organizational aspects of the job that require sustained physical or mental effort and are therefore associated with certain physiological and psychological costs" (Demerouti et al., 2001, p. 501). The Job demands-resources model (see for example Schaufeli & Taris, 2014) assumes that when job demands are high, additional effort must be made to achieve the work goals and to prevent decreasing performance. This obviously comes with physical and psychological costs, such as fatigue, irritability, and burnout. Given the fact that both of those constructs have an established relationship with employee burnout, work engagement and well-being in general, we wanted to use them as covariates in order to "extract" the "pure" effect of greenery, regardless of the tendency of the person to perceive nature's beauty, or his/her level of job demands.

The present study

Although the links between urban greenery and employees' physical and psychological health have been summarised in many publications, there are still significant gaps in the knowledge and inconsistencies in the findings. For example, little is known about how natural elements affect indicators of well-being at work, such as burnout or work engagement. In fact, to the best of our knowledge, there are very few studies that have explored the effects of natural elements on burnout (Cordoza et al., 2018; Hyvönen et al., 2018; Thompson & Bruk-Lee, 2019). They demonstrated significant correlation between workplace nature exposure, directed attention, and strain outcomes (burnout, among others). Likewise, there is little information available regarding the seasonality of the effect of the workplace greenery. Does establishing green resources in the workplace affect employee well-being only during the spring – blooming season for most species – or can employees benefit throughout the whole year? Finally, many studies have yielded a weak association between natural elements and health (e.g., Houlden et al., 2017; see also Bringslimark et al., 2009) or have been criticised for having poor design, failing to exclude confounding effects or reverse causality (Lee & Maheswaran, 2011; see also Grinde & Patil, 2009; Yeo et al., 2020).

Our study addressed the above-mentioned gaps and raised the following research questions: 1) Does workplace greenery affect work burnout and work engagement, when job demands and engagement with natural beauty are controlled for? 2) Do the effects of workplace greenery on work burnout and work engagement differ across different seasons (winter vs spring) and samples (Ljubljana, Slovenia vs Novi Sad, Serbia)? 3) Does gender moderate the effects of workplace greenery on work burnout and work engagement?

In order to address those questions, we performed two independent studies, one in Ljubljana (Slovenia), and the second in Novi Sad (Serbia). In line with the Attention Restoration Theory and Stress Recovery Theory, and

building upon research findings that reveal restorative power of the natural elements and its effects on stress reduction (Hyvönen et al., 2018; Lottrup et al., 2013), following hypotheses were formulated:

H1: There are statistically significant differences in the levels of work burnout (WB) and work engagement (WE) with respect to the possibility of viewing a green outdoor environment through the window during work (WG1). We expect more work burnout and less engagement in employees lacking the possibility of viewing a green outdoor environment through the window during the working day.

H2: There are statistically significant differences in the levels of work burnout (WB) and work engagement (WE) with respect to the possibility of taking a break in a garden, park or other natural environment during the working day (WG2). We expect more work burnout and less engagement in employees having less possibility of taking a break in a garden, park or other natural environment during the working day.

H3: There are statistically significant differences in the levels of work burnout (WB) and work engagement (WE) regarding the presence/absence of workplace indoor plants (WG3). We expect more work burnout and less engagement in employees who work in offices without indoor plants.

Also, even though literature review did not paint the whole picture, most studies (Astell-Burt et al., 2014; Fischer et al., 2018; Jiang et al., 2014; Lottrup et al., 2013) showed that there are some gender differences in responding to urban greenery. Therefore, the fourth hypothesis was as follows:

H4: Male and female employees will respond differently in contact with both outdoor and indoor greenery.

Study 1 (Ljubljana, Slovenia)

Method

Instruments

The workplace greenery index (WGI) measures employees' nature exposure in the workplace. The questionnaire consists of three items, two of which have been taken from Lottrup et al. (2013), and the third has been added to measure the presence of indoor plants. The first item (WG1) measures the visibility of outdoor greenery through the window ("Do you have the possibility of viewing a green outdoor environment through the window while you are working?", with the response categories being "*Yes*" or "*No*"). The second question (WG2) relates to actual access to the outdoor environment during work ("Do you have the possibility of taking a break in a garden, park or other natural environment during your working day?", with the response categories being "*No*", "*Yes, but I never use it*", "*Yes, and I use it sometimes*", "*Yes, and I use it often*"). Finally, with the third question (WG3) participants were asked about the presence of indoor greenery in their workplaces ("Do you have indoor plants in your workplace?", the possible responses being "*Yes*" or "*No*").

The Work Burnout scale from the Copenhagen burnout inventory (Kristensen et al., 2005; Serbian version: Berat et al., 2016). WB assesses the degree of physical and mental fatigue and exhaustion that is perceived by the person, related to his/her work. It consists of seven items (e.g., "Is your work emotionally exhausting?"), with a five–point Likert scale (from "*Never/almost never*" to "*Always*"). The original scoring was used (the response "*Never/almost never*" is calculated as 0, "*Seldom*" as 25, "*Sometimes*" as 50, "*Often*" as 75 and "*Always*" as 100), and the total score is obtained as arithmetic means of the responses to all seven items. Earlier studies have shown that the scale is valid and reliable (Milfont et al., 2008). Verification on the samples of employees in Serbia has shown that it has a high internal consistency ($\alpha = .88$), as well as correlating to distress in the expected manner

(*r* = .54), turnover intentions (*r* ranging from .36 to .40) and job satisfaction (*r* ranging from .29 to -.44; Berat et al., 2016). The α coefficient obtained for the scale was .88 in Study 1 (Ljubljana), and .90 in Study 2 (Novi Sad), respectively.

WE was measured using the short version of *the Utrecht Work Engagement Scale* (UWES-9; Schaufeli et al., 2006; both Serbian and Slovenian versions of the scale can be obtained from: https://www.wilmarschaufeli.nl/tests/#engagement). The measure consists of three subscales: Vigor (VI), Dedication (DE), and Absorption (AB). Each scale consists of three items. Each item needs to be assessed on a 7-point scale (*1 = never* to *7 = always*). Item samples: "At my job, I feel strong and vigorous." (VI), "I am proud of the work that I do" (DE) and "I get carried away when I'm working" (AB). The α coefficient obtained in this study for the scale was .93.

The Job Demands Scale from the Job demands-resources questionnaire (Boyar et al., 2007), translated into Slovenian and adapted by Tement et al., (2010). The Job Demands Scale contains five items (sample item: "My work requires all of my attention"). The α coefficient obtained in this study for the scale was .89.

Appreciation of natural beauty was measured with the natural beauty dimension from the *Engagement with Beauty Scale* (EBS; Diessner et al., 2008). The scale measures the different psychological processes by which humans encounter beauty in the natural environment: perception or cognition ("I notice beauty in one or more aspects of nature"), physiological arousal ("When perceiving beauty in nature I feel changes in my body, such as a lump in my throat, an expansion in my chest, faster heartbeat, or other bodily responses"), conscious emotion ("When perceiving beauty in nature I feel emotional, it 'moves me,' such as feeling a sense of awe, or wonder or excitement or admiration or upliftment."), and transcendence or spirituality ("When perceiving beauty in nature I feel something like a spiritual experience, perhaps a sense of oneness, or being united with the universe, or a love of the entire world."). Each of the four items was assessed on a 7-point scale (*1 = very unlike me* to *7 = very much like me*). The α coefficient obtained in this study for the scale was .80.

Several items of socio-demographic data were collected as well, including gender, age, education level, tenure with current employer and the number of working hours per week.

Sample and procedure

The original idea behind the study was to implement a two-wave panel design, with a time lag of 4 months between Time 1 and Time 2. With that in mind, data for Time 1 were collected in January 2016. The invitation to participate in the survey was sent to five different departments of the public administration of the City of Ljubljana, whose offices are situated in the urban part of the city of Ljubljana. Participants were invited by their superiors via company email, were supplied with general information about the aim of the study, and informed that data obtained in the study would be considered strictly confidential and used only for research purposes. The questionnaires were administered individually, and it took the respondents approximately 10 minutes to complete them. Participation was anonymous and voluntary.

A total of 132 employees filled in a Google Form, out of approximately 300 full-time employees – response rate 43%. Two cases (1.5%) were initially deleted because they did not fill out at least one whole questionnaire, so the final sample consists of 130 respondents with usable data (38 males [29.2%], 92 females [70.8%]; average age 44 [range 27-66], 13 participants with no response), with nearly 42 working hours per week on average. 13 employees (10.2%) had finished secondary school, 100 employees had either bachelor's or master's degrees (78.1%), and 15 participants (11.7%) had either a magister degree or PhD (two participants with no response). Missing data were present in sparse numbers (all but one respondent with less than 10% missing data scattered across the dataset), so we decided to replace missing data using an Expectation-Maximisation algorithm.

Time 2 was carried out during the spring season, May-June 2016. However, due to attrition, at Time 2 data were collected from only 61 participants. In addition, a large number of participants did not use the same code for both measurements and thus data could not be reconciled.

Therefore, in this paper only data from Time 1 have been reported. However, data from Time 2 can be obtained from the first author upon request.

The SPSS 23.0 package was used to process the results. Descriptive statistics techniques were used to describe the data, and ANCOVA for testing the hypotheses.

Results

All variables in the study show a distribution that does not significantly deviate from normal. Also, α coefficients of internal consistency show acceptable levels for all the study variables. However, in two subjects' values were found for the work engagement which qualified them as outliers, namely z = 3.33 and 3.49 (Tabachnik & Fidell, 2013). Thus, Table 1 shows descriptive statistics for the variables used in the study, as well as their interrelationships, after the two outliers have been dropped.

Table 1

Study 1 - mean, standard deviation, skewness, kurtosis, and α coefficients of variable and correlations between the variables.

	M	*SD*	*Sk*	*K*	α	2	3	4
1. Work Burnout (WB)	42.24	22.04	.16	-.43	.88	-.43**	.22**	.05
2. Work Engagement (WE)	37.38	9.62	-.61	-.05	.95		.35**	.06
3. Job demands (JD)	20.68	3.79	-.90	.41	.89			.09
4. Engagement with Natural Beauty (EWNB)	18.52	4.60	-.77	.00	.83			

Notes. N = 128. M – mean, SD – standard deviation, Sk – Skewness, K – Kurtosis, α – coefficient of internal consistency.

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

Table 2 shows the results of the ANCOVA analysis, where the three WG questions served as independent variables, burnout and work engagement as dependent variables, and job demands and engagement with natural beauty as covariates. Since dropping outliers from the dataset

significantly impacted results, Table 2 presents the ANCOVA results with and without outliers. As shown in Table 2, there were no significant differences in well-being indicators (burnout and work engagement) for either the window view (WG1) or the possibility of taking a break in a garden, park, or other natural environments during the working day (WG2). Thus, *hypotheses H1 and H2 were rejected*. There were significant differences in the level of work engagement in the case of the presence/absence of indoor plants at the workplace. It was shown that employees with indoor plants in their offices scored higher on work engagement (estimated M = 37.42, 95% *CI*: 35.83, 39.89; SE = 1.02) in comparison to those who did not (estimated M = 33.42, 95% *CI*: 29.70, 37.13; SE = 1.88). However, after two outliers had been removed, significant differences disappeared. Thus, *hypothesis H3 was rejected* as well. The fourth hypothesis was tested in the same way as the third, with gender introduced as the second categorical independent variable. Firstly, we wanted to test whether there were gender differences in burnout and work engagement. Results of t-test showed that there were no gender differences in either burnout, $t(126) = -0.28$, $p = .79$ or work engagement, $t(126) = -0.33$, $p = .74$. Furthermore, the results showed that gender moderated the effect of indoor plants on work engagement, WG3 x Gender, $F(1,124) = 4.07$, $p = .046$, $\eta^2 = .03$ while such an effect on burnout proved to be marginally significant, WG3 x Gender, $F(1,124)=3.71$, $p = .056$, $\eta^2=.03$. However, after two outliers had been removed from the dataset, significant differences disappeared for both engagement, WG3 x Gender, $F(1,122) = .33$, $p = .56$, $\eta^2 = .00$ and burnout, WG3 x Gender, $F(1,122) = 1.39$, $p = .24$, $\eta^2 = .01$. Thus, *hypothesis H4 was rejected*. Finally, further analysis revealed that women reported having indoor plants more frequently than men, $\chi^2(1) = 21.34$, $p = .00$, so the results regarding the moderation effect of gender in the relationship between indoor greenery and work engagement should be interpreted with caution.

Table 2

Study 1 - ANCOVA models testing differences in WB and WE with and without outliers

		With outliers (N = 130)					Without outliers (N = 128)				
Source	Dependent Variable	*df*	Error *df*	*F*	*p*	η^2	*df*	Error *df*	*F*	*p*	η^2
WG1	Work Burnout	1	126	.17	.68	.00	1	124	.01	.94	.00
	Work Engagement	1	126	.39	.54	.00	1	124	2.22	.14	.02
WG2	Work Burnout	3	124	1.05	.37	.03	3	122	1.20	.32	.03
	Work Engagement	3	124	1.14	.33	.03	3	122	1.78	.16	.04
WG3	Work Burnout	1	126	.01	.93	.00	1	124	.50	.48	.00
	Work Engagement	1	126	4.27	.04	.03	1	124	.97	.33	.01

Note. In the cases of all three WG variables, JD and EWNB served as covariates. JD was a significant covariate in the case of all three independent variables, whereas EWNB was not in any of the three. See text for more details.

Discussion

The results have shown that outdoor greenery (accessed either visually or physically) did not exert a significant effect on employee well-being as measured by burnout and work engagement. It is worth noting though that hypotheses regarding the effects of outdoor greenery were postulated based on the initial study design that should have included two measurements (winter/spring). However, due to the sample attrition, the results here were based only on the Time 1 measurement (winter). Given the fact that in wintertime outdoor greenery is scarce, this result is not surprising. Gender differences in the context of workplace greenery are also worthy of discussion. It appears that male and female workers react differently to the presence/absence of indoor plants. However, after two outliers were dropped from the dataset, the moderating effect of gender disappeared, thus rendering the results inconclusive. What is puzzling is the finding that women report having indoor plants in their offices more frequently than men. It appears that they choose to plant them for reasons other than restorative

(for example, aesthetic or out of cultural habit). Overall, there is insufficient data to support any of the formulated hypotheses and more research is needed on this topic.

Study 2 (Novi Sad, Serbia)

Method

Instruments

In Study 2, the instruments used to assess the presence/absence of greenery, engagement with natural beauty, work engagement and work burnout were the same as those used in Study 1.

The job characteristics questionnaire (Popov et al., 2022) was used to measure job demands. More specifically, three subscales measuring three fundamental job demands were used, those being: quantitative job demands (sample item: "Your job requires you to work very fast"), cognitive job demands (sample item: "Your job requires you to be very focused on and to pay full attention to the task at hand."), and emotional job demands (sample item: "Your job requires you to behave kindly (in conversation with customers, clients, etc.) even when you don't feel that way."). The dimensions were represented with three items each, so that the total number of items used in the study amounted to 9. The α coefficient obtained in this study for the scale was .64.

Sample and procedure

A total sample of 191 full-time employees from various industries (both private and public sector companies) participated in the study (101 females, 52.9%). The average age of participants at the time of the data collection was 37 (range 18-64). Average tenure for participants was 9 years, with 35 working hours per week on average. 42 employees (18.2%) had finished secondary school, 19 employees had bachelor's degrees (8.2%), 124 (53.7%) had master's degrees, and 5 participants had PhDs (2.2%).

Data was collected via Google Forms during May 2019. As was the case in Study 1, the first page of the questionnaire provided respondents with basic information on the main aim of the study. Respondents were informed that participation in the study was voluntary and were offered the opportunity to contact the authors. The questionnaires were completed individually and took the participants approximately 10 minutes to complete. Due to settings of Google Forms (all questions were obligatory to answer), there were no missing values or incomplete data.

Results

Table 3 shows descriptive values for the variables used in the study. It is evident that all variables show a fairly normal distribution. Also, the α coefficient of internal consistency is at least acceptable for all the variables except for job demands (JD), which evidenced a somewhat lower level of internal consistency. Given the fact that we used a composite score of job demands consisting of three relatively independent job demands indicators (quantitative, qualitative, and emotional), this result is not unexpected. No outliers were detected for any of the studied variables.

Table 3

Study 2 - mean, standard deviation, skewness, kurtosis, and α coefficients of variable and correlations between the variables

	M	*SD*	*Sk*	*K*	α	2	3	4
1. Work Burnout (WB)	47.88	24.79	.04	-.67	.90	-.52**	.39**	.17*
2. Work Engagement (WE)	34.77	10.14	-.46	-.05	.93		.00	.09
3. Job demands (JD)	32.61	5.22	-.18	-.13	.64			.19**
4. Engagement with Natural Beauty (EWNB)	13.81	3.96	-.39	-.42	.80			

Notes. $N = 191$. *M* – mean, *SD* – standard deviation, *Sk* – Skewness, *K* – Kurtosis, α – coefficient of internal consistency.

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

Table 4 shows results of the ANCOVA analysis, where the three WG questions served as independent variables, burnout (WB) and work engagement (WE) as dependent variables, and job demands (JD) and engagement with natural beauty (EWNB) as covariates. As is shown in Table 4 and Figure 1, respondents with the possibility of viewing a green outdoor environment through their window during work (WG1) reported significantly less burnout (estimated *M* = 45.32, 95%*CI*: 47.49, 58.67, *SE* = 2.02), in comparison to those without that possibility (estimated *M* = 53.08, 95%*CI*: 47.49, 58.67, *SE* = 2.83). Also, respondents with the possibility of viewing a green outdoor environment through the window reported higher levels of WE (estimated *M* = 36.38, 95%*CI*: 34.62, 38.14; *SE* = .89), compared to those without (estimated *M* = 31.67, 95%*CI*: 29.21, 34.13; *SE* = 1.25). Thus, *hypothesis H1 was supported*. In the case of the possibility of taking a break in a garden, park, or other natural environments during the working day (WG2) the differences were also significant (Figure 2), thus *hypothesis H2 was supported*. Bonferroni pairwise comparisons showed significant differences in the level of burnout between the group of respondents who did not have access to outdoor greenery (estimated *M* = 55.27, 95%*CI*: 49.50, 61.04; *SE* = 2.92) as compared to both those who did have access and used it sometimes (estimated *M* = 43.80, 95%*CI*: 38.76, 48.85; *SE* = 2.56) and who did have access and used it often (estimated *M* = 41.00, 95%*CI*: 33.98, 48.02; *SE* = 3.56). Additionally, significant differences in the level of work engagement were registered between the group of respondents who did not have access to outdoor greenery (estimated *M* = 30.96, 95%*CI*: 28.37, 33.54; *SE* = 1.31) as compared to both those who had access and used it sometimes (estimated *M* = 36.26, 95%*CI*: 34.00, 38.52; *SE* = 1.15) and those who had access and used it often (estimated *M* = 37.08, 95%*CI*: 33.94, 40.23; *SE* = 1.59). Finally, there were no significant differences in levels of either WB or WE in the third condition (WG3 – "Do you have indoor plants at your workplace?"). Thus, *hypothesis H3 was rejected*. To test whether the effects of workplace greenery depend on gender, we performed additional analysis for the independent variables WG1 and WG2. Prior to the hypothesis testing, we wanted to examine whether

there were gender differences in burnout and work engagement. Results of t-test showed that there were no gender differences in burnout, *t*(189)= -1.33, *p* = .18 and work engagement, *t*(189)= .40, *p* = .69. Moderation analysis showed that gender did not moderate the relationship between either of the two greenery variables and employee well-being. In the case of the possibility of viewing green outdoors through the window: WG1 x Gender, *F*(1,185)= .87, *p* = .35, η^2 = .01 for work burnout; WG1 x Gender, *F*(1,185) = .81, *p* = .78, η^2 = .00 for WE. For the second independent variable, the possibility of taking a break in a garden, park or other natural environments during the working day: WG2 x Gender, *F*(3,181)= 1.79, *p* = .15, η^2 = .03 for work burnout, and WG2 x Gender, *F*(3,181) = 1.50, *p* = .22, η^2 = .02. Thus, *hypothesis H4 was rejected*.

Table 4

Study 2 - ANCOVA models testing differences in WB and WE

Source	Dependent Variable	*df*	Error *df*	*F*	*p*	η^2
WG1	Work Burnout	1	187	5.04	.03	.03
	Work Engagement	1	187	9.30	.00	.05
WG2	Work Burnout	3	185	5.00	.00	.08
	Work Engagement	3	185	4.13	.01	.06
WG3	Work Burnout	1	187	1.32	.25	.01
	Work Engagement	1	187	0.85	.36	.01

Notes. *N* = 191. In the cases of all three WG variables, JD and EWNB served as covariates. JD proved to be a significant covariate in the case of all three independent variables, but only when predicting WB. EWNB was not significant in any of the cases. See text for more details.

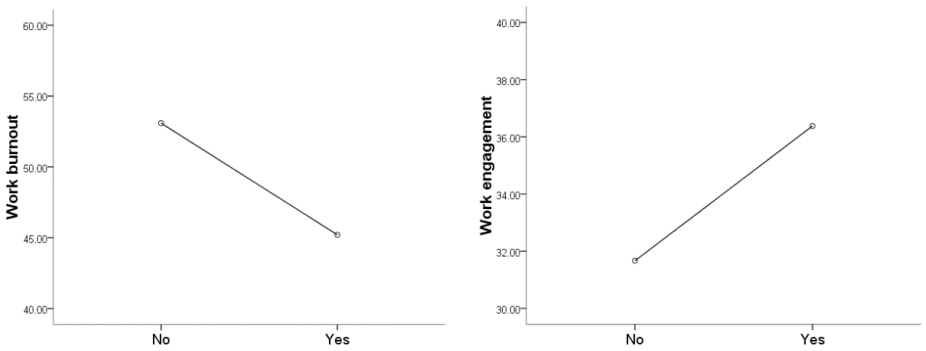

Figure 1. Main effect of WG1 ("Do you have the possibility of viewing a green outdoor environment through the window while you are working?") on Work burnout and Work engagement respectively

Figure 1. Main effect of WG1 on Work burnout and Work engagement.

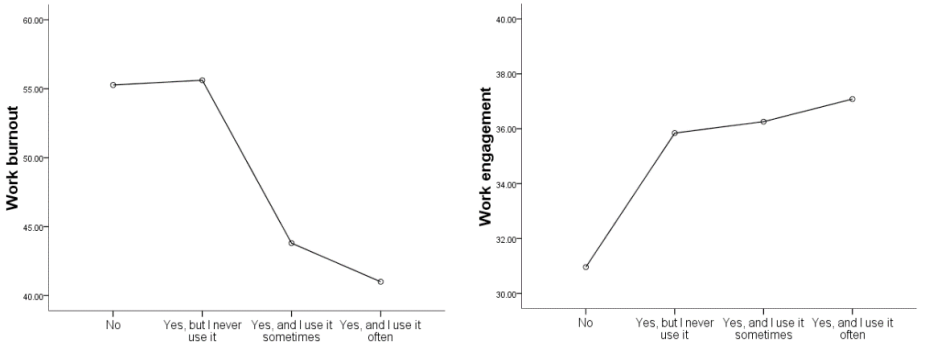

Figure 2. Main effect of WG2 ("Do you have the possibility of taking a break in a garden, park or other natural environments during your working day?") on Work burnout and Work engagement

Figure 2. Main effect of WG2 on Work burnout and Work engagement.

Discussion

It is interesting to notice the almost perfectly inverse pattern of results in the second study (Novi Sad, spring) in comparison to the first one (Ljubljana, winter). Specifically, the results of Study 2 have generally shown that outdoor natural elements can have restorative power and promote well-being in employees. In contrast, indoor plants were not found to be effective, nor did their interaction with gender have any influence on the well-being of employees. It can be concluded that, if available, outdoor greenery (the

possibility of visual and physical access) has a greater influence on the well-being of employees than indoor plants.

General discussion and conclusions

To date, many researchers have attempted to understand how people respond to the natural environment (Elings, 2006; Ulrich, 1984; see Lee & Maheswaran, 2011 for a detailed review). Even though they are not unambiguous in that regard, the results of previous research have generally shown that human beings tend to respond positively in contact with nature, in terms of affective response, psychological well-being, stress recovery etc. (Barton et al., 2009; Bowler et al., 2010; Elings, 2006; Houlden et al., 2017; Lottrup et al., 2013; Sanchez et al., 2018; Smith & Pitt, 2009; Wang et al., 2019). However, to date, very few studies have explored the effects of natural elements at work on employee well-being (Cordoza et al., 2018; Hyvönen et al., 2018; Thompson & Bruk-Lee, 2019). Moreover, to the best of our knowledge none of them have focused on the effect of greenery on one of the positive indicators of employee well-being, namely work engagement. To fill that gap in knowledge, we conducted two independent cross-sectional studies, the first in Ljubljana (Slovenia), and the second in Novi Sad (Serbia). The main aim of said studies was to explore whether workplace greenery affects employee well-being (both positive and negative aspects of well-being – work engagement and burnout, respectively), when job demands and engagement with natural beauty are controlled for.

There were several important findings. Firstly, it has been shown that when present, outdoor natural elements exert a significant effect on work burnout and engagement and that effect is in the expected direction. The same goes for both visual access (the possibility of viewing a green outdoor environment through a window), and physical access (the possibility of taking a break in a natural environment). As has been already discussed, this effect was significant only in the second study that was carried out in spring, and not in the first one, which took place during the winter season. Thus, it can be hypothesized that the effects of outdoor workplace greenery (both visual

and physical access) may be greater in spring than in winter, but the exact conclusion requires more strict research design (see Limitation section for more details). Furthermore, in the case of physical access to outdoor greenery during the working day, it is necessary to control for the possible confounding effect of physical exercise (i.e., walking), and thus to extract the "pure" effect of greenery. Indeed, it is possible that access to outdoor greenery during the working day boosts employees' physical activity, which in turn contributes to improving their well-being as shown earlier (see for example, Mitchell, 2013; Rebar et al., 2015; Weng & Chiang, 2014). That being said, the important finding of the current research is that the effect of making use of physical access to nature was significant in Study 2 (physical activity in a "greener" outdoor environment), while in Study 1 (physical activity in a "less green" environment) was not. This finding could contribute to the growing body of knowledge that considers greenery as being a key agent that boosts employee well-being.

The effect of indoor greenery is less clear. The results showed that, after two outliers had been removed in Study 1, indoor plants did not exert a significant effect on either burnout or engagement. It appears that, when present, outdoor natural elements have a greater effect on humans than indoor plants, as has been shown earlier (Chang & Chen, 2005; Korpela et al., 2017). One possible explanation is the confounding effect of depth of perception or daylight that enters the room through a window. For example, the classic work from Ulrich (1984) received criticism that the reported health benefits of looking through a hospital window was not due to natural elements but to the effects of a window regardless of the scenery. Yet, in later studies (Ulrich et al., 1991) it was shown that viewing a natural environment was superior in terms of physiological restoration after stressful events. However, when visual contact with both indoor and outdoor natural elements is present (as in Study 2), it is possible that depth perception (i.e., the view through the window onto the broader natural landscape or into the distance) may have more restorative power than a lack thereof (i.e., the view of "just" an indoor plant) and serve as the key component in boosting

employee well-being. Therefore, future studies should be focused on extracting the "true" effect of greenery.

It has been found previously that job stress affects male and female employees in different ways, as Lottrup et al. (2013) discussed in more detail. Additionally, in previous studies it has been shown that males and females react to exposure to nature differently (see for example, Astell-Burt et al., 2014; Shibata & Suzuki, 2002). The results from our studies do not support the hypothesis of gender as a moderating variable in the relationship between nature and employee well-being. Although women reported a greater number of plants in their offices as compared to men, there is no evidence that this greenery played a protective role in recovery from stress, nor that it helped in boosting work engagement. One way to understand such inconsistencies in findings in comparison to earlier studies is to take into account the different types of outcome measure. In different studies different indicators of response to the natural elements have been utilised (somatic outcomes, physiological response or reported perceptions of (mental) health levels; see for example Bos et al., 2016; Korpela et al., 2017), which makes it difficult to compare them.

Finally, the literature on environmental preference and restoration has to date been guided mostly by Stress Recovery Theory (Ulrich et al., 1991) and Attention Restoration Theory (Kaplan & Kaplan, 1989). This paper shows that theoretical perspectives for understanding workplace greenery could be broadened. Specifically, the results from this study could be discussed in the context of the job demands-resources theory (Demerouti et al., 2001), as suggested in one recent paper (Thompson & Bruk-Lee, 2019). In accordance with the job demands-resources theory, workplace greenery could be understood as one of the job resources that has restorative potential that is available to the employees to use (along with other work environment resources, such as ergonomically well-designed workplaces, social support, feedback, etc.). Naturally, a true and unambiguous support for the biophilia hypothesis is yet to be provided.

Limitations and implications for future research

The results of the study should be taken with some caution due to its limitations. *Firstly*, in Study 1 a relatively small number of employees took part in the study. Percentagewise, the response rate in the first wave was not small, given the fact that 130 out of (approximately) 300 full-time Ljubljana public administration employees took part in the study. However, in absolute numbers, it is still a relatively small sample. *Secondly*, as has been discussed earlier in the paper, two studies (Novi Sad and Ljubljana) were not carried out in the same period of the year. More precisely, Study 1 (Ljubljana) was designed as a short prospective study with two measurement times, but due to the high attrition in the second measurement, only data from measurement at Time 1 were usable. Therefore, the obtained results from the two studies are comparable only in the context of indoor greenery. Furthermore, the comparability of the two studies is further reduced due to the fact that they are two different samples (sample of employed from one public institution in study 1 vs heterogeneous sample of employees from various private companies in study 2). *Thirdly*, self-selection bias was not controlled for, and it cannot be ruled out a possibility that those who are most interested in this topic participated in the study and thus made the results biased in a sense that they reported an even greater effect of greenery than it would have been in the general population (see for example Bethlehem, 2010; Schaurer & Weiss, 2020).

Finally, we would like to invite other researchers to replicate this study particularly in other countries, as it would be very useful for establishing cross-cultural validity of the used measures and obtained results in general.

Conclusions

This study is among the first to examine the relationships between exposure to nature at work and employees' reported work burnout and work engagement. In general, the results give limited support to the growing body of evidence to suggest that natural elements can boost employee well-being

and mental health in general. It could be concluded that natural elements play a significant role in explaining employee well-being levels, even though that role is not always clear. Additionally, it seems that, when present, outdoor greenery exerts a greater impact on employee well-being than indoor plants. In the absence of outdoor elements (winter season in Study 1), results on the effect of indoor plants on work burnout and engagement are inconclusive. Finally, the moderating effect of gender is not clear and further studies on this topic are needed.

Conflict of interest

We have no conflicts of interest to disclose.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding

This work was supported by the GREEN SURGE, EU FP7 collaborative project, FP7-ENV.2013.6.2-5-603567.

References

Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. A longitudinal study. *Journal of epidemiology and community health*, *68*(6), 578–583. https://doi.org/10.1136/jech-2013-203767

Barton, J., Hine, R., & Pretty, J. (2009). The health benefits of walking in green spaces of high natural and heritage value. *Journal of Integrative Environmental Sciences*, *6*(4), 261–278. https://doi.org/10.1080/19438150903378425

Berat, N., Jelić, D., & Popov, B. (2016). Serbian version of the work burnout scale from the Copenhagen burnout inventory: Adaptation and psychometric properties. *Primenjena psihologija*, *9*(2), 177–198. https://doi.org/10.19090/pp.2016.2.177-198

Bethlehem, J. (2010). Selection bias in web surveys. *International statistical review, 78*(2), 161–188. https://doi.org/10.1111/j.1751-5823.2010.00112.x

Bogerd, N. van den, Dijkstra, S. C., Seidell, J. C., & Maas, J. (2018). Greenery in the university environment: Students' preferences and perceived restoration likelihood. *Plos one*, *13*(2), e0192429. https://doi.org/10.1371/journal.pone.0192429

Bos, E. H., Van der Meulen, L., Wichers, M., & Jeronimus, B. F. (2016). A Primrose Path? Moderating Effects of Age and Gender in the Association between Green Space and Mental Health. *International Journal of Environmental Research and Public Health*, *13*(5), 492–500. https://doi.org/10.3390/ijerph13050492

Bowler, D. E., Buyung-Ali, L. M., Knight, T. M., & Pullin, A. S. (2010). A systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Public Health*, *10*(1), 456–466. https://doi.org/10.1186/1471-2458-10-456

Boyar, S. L., Carr, J. C., Mosley, D. C., & Carson, C. M. (2007). The Development and Validation of Scores on Perceived Work and Family Demand Scales. *Educational and Psychological Measurement*, *67*(1), 100–115. https://doi.org/10.1177/0013164406288173

Bringslimark, T., Hartig, T., & Patil, G. G. (2009). The psychological benefits of indoor plants: A critical review of experimental literature. *Journal of Environmental Psychology*, *29*(4), 422–433. https://doi.org/10.1016/j.jenvp.2009.05.001

Capaldi, C. A., Passmore, H. A., Ishii, R., Chistopolskaya, K. A., Vowinckel, J., Nikolaev, E. L., & Semikin, G. I. (2017). Engaging with natural beauty may be related to well-being because it connects people to nature: Evidence from three cultures. *Ecopsychology*, *9*(4), 199–211. https://doi.org/10.1089/eco.2017.0008

Chang, C.-Y., & Chen, P.-K. (2005). Human Response to Window Views and Indoor Plants in the Workplace. *HortScience*, *40*(5), 1354–1359. https://doi.org/10.21273/HORTSCI.40.5.1354

Cordoza, M., Ulrich, R. S., Manulik, B. J., Gardiner, S. K., Fitzpatrick, P. S., Hazen, T. M., Mirka, A., & Perkins, R. S. (2018). Impact of Nurses Taking Daily Work Breaks in a Hospital Garden on Burnout. *American Journal of Critical Care*, *27*(6), 508–512. https://doi.org/10.4037/ajcc2018131

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*(3), 499–512. https://doi.org/10.1037/0021-9010.86.3.499

Diessner, R., Solom, R. D., Frost, N. K., Parsons, L., & Davidson, J. (2008). Engagement With Beauty: Appreciating Natural, Artistic, and Moral Beauty. *The Journal of Psychology*, *142*(3), 303–332. https://doi.org/10.3200/JRLP.142.3.303-332

Dravigne, A., Waliczek, T. M., Lineberger, R. D., & Zajicek, J. M. (2008). The effect of live plants and window views of green spaces on employee perceptions of job satisfaction. *HortScience*, *43*(1), 183–187. https://doi.org/10.21273/HORTSCI.43.1.183

Elings, M. (2006). People-plant interaction: The physiological, psychological and sociological effects of plants on people. In J. Hassink & M. Van Dijk (Eds.), *Farming for health* (pp. 43–55). Springer, Netherlands. https://doi.org/10.1007/1-4020-4541-7_4

Fischer, L. K., Honold, J., Cvejić, R., Delshammar, T., Hilbert, S., Lafortezza, R., Nastran, M., Nielsen, A. B., Pintar, M., van der Jagt, A. P. N., & Kowarik, I. (2018). Beyond green: Broad support for biodiversity in multicultural European cities. *Global Environmental Change, 49*, 35–45. https://doi.org/10.1016/j.gloenvcha.2018.02.001

González-Romá, V., Schaufeli, W. B., Bakker, A. B., & Lloret, S. (2006). Burnout and work engagement: Independent factors or opposite poles? *Journal of Vocational Behavior*, *68*(1), 165–174. https://doi.org/10.1016/j.jvb.2005.01.003

Grinde, B., & Patil, G. G. (2009). Biophilia: Does Visual Contact with Nature Impact on Health and Well-Being? *International Journal of Environmental Research and Public Health*, *6*(9), 2332–2343. https://doi.org/10.3390/ijerph6092332

Houlden, V., Weich, S., & Jarvis, S. (2017). A cross-sectional analysis of green space prevalence and mental wellbeing in England. *BMC Public Health*, *17*(1), 460–469. https://doi.org/10.1186/s12889-017-4401-x

Hyvönen, K., Törnroos, K., Salonen, K., Korpela, K., Feldt, T., & Kinnunen, U. (2018). Profiles of Nature Exposure and Outdoor Activities Associated with Occupational Well-Being among Employees. *Frontiers in Psychology*, *9*(754), 1–13. https://www.frontiersin.org/article/10.3389/fpsyg.2018.00754

Jiang, B., Chang, C.-Y., & Sullivan, W. C. (2014). A dose of nature: Tree cover, stress reduction, and gender differences. *Landscape and Urban Planning*, *132*, 26–36. https://doi.org/10.1016/j.landurbplan.2014.08.005

Kaplan, R., & Kaplan, S. (1989). *The Experience of Nature: A Psychological Perspective*. Cambridge University Press.

Korpela, K., De Bloom, J., Sianoja, M., Pasanen, T., & Kinnunen, U. (2017). Nature at home and at work: Naturally good? Links between window views, indoor

plants, outdoor activities and employee well-being over one year. *Landscape and Urban Planning*, *160*, 38–47. https://doi.org/10.1016/j.landurbplan.2016.12.005

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Leather, P., Pyrgas, M., Beale, D., & Lawrence, C. (1998). Windows in the Workplace Sunlight, View, and Occupational Stress. *Environment and Behavior*, *30*(6), 739–762. https://doi.org/10.1177/001391659803000601

Lee, A. C. K., & Maheswaran, R. (2011). The health benefits of urban green spaces: A review of the evidence. *Journal of Public Health*, *33*(2), 212–222. https://doi.org/10.1093/pubmed/fdq068

Lottrup, L., Grahn, P., & Stigsdotter, U. K. (2013). Workplace greenery and perceived level of stress: Benefits of access to a green outdoor environment at the workplace. *Landscape and Urban Planning*, *110*, 5–11. https://doi.org/10.1016/j.landurbplan.2012.09.002

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job Burnout. *Annual Review of Psychology*, *52*(1), 397–422. https://doi.org/10.1146/annurev.psych.52.1.397

Milfont, T. L., Denny, S., Ameratunga, S., Robinson, E., & Merry, S. (2008). Burnout and wellbeing: Testing the Copenhagen Burnout Inventory in New Zealand teachers. *Social Indicators Research*, *89*, 169–177. https://doi.org/10.1007/s11205-007-9229-9

Mitchell, R. (2013). Is physical activity in natural environments better for mental health than physical activity in other environments? *Social Science & Medicine*, *91*, 130–134. https://doi.org/10.1016/j.socscimed.2012.04.012

Pines, A., & Aronson, E. (1988). *Career burnout: Causes and cures*. New York: Free Press.

Popov, B., Jelić, D., Milinković, I., & Dinić, B. (2022). *Psychometric properties of the Job characteristics questionnaire*. *Manuscript in preparation*.

Raanaas, R. K., Evensen, K. H., Rich, D., Sjøstrøm, G., & Patil, G. (2011). Benefits of indoor plants on attention capacity in an office setting. *Journal of Environmental Psychology*, *31*(1), 99–105. https://doi.org/10.1016/j.jenvp.2010.11.005

Rebar, A. L., Stanton, R., Geard, D., Short, C., Duncan, M. J., & Vandelanotte, C. (2015). A meta-meta-analysis of the effect of physical activity on depression and

anxiety in non-clinical adult populations. *Health Psychology Review*, *9*(3), 366–378. https://doi.org/10.1080/17437199.2015.1022901

Repke, M. A., Berry, M. S., Conway, L. G., Metcalf, A., Hensen, R. M., & Phelan, C. (2018). How does nature exposure make people healthier? Evidence for the role of impulsivity and expanded space perception. *PLoS ONE*, *13*(8), e1–e20. https://doi.org/10.1371/journal.pone.0202246

Sanchez, J. A., Sanchez, S. V., Ikaga, T., Ichihara, M., & Harimoto, K. (2018). The impact of greenery and daylight on productivity and well-being at the workplace: An experimental case study. *Journal for Facility Management*, *1*(15), 20–32. https://doi.org/10.34749/jfm.2017.2647

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior*, *25*(3), 293–315. https://doi.org/10.1002/job.248

Schaufeli, W. B., & Greenglass, E. R. (2001). Introduction to special issue on burnout and health. *Psychology & Health*, *16*(5), 501–510. https://doi.org/10.1080/08870440108405523

Schaufeli, W. B., & Taris, T. W. (2014). A critical review of the Job Demands-Resources Model: Implications for improving work and health. In G. Bauer & O. Hämmig (Eds.), *Bridging occupational, organizational and public health* (pp. 43–68). Dordrecht, the Netherlands: Springer. http://dx.doi.org 10.1007/978-94-007-5640-3_4

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a Short Questionnaire: A cross-national study. *Educational and Psychological Measurement*, *66*(4), 701–716. https://doi.org/10.1177/0013164405282471

Schaufeli, W. B., Salanova, M., González-romá, V., & Bakker, A. B. (2002). The Measurement of Engagement and Burnout: A Two Sample Confirmatory Factor Analytic Approach. *Journal of Happiness Studies*, *3*(1), 71–92. https://doi.org/10.1023/A:1015630930326

Schaurer, I., & Weiß, B. (2020). Investigating selection bias of online surveys on coronavirus-related behavioral outcomes. *Survey Research Methods, 14*(2), 103–108. https://doi.org/10.18148/srm/2020.v14i2.7751

Shibata, S., & Suzuki, N. (2002). Effects of the foliage plant on task performance and mood. *Journal of Environmental Psychology*, *22*(3), 265–272. https://doi.org/10.1006/jevp.2002.0232

Shin, W. S. (2007). The influence of forest view through a window on job satisfaction and job stress. *Scandinavian Journal of Forest Research*, *22*(3), 248–253. https://doi.org/10.1080/02827580701262733

Shirom, A. (1989). Burnout in work organizations. In C. L. Cooper & I. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 26–48). New York: Wiley.

Smith, A., & Pitt, M. (2009). Sustainable workplaces: Improving staff health and well-being using plants. *Journal of Corporate Real Estate*, *11*(1), 52–63. https://doi.org/10.1108/14630010910940552

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics (6th ed.)*. Boston, MA: Pearson.

Tement, S., Korunka, C., & Pfifer, A. (2010). Toward the assessment of the work–family interface: Validation of the Slovenian versions of work–family conflict and work–family enrichment scales. *Psihološka Obzorja*, *19*(3), 53–74.

Thompson, A., & Bruk-Lee, V. (2019). Naturally! Examining Nature's Role in Workplace Strain Reduction. *Occupational Health Science*, *3*(1), 23–43. https://doi.org/10.1007/s41542-019-00033-5

Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science (New York, N.Y.)*, *224*(4647), 420–421. https://doi.org/10.1126/science.6143402

Ulrich, R.S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*(3), 201–230. https://doi.org/10.1016/S0272-4944(05)80184-7

United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. New York: United Nations.

Wang, R., Helbich, M., Yao, Y., Zhang, J., Liu, P., Yuan, Y., & Liu, Y. (2019). Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures. *Environmental Research*, *176*, 108535. https://doi.org/10.1016/j.envres.2019.108535

Weng, P.-Y., & Chiang, Y.-C. (2014). Psychological Restoration through Indoor and Outdoor Leisure Activities.*Journal of Leisure Research*, *46*(2), 203–217. https://doi.org/10.1080/00222216.2014.11950320

Yeo, N. L., Elliott, L. R., Bethel, A., White, M. P., Dean, S. G., & Garside, R. (2020). Indoor Nature Interventions for Health and Wellbeing of Older Adults in Residential Settings: A Systematic Review. *The Gerontologist*, *60*(3), e184–e199. https://doi.org/10.1093/geront/gnz019

Zhang, J. W., Howell, R. T., & Iyer, R. (2014). Engagement with natural beauty moderates the positive relation between connectedness with nature and psychological well-being. *Journal of Environmental Psychology*, *38*, 55–63. https://doi.org/10.1016/j.jenvp.2013.12.013

Primenjena psihologija
Vol. 16, No. 1, pp. 59-81, 2023

Research Article

„You Can Run but You Can't Hide "– The Role of Avoidant Coping in Mental Health of Athletes during COVID-19 Pandemic

Vesna Barzut [1]✉, Dušanka Đurović [1] and Ivana Novakov [1]

[1] *Department of psychology, Faculty of sport and psychology, Educons University, Novi Sad, Serbia*

ABSTRACT

This study aimed to investigate the psychological functioning of athletes in Serbia during the second wave of the COVID-19 pandemic. We investigated the relationships of different coping strategies (problem-focused, emotion-focused, and avoidant coping), intolerance of uncertainty, and perceived levels of depression, anxiety, and stress. A total of 117 athletes took part in this study (62.40% male, average age 29.95). The results showed that the use of avoidant coping and problem-focused strategies were associated with poorer indicators of mental health. Furthermore, intolerance of uncertainty had both direct and indirect (via avoidant coping strategies) effects on experienced levels of depression, anxiety, and stress.

Keywords: coping strategies, COVID-19, athletes, mental health, intolerance of uncertainty

UDK: 796.071-056.32:578.834
DOI: 10.19090/pp.v16i1.2411
Received: 28.05.2022.
Revised: 07.09.2022.
Accepted: 19.09.2022.

✉ Corresponding author email: vesna.barzut@tims.edu.rs

Introduction

COVID-19 outbreak completely changed our daily lives, affecting not only health but all domains of human functioning. The World Health Organization (WHO, 2020) declared a global pandemic and many countries across the world introduced various restrictive measures (e.g., social distancing, self-isolation, and quarantine of those who were exposed to COVID-19) in order to get control over the spreading of coronavirus disease. Many sports events (including Olympic games) were postponed or completely canceled, which affected the professional and private goals of many professional athletes. COVID-19 pandemic also greatly affected mental health of people across the globe. Participants from general population reported higher level of anxiety and distress (Keeter, 2020; Liu et al. 2020), insomnia (Voitsidis et al., 2020) or even symptoms of posttraumatic stress (Liu et al. 2020; González-Sanguino et al., 2020). While some studies, conducted before COVID-19 pandemic, suggest that athletes are showing fewer signs of depression compared with non-athletes (Brand et al., 2013) there is also evidence that professional athletes are also experiencing various mental health symptoms and disorders at the same or even higher extent compared to non-athletes (e.g., Readon et al. 2019).

Mental health of athletes during COVID-19 pandemic

In the context of global pandemic, athletes were facing direct effects of the pandemic such as potential illness, death of loved ones or fear of their own mortality, social isolation, and loss of personal freedom; and sport-specific effects such as losing job and career opportunity due to delay or cancellation of important sport events, losing or lack of sponsorship that affected financial status and living conditions, etc. Additionally, many countries banned sports activities, which could also be one of the factors that contribute to the mental and physical health of athletes and physically active people, as there is a well-documented relationship between exercise and mental health. For example, it was shown that physical activity is one of the factors that could reduce levels of post-traumatic stress, depression, and anxiety (Leard Mann

et al., 2011; Mammen & Faulkner, 2013; Schuch et al., 2019). Finally, people could use different sports activities to better cope with ongoing pandemic and related stress and in situations where sports activities were banned, they had to come up with different alternative strategies.

In one of the studies that investigated the psychological functioning of athletes during the COVID-19 outbreak (Şenışık et al., 2021) it was shown that the general distress (measured as total score on The Depression, Anxiety, and Stress Scale [DASS-21; Lovibond & Lovibond, 1995]) was significantly lower among athletes compared to non-athletes. Another study conducted in Serbia (Sokić et al., 2021) showed negative association between training and distress. Namely, professional athletes, as well as those with a high level of physical activities, experienced less distress during the outbreak of the pandemic and curfew. Interestingly, it was shown that professional athletes who changed their training routine experienced lower levels of anxiety compared to recreational athletes who did not change their training routine and recreational athletes who did change it (Sokić et al, 2021). The authors of this study explained that in the time of a global pandemic it seems that physical activity does not provide additional benefits for mental health (Sokić et al, 2021). On the other hand, elite athletes might be more resilient compared to other groups and that could help them to better adjust even during crisis.

Coping strategies and mental health during COVID-19 pandemic

As we already mentioned, exercise has many positive effects on mental health (Brand et al., 2013; Mikkelsen et al., 2017). We are learning more and more about the psychological, but also biochemical mechanisms, through which exercise and physical activity have a positive effect on mental health and mood. However, when dealing with stressful events such as pandemic, our coping strategies could probably explain why some people handle stressful events better than the others. The relationship between the use of coping strategies and psychological functioning was also investigated in the context of the COVID-19 pandemic. Nevertheless, these studies showed

mixed results regarding the relationship between coping strategies and mental health. For example, some studies showed (e.g., Guo et al., 2020) that problem-focused strategies are associated with fewer mental health issues while the opposite trend was observed for emotion-focused coping strategies. Contrary, another study (Li, 2020) suggested that using both problem-focused and emotion-focused strategies is associated with better psychological status while only usage of problem-focused strategies was linked with higher PTSD symptoms. Furthermore, it was shown that positive thinking, active coping, and social support are positively associated with well-being and negatively with stress, anxiety, and depression (Budimir et al., 2021). Also, avoiding thinking about issues or struggling to cope was associated with higher levels of depression and anxiety (Kar et al., 2021). The study conducted with the athlete population in Serbia at the beginning of the pandemic showed that emotion-focused and avoidant coping strategies were associated with worse mental health while problem-focused strategies were negatively associated with depression and anxiety (Popov et al., 2021). However, it is important to note that all these studies were conducted during the first wave of the pandemic. One study which was conducted during the second wave of the pandemic in Japan (Fukase et al., 2021) showed that certain coping strategies were associated with depression. Those strategies were: planning, use of instrumental support, denial, behavioral disengagement, and self-blame. Planning and instrumental support could be classified as problem-focused strategies, while the rest belong to the group of avoidant coping strategies.

Our study was conducted in the period from December 2020 to the first week of March 2021. At that time, many athletes continued with their regular activities. Some of the competitions were even re-established, although there was still a lot of uncertainty (e.g., it was not yet sure if the Olympic games would be held or not), since it was unknown in which direction the pandemic would go. In such a situation, athletes had limited control over their professional future, as they could only take precautionary measures and keep training, while decisions about holding sports events were out of their control. When it seems hard or impossible to control a

problem in a highly uncertain situation, use of different coping strategies could be more or less effective. Although it is often assumed that using problem-focused strategies is associated with better mental health in some extreme situation (such as global pandemic) that is not necessarily true. For example, a study that investigated the association between coping strategies and adjustment in the low-control situation (e.g., Terry & Haynes, 1998) showed that problem-focused management and avoidant coping strategies were associated with poorer outcomes while emotion-focused ones were related to more positive outcomes.

Intolerance of uncertainty, coping strategies, and mental health

In the times of global crises such as pandemics, which generally bring heightened levels of fear and uncertainty, one of the constructs which can facilitate understanding of the human capacity to deal with such circumstances is intolerance of uncertainty (IU). This concept could be defined as "an individual's dispositional incapacity to endure the aversive response triggered by the perceived absence of salient, key, or sufficient information, and sustained by the associated perception of uncertainty" (Carleton, 2015, p. 31). Previous studies confirmed the relation between high IU and anxiety and depression (e.g., Boelen & Reijntjes, 2009; Gentes & Ruscio, 2011). Furthermore, it was shown that intolerance of uncertainty is a useful concept for understanding psychological functioning during pandemics, both H1N1 (Taha et al., 2014) and current, COVID-19 (e.g., Blanuša et al., 2020; Blanuša et al., 2021; Ferreira et al., 2020; Satici et al., 2020). Also, the association of intolerance of uncertainty with particular coping strategies was confirmed. For example, IU had not only direct effects on H1N1 anxiety, but also indirect ones via emotion-based coping strategies (Taha et al., 2016).

The present study

In this study we aimed to examine the psychological functioning of athletes in the situation of prolonged stress caused by the second wave of the COVID-19 pandemic. In order to do so we formulated several hypotheses:

H_1: Emotion-focused strategies would be associated with lower symptoms of depression, anxiety, and stress. At the time of data collection, coronavirus was not a completely new threat anymore, the medical treatment protocols were improved, and during December of 2020, Serbia got the first coronavirus vaccine. Therefore, we assumed that athletes` main concerns were training and preparing for competitions (and facing sports-related uncertainties).
H_2: Problem-focused strategies, as well as avoidant coping strategies, will be associated with elevated symptoms of depression, anxiety, and stress. Avoiding dealing with unpleasant events is in most situations a dysfunctional strategy. On the other hand, if the main source of stress is out of someone's control (e.g., uncertainty about future sports events), we assume that problem-focused strategies could only increase distress.
H_3: The intolerance of uncertainty would predict levels of experienced depression, anxiety, and stress among athletes, both directly and indirectly via the usage of coping strategies (problem-focused, emotion-focused, or avoidant).

Method

Sample and procedure

A total of 117 athletes (77 professional and 40 recreational athletes) took part in this study, 62.4% of which were male. The average age of study participants was 29.95 (*SD* = 11.61). According to Serbian Law on sports (Zakon o sportu, 2016), recreational sport includes engaging in sports activities for the purpose of leisure and recreation, improving health or improving one's own results, in all segments of the population. On the other hand, professional athletes are those whose primary or only occupation is sport (i.e., they are paid to train and to compete). The status of athletes in our study was determined based on their self-report. In Table 1 more detailed information about our sample is given.

Table 1

Demographic characteristics of the sample

	professional athletes	recreational athletes
male		
athletes without disability	33	21
athletes with disability	10	9
female		
athletes without disability	28	10
athletes with disability	6	0

Because the group of athletes with disabilities was relatively small and they did not differ (based on Mann-Whitney U test) from the group of athletes without disabilities regarding their intolerance of uncertainty, anxiety, depression, or stress score as well as regarding usage of problem-focused, emotion-focused and avoidant coping strategies, we did not use this factor in further analysis. Our initial idea was to compare athletes with and without disabilities since we assumed that athletes with disabilities might developed some additional coping strategies and resilience that could be also useful in the context of pandemic. However, in this study those differences were not obtained.

Participants voluntarily took part in this online study (the survey was administered using the *Google forms* software). We used the snowball sampling method for recruiting participants. Before filling out the questionnaire, participants were informed about the aim of the research and by clicking the button "I agree to take part in this study" they gave consent.

Instruments

Brief COPE (Carver, 1997: Serbian translation and adaptation: Živanović & Vukčević-Marković, 2019)

We used Brief COPE (Carver, 1997: Serbian translation and adaptation: Živanović & Vukčević-Marković, 2019) to measure 14 coping strategies: self-distraction, active coping, denial, substance use, emotional support, instrumental support, behavioral disengagement, venting, positive reframing, planning, humor, acceptance, religion, self-blame. For analysis and more comprehensive interpretation we grouped these 14 coping strategies into 3 broader groups (like it was done in Popov et al., 2021): *problem-focused* (active coping, instrumental support, and planning), *emotion-focused* (acceptance, positive reframing, emotional support, religion, and humor) and *avoidant coping strategies* (denial, self-distraction, self-blame, behavioral disengagement, venting and substance use). The Cronbach alpha coefficients were: .81 for problem-focused strategies (6 items), .70 for emotion-focused (10 items), and .73 for avoidant strategies (12 items). Participants rated using a 4-point scale (from 1 to 4) how each of described behaviors is typical for them in a highly stressful situation.

The Depression Anxiety Stress Scales-21 (DASS-21; Lovibond & Lovibond, 1995; Serbian translation: Jovanović et al., 2014)

Furthermore, for evaluating levels of anxiety, stress, and depression we used The Depression Anxiety Stress Scales-21 (DASS-21; Lovibond & Lovibond, 1995; Serbian translation: Jovanović et al., 2014). DASS-21 consists of 21 items with a 4-point scale (from 0 to 3) and 3 subscales: depression (7 items, α = .87), anxiety (7 items, α = .79), and stress (7 items, α = .87). Participants were instructed to evaluate how they felt in the last 7 days.

Intolerance of Uncertainty Scale (IUS-11; Mihić et al., 2014)

For measuring intolerance of uncertainty, we used the Serbian version of the short Intolerance of Uncertainty Scale (IUS-11; Mihić et al., 2014). This instrument consists of 11 items with a 5-point scale (α = .85). The items

examine a person's tendency to react negatively in uncertain and ambiguous situations, evaluating cognitive, behavioral, and emotional aspects of those reactions.

Data analytic plan

JASP version 0.16.3 (JASP team, 2022) was used for the analysis. The Cronbach's alpha was estimated to examine the reliability of used instruments. Descriptive statistics were calculated, and we used Pearson's coefficient correlation analysis for testing the relationships between the variables. Finally, 9 simple mediation analysis [1 predictor x 3 mediators x 3 outcomes] were performed taking intolerance of uncertainty as the predictor, coping strategies (problem-focused, emotion-focused and avoidant) as mediators, and the depression, anxiety, and stress as outcome variables. Based on Hair's recommendation (Hair et al., 2013), the minimum sample size required to detect a minimum R^2 value of 0.10 in any endogenous construct in a structural model with two independent variables for a significance of 5% (i.e., .05 alpha level) and assuming statistical power of 80% is 110 (in our study 117 subjects participated). Additionally, since we tested 9 models, we used Bonferroni correction. The adjusted alpha level was .006.

Results

Descriptive statistics are presented in Table 2. Skewness value was smaller than 3 (ranged from -.85 to 1.99) and kurtosis value was smaller than 10 (ranged from -.04 to 3.82), which is considered appropriate for structural equation modeling (Kline, 2005).

Table 2

Descriptive statistics for variables included in this report

	Theoretical range	Achieved range	*M*	*SD*	*Skewness*	*Kurtosis*
Avoidant coping (COPE)	12-24	12-39	22.77	4.95	.22	-.04
Emotion-focused (COPE)	10-40	11-38	27.39	5.09	-.21	.15
Problem-focused (COPE)	6-24	6-24	18.37	3.92	-.85	.38
Depression (DASS-21)	0-21	0-17	2.96	3.87	2.00	3.82
Anxiety (DASS-21)	0-21	0-16	2.66	3.24	1.50	2.15
Stress (DASS-21)	0-21	0-20	5.87	4.66	.88	.59
Intolerance of uncertainty (IUS-11)	11-55	11-50	22.04	7.40	1.09	1.82

Perceived levels of anxiety, stress and depression during second wave of coronavirus pandemic

The perceived levels of depression, anxiety and stress in our sample are comparable with the results from the validation study of the instrument on the Serbian student population (Jovanović et al., 2014) and the study that investigated athletes during the COVID-19 outbreak in Serbia (Sokić et al., 2021). This indicated that the majority of athletes did not report higher levels of symptoms during the second wave of the coronavirus pandemic. Additionally, the scores obtained in our study are lower compared with the study conducted in Austria in general population at the beginning of the COVID-19 pandemic (Traunmüller et al., 2020)[1], and multiple times lower compared to results from Spain[2] (González-Hernández et al., 2021) in athletic population. In order to estimate the severity of perceived levels of depression,

[1] $M_{depression}$ = 5.42 (SD = 8.38), $M_{anxiety}$ = 8.88 (SD = 10.26), M_{stress} = 10.58 (SD = 10.85), $M_{total\ DASS\text{-}21}$ = 24.87 (SD = 26.97)

[2] $M_{depression}$ = 13.15 (SD = 3.11), $M_{anxiety}$ = 15.73(SD = 3.61), M_{stress} = 16.38 (SD = 3.04)

anxiety and stress we compared our results with norms that are proposed by the authors of this scale (Lovibond & Lovibond, 1995). Since we used short version of the scale DASS-21, cut off scores are calculated by double multiplying the scores on each of subscales. As we can see from the Table 3, the majority of athletes' scores on all three scales are within the range that is considered normal. However, it is important to notice that 12.8% of athletes experienced depressive symptoms (from moderate to extremely severe), 23.1% experienced moderate-to-extreme anxiety, and 17.9% experienced moderate-to-extremely severe symptoms of stress.

Table 3

Distribution of DASS-21 scores

	Normal	Mild	Moderate	Severe	Extremely severe
Depression	79.5%	7.7%	5.1%	3.4%	4.3%
Anxiety	71.8%	5.1%	13.7%	5.1%	4.3%
Stress	68.4%	13.7%	7.6%	6.9%	3.4%

The relationship between intolerance of uncertainty, coping strategies and depression, anxiety, and stress

First, we checked the correlation between variables (Table 4). It is interesting to note that problem-focused strategies are not significantly associated with symptoms of depression, anxiety and stress but they are positively associated with intolerance of uncertainty. Furthermore, emotion-focused strategies are not significantly correlated with symptoms of anxiety but there are positive correlations with depression, stress, and intolerance of uncertainty. Avoidant coping strategies showed the highest positive correlations with symptoms of depression, anxiety, and stress. Finally, as expected, all 3 groups of coping strategies were intercorrelated.

Table 4

Intercorrelations among the variables

	1	2	3	4	5	6	7
Emotion-focused COPE							
Problem-focused COPE	.65**						
Avoidant COPE	.45**	.38**					
Depression DASS-21	.18*	.02	.51**				
Anxiety DASS-21	.15	-.02	.40**	.58**			
Stress DASS-21	.19*	.17	.55**	.77**	.64**		
IUS-11 total	.26**	.22*	.50**	.45**	.41**	.46**	

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

Next, we tested mediation models using JASP, version 0.16.3 (JASP team, 2022). Results showed 3 significant mediation models. It was shown that IU has not only a direct effect on experienced levels of depression among athletes but also an indirect effect via avoidant coping strategies. The same trend was observed regarding anxiety and stress. Avoidant coping strategies had a mediating role (Table 5) between IU and anxiety, depression, and stress (Figure 1, 2 and 3).

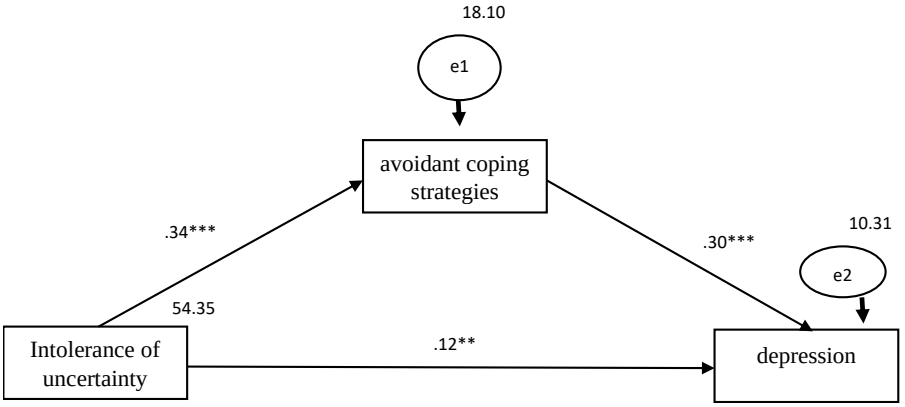

Figure 1. Mediation model for depression

Notes. Unstandardized beta coefficients are noted in the diagram. All pathways are significant.

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

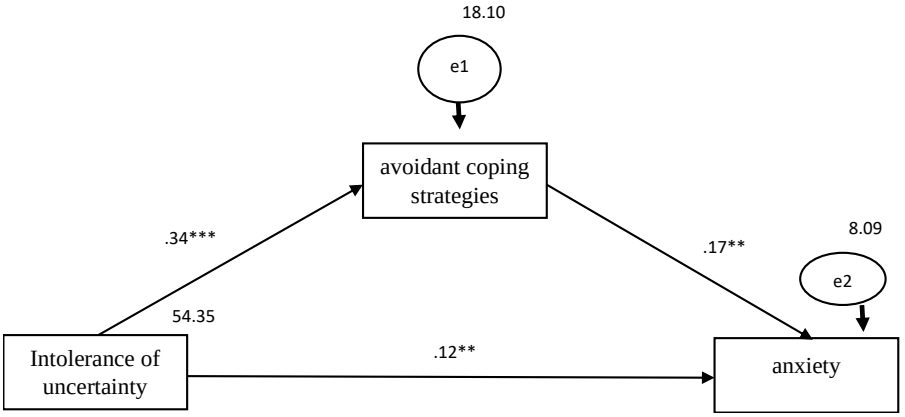

Figure 2. Mediation model for anxiety

Notes. Unstandardized beta coefficients are noted in the diagram. All pathways are significant.

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

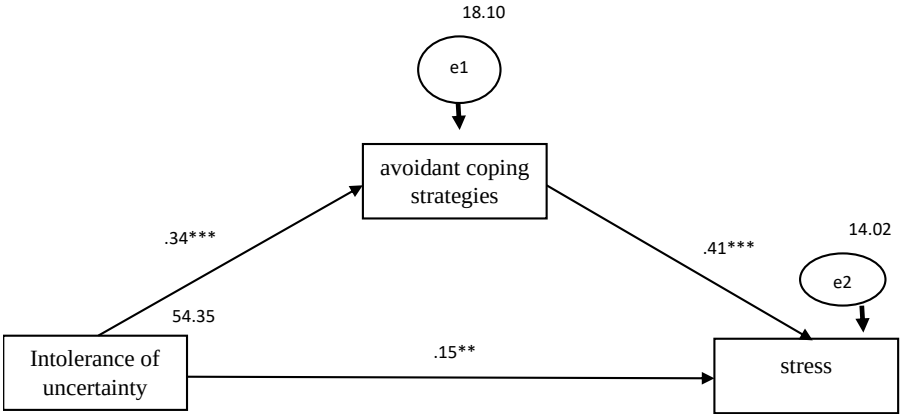

Figure 3. Mediation model for anxiety

Notes. Unstandardized beta coefficients are noted in the diagram. All pathways are significant.

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

Table 5

Mediation analysis - effects

Predictor	Mediator	Outcome	effects	estimate	*S.E.*	*z*-value	*p*	95% Confidence interval		result
								lower	*upper*	
IUS-11	Problem-focused COPE	Depression	Direct	0.24	0.04	5.49	<.001	0.16	0.33	Not significant
			Indirect	-0.01	0.01	-0.93	.350	-0.03	0.01	
			Total	0.23	0.04	5.38	<.001	0.15	0.32	
		Anxiety	Direct	0.19	0.04	5.16	<.001	0.12	0.27	Not significant
			Indirect	-0.01	0.01	-1.25	.212	-0.03	0.01	
			Total	0.18	0.04	4.92	<.001	0.11	0.25	
		Stress	Direct	0.28	0.05	5.28	<.001	0.17	0.38	Not significant
			Indirect	0.01	0.01	0.79	.431	-0.02	0.03	
			Total	0.29	0.05	5.59	<.001	0.19	0.39	
	Emotion-focused COPE	Depression	Direct	0.22	0.05	4.99	<.001	0.14	0.31	Not significant
			Indirect	0.01	0.01	0.81	.418	-0.01	0.03	
			Total	0.23	0.04	5.38	<.001	0.15	0.32	
		Anxiety	Direct	0.18	0.04	4.63	<.001	0.10	0.25	Not significant
			Indirect	0.01	0.01	0.48	.629	-0.02	0.02	
			Total	0.18	0.04	4.92	<.001	0.11	0.25	
		Stress	Direct	0.28	0.05	5.18	<.001	0.17	0.38	Not significant
			Indirect	0.01	0.01	0.87	.385	-0.02	0.04	
			Total	0.29	0.05	5.59	<.001	0.19	0.39	
	Avoidant COPE	Depression	Direct	0.13	0.05	2.83	.005	0.04	0.23	Significant
			Indirect	0.10	0.03	3.55	<.001	0.05	0.16	
			Total	0.23	0.04	5.38	<.001	0.15	0.32	
		Anxiety	Direct	0.12	0.041	2.995	.003	0.04	0.21	Significant
			Indirect	0.06	0.023	2.525	.012	0.01	0.10	
			Total	0.18	0.037	4.921	<.001	0.11	0.25	
		Stress	Direct	0.15	0.054	2.788	.005	0.05	0.26	Significant
			Indirect	0.14	0.035	3.925	<.001	0.07	0.21	
			Total	0.29	0.052	5.585	<.001	0.19	0.39	

Discussion

This study aimed to evaluate the psychological functioning of athletes during the second wave of the COVID-19 pandemic. Symptoms of anxiety, depression, and stress were used as indicators of psychological functioning. The studies from the beginning of the pandemic (e.g., Şenışık et al., 2021) showed that the athletes reported better functioning compared to

non-athletes, but we were interested in their mental health during the second wave of the pandemic. For the COVID-19 outbreak, it was typical that we all were facing the same uncertainty regarding our health and everyday life, and we all faced the same restrictions (at least when we talk about the same country). However, during the second wave, the sport remained one of the areas with high uncertainty regarding the future situation.

Reported levels of depression, anxiety, and stress symptoms in our study are lower compared to studies conducted during the first wave (e.g., González-Hernández et al., 2021; Traunmüller et al., 2020; González-Hernández et al., 2021) which could indicate some sort of adjustment. However, it is important to note that the percentage of athletes with moderate-to-extremely severe symptoms were high. Namely, 12.8% of athletes experienced depressive symptoms, 23.1% experienced moderate-to-extremely severe anxiety, and 17.9% experienced moderate-to-extremely severe symptoms of stress. In other words, even though it seems that the population of athletes is coping well with the pandemic, some individuals experienced a significant disturbance in their psychological functioning. However, since this was a cross-sectional study, it is not possible to check whether these athletes were experiencing higher distress even before the pandemic. That could be the case for those who are ending their careers, dealing with changes, injuries, etc.

Furthermore, we also detected the association between the use of particular coping strategies and experienced levels of depression, anxiety, and stress. Usage of avoidant coping strategies was associated with higher levels of anxiety, stress, and depression. This is in line with the results of previous studies (Fukase et al., 2021; Popov et al., 2021). The negative impact of avoidant coping strategies on mental health is not surprising. Living in a time of pandemic means that we will be constantly reminded about the ongoing global crisis, so to avoid unpleasant thoughts and feelings, an individual must put the additional effort that will increase the burden and experienced levels of distress (since we cannot hide from the unpleasant reality of pandemic). Furthermore, emotion-focused strategies were associated with higher levels of depressive and stress symptoms but not with

anxiety. Interestingly, at the beginning of the pandemic, the use of emotion-focused strategies was associated with higher levels of depressive and anxiety symptoms (Popov et al., 2021). This is an interesting result since it might seem logical that in a situation with high uncertainty, working on acceptance and emotion regulation will be beneficial. Finally, in our study, problem-focused strategies were not associated with experienced levels of depression, anxiety, and stress. At the beginning of the pandemic, using problem-focused strategies was associated with lower levels of depression, anxiety, and stress (Popov et al., 2021). In other words, when uncertainty regarding the virus and disease was high, focusing on the problem and taking precautions measures could give someone a sense of control and hope, which could be a protective factor for mental health. However, at the time when we collected our data, the main concerns for athletes were what will be decided regarding the future competition. Since they could not directly influence these decisions, perhaps using the problem-focused strategies did not provide additional benefits for their mental health.

This study also showed the association between someone's intolerance of uncertainty, coping strategies, and the experienced level of stress, anxiety, and depression. Intolerance of uncertainty was a predictor of the experienced level of depression, anxiety, and stress which is a common result in literature (Boelen & Reijntjes, 2009; Gentes & Ruscio, 2011). This is also in accordance with the idea that intolerance of uncertainty is a transdiagnostic risk factor for emotional disorders (e.g., Boelen & Reijntjes, 2009; Gentes & Ruscio, 2011). Intolerance of uncertainty also was associated with the use of different coping strategies (avoidant, emotion-focused and problem-focused). Such a result is in line with previous studies (e.g., Sankar et al., 2017) where behavioral manifestations of Intolerance of uncertainty were classified into five categories: under-engagement, over-engagement, impulsive decision making, flip-flop, and dither. Under-engagement describes avoidant behavior while over-engagement is described as approach behavior i.e., there is a clear parallel with avoidant coping strategies and problem-focused strategies.

Furthermore, it is interesting to mention the indirect effect of intolerance of uncertainty on depression, stress, and anxiety via the avoidant coping mechanism. Higher intolerance of uncertainty was associated with the usage of avoidant coping strategies that are shown as malfunctioning coping mechanisms. In other words, an athlete who cannot tolerate uncertainty might get involved in dysfunctional patterns of behavior and thinking, such as substance use, venting, denial, self-distraction, self-blame, and behavioral disengagement, as a form of self-soothing and self-distraction, which in turn will lead again to the increased levels of stress, anxiety, and depression. Improving communications and providing clear and timely information to the athletes perhaps could decrease the uncertainty and related dysfunctional behavior during the crisis. However, since uncertainty is an inevitable part of life, athletes could also benefit from the usage of brief cognitive intervention that would target intolerance of uncertainty and help them cope better (Oglesby et al., 2017). Furthermore, the obtained results indicate the need to promote resilience in athletes in times of crisis, which could be achieved through counseling and professional support to reduce the harmful psychological and emotional effects of the situation caused by the pandemic.

Limitations

This study has several limitations. The main issue was the fact that the sample was convenient and unrepresentative. Even though we had a category of professional athletes, the number of elite athletes (such as Olympic athletes) was small, as well as the number of athletes with disabilities. Further studies should include a larger, more representative sample. Also, it would be interesting to compare the athletes with a disability and those without it. Perhaps athletes with a disability are already using some additional coping strategies that might increase their resilience even in times of global crisis. Potentially, we did not detect differences due to the small sample size. Furthermore, this was a cross-sectional study, and therefore we do not have the previous indicators of mental health measures of the athletes who participated. It would be interesting to follow possible changes in the

usage of particular coping strategies as well as the experienced levels of depression, anxiety, and stress. Finally, since our study was correlational, the assumed causal relationship between the variables needs to be further explored.

Acknowledgement

Part of the data were previously presented on conference Scientific Trends in Psychology, 2021, October 2021, Faculty of Philosophy, University of Novi Sad, Serbia.

Conflict of Interests

The authors declare that they have no conflict of interests.

Data availability statement

Data files are available upon request.

References

Blanuša, J., Barzut, V. & Knežević, J. (2020). Direct and indirect effect of intolerance of uncertainty on distress during the COVID-19 pandemic. *Primenjena Psihologija, 13*(4), 473–487. https://doi.org/10.19090/pp.20.4.473-487

Blanuša, J., Barzut, V. & Knežević, J. (2021). Intolerance of Uncertainty and Fear of COVID-19 Moderating Role in Relationship Between Job Insecurity and Work-Related Distress in the Republic of Serbia. *Frontiers in Psychology*, 2170. https://doi.org/10.3389/fpsyg.2021.647972

Boelen, P. A., & Reijntjes, A. (2009). Intolerance of uncertainty and social anxiety. *Journal of Anxiety Disorders*, *23*(1), 130–135. https://doi.org/10.1016/j.janxdis.2008.04.007

Brand, R., Wolff, W., & Hoyer, J. (2013). Psychological symptoms and chronic mood in representative samples of elite student-athletes, deselected student-athletes and comparison students. *School Mental Health, 5*, 166–174. https://doi.org/10.1007/s12310-012-9095-8

Budimir, S., Probst, T., & Pieh, C. (2021). Coping strategies and mental health during COVID-19 lockdown. *Journal of Mental Health*, *30*(2), 156–163. https://doi.org/10.1080/09638237.2021.1875412

Carleton, R. N. (2016). Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders*, *39*, 30–43. https://doi.org/10.1016/j.janxdis.2016.02.007

Carver, C. S. (1997). You want to Measure Coping But Your Protocol's Too Long: Consider the Brief COPE. *International Journal of Behavioral Medicine, 4*, 92–100. https://doi.org/10.1207/s15327558ijbm0401_6

Ferreira, D. C. S., Oliveira, W. L., Delabrida, Z. N. C., Faro, A., & Cerqueira-Santos, E. (2020). Intolerance of uncertainty and mental health in Brazil during the Covid-19 pandemic. *Suma Psicológica, 27*(1), 62–69. https://doi.org/10.14349/sumapsi.2020.v27.n1.8

Fukase, Y., Ichikura, K., Murase, H., & Tagaya, H. (2021). Depression, risk factors, and coping strategies in the context of social dislocations resulting from the second wave of COVID-19 in Japan. *BMC Psychiatry*, *21*(1), 1–9. https://doi.org/10.1186/s12888-021-03047-y

Gentes, E. L., & Ruscio, A. M. (2011). A meta-analysis of the relation of intolerance of uncertainty to symptoms of generalized anxiety disorder, major depressive disorder, and obsessive–compulsive disorder. *Clinical Psychology Review*, *31*(6), 923–933. https://doi.org/10.1016/j.cpr.2011.05.001

González-Hernández, J., López-Mora, C., Yüce, A., Nogueira-López, A., & Tovar-Gálvez, M. I. (2021). "Oh, My God! My Season Is Over!" COVID-19 and Regulation of the Psychological Response in Spanish High-Performance Athletes. *Frontiers in Psychology*, *12*, 622529. https://doi.org/10.3389/fpsyg.2021.622529

González-Sanguino, C., Ausín, B., Castellanos, M. Á., Saiz, J., López-Gómez, A., Ugidos, C., & Muñoz, M. (2020). Mental health consequences during the initial stage of the 2020 Coronavirus pandemic (COVID-19) in Spain. *Brain, Behavior, and Immunity*, *87*, 172–176. 10.1016/j.bbi.2020.05.040

Guo, J., Feng, X. L., Wang, X. H., & van IJzendoorn, M. H. (2020). Coping with COVID-19: exposure to COVID-19 and negative impact on livelihood predict elevated mental health problems in Chinese adults. *International Journal of Environmental Research and Public Health*, *17*(11), 3857. https://doi.org/10.3390/ijerph17113857

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2013). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: Sage Publications.

JASP Team. (2022). JASP (Version 0.16.1)[Computer software]. Retrieved from https://jasp-stats.org/.

Jovanović, V., Gavrilov-Jerković, V., Žuljević, D., & Brdarić, D. (2014). Psychometric evaluation of the depression anxiety stress scales-21 (DASS-21) in a Serbian student sample. *Psihologija, 47*(1), 93–112. https://doi.org/10.2298/psi1401093j

Kar, N., Kar, B., & Kar, S. (2021). Stress and coping during COVID-19 pandemic: Result of an online survey. *Psychiatry Research*, *295*, 113598. https://doi.org/10.1016/j.psychres.2020.113598

Keeter, S. (2022, June 30). People financially affected by coronavirus outbreak are experiencing more psychological distress than others. Pew Research Center. Retrieved July 13, 2022, from https://www.pewresearch.org/fact-tank/2020/03/30/people-financially-affected-by-covid-19-outbreak-are-experiencing-more-psychological-distress-than-others/

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. 2nd Ed. New York: Guilford Press.

LeardMann, C. A., Kelton, M. L., Smith, B., Littman, A. J., Boyko, E. J., Wells, T. S., ... & Millennium Cohort Study Team. (2011). Prospectively assessed posttraumatic stress disorder and associated physical activity. *Public Health Reports*, *126*(3), 371–383. https://doi.org/10.1177/003335491112600311

Li, Q. (2020). Psychosocial and coping responses toward 2019 coronavirus diseases (COVID-19): a cross-sectional study within the Chinese general population. *QJM: An International Journal of Medicine*, *113*(10), 731–738. https://doi.org/10.1093/qjmed/hcaa226:hcaa226

Liu, D., Ren, Y., Yan, F., Li, Y., Xu, X., Yu, X., Qu, W., Wang, Z., Tian, B., Yang, F., Yao, Y., Tan, Y., Jiang, R., & Tan, S. (2020). Psychological impact and predisposing factors of the coronavirus disease 2019 (covid-19) pandemic on general public in China. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3551415

Liu, N., Zhang, F., Wei, C., Jia, Y., Shang, Z., Sun, L., ... & Liu, W. (2020). Prevalence and predictors of PTSS during COVID-19 outbreak in China hardest-hit areas: Gender differences matter. *Psychiatry Research, 287,* 112921. 10.1016/j.psychres.2020.112921

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety & Stress Scales. (2nd Ed.).* Psychology Foundation of Australia.

Mammen, G., & Faulkner, G. (2013). Physical activity and the prevention of depression: a systematic review of prospective studies. *American Journal of*

Preventive Medicine, *45*(5), 649–657. https://doi.org/10.1016/j.amepre.2013.08.001

Mihić, L., Sokić, J., Samac, N., & Ignjatović, I. (2014). Serbian adaptation and validation of the intolerance of uncertainty scale. *Primenjena Psihologija, 7*(3-1), 347–370. https://doi.org/10.19090/pp.2014.3-1.347-370

Mikkelsen, K., Stojanovska, L., Polenakovic, M., Bosevski, M., & Apostolopoulos, V. (2017). Exercise and mental health. *Maturitas*, *106*, 48–56. https://doi.org/10.1016/j.maturitas.2017.09.003

Oglesby, M. E., Allan, N. P., & Schmidt, N. B. (2017). Randomized control trial investigating the efficacy of a computer-based intolerance of uncertainty intervention. *Behaviour research and therapy*, *95*, 50–57. https://doi.org/10.1016/j.brat.2017.05.007

Popov, S., Sokić, J., & Stupar, D. (2021). Activity matters: Physical exercise and stress coping during the 2020 COVID-19 state of emergency. *Psihologija*, *54*(3), 307–322. https://doi.org/10.2298/PSI200804002P

Reardon, C. L., Hainline, B., Aron, C. M., Baron, D., Baum, A. L., Bindra, A., Budgett, R., Campriani, N., Castaldelli-Maia, J. M., Currie, A., Derevensky, J. L., Glick, I. D., Gorczynski, P., Gouttebarge, V., Grandner, M. A., Han, D. H., McDuff, D., Mountjoy, M., Polat, A., Purcell, R., Putukian, M., Rice, S., Sills, A., Stull, T., Swartz, L., Zhu, L.J,, Engebretsen, L. (2019). Mental health in elite athletes: International Olympic Committee Consensus statement (2019). *British Journal of Sports Medicine, 53*(11), 667–699. https://doi.org/10.1136/bjsports-2019-100715

Sankar, R., Robinson, L., Honey, E., & Freeston, M. (2017). 'We know intolerance of uncertainty is a transdiagnostic factor, but we don't know what it looks like in everyday life': A systematic review of intolerance of uncertainty behaviours. *Clinical Psychology Forum, 296,* 10–15.

Satici, B., Saricali, M., Satici, S. A., & Griffiths, M. D. (2020). Intolerance of uncertainty and mental wellbeing: serial mediation by rumination and fear of COVID-19. *International Journal of Mental Health and Addiction, 20*, 2731–2742. https://doi.org/10.1007/s11469-020-00305-0

Savitsky, B., Findling, Y., Ereli, A., & Hendel, T. (2020). Anxiety and coping strategies among nursing students during the covid-19 pandemic. *Nurse Education in Practice*, *46*, 102809. https://doi.org/10.1016/j.nepr.2020.102809

Schuch, F. B., Stubbs, B., Meyer, J., Heissel, A., Zech, P., Vancampfort, D., ... & Hiles, S. A. (2019). Physical activity protects from incident anxiety: A meta-analysis of

prospective cohort studies. *Depression and Anxiety, 36*(9), 846–858. https://doi.org/10.1002/da.22915

Şenışık, S., Denerel, N., Köyağasıoğlu, O., & Tunç, S. (2021). The effect of isolation on athletes' mental health during the COVID-19 pandemic. *The Physician and Sports Medicine*, *49*(2), 187–193. https://doi.org/10.1080/00913847.2020.1807297

Sokić, J., Popov, S., Dinić, B. M., & Rastović, J. (2021). Effects of Physical Activity and Training Routine on Mental Health During the COVID-19 Pandemic and Curfew. *Frontiers in Psychology*, *12*, 624035. https://doi.org/10.3389/fpsyg.2021.624035

Taha, S., Matheson, K., Cronin, T., & Anisman, H. (2014). Intolerance of uncertainty, appraisals, coping, and anxiety: The case of the 2009 H 1 N 1 pandemic. *British Journal of Health Psychology, 19*(3), 592–605. https://doi.org/10.1111/bjhp.12058

Terry, D. J., & Hynes, G. J. (1998). Adjustment to a low-control situation: Reexamining the role of coping responses. *Journal of Personality and Social Psychology*, *74*(4), 1078–1092. https://doi.org/10.1037/0022-3514.74.4.1078

Traunmüller, C., Stefitz, R., Gaisbachgrabner, K., & Schwerdtfeger, A. (2020). Psychological correlates of COVID-19 pandemic in the Austrian population. *BMC Public Health, 20*(1), 1–16. https://doi.org/10.1186/s12889-020-09489-5

Voitsidis, P., Gliatas, I., Bairachtari, V., Papadopoulou, K., Papageorgiou, G., Parlapani, E., Syngelakis, M., Holeva, V. & Diakogiannis, I. (2020). Insomnia during the COVID-19 pandemic in a Greek population. *Psychiatry Research, 289*, 113076. https://doi.org/10.1016/j.psychres.2020.113076

World Health Organization. (2020, March 18). Mental health and psychosocial considerations during the covid-19 Outbreak. World Health Organization. Retrieved July 13, 2022, from https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-MentalHealth-2020.1

Zakon o sportu [Law on sports] 2016 ("Sl. glasnik RS", br. 10/2016) (RS) retrieved from: https://www.paragraf.rs/propisi/zakon_o_sportu.html

Živanović, M. & Vukčević-Marković, M. (2019). Strategije prevladavanja stresa - kratka forma [Brief COPE]. https://osf.io/7huz8

Primenjena psihologija
Vol. 16, No. 1, pp. 83-106, 2023

Research Article

Theoretical models of Short Dark Tetrad (SD4) and its relationship with socially desirable responding: Findings on the Croatian version

Anja Wertag [1]✉, Maja Ribar [1] and Ines Sučić [1]

[1] *Institute of Social Sciences Ivo Pilar, Zagreb, Croatia*

ABSTRACT

The interest in the so-called dark traits in the area of individual differences is rising, and there are several instruments assessing currently the most prominent dark traits constellation, the Dark Tetrad. The first aim of this study was to examine the structure of the newly developed Short Dark Tetrad (SD4) measure in the Croatian context using latent variable modeling and testing two competing models: the confirmatory factor analytic model with four interrelated factors where items load only on their respective trait or factor, and the bifactor model with both the general "dark" factor where all the items load on and four specific orthogonal factors on which only their respective items load. The second aim of this study was to investigate the relationship between the Dark Tetrad traits and egoistic and moralistic socially desirable responding. Data were collected on a convenience sample of 439 participants (81% female) in Croatia. The results showed that, although some of the fit statistics of both tested models fell somewhat below the conventional acceptable fit threshold, their values were comparable to those from the original validation study, indicating that the Croatian version of SD4 is valid and can be used to assess the Dark Tetrad traits. Moreover, relations between the Dark Tetrad traits and egoistic and moralistic socially desirable responding point to the distinctiveness of narcissism from the rest of the Dark Tetrad.

Keywords: Dark Tetrad, structural models, socially desirable responding

UDK: 159.923.072(497.5)
DOI: 10.19090/pp.v16i1.2426
Received: 06.09.2022.
Revised: 26.10.2022.
Accepted: 01.11.2022.

✉ Corresponding author email: anja.wertag@pilar.hr

Introduction

After the introduction of the Dark Triad of personality (Paulhus & Willaims, 2002), the popularity of the dark traits in the area of individual differences has been exponentially rising. The Dark Triad consists of three distinct, but conceptually and empirically overlapping traits: Machiavellianism, narcissism, and psychopathy, and relatively recently, sadism was added to this constellation, forming the Dark Tetrad (Paulhus, 2014). These traits are socially offensive and aversive, but still within the normal or everyday range. Considering their relationship with basic personality traits, the most notable are negative relationships with agreeableness (e.g., Muris et al., 2017) and honesty-humility (e.g., Međedović & Petrović, 2015). Moreover, although overlapping, each Dark Tetrad trait has its distinguishing features. The key features of Machiavellianism are callousness and manipulation, of narcissism grandiosity and attention craving, of psychopathy impulsivity, and callous thrill-seeking, while enjoyment of cruelty is a distinctive feature of sadism (Paulhus, 2014). Although there are several instruments depicting each of the dark traits (see Dinić et al., 2020; Furnham et al., 2013), the most promising instrument capturing the Dark Tetrad traits seems to be the newly developed Short Dark Tetrad (SD4; Paulhus et al., 2020). As the Croatian version of this instrument was not previously evaluated, the question of the viability of the theoretical structure of this version arises. Moreover, as the common features of dark traits reflect socially aversive character prone to manipulative behavior (see Furnham et al., 2013; Paulhus, 2014), another interesting question is the relationship between the Dark Tetrad traits (as captured by SD4) and socially desirable responding.

Previous research demonstrated that SD4 has a good construct validity and is valid for the assessment of the Dark Tetrad traits in different cultural contexts (see Blötner, Ziegler, et al., 2022; Pechorro et al., 2022), and that the factor structure of SD4 can be meaningfully compared between cultures (Blötner et al., 2022). Thus, the aim of this study was to further examine the structure of the SD4 instrument in the Croatian context using

latent variable modeling. We opted to test the two competing theoretical models from the literature on the measures of the Dark Triad (e.g., Chiorri et al., 2019; Dinić et al., 2018; Persson et al., 2019), which were also previously tested in modeling SD4 (see Neumann et al., 2021). The first model was the confirmatory factor analytic (CFA) model in which four dark traits are defined as correlated, but distinct factors. In other words, this model included four interrelated factors where items loaded only on their respective trait or factor. The second one was the bifactor model, in which each item loads on its respective trait or factor, but also on a general factor, with all factors being orthogonal, thus allowing insight into the extent to which items are influenced by one common factor compared to the extent to which they are influenced by specific factors (Riese et al., 2010). In the context of the Dark Tetrad, this model included the general "dark" factor on which all the items loaded and four specific orthogonal factors on which only their respective items loaded (Figure 1).

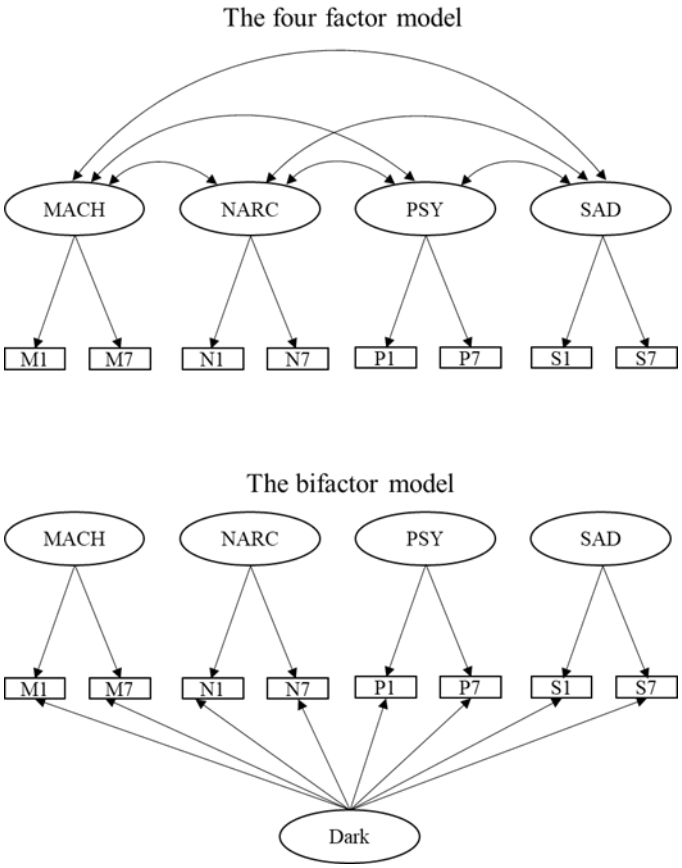

Figure 1. Two Tested Theoretical Models of the Dark Tetrad (Simplified Versions)
Note. MACH = Machiavellianism; NAR = narcissism; PSY = psychopathy; SAD = sadism.

Furthermore, given that the Dark Tetrad traits are socially aversive and characterized by manipulativeness, the second aim of this study was to investigate the relationship between the socially desirable responding and Dark Tetrad traits captured by SD4. Socially desirable responding can be defined as a tendency of giving overly positive self-descriptions (e.g., Paulhus, 2002), and whilst most conceptualizations consider it as a unidimensional construct, there is also evidence of its more complex structure (Holden & Passey, 2009). When it comes to the conceptualization of social desirability

as a unidimensional construct and its relationship with the Dark Triad traits, the results of previous research are inconsistent, and partly depend on the measure of dark traits. For example, some studies using full measures of each of the Dark Triad traits showed contradictory results for the relationship between narcissism and socially desirable responding, but more consistent results for Machiavellianism and psychopathy, indicating a negative relationship (Kowalski et al., 2016; 2018). Another study using two different short measures of the Dark Triad traits showed a consistent positive relationship between socially desirable responding and Machiavellianism, but inconsistent results for narcissism and psychopathy (Pineda et al., 2020). More specifically, socially desirable responding was positively related to psychopathy captured exclusively by Short Dark Triad (SD3; Jones & Paulhus, 2014), and narcissism captured exclusively by Dirty Dozen (DD; Jonason & Webster, 2010).

Paulhus (1984) proposed a two-component model of socially desirable responding, which includes self-deception and impression management. While self-deception is characterized by a person's belief in his or her positive self-report, impression management is characterized by deliberate positivity in self-reports. Research on the relationship between these two components of socially desirable responding and the Dark Triad traits (i.e., Gamache et al., 2018; Savard et al., 2017) primarily shows that the Dark Triad traits are not related to self-deception (with the exception of DD psychopathy; Gamache et al., 2018), and are negatively related to impression management, with the relationship being lowest in case of narcissism. Research including the Dark Tetrad traits also indicated they have a stronger negative relationship with impression management than with self-enhancement, except for narcissism which is positively related to self-enhancement (Womick et al., 2019).

In his later work, Paulhus (2002) proposed a new model according to which socially desirable responding can be classified by the levels of consciousness (i.e., conscious-unconscious) and content of self-presentation (i.e., egoistic-moralistic). More specifically, unconscious self-enhancement refers to positively biased self-descriptions that one believes to be true, while

conscious impression management represents a deliberate attempt to create a favorable self-image, depending on the characteristics of the situation. Regarding the content, egoistic bias manifests as a tendency of exaggerating social and intellectual competencies, and moralistic bias manifests as overemphasizing moral qualities and respect for social rules (Paulhus & John, 1998). Previous studies investigating relations between egoistic and moralistic socially desirable responding and basic personality traits indicated that egoistic socially desirable responding is more related to agentic traits, such as openness to experience/intellect and extraversion, while moralistic socially desirable responding is more related to communal traits, such as conscientiousness and agreeableness (e.g., Parmač Kovačić et al., 2014; Paulhus, 2002). Since previously described research indicated the stronger relationship of dark traits with deliberate impression management, or self-presentation, which can be conceptualized as egoistic and moralistic (Paulhus, 2002), the question of the relationship between dark traits and these two types of socially desirable responding arises. To the best of our knowledge, the relationship between these two types of socially desirable responding and the Dark Tetrad traits has not yet been examined.

Given that previous research comparing different models in diverse cultural contexts showed mixed findings on the specific model which fits the data the best (e.g., Neumann et al., 2021; Pechorro et al., 2022), we approached this research question as exploratory. In line with narcissistic grandiosity and specific characteristics of the other three Dark Tetrad traits (Paulhus, 2014), as well as the previous findings on the relationship of egoistic and moralistic socially desirable responding and basic personality traits (e.g., Parmač Kovačić et al., 2014; Paulhus, 2002) and relationship of the Dark Triad with basic personality traits (i.e., negative relationship of all three Dark Triad traits to agreeableness, negative relationship of psychopathy and Machiavellianism to conscientiousness, and positive relationship of narcissism and openness; Muris et al., 2017), we expected that narcissism would be primarily positively related to egoistic socially desirable responding, while psychopathy, Machiavellianism, and sadism (which, captured by SD4, is strongly correlated to and has a quite similar nomological network as psychopathy; Blötner,

Ziegler, et al., 2022) would be primarily negatively related to moralistic socially desirable responding. In sum, the present study aimed at testing the structure of SD4 using two latent variable models (i.e., four factor and bifactor model), and examining the Dark Tetrad traits' relations with two types of socially desirable responding (i.e., moralistic and egoistic).

Method

Participants and procedure

Data were collected online, on a convenience sample of a total of 439 participants (81 male; age range: 18-64, M_{age} = 25.42, SD_{age} = 7.94) in Croatia. Data were collected online, using the 1KA platform (https://www.1ka.si/), and by sharing the questionnaire link through personal contacts and social networks. Thus, a convenient sample of participants was obtained, and collected by a non-probabilistic sampling method. All the aspects of the study were approved by the institutional Ethical Board (Protocol No. 11-73/21-1199).

At the beginning of the questionnaire, the goal and purpose of the research were explained to the participants. Data confidentiality and anonymization were guaranteed. It was pointed out that by agreeing to fill out the questionnaire, they confirm their voluntary participation, and they are aged ≥18. In the first part of the questionnaire, socio-demographic data related to age and gender were collected from the participants. Then the participants expressed a degree of (non)agreement with the statements from the measures described in the next section. Participants had the option of skipping the questions they did not want to answer, and the option of withdrawing from the questionnaire without data recording and any consequences. Due to this, the number of participants that completed the SD4 was 439, while it was somewhat lower for socially desirable responding. At the end of the questionnaire, the participants were thanked for their participation, and the researcher's contact information for the participants' inquiries was indicated.

Measures

Short Dark Tetrad (Paulhus et al., 2020)

Short Dark Tetrad (Paulhus et al., 2020) measures dark traits with seven items per trait, with a response scale ranging from 1 (*not at all*) to 5 (*very much*). Two researchers independently translated the SD4 items from English to Croatian, reaching a consensus for the translations in cases of disagreement. Croatian translation of the scale is shown in Appendix A. Reliabilities in this study, measured by Cronbach's alpha were Machiavellianism = .68, narcissism = .73, psychopathy = .77, sadism = .74.

Social Desirability Scale (Parmač Kovačić et al., 2014)

Social Desirability Scale (Parmač Kovačić et al., 2014) consists of two subscales: egoistic and moralistic socially desirable responding (10 items per subscale) with a response scale ranging from 1 (*completely false*) to 7 (*completely true*). In line with the recommended scoring procedure (Paulhus 2002), only responses of 6 and 7 on positively oriented or 1 and 2 on negatively oriented items were counted as socially desirable. Reliability in this study, measured by Cronbach's alpha were .72 and .69 for egoistic and moralistic socially desirable responding, respectively.

Data analytic approach

Latent variable modeling was conducted in lavaan package, version 0.6-9 (Rosseel, 2012) in the R statistical environment (R Core Team, 2021), using robust weighted least squares estimation (weighted least square mean and variance adjusted; WLSMV estimator, where manifest variables were treated as categorical using ordered argument), and variances of all factors were fixed to 1. Model fit was assessed using the following indexes and guidelines: Tucker-Lewis index (TLI) and the comparative fit index (CLI) with values close to .95 indicating acceptable model fit, standardized root mean square residual (SRMR) with a value close to .08 and root mean square error of approximation (RMSEA) with a value close to .06 indicating acceptable model fit (Hu & Bentler, 1999), bearing in mind that these are only rough guidelines that may

be hard to reach in practice (Marsh et al., 2004). Since χ^2 difference test, like model χ^2, is sensitive to sample size, the differences in CFI < .01 and in RMSEA < .015 were treated as indicators of non-difference between the models (Chen, 2007).

The relationship between dark traits and socially desirable responding was examined at the level of bivariate relationships, but also by means of multiple regression analyses with two types of social desirability as criteria, and dark traits as predictors. Since dark traits were expected to be intercorrelated, multiple regression was conducted in order to examine their independent predictive contribution to socially desirable responding (Furnham et al., 2013). Prior to regression analyses, normal q-q plots of standardized residuals were examined and they did not indicate deviations of residual errors from normality.

Results

The structure of SD4

Means, standard deviations, and intercorrelations between SD4 items are shown in Appendix B. The bifactor model had somewhat better fit indices values (*RMSEA* = .06, *SRMR* = .07, *CFI* = .90, *TLI* = .88) than four factor model (*RMSEA* = .07, *SRMR* = .08, *CFI* = .85, *TLI* = .83). Considering acceptable model fit guidelines (Hu & Bentler, 1999), in the case of four factor model, only SRMR value indicated acceptable fit, although RMSEA was relatively close to be acceptable, while in case of bifactor model both SRMR and RMSEA values indicated acceptable model fit. However, considering the differences in *CFI* < .01 and in *RMSEA* < .015 as criteria for model comparison (Chen, 2007), it can be concluded that the four-factor model fitted the data equally well as the bifactor model. The latent correlation between the dark traits in the four-factor model was smallest in the case of narcissism and sadism, and largest in the case of psychopathy and sadism (Table 1).

Table 1

Latent Factor Intercorrelations Between the Dark Tetrad Traits in Four factor Model

Factor (trait)	1	2	3	4
1. Machiavellianism	-			
2. Narcissism	.35	-		
3. Psychopathy	.30	.46	-	
4. Sadism	.35	.25	.60	-

Note. All coefficients are significant at $p < .001$.

Loadings for each of the tested models are shown in Table 2. All factor loadings in four factor model were greater than .30, and most of them were greater than .50. In the bifactor model, factor loadings for the Machiavellianism and narcissism items were generally stronger in case of their respective factor compared to the general factor, while some psychopathy and sadism items loaded on general factor to a greater degree than on their respective factor.

Table 2

Standardized Factor Loadings for Each of the Tested Models

Short Dark Tetrad (SD4) dimensions and respective items	Four factor model	Bifactor model	
		Specific	General
Machiavellianism			
It's not wise to let people know your secrets.	**.35**	**.42**	.11
Whatever it takes, you must get the important people on your side.	**.64**	**.50**	**.35**
Avoid direct conflict with others because they may be useful in the future.	**.51**	**.60**	.15
Keep a low profile if you want to get your way.	**.61**	**.55**	**.31**
Manipulating the situation takes planning.	**.32**	**.38**	*.08*
Flattery is a good way to get people on your side.	**.52**	**.46**	.28
I love it when a tricky plan succeeds.	**.70**	**.39**	**.43**

Narcissism			
People see me as a natural leader.	**.69**	**.59**	**.37**
I have a unique talent for persuading people.	**.70**	**.50**	**.44**
Group activities tend to be dull without me.	**.61**	**.40**	**.39**
I know that I am special because people keep telling me so.	**.63**	**.49**	**.36**
I have some exceptional qualities.	**.57**	**.72**	.12
I'm likely to become a future star in some area.	**.62**	**.69**	.21
I like to show off every now and then.	**.34**	.13	.30
Psychopathy			
People often say I'm out of control.	**.71**	**.44**	**.57**
I tend to fight against authorities and their rules.	**.60**	.29	**.50**
I've been in more fights than most people of my age and gender.	**.66**	**.36**	**.54**
I tend to dive in, then ask questions later.	**.54**	**.47**	**.39**
I've been in trouble with the law.	**.66**	**.59**	**.45**
I sometimes get into dangerous situations.	**.73**	**.55**	**.54**
People who mess with me always regret it.	**.66**	*-.04*	**.71**
Sadism			
Watching a fist-fight excites me.	**.88**	**.66**	**.57**
I really enjoy violent films and video games.	**.74**	**.78**	**.35**
It's funny when idiots fall flat on their face.	**.64**	**.47**	**.44**
I enjoy watching violent sports.	**.79**	**.73**	**.42**
Some people deserve to suffer.	**.57**	.19	**.49**
Just for kicks, I've said mean things on social media.	**.43**	*.15*	**.38**
I know how to hurt someone with words alone.	**.62**	.13	**.58**

Note. Loadings greater than .30 are typed in bold, and loadings not significant at $p < .05$ are typed in italics. In the case of the bifactor model, loadings that were greater in the case of general compared to specific factors are underlined.

Although both tested models fitted the data equally well, it should be kept in mind that substantial, or non-trivial loadings of all items on general and specified specific factors would imply the validity of the bifactor model (Riese et al. 2010; Watts et al., 2019). Since that was not the case here, for further analyses we opted to include the originally proposed four correlated

factors model. An additional reason for choosing four factor model was following the recommendation that it might be best for uncovering how etiological factors are linked to the expression of each dark domain (Neumann et al., 2021) and with the intention to facilitate the interpretability and comparability of our findings with previous and future ones.

The relationship between the Dark Tetrad traits and egoistic and moralistic socially desirable responding

Egoistic socially desirable responding was strongly and positively correlated with narcissism and weakly and positively with psychopathy and sadism, while moralistic socially desirable responding was negatively correlated with all four Dark Tetrad traits (Table 3).

Table 3

Descriptive Statistics and Intercorrelations Between Study Variables

Variable	*M*	*SD*	1	2	3	4	5	6
1. Egoistic socially desirable responding	2.19	2.14	-					
2. Moralistic socially desirable responding	3.23	2.27	.10*	-				
3. Machiavellianism	3.27	0.57	.04	−.34**	-			
4. Narcissism	2.99	0.62	.48**	−.14**	.24**	-		
5. Psychopathy	2.03	0.67	.15**	−.30**	.21**	.37**	-	
6. Sadism	1.98	0.69	.14**	−.46**	.27**	.20**	.47**	-

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

Regression analyses were conducted to get a better insight into the relationship between the Dark Tetrad and socially desirable responding, when the shared variance between dark traits is taken into account (Table 4). The Dark Tetrad traits explained 25% of the variance of egoistic socially desirable responding, $R^2 = .246$, $F(4, 421) = 34.29$, $p < .001$, and 27% of the variance of moralistic socially desirable responding, $R^2 = .272$, $F(4, 419) = 39.16$, $p < .001$). Higher narcissism was significantly associated with a higher level of egoistic

socially desirable responding, while the contribution of Machiavellianism and sadism was negative and marginally significant. Both lower Machiavellianism and sadism were significantly associated with a higher level of moralistic socially desirable responding, with the contribution of psychopathy being negative and marginally significant.

Table 4

Standardized Regression Coefficients of the Dark Tetrad Traits on Each Type of Socially Desirable Responding

	Egoistic (*N* = 426)		Moralistic (*N* = 424)	
Predictor	β	*p*	β	*p*
Machiavellianism	−.09	.048	−.24	< .001
Narcissism	.51	< .001	.03	.472
Psychopathy	−.06	.243	−.10	.051
Sadism	.10	.051	−.36	< .001

Discussion

This study aimed to examine the structure of the Croatian version of SD4, as well as the relationship of the Dark Tetrad traits (measured with SD4) with two types of deliberate socially desirable responding – egoistic and moralistic. The structure of SD4, although not perfect in comparison with rough acceptable model fit guidelines (Hu & Bentler, 1999), an inspection of factor loadings and their comparison with those from other studies indicated that the Croatian version has an acceptable structure. Its relationship with the two examined types of socially desirable responding was in line with expectations based on previous research and those that were theoretically driven.

The first aim of this study was to examine the structure of SD4 in the Croatian context using latent variable modeling, testing two competing models: the four-factor model and the bifactor model. The results showed that some of the fit statistics fell somewhat below the conventionally acceptable fit thresholds (Hu & Bentler, 1999); however, as previously noted,

these thresholds may be hard to reach in practice (Marsh et al., 2004). Many widely used personality inventories that measure multiple dimensions fail to reach conventional adequate fit criteria, so results should be interpreted in the context of previous validation studies since they could be more reasonable criteria for evaluation than conventional rules of a thumb (Hopwood & Donnellan, 2010). The values of fit statistics of both models tested in this study were comparable to those from the original study testing these models (see Neumann et al., 2021), along with the pattern of factor loadings for tested models. Although the values of four correlated factors model fit statistics were lower than those obtained in the Portuguese sample (Pechorro et al., 2022), they were comparable to those obtained in the German sample (Blötner, Ziegler, et al., 2022). The finding that some psychopathy and sadism items in the bifactor model loaded on general factor to a greater degree than to their respective factors is in line with previous findings indicating that general factor mostly reflected item content of these two dark traits (Neumann et al., 2021). The plausible reason for the representation of both psychopathy and sadism in the general factor could be a high correlation between them; however, such a high correlation is in line with the results of the recent meta-analysis (Bonfá-Araujo et al., 2022). Moreover, there are findings indicating that the conceptualization of sadism in SD4 is problematic and should be revised (Blötner & Beisemann, 2022), so this question remains open for further research on the Croatian version of SD4.

Although the bifactor model had somewhat better fit statistics than the four correlated factors model, the difference between the fit statistics was not large enough to conclude that the bifactor model was superior. Moreover, the pattern of factor loadings in each model was used as an additional criterion to evaluate the models' viability. First, since in the bifactor model general and group factors are orthogonal, it can be used as guidance whether the scale is multidimensional or unidimensional, and large loadings on general and low loadings on group factors would imply unidimensionality (Riese et al., 2010), which was not the case in our model. Second, in order for the bifactor model to be a valid representation of data, it should have non-

trivial loadings on all factors (specific and general; Riese et al. 2010; Watts et al., 2019), which was also not the case in our data. Therefore, the four correlated factors model was chosen as a better representation of SD4 and used for further analyses. The additional reason was to facilitate the interpretability and comparability of our findings with previous and future ones, following the recommendations and findings that the choice of the SD4 analytic model may not matter substantially in examining the external correlates of dark traits (see Neumann et al., 2021).

In the four correlated factors model all loadings were satisfactory in size, with most of them greater than .50. Reliabilities of SD4 scales measured by Cronbach's alpha were adequate (i.e., around .70 or above), and although they were somewhat lower compared to the original validation samples (Paulhus et al., 2021) or Portuguese version (Pechorro et al., 2022), they were comparable to German version (Blötner, Ziegler, et al., 2022). If we interpret the sizes of Pearson correlations of .10, .20. .30, and .40 as small, medium, large, and very large effect sizes, respectively (Funder & Ozer, 2019; Gignac & Szodorai, 2016), intercorrelations between SD4 subscales in our research ranged from medium to very large. Obtained effect sizes are in the range of those from previous studies (Blötner, Ziegler, et al., 2022; Paulhus et al., 2021; Pechorro et al., 2022). The patterns of correlations most closely followed those obtained in the German sample (Blötner, Ziegler, et al., 2022), with the highest correlation between psychopathy and sadism, followed by the correlation between psychopathy and narcissism. Furthermore, the narcissism item regarding showing off every now and then had the lowest loadings in four factor model both in Croatian, German and Portuguese contexts (see Blötner, Ziegler, et al., 2022; Pechorro et al., 2022). Taken together, our findings indicate that the Croatian version of SD4 is valid and comparable to the original version.

The second aim of this study was to investigate the relationship between the Dark Tetrad traits and egoistic and moralistic socially desirable responding. In line with expectations, narcissism was positively related to egoistic socially desirable responding, while psychopathy, Machiavellianism, and sadism were negatively related to moralistic socially desirable

responding. The positive relationship between egoistic socially desirable responding and primarily narcissism was expected, given that egoistic socially desirable responding reflects a tendency of exaggerating one's competencies, which resembles grandiosity, a key feature of narcissism (Paulhus, 2014), and is in line with previous findings that narcissism, compared to other dark traits is related to exhibiting self-enhancement the most (Blötner, Ziegler, et al., 2022; Paulhus & Williams, 2002). Furthermore, the negative relationship between moralistic socially desirable responding and the other three Dark Tetrad traits reflects both theoretical assumptions and empirical findings (see Parmač Kovačić et al., 2014) indicating that moralistic socially desirable responding correlates higher with communal traits (i.e., agreeableness and conscientiousness). Namely, among the Dark Triad traits, narcissism has the lowest correlations with agreeableness and is not related to conscientiousness, while the correlations of psychopathy and Machiavellianism with these traits are similar (see Muris et al., 2017). Finally, the relations of the Dark Tetrad traits and egoistic and moralistic socially desirable responding point to the distinctiveness of narcissism from the rest of the Dark Tetrad, contributing to findings that narcissism is different from the other Dark Tetrad traits (e.g., Book et al., 2016; Dinić et al., 2020).

There are several limitations imposed on the generalizability of the findings due to restriction to self-report questionnaires, cross-sectional study design, and sample characteristics – a convenience sample of mainly younger females. Primarily, as our sample consisted of more than 80% females, invariance across gender could not be tested. To overcome this shortcoming, we encourage future research to investigate the psychometric properties of the SD4 scales in more gender-balanced samples. Another proof of the scale's construct validity would be to examine its nomological network and compare obtained relationships to those from other studies (see Blötner, Ziegler, et al, 2022; Paulhus et al, 2021; Pechorro et al., 2022), as well as its longitudinal stability or invariance. To corroborate the associations established by the present research between SD4 and socially desirable responding it would be useful to explore the relationships among these constructs using behavioral and observational paradigms (e.g., *fake-good* or *fake-bad*), as the recent

meta-analysis showed that the Dark Triad traits are fakable, with similar effect sizes to previous meta-analyses of Big Five measures (Walker et al., 2022). Therefore, in future research emphasis should be given to the exploration of the relationship between SD4 and socially desirable responding in different social situations and in longitudinal designs to better understand cross-situational stability and the temporal order of these relationships.

In conclusion, the results of our study indicate that the Croatian version of SD4 is valid and can be used to assess the Dark Tetrad traits. Concerning the two examined models, our results did not indicate the superiority of the bifactor model compared to the four-factor model, and the general factor did not reflect all, nor most of the SD4 items. Thus, we did not find support for the use of the general "dark" factor. The relations of the Dark Tetrad traits and egoistic and moralistic socially desirable responding corroborated the distinctiveness of narcissism from the rest of the Dark Tetrad since narcissism was positively related to egoistic socially desirable responding, while psychopathy, Machiavellianism, and sadism were negatively related to moralistic socially desirable responding. Moreover, since these relationships were in line with theoretically and empirically based expectations, they can be viewed as an indicator of the usefulness of the Croatian version of SD4. Although our study design did not allow us a more comprehensive insight into its validity (e.g., measurement invariance, nomological network), we hope that this study will provide the first step toward more comprehensive future tests of its validity.

Acknowledgement

We thank Maja Đuras for help during the data collection process.

Conflict of interest

We have no conflicts of interest to disclose.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Blötner, C., & Beisemann, M. (2022). The Dark Triad is dead, long live the Dark Triad: An item-response theoretical examination of the Short Dark Tetrad. *Personality and Individual Differences*, *199*, 111858. https://doi.org/10.1016/j.paid.2022.111858

Blötner, C., Webster, G. D., & Wongsomboon, V. (2022). Measurement invariance of the Short Dark Tetrad across cultures and genders. *European Journal of Psychological Assessment*. Advance online publication. https://doi.org/10.1027/1015-5759/a000715

Blötner, C., Ziegler, M., Wehner, C., Back, M. D., & Grosz, M. P. (2022). The nomological network of the Short Dark Tetrad scale (SD4). *European Journal of Psychological Assessment, 38*(3), 187–197. https://doi.org/10.1027/1015-5759/a000655

Bonfá-Araujo, B., Lima-Costa, A. R., Hauck-Filho, N., & Jonason, P. K. (2022). Considering sadism in the shadow of the Dark Triad traits: A meta-analytic review of the Dark Tetrad. *Personality and Individual Differences*, *197*, 111767. https://doi.org/10.1016/j.paid.2022.111767

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chiorri, C., Garofalo, C., & Velotti, P. (2019). Does the Dark Triad manifest similarly in men and women? Measurement invariance of the Dirty Dozen across sex. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 38*(3), 659–675. https://doi.org/10.1007/s12144-017-9641-5

Dinić, B. M., Bulut Allred, T., Petrović, B., & Wertag, A. (2020). A test of three sadism measures. *Journal of Individual Differences*, *41*(4), 219–227. https://doi.org/10.1027/1614-0001/a000319

Dinić, B. M., Petrović, B., & Jonason, P. K. (2018). Serbian adaptations of the Dark Triad Dirty Dozen (DTDD) and Short Dark Triad (SD3). *Personality and Individual Differences, 134*, 321–328. https://doi.org/10.1016/j.paid.2018.06.018

Dinić, B. M., Wertag, A., Tomašević, A., & Sokolovska, V. (2020). Centrality and redundancy of the Dark Tetrad traits. *Personality and Individual Differences*, *155*, 109621. https://doi.org/10.1016/j.paid.2019.109621

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10 year review. *Social and Personality Psychology Compass, 7*(3), 199–216. https://doi.org/10.1111/spc3.12018

Gamache, D., Savard, C., & Maheux-Caron, V. (2018). French adaptation of the Short Dark Triad: Psychometric properties and a head-to-head comparison with the Dirty Dozen. *Personality and Individual Differences, 122*, 164–170. https://doi.org/10.1016/j.paid.2017.10.027

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Holden, R. R., & Passey, J. (2009). Social desirability. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 441–454). The Guilford Press.

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*(3), 332–346. https://doi.org/10.1177/1088868310361240

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment, 22*(2), 420–432. https://doi.org/10.1037/a0019265

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3): A brief measure of dark personality traits. *Assessment, 21*(1), 28–41. https://doi.org/10.1177/1073191113514105

Kowalski, C. M., Rogoza, R., Vernon, P. A., & Schermer, J. A. (2018). The Dark Triad and the self-presentation variables of socially desirable responding and self-monitoring. *Personality and Individual Differences, 120*, 234–237. https://doi.org/10.1016/j.paid.2017.09.007

Kowalski, C. M., Vernon, P. A., & Schermer, J. A. (2016). The general factor of personality: The relationship between the big one and the dark triad. *Personality and Individual Differences, 88*, 256–260. https://doi.org/10.1016/j.paid.2015.09.028

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Međedović, J., & Petrović, B. (2015). The Dark Tetrad: Structural properties and location in the personality space. *Journal of Individual Differences, 36*(4), 228–236. https://doi.org/10.1027/1614-0001/a000179

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science*, *12*(2), 183–204. https://doi.org/10.1177/1745691616666070

Neumann, C. S., Jones, D. N., & Paulhus, D. L. (2021). Examining the Short Dark Tetrad (SD4) across models, correlates, and gender. *Assessment*, 1073191120986624. https://doi.org/10.1177/1073191120986624

Parmač Kovačić, M., Galić, Z., & Jerneić, Ž. (2014). Social desirability scales as indicators of self-enhancement and impression management. *Journal of Personality Assessment, 96*(5), 532–543. https://doi.org/10.1080/00223891.2014.916714

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598–609. https://doi.org/10.1037/0022-3514.46.3.598

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Lawrence Erlbaum Associates Publishers.

Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science*, *23*(6), 421–426. https://doi.org/10.1177/0963721414547737

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic bias in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality, 66*, 1025–1060. https://doi.org/10.1111/1467-6494.00041

Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*(6), 556–563. https://doi.org/10.1016/S0092-6566(02)00505-6

Paulhus, D. L., Buckels, E. E., Trapnell, P. D., & Jones, D. N. (2021). Screening for dark personalities. *European Journal of Psychological Assessment, 7*(3), 208–222. https://doi.org/10.1027/1015-5759/a000602

Pechorro, P., Karandikar, S., Carvalho, B., DeLisi, M., & Jones, D. N. (2022). Screening for dark personalities in Portugal: Intra-and interpersonal correlates, reliability and invariance of the Short Dark Tetrad Portuguese version. *Deviant Behavior*, 1–16. https://doi.org/10.1080/01639625.2022.2071655

Persson, B. N., Kajonius, P. J., & Garcia, D. (2019). Revisiting the structure of the Short Dark Triad. *Assessment, 26*(1), 3–16. https://doi.org/10.1177/1073191117701192

Pineda, D., Sandín, B., & Muris, P. (2020). Psychometrics properties of the Spanish version of two Dark Triad Scales: The Dirty Dozen and the Short Dark Triad. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 39*(5), 1873–1881. https://doi.org/10.1007/s12144-018-9888-5

R Core Team. (2021). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*(6), 544–559. http://dx.doi.org/10.1080/00223891.2010.496477

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Savard, C., Simard, C., & Jonason, P. K. (2017). Psychometric properties of the French-Canadian version of the Dark Triad Dirty Dozen. *Personality and Individual Differences, 119*, 122–128. https://doi.org/10.1016/j.paid.2017.06.044

Walker, S. A., Double, K. S., Birney, D. P., & MacCann, C. (2022). How much can people fake on the dark triad? A meta-analysis and systematic review of instructed faking. *Personality and Individual Differences, 193*, 111622. https://doi.org/10.1016/j.paid.2022.111622

Watts, A. L., Poore, H. E., & Waldman, I. D. (2019). Riskier tests of the validity of the bifactor model of psychopathology. *Clinical Psychological Science, 7*(6), 1285–1303. https://doi.org/10.1177/2167702619855035

Womick, J., Foltz, R. M., & King, L. A. (2019). "Releasing the beast within"? Authenticity, well-being, and the Dark Tetrad. *Personality and Individual Differences, 137*, 115–125. https://doi.org/10.1016/j.paid.2018.08.022

Appendix A

Croatian Translation of SD4 Items

Item label	Croatian item content
Mach_1	Nije pametno odavati svoje tajne.
Mach_2	Bez obzira na cijenu, neophodno je pridobiti važne ljude na svoju stranu.
Mach_3	Dobro je izbjegavati izravne sukobe s drugima jer mogu biti od koristi u budućnosti.
Mach_4	Potrebno se pritajiti da bi dobio ono što želiš.
Mach_5	Za kontrolu nad situacijom potrebno je planiranje.
Mach_6	Laskanje je efikasno u pridobivanju ljudi na svoju stranu.
Mach_7	Volim kad domišljat plan uspije.
Narc_1	Ljudi me vide kao rođenog vođu.
Narc_2	Talentiran/a sam za uvjeravanje drugih.
Narc_3	Grupne aktivnosti znaju biti dosadne bez mene.
Narc_4	Znam da sam poseban/na jer mi drugi to stalno govore.
Narc_5	Imam puno izvrsnih kvaliteta.
Narc_6	Vjerojatno ću postati iznimno uspješan/na u nečem.
Narc_7	Povremeno se volim hvaliti.
Psy_1	Često za mene kažu da sam van kontrole.
Psy_2	Sklon/a sam suprotstavljati se autoritetima i njihovim pravilima.
Psy_3	Češće sam se tukao/la od većine ljudi moje dobi i spola.
Psy_4	Prvo uletim u nešto, a tek onda postavljam pitanja.
Psy_5	Imao/la sam problema sa zakonom.
Psy_6	Ponekad upadnem u opasne situacije.
Psy_7	Ljudi koji mi se zamjere uvijek požale.
Sad_1	Gledanje tučnjave me uzbuđuje.
Sad_2	Uživam u nasilnim filmovima i video-igrama.
Sad_3	Zabavno je kad neka budala padne i razbije se.
Sad_4	Volim gledati nasilne sportove.
Sad_5	Neki ljudi zaslužuju patnju.
Sad_6	Radi zabave, pisao/la sam zlobne stvari po društvenim medijima.
Sad_7	Znam kako povrijediti nekoga samo riječima.

Note. Mach = Machiavellianism; Narc = Narcissism; Psy = Psychopathy; Sad = Sadism.

Appendix B

Means, Standard Deviations, and Intercorrelations Between SD4 Items

Item label	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1. Mach_1	3.70	0.88	-																											
2. Mach_2	2.44	1.03	**.20**	-																										
3. Mach_3	3.13	1.09	**.17**	**.36**	-																									
4. Mach_4	2.47	1.00	**.31**	**.35**	**.32**	-																								
5. Mach_5	3.67	0.91	**.15**	**.18**	**.13**	**.19**	-																							
6. Mach_6	3.32	1.07	**.17**	**.22**	**.31**	**.27**	**.19**	-																						
7. Mach_7	4.18	0.78	**.17**	**.24**	**.25**	**.24**	**.26**	**.24**	-																					
8. Narc_1	2.72	1.09	.10	.11	−.05	.05	.09	.01	**.23**	-																				
9. Narc_2	3.15	1.04	.03	**.20**	**.13**	**.12**	.07	.10	**.32**	**.53**	-																			
10. Narc_3	2.46	1.02	.11	**.20**	.10	.09	.07	.10	**.17**	**.33**	**.31**	-																		
11. Narc_4	2.54	1.08	−.02	**.14**	.11	.09	.12	.08	**.17**	**.34**	**.31**	**.45**	-																	
12. Narc_5	3.68	0.77	−.02	.07	.08	.00	−.01	−.03	**.16**	**.33**	**.28**	**.25**	**.35**	-																
13. Narc_6	3.44	0.84	.00	.11	.04	.07	.02	−.06	**.26**	**.40**	**.33**	**.23**	**.31**	**.53**	-															
14. Narc_7	2.96	1.13	−.03	**.21**	.09	.09	−.02	**.15**	.09	**.12**	.10	.11	.11	**.14**	**.19**	-														
15. Psy_1	2.03	1.03	.05	**.18**	.04	**.21**	.04	**.17**	.08	**.15**	**.18**	**.28**	**.24**	.05	.07	**.19**	-													
16. Psy_2	2.83	1.19	.04	.08	−.06	**.15**	.00	.07	.09	**.29**	**.24**	**.18**	**.18**	**.15**	**.19**	.11	**.44**	-												
17. Psy_3	1.57	0.98	.05	.08	.07	**.17**	.00	**.15**	.09	**.11**	**.15**	.09	.05	.00	−.03	.09	**.38**	**.32**	-											
18. Psy_4	2.40	1.12	−.01	**.13**	.00	**.16**	−.08	.12	.08	.08	.08	**.22**	**.17**	.07	.08	**.12**	**.48**	**.30**	**.26**	-										
19. Psy_5	1.37	0.78	.03	.09	−.03	.12	−.04	**.14**	.09	.10	.11	**.14**	.05	.04	.05	.08	**.27**	**.23**	**.35**	**.22**	-									
20. Psy_6	1.97	1.10	0.00	.03	.02	0.08	−.04	**.16**	.10	.12	**.22**	**.15**	**.19**	.11	.11	**.16**	**.41**	**.31**	**.33**	**.39**	**.49**	-								
21. Psy_7	2.05	1.01	**.20**	**.31**	.07	**.17**	.05	.12	**.20**	**.28**	**.24**	**.27**	**.26**	**.14**	**.24**	**.16**	**.32**	**.31**	**.29**	**.20**	**.22**	**.28**	-							
22. Sad_1	1.50	0.92	.01	**.16**	.08	.12	−.05	.09	**.18**	.09	**.15**	.12	.07	.03	**.12**	.06	**.30**	**.21**	**.32**	**.20**	**.20**	**.33**	**.27**	-						
23. Sad_2	1.73	1.09	−.03	.10	.10	.04	−.04	.03	**.20**	.09	**.15**	.01	.02	.03	.04	.10	.11	**.14**	**.25**	.09	.08	**.18**	**.16**	**.57**	-					
24. Sad_3	2.01	1.20	.05	**.20**	.07	**.14**	.09	**.16**	**.20**	.04	.06	.10	.10	.01	.02	.12	**.23**	**.14**	**.21**	**.18**	.11	**.16**	**.22**	**.41**	**.43**	-				
25. Sad_4	1.76	1.10	−.06	.10	.01	.06	−.02	.03	**.19**	.09	.11	.09	.06	.06	**.14**	.04	**.20**	**.19**	**.23**	**.14**	.10	**.31**	**.19**	**.62**	**.61**	**.35**	-			
26. Sad_5	2.26	1.33	**.13**	**.21**	**.20**	**.23**	.07	**.16**	**.24**	.06	.11	.11	.02	−.03	.01	.05	**.18**	**.17**	**.18**	**.13**	**.21**	**.24**	**.31**	**.24**	**.23**	**.35**	**.17**	-		
27. Sad_6	1.29	0.69	−.01	.10	.10	**.12**	.07	.12	.00	−.01	−.02	.03	.09	−.03	−.03	**.14**	**.16**	**.12**	.10	.07	**.15**	**.13**	**.17**	.12	.11	**.17**	.08	**.16**	-	
28. Sad_7	3.32	1.20	.08	.05	.07	**.16**	.07	**.14**	**.22**	**.28**	**.28**	.07	**.13**	.07	.09	.11*	**.21**	**.28**	**.19**	**.18**	**.20**	**.31**	**.27**	**.26**	**.23**	**.26**	**.24**	**.35**	**.23**	-

Notes. Mach = Machiavellianism; Narc = Narcissism; Psy = Psychopathy; Sad = Sadism. Correlation coefficients significant at $p < .01$ are typed bold and underlined, and correlation coefficients significant at $p < .05$ are underlined.

Primenjena psihologija
Vol. 16, No. 1, pp. 107-147, 2023

Research Article

Linking Metacognition, Workplace Cognitive Competencies and Performance: An Integrative Review-Based Conceptual Framework

Shilpa Jain [1], Bhavna Bajaj [2] and Aarushi Singh [1]✉

[1] *University School of Management Studies, Guru Gobind Singh Indraprastha University, New Delhi, India*

[2] *School of Business and Management, Christ University (Deemed to be University), Bangalore, India*

ABSTRACT

A key driver of workplace cognitive competencies is metacognition which has been shown to impact performance among nurses, teachers and firefighters, however, it is scarcely studied among managerial employees. The research investigating this relationship is also scattered across multiple domains limiting its' utility for researchers and practitioners. This paper, therefore, presents an integrative review of the existing empirical literature from the Web of Science and Scopus database to trace the linkages of metacognition, workplace cognitive competencies and performance at work. The identified linkages are then formulated into a conceptual framework clarifying how various workplace cognitive competencies and performance may be linked to metacognition. The findings indicate linkages between metacognition and various workplace cognitive competencies such as problem-solving, decision-making, innovation, creativity and knowledge acquisition. The present research also establishes the link of metacognition and cognitive competencies with learning, individual and firm performance. The review paves way for metacognition to be considered as a distinct construct in the workplace, identifies gaps and provides direction for future research.

Keywords: Metacognition, Metacognitive Ability, Conceptual framework, Integrative Review, Employee Performance

UDK: 331.105.2-057.16:159.95
DOI: 10.19090/pp.v16i1.2427
Received: 07.09.2022.
Revised: 10.02.2023.
Accepted: 20.02.2023.

✉ Corresponding author email: aarushis1996@gmail.com

Introduction

Modern jobs with high usage of technology place higher demand for cognitive competencies on workers (Hardy et al., 2019; Torraco, 2002). Research on the workplace cognitive competencies (such as decision-making, problem-solving, innovation or creativity, knowledge acquisition and learning tasks) over the last four decades has shown the validity and utility of them in predicting workplace performance in different settings (Boyatzis, 1991; Boyatzis et al., 2017; Rozhkov et al., 2017; Ryan et al., 2009). Hacker et al (2009) have mentioned that metacognition is the key driver of competencies like learning, decision-making, problem-solving and collaboration impacting performance at work and this has been scarcely studied by organisational psychologists.

A few studies relating metacognition to cognitive competencies are available in niche areas such as nursing (Oh, 2016), teaching (Duman, 2018) and fire-fighting (Frye & Wearing, 2016), but studies involving practising managers are virtually limited. The studies in these areas have examined the linkage of metacognition with cognitive competencies such as decision-making (Mattingly et al., 2016; Shepherd et al., 2010), problem-solving (Liu & Liu, 2020a; Urban & Wood, 2017), innovation or creativity (Berraies, 2020) and knowledge acquisition (Zumbach et al., 2020). Recently, researchers (Cho & Linderman, 2019; Lyons & Bandura, 2019) in the management field have demonstrated that metacognition is linked to managerial performance, thereby suggesting that the metacognition of an individual is a likely determinant of his/her performance in the workplace. The logical linkage of metacognition to cognitive competencies and performance prompted us to identify the linkages studied in the literature. We focused only on empirical work which provides evidence of relationships among the variables studied.

The current work is an attempt to provide such a conceptual framework relating metacognition to various cognitive competencies impacting performance in the workplace. We arranged and integrated the diverse body of empirical literature relating to metacognition among

professionals to trace these linkages. Furthermore, we identify themes based on various cognitive competencies and then formulate and present an integrative conceptual framework linking metacognition to the cognitive competencies impacting employee performance. Finally, the available insights from the existing body of knowledge help to identify research gaps and recommend directions for future research in the field of organisational psychology and management.

This study attempts to reflect upon the following research objectives;
RO1: To identify and integrate empirical literature linking metacognition and cognitive competencies impacting performance in the workplace.
RO2: To formulate an integrative conceptual framework based on the identified empirical linkages using relational and thematic analyses.
RO3: To discuss the consolidated insights from the metacognition literature and make propositions regarding the linkages between components and sub-components of metacognition with the cognitive competencies at work that facilitate individual performance.
RO4: To identify research gaps in the literature for providing directions for future research and implications for researchers and practitioners.

Theoretical Background

Metacognition

The term metacognition was first coined by John Flavell in 1976 in his seminal work "Metacognitive Aspects of Problem Solving" (Flavell, 1976). Metacognition simply means 'cognition about cognition, 'knowing about knowing' and 'thinking about thinking' (Alter & Oppenheimer, 2009). The research on metacognition has been drawn over the years from three distinct theoretical paradigms of cognitive development psychology (Piaget, 1950), cognitive psychology (Hart, 1965) and social development psychology (Vygotsky, 1962). The definition of metacognition evolved with Gavelek and Raphael (1985, pp.22-23) defining it as "the abilities of individuals to adjust their cognitive activity to promote more effective comprehension " thereby suggesting that control or adjustment of cognitive activity was an inherent

part of metacognition. Eventually, it was defined as the information individuals possess while completing a task including a deliberate organization in cognitive processes (Brown et al., 1983). Another definition by Paris & Winograd (1990) included two essential features of metacognition: "self-appraisal and self-management of cognition". Matsaggouras (1994) further specified that metacognition includes awareness and control of emotions that accompany the cognitive processes and the person's ability to monitor them. Livingstone (1997) suggested that whereas cognition is the set of all mental abilities and processes related to knowledge, metacognition involves overseeing cognitive goals to ensure they have been met. The commonality in the evolving definitions of metacognition was that it included monitoring strategies for the learning process (Bonner, 1998). Eventually, researchers began to see metacognition as a master that coordinates the smooth operation of all other cognitive processes (Hacker et al., 1998).

In essence, metacognition represents the control that the individual has over their cognition as a function of an ability that differs within individuals, to consider alternative cognitive strategies to cope with a changing environment (Haynie et al., 2010). The term 'Metacognition' has been used interchangeably in literature with Metacognitive Ability by various researchers (Jia et al., 2019). Over the years, there has been the development of various theoretical models of metacognition discussing its' various components and sub-components. The sequential examination of these theoretical models suggests that the conceptualisation of metacognition started with the model proposed by Flavell (1981). He described metacognition as consisting of four subcomponents - knowledge, experiences, goals-tasks, and strategies. Further, Brown (1987) clarified the two-component model consisting of two main aspects of metacognition i.e. metacognitive knowledge (also referred to as metacognitive awareness) and metacognitive regulation (or metacognitive skill). Although, several authors have put forth various subcomponents of Metacognitive Ability, nearly all of the relevant research in the last twenty years on metacognition has

confirmed Brown's two-component model of metacognition (Harrison & Vallin, 2018; Kuhn & Dean, 2004; Schraw et al., 2006).

Metacognitive knowledge refers to an individual's knowledge of his/ her cognitive structure and process (Flavell, 1979). It is a part of the individual's evaluation of their belief system but may be inaccurate such that individuals may overestimate or underestimate their competencies (called metacognitive accuracy) (Veenman et al., 2006). This ability to be metacognitively accurate can be quantified and increased with training and improvement techniques (Knox et al., 2017). *Metacognitive regulation* on the other hand refers to an individual's ability to monitor and evaluate his/ her cognitive activity. It is concerned with planning, critical evaluation and conscious execution of appropriate actions to achieve a particular goal (Martinez, 2006; Schraw et al., 2006).

Metacognitive Ability thus enables individuals to engage in self-awareness (i.e. knowledge) and regulation of cognitive processes. These two components of metacognition have been further subdivided into sub-components by various authors (Schraw et al., 2006) as briefed in Appendix A. These components and sub-components of metacognition have been used interchangeably in the literature in various professional domains.

Cognitive Competencies

Competency is defined as his/her "capability or ability" associated with a motive that impacts the achievement of goals and objectives. These competencies are something that an employee "must know" and "able to do" to perform a task effectively. Traditionally, Spencer and Spencer (1993) defined competencies as "the underlying characteristics of the person that lead to or cause effective or superior performance". The increment in the dissatisfaction associated with the traditional measures of cognitive intelligence has led to the emergence of scholarly interest in the concept of competencies (Boyatzis et al., 2017). These competencies account for a large amount of variance in performance at work, especially among studies examining the performance of professionals (Ryan et al., 2009). Spencer and Spencer (1993)integrate a parsimonious framework of competencies into

emotional, social, and cognitive. Recently, there has been a growing interest in cognitive competencies in the workplace with an increment in the requirement of analytical and conceptual thinking while performing a task at work. Cognitive competencies are said to help individuals in analysing information and situation at work. They help individuals to approach tasks by looking at them as an element of a larger system instead of viewing them as small components (Ackoff & Addison, 2010). These cognitive competencies are not restricted to personal traits of individuals, rather they can be developed through individual interventions such as learning, monitoring and regulation (Bonesso et al., 2018). Given the role of metacognition in cognitive monitoring, cognitive regulation, success on learning tasks, and the high cognitive demand of present-day jobs prompted us to further explore the literature to find linkages between metacognition, cognitive competencies, and performance at work.

A quick review of the literature revealed the presence of various studies in other professional domains except for the managerial workspace. This called for an integrative literature review.

Method

An integrative review is a form of literature study that reviews, synthesizes and critiques literature related to a subject matter comprehensively to formulate a new framework or viewpoint on the area of review (Torraco, 2005). This form of review is used to address both a mature topic or a new or emerging topic such as the one in this work (Snyder, 2019). According to Shahbaz and Parker (2021), a replicable integrative review must begin with defining the research objectives and proceed to formulate a conceptual framework through discrete steps (Fornes et al., 2008; Wollard & Shuck, 2011). In line with their recommendations, the current work flows through five steps (1) Defining the research objectives (mentioned in the introduction section) (2) Identifying sources, database and keywords (3) Selecting and evaluating the studies (4) Formulation of a conceptual framework (5) Reporting of discussions.

Identifying sources, databases and keywords

The multidisciplinary databases Web of Science (WoS) and Scopus were accessed to maintain the quality and consistency of the articles reviewed for the literature (Aghaei Chadegani et al., 2013). A systematic search was conducted (on 25th February 2022) on these databases to identify all peer-review papers published on the topic in the last 31 years, i.e., 1990-2022. Since the first objective of the study was to identify the linkage of metacognition with various cognitive competencies impacting performance in the workplace. A preliminary search of the databases clarified that very few relevant results could be elicited in the context of the workplace by using a combination of the following broad keywords "Metacognition", "performance", "work performance" and "employee performance" in the title, abstract and keyword. The search query was thus expanded to identify the linkage of metacognition with various workplace cognitive competencies such as problem-solving, decision-making, innovation and creativity along with performance (Rozhkov et al., 2017; Ryan et al., 2009) among professionals from diverse fields. Since these competencies impact individual and overall firm performance (Cho & Linderman, 2019; Lyons & Bandura, 2019) which has been recently explored with metacognition in a few studies in the field of management. These key terms recognised for metacognition and workplace cognitive competencies took account of the diverse terminology for metacognition, performance and dimensions of cognitive competencies. This ensured the complete coverage of the broad and scattered literature on metacognition relevant to the field of management.

A BOOLEAN search criterion (title, abstract and keyword) was performed in this review as used across disciplines to structure query (Pohl et al., 2010) in the databases as follows:

TS (Metacognition OR Metacognitive Skills OR Metacognitive Ability) AND TS (Performance OR Decision making OR Problem Solving OR Innovation OR Creativity)

The search query resulted in 7,127 results (included only articles in; English language, published post 1990 and journal & review articles). We

removed the articles (n=6,762) relating to the domain of neurology, animal metacognition, child development and clinical metacognition. Post-data extraction we moved on to the screening stage of the integrative review process with 365 studies.

Selecting and evaluating the studies

The articles obtained were then examined to qualify for the study's inclusion and exclusion criteria. As recommended by Torraco (2005) initial primary screening was performed by reviewing the title, abstract and keywords of the articles. These studies were either included in OR excluded for further full-text analysis.

We included studies that -

1. Explicitly used the components and sub-components of metacognition; metacognitive knowledge and metacognitive regulation (refer to Appendix A) in their title, abstract and keywords. Studies that didn't directly investigate metacognition and its' sub-components were thus excluded.
2. Explored metacognition in the context of cognitive tasks/managerial work/cognitive competencies such as *decision-making, problem-solving, innovation or creativity, knowledge acquisition, learning tasks and individual, team & firm performance* in the context of various professional domains relevant for managerial work.
3. Were published in journals in the domain relevant for the field of management such as *Business; Business Finance; Psychology Applied; Industrial Relations; Labor Psychology; Economics; Management, Social Sciences and Interdisciplinary categories.*
4. Belonged to peer-reviewed journal papers so that they present scientifically validated knowledge.

All duplicate articles (n=35) were eliminated. A total of 253 articles were then excluded for not meeting the inclusion criteria. The remaining 77 articles were then reviewed through full-text examination (Refer: Figure 1).

Formulation of a Conceptual Framework

After the full-text examination of the 77 articles, 39 were found to be relevant to the objective of the study (Refer: Figure 1). In the full-text screening of the articles the authors examined the metacognition construct investigated and measured in the study. The studies which didn't specifically investigate the knowledge and regulation components and sub-components were excluded from the review. Also, the studies that investigated metacognition specifically with cognitive competencies/tasks/work relevant for workplace were included in the review.

The selected studies were then examined through the technique of relational and thematic analysis (Wollard & Shuck, 2011) to detect the relationship of components and sub-components of metacognition with various workplace cognitive competencies. The following themes were identified from the included articles: *Performance on academic and learning tasks; Decision Making; Problem Solving; Innovation Performance; Individual Performance & Firm Performance*. These categories were then processed by developing a conceptual framework, presenting the visual link between concepts and constructs (Wollard & Shuck, 2011).

The following diagram illustrates the search and selection process of the review (Figure 1).

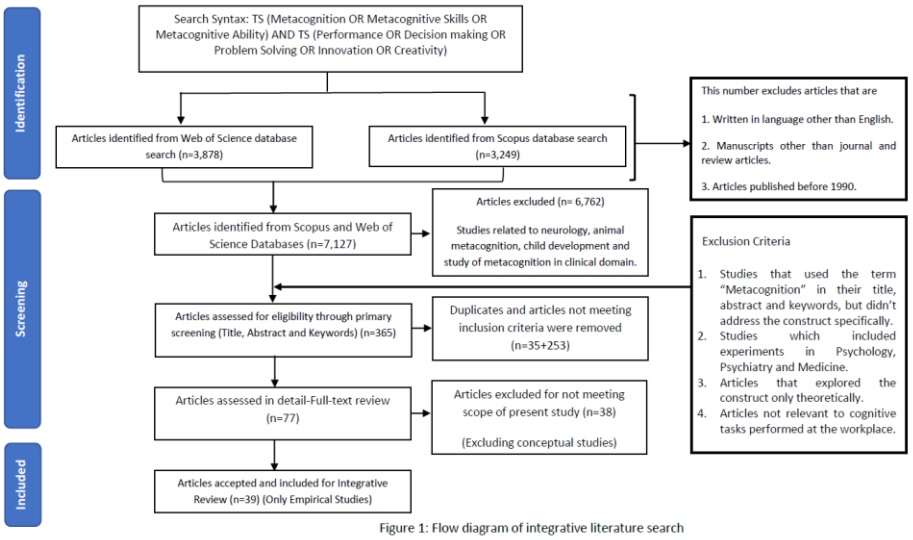

Figure 1. Flow diagram of integrative literature search

General Discussion

Researchers' interest in exploring linkages between metacognition and various cognitive competencies to improve performance at work has been steadily increasing since the 2000s. Predictably, most of the research remains confined to journals of Psychology/ Applied Psychology which carry research that is relevant to the domain of Organisational Psychology & Behaviour, Human Resource Management and Performance Management. Out of 77 relevant articles, 39 articles had direct implications for the field of management. Further, it became evident from the review that the researchers in the field of Computer-Human Interaction are actively investigating metacognition to understand performance in learning environments as well as work contexts. The analysis also depicts that research in the context of sports, entrepreneurship, security forces, fire workers and managerial professionals have pointers on linkages between metacognition and cognitive competencies at the workplace.

A large proportion of the 39 relevant studies have their origins in the USA and Australia however, researchers from 15 other countries (Austria,

France, Ireland, Germany, Japan, Macau, and others) too have begun work on the subject. The review highlighted the scarcity of literature on management professionals in the field of metacognition as the majority of the studies are restricted to professional domains other than management. The findings indicate the presence of three studies investigating the positive impact of metacognition on employee performance among management professionals and the glaring need to explore it as an essential construct at work.

The identified key features of the included studies (latest 6 years) are tabulated in Table 1 (Refer to Appendix A for the features of 39 studies). The majority of the included studies have adopted an experimental research design (23) for investigating the relationships among the constructs. Eight studies have used the survey method, seven studies have utilised questionnaires for data collection and one has used the interview method. The components of metacognition have also been studied as intervening variables with academic and managerial cognitive competencies. Four studies have used metacognitive intervention as a mediating variable and two studies have considered the moderating role of metacognition with workplace competencies.

The relational analysis of the 39 studies included in the review (tabulated in Table 1 and Appendix B) clarified the various linkages of metacognition (its' components and sub-components) with various cognitive competencies in the workplace. Further, there was a need to identify the major themes with which the various components of metacognition are investigated in the literature.

Table 1

Key features of studies selected for integrative review (Latest 6 years)

No.	Author (Year)	Subject	Field/Industry	Sample Size	Research Design	Study Design			
						Independent Variable	Dependent Variable	Moderating Variable	Mediating Variable
1	Zumbach et al. (2020)*	Students	Gaming and Technology	131	Experimental Study	Metacognitive Strategies (Training and Prompting)	Knowledge Acquisition#** Cognitive Load***		
2	Liu and Liu (2020b)*	Students	Gaming and Technology	159	Experimental Study	Metacognition Goal-orientation	Problem-Solving#** Learning Performance#**		
3	Zhao and Ye (2020)*	Undergraduate Students	Students	230	Experimental Study	Metacognitive Calibration	Performance on learning task#*	Time Spent on task	
4	Cho and Linderman (2019)	Managers	Operational Management	235	Survey Study	Managerial Metacognition	Firm Performance#** Performance#*		Social Process Improvement Practices*** Technical Process Improvement Practices***
5	Najmaei and Sadeghinejad (2019)	CEO and Top-management team members	Entrepreneurs (SME)	105	Questionnaire Study	Teams' metacognitive ability diversity	Firm Performance#		Innovativeness* Pro-activeness Risk-Taking
6	Berraies (2020)	Manager	Diverse fields	186	Survey Study	Metacognitive CQ* Cognitive CQ Behavioural CQ Motivational CQ	Innovation Performance#*	Collaborative Climate	Knowledge Sharing*
7	Yoo et al. (2018)	Staff, Managers, Executives	Electronics, Automobile, Construction	415	Survey Study	Technological Innovation	Business Performance#***		Metacognition at

			, Textile, Food & Beverages				Organisation Effectiveness***		Organisational Level
8	Rhodes et al. (2018)	CEO and Senior executives	Entrepreneurs	140	Quantitative Study	Teams' metacognitive knowledge diversity Teams' metacognitive experience diversity	Firm Performance#**		Top Management Teams Entrepreneurial Orientation
9	Pan and Sun (2018)	Professionals	Marketing, Service, R&D, and Technology	361 & 62	Questionnaire Study	Metacognition	Employee adaptive performance#*	Job Complexity***	Emotional Control***
10	Hidayat et al. (2018)	Students	Mathematics	538	Survey Study	Metacognition	Mathematical modelling competencies#**		Cognitive Strategy** Self-Checking**
11	Elferink-Gemser et al. (2018)	Athletes	Table Tennis	60	Psychological Test	Metacognition and Executive functions	Performance#**		
12	Kim (2018)	Students	University Students	30	Experimental Study	Metacognitive monitoring feedback	Performance in Computer-based training#*		
13	Urban and Wood (2017)	Entrepreneurs	Entrepreneurs	784	Questionnaire Study	Entrepreneurial Metacognition*** Corporate building blocks*** Entrepreneurial alertness	Corporate entrepreneurial activity#	Firm age and size, sector, gender, education, work tenure	
14	Kiso and Hershey (2017)	Professionals	Diverse Fields	90	Questionnaire Study	Self-regulated learning Financial Knowledge Financial worry Retirement metacognition*	Problem-Solving# Financial Planning activities#		

Note. For the studies with significant results *p<0.05, **p<0.01, ***p<0.001
For the variables included in the conceptual framework.

Notes. For the studies with significant results *p<0.05, **p<0.01, ***p<0.001.
For the variables included in the conceptual framework.

Identification of Themes

The strategic diagram visualization instrument was used to detect the themes of various cognitive competencies at work investigated with metacognition through the R software, bibliometrix package (Refer: Figure 2). The strategic diagram presents the themes using a co-word analysis where the nodes in the network represent the keywords and the linkage between these nodes is the number of times the same articles mention the linked keywords (Chen et al., 2016). The clusters of keywords represented by node shape in the network are identified using the community detection algorithm (Newman & Girvan, 2004). The keywords are indicative of the themes in the knowledge domain and the co-occurrence of these in a document is indicative of the linkage of the relationship between the themes of the document. The strategic diagram in this study constituted of authors' keywords used in the 39 included studies. We considered 200 keywords out of a total of 455 keywords using bibliometrix package by categorizing various themes based on two measures; centrality and density (Cobo et al., 2015). Centrality (Relevance degree) indicates the degree of interaction between the themes (in this case interaction between metacognition and the various cognitive competencies). The size of the nodes is representative of the degree of centrality of each keyword. Density (Development degree) measures the strength of internal ties within a theme i.e. the strength of relationships within metacognition and cognitive competencies.

The four quadrants presented in the strategic diagram represent the four types of themes – the upper right quadrant is the "Motor Theme" indicating high centrality and high density i.e. themes which have been well developed and important in the research field; the upper left quadrant is the "Niche Theme" with high density and low centrality, themes which have marginal relevance for research; "Emerging or Declining theme" is represented in the lower left quadrant of the diagram depicting themes which are marginal and weakly developed with low density and low centrality. At last the "Basic or General theme" is represented in the lower right quadrant, these

are found to be interconnected with other themes however are not well developed (low density and high centrality) (Bamel et al., 2022).

We identified nine themes, spread across the four quadrants. The "Motor Theme" quadrant depicts 'Academic and Learning Performance', 'Decision Making' and 'Problem Solving'. This quadrant confirms the high development of these themes with metacognition in the academic domain among students. The 'Academic and Learning Performance' includes sub-themes namely, knowledge acquisition, to name a few. While the term 'decision making' includes sub-themes related to strategic, financial and team decisions. The theme of problem-solving included the sub-themes such as; learning performance, goal-orientation and cognitive strategy. However, a portion of these themes fall under "Basic Themes" and there is a need to further explore the relationships of these themes with metacognition with organisational perspective.

The "Basic or General theme" includes 'Individual Performance', 'Team Performance' and 'Firm Performance'. This indicates that these themes are interconnected with metacognition via other themes and need further investigation of having a direct causal relationship. The high centrality of these themes is indicative of the strength of its ties with the other themes over the years. Furthermore, the size of the circles shows that these themes are less researched in the area of management however, constitute keywords related to the managerial workplace indicating the presence of metacognition and performance at the workplace.

The "Emerging theme or declining theme" quadrant depicts the following themes; 'Creativity', a small part of 'Entrepreneurial Cognition' and 'Innovation Performance'. The sub-themes discussed sporadically in this section are related to culture, entrepreneurship, collaboration, knowledge sharing and cooperation. Innovation, creativity and entrepreneurship are gaining traction in recent years due to their relevance in the modern and expanding workplace and therefore, require further investigation. The "Niche themes" quadrant does not depict any themes in the context of metacognition at the workplace which has marginal relevance for research.

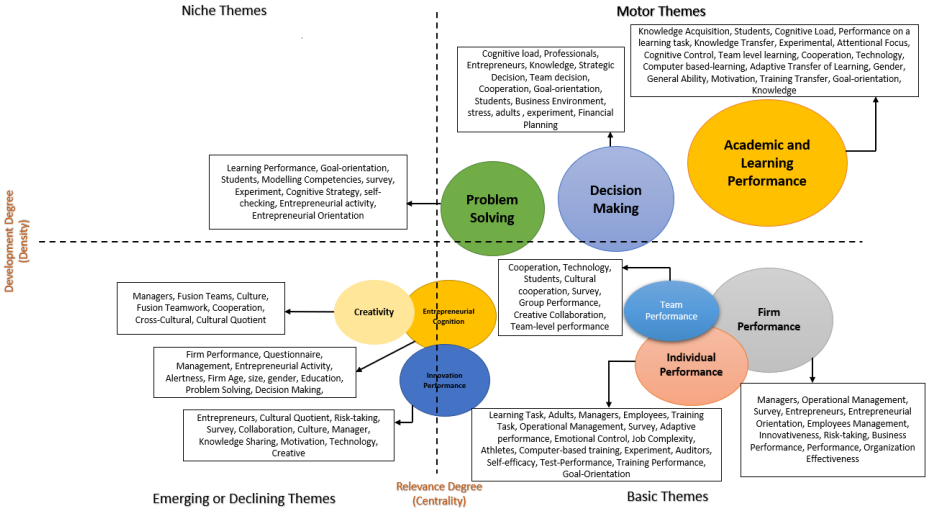

Figure 2. Strategic Diagram

Further, a detailed analysis of the identified themes under each quadrant of the strategic diagram is conducted. A linkage description number is assigned to every established relationship between components of metacognition and the above-identified themes based on the empirical evidence from the extant literature (Refer to Table 2). These themes were then presented through one-way, two-way and expected links which are represented through arrows as visualised in the framework (Figure 3). The empirical studies investigating metacognition with various cognitive competencies and performance are limited to fields other than management such as nursing, fire-fighting, teaching and sports. Therefore, based on empirical linkages explored in these niche domains the authors make some expected propositions linking components and sub-components of metacognition with the cognitive competencies that could facilitate individual and firm performance in the field of management.

Academic and Learning Performance

Metacognition and its sub-components have been majorly empirically explored in the context of educational learning. The linkage of academic and learning performance with metacognitive skilfulness and metacognitive accuracy provides evidence for the workplace in terms of performance on training tasks and knowledge acquisition. The accuracy of the metacognitive judgement is considered relevant for successful learning of the task (Miller & Geraci, 2014). The studies emphasize the role of metacognitive calibration and monitoring accuracy in the enhancement of individuals' knowledge and application of learning strategies. This has also confirmed the importance of the application of metacognitive training and learning strategies to promote individual metacognition among employees (Zhao & Ye, 2020). These components are also found helpful in identifying the difficulties of performing a task and thereby behavioural changes required to develop greater knowledge, performance strategies and confidence of individuals in the task to be performed (Ford et al., 1998; Schmidt & Ford, 2003). The metacognitive strategy component such as planning, monitoring, reviewing & evaluating also tend to influence the attention and cognitive control exerted by individuals while training on a task (Brick et al., 2015). Metacognitive training is therefore found helpful for trainees in carrying out corrective actions in a learning or training task (Kim, 2018) as it increases the capabilities of an individual to gain new knowledge by continuous awareness and monitoring of their strengths, weakness and learning strategies (Zhao & Ye, 2020). Metacognition acts as a directive for better test preparation and information processing required in a job application test (Clause et al., 2001). Various interventions such as self-directed prompts are found to be beneficial in the reviewed studies (Zumbach et al., 2020) and are found to cultivate strategic learning activities amongst learners resulting in better learning performance (Bannert et al., 2015; Schmidt & Ford, 2003). These linkages have also been confirmed amongst dyads and teams (Dierdorff & Ellington, 2012; Norman & Furnes, 2016). Prior research evidence also confirms how improvement in the learning and training performance of the employees

further raises their work performance through the acquisition of knowledge and skill set required for the job to be undertaken (Greco et al., 2018; Guan & Frenkel, 2019). Hence, we propose;

Proposition 1: The Components and Sub-Components of metacognition of an individual are related to Performance on a learning, knowledge acquisition or training task at the workplace and thus with the work performance (1 →2, 3→2, 4→5→6; Figure 3).

Decision Making

Decision making an important workplace cognitive competency is found to be linked to components and sub-components of metacognition (Kiso & Hershey, 2017). An individual who has a strong awareness of his or her knowledge tends to further perceive lower cognitive difficulties in engaging in the cognitive task. The task-specific metacognitive perception (Kiso & Hershey, 2017) also helps an individual in determining the uncertainty and opportunities in the environment and avoiding errors associated with it (Frye & Wearing, 2016; Mattingly et al., 2016). It is therefore argued that individuals who are high on metacognition tend to exhibit less erratic strategic decisions while operating in a dynamic environment (Mitchell et al., 2011). The above studies indicate the role of individuals' metacognition in their decision-making. The literature further confirms how these individual differences in decision-making contribute to the work performance of employees engaged in a cognitive task (Ceschi et al., 2017). Based on the above discussion, we thus propose;

Proposition 2: The Components and Sub-Components of metacognition of an individual are related to his/her decision-making and thus to the work performance (7→8; Figure 3).

Problem Solving

The findings of the review confirm the link between individuals' metacognition and problem-solving also. The components of metacognition are found to help learners strategically analyse the problem, formulate a mental structure, select an appropriate strategy and identify the obstacles that may hamper the problem-solving process (Hidayat et al., 2018; Liu & Liu, 2020b). A study examining the impact of self-questioning techniques on problem-solving performance and metacognition of employees (Ng et al., 2011), confirmed a significant positive effect of self-questioning intervention techniques on problem-solving. The authors further suggested that the long-term application of these techniques can enhance the metacognition of the individual. The above studies, therefore, indicate that metacognitive interventions and training programmes can help improve the problem-solving techniques of an individual and thus work performance (Giampaoli et al., 2017; Ng et al., 2011). We therefore propose;

Proposition 3: The Components and Sub-Components of metacognition of an individual are related to his/her problem-solving and thus to the work performance (12→13; Figure 3).

Cultural Metacognition and Innovative Performance

Management researchers have found linkages between Metacognitive and Cultural Quotient among employees (Chua et al., 2012; Mor et al., 2013). A higher metacognitive cultural quotient is linked with higher knowledge sharing and further to the higher innovative performance of the firm (Berraies, 2020). The literature analysis further shows how individuals high on cultural metacognition work well in fusion teams, promoting creativity in multicultural teams (Crotty & Brett, 2012). The development of cultural metacognition also acts as a crucial dimension in strengthening knowledge sharing, creativity, innovative performance, cooperation and coordination at the workplace (Najmaei & Sadeghinejad, 2019). This improvement in innovation performance further contributes to the increment in performance at work. We therefore propose;

Proposition 4: The Components and Sub-Components of metacognition of an individual are related to his/her Innovative Performance and Cultural Metacognition and thus to the work performance (9→10→11; Figure 3).

Individual and Firm Performance

Recent empirical studies confirm the contribution of managerial metacognition in helping and identifying sources of errors and effective improvement strategies contributing to individual and firm performance (Cho & Linderman, 2019). These results imply the relevance of managerial metacognition at both individual and organisational levels in improving individual and business performance. Yoo et al. (2018) along similar lines investigated the positive mediating role of metacognition on the relationship between technological innovation capabilities, business performance & organisational effectiveness in the area of sustainable management. The results of the study provide evidence of how the positive role of metacognition at the organisational level helps SME managers in the problem-solving process and the successful improvement of business performance and organisational effectiveness. Rhodes et al. (2018) also established the importance of metacognitive knowledge and experience among managers for improvement in the performance of SMEs. These metacognitive components are found to be helpful for individuals to understand their range of knowledge, skills and decision-making thereby, resulting in positive firm performance. Along similar lines results of a doctoral study of 1216 working professionals from diverse fields by Bajaj and Jain (2020) also confirm the positive relationship between metacognitive scores and the performance of employees at the workplace.

Proposition 5: The Components and Sub-Components of metacognition of an individual are related to his/her Individual, Team and Firm Performance (14→19→20, 15→16→19→20, 15→17→18; Figure 3).

Table 2
Known linkages of metacognition and cognitive competencies in the workplace

Metacognition	Workplace Cognitive Competencies	Author (Year)	Linkage description no.
Metacognitive Accuracy (Metacognitive Monitoring; Accurate Calibration)	Performance on a learning task	Bannert et al. (2015); Zhao and Ye (2020)	1
	Performance on a learning task	Liu and Liu (2020b); Kim (2018); Norman and Furnes (2016); Brick et al. (2015); Dierdorff and Ellington (2012)#; Crook and Beier (2010); Schmidt and Ford (2003)*; Tempelaar (2006); Clause et al. (2001)*; Fiore et al. (2002)	3
Metacognitive Skilfulness (Task-Specific Metacognitive Activity; Use of Metacognitive Strategy)	Transfer of Learning	Keith and Frese (2005)*	4→5
	Knowledge Acquisition	Zumbach et al. (2020); Schmidt and Ford (2003)*; Ford et al. (1998)#; Fiore et al. (2002)	4
	Decision Making	Kiso and Hershey (2017); Frye and Wearing (2016); Mattingly et al. (2016); Robert Mitchell et al. (2011); Batha and Carroll (2007); Dierdorff and Ellington (2012)#	7

Metacognition (Metacognitive Awareness Metacognitive Ability Metacognitive Predisposition)	Problem Solving	Liu and Liu (2020b); ; Hidayat et al. (2018); Ng et al. (2011); Brand et al. (2003)	12
Managerial Metacognition	Firm Performance	Cho and Linderman (2019); Yoo et al. (2018)*; Rhodes et al. (2018)	16→19→20 17→18
	Performance	Cho and Linderman (2019); Pan and Sun (2018)	16
Cultural Metacognition	Innovative Performance	Berraies (2020); Crotty and Brett (2012)	10
Cultural Metacognition	Firm Performance	Mor et al. (2013); Chua et al. (2012)	10→11→19→20
Metacognition (Metacognitive Awareness Metacognitive Ability Metacognitive Predisposition)	Performance	Elferink-Gemser et al. (2018); Nietfeld (2003); Najmaei and Sadeghinejad (2019); Cho and Jung (2014)	14
Metacognitive Skilfulness (Task-Specific Metacognitive Activity; Use of Metacognitive Strategy)	Performance	Plumlee et al. (2015)	14

Metacognitive Activity	Collaboration in groups	Nonose et al. (2014); Kwon et al. (2013)	14→19

Notes. Performance = Individual Performance at the workplace.
*: Indicates that metacognition is a mediating variable in the study;
#: Indicates that metacognition is a moderating variable in the study;
( ): Terms used interchangeably in literature with Primary Construct.

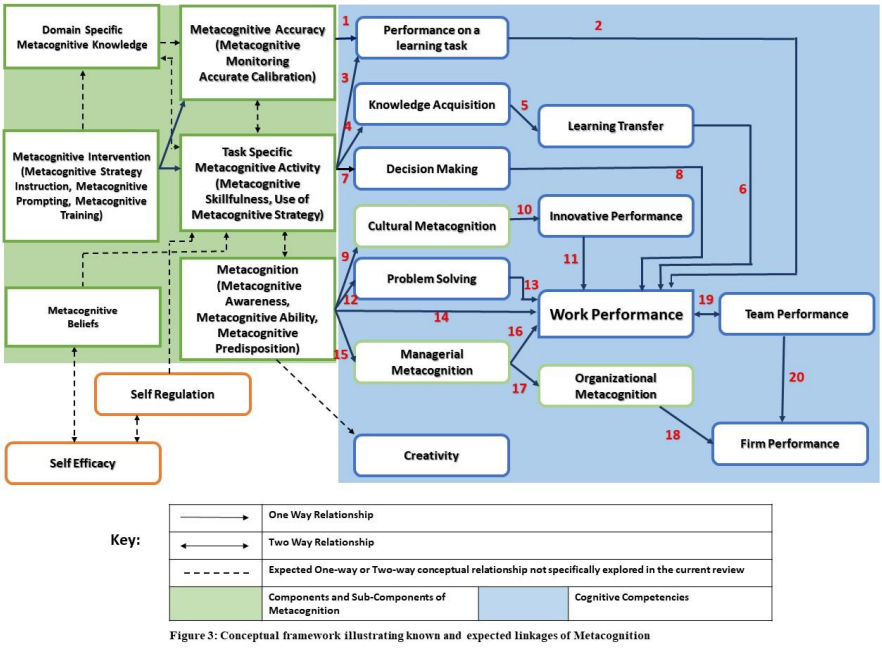

Figure 3: Conceptual framework illustrating known and expected linkages of Metacognition

Figure 3. Conceptual framework illustrating known and expected linkages of Metacognition

Implications and Research gaps

The previous sections have addressed two foremost objectives of the study. Firstly, the identification and integration of linkages between

metacognition and workplace cognitive competencies and secondly, the formulation and presentation of these relationships into a comprehensive conceptual framework. The consolidated framework can be used by scholars and practitioners in the field of human resource management to study and develop the role of metacognition in the context of the workplace.

The 39 identified studies provide various workplace cognitive competencies such as learning performance, decision-making, problem-solving, and innovative performance apart from individual and firm performance linked with metacognition and its components. However, the review also confirms the sporadicity of this empirical evidence and limited literature directly investigating the link between managerial metacognition and work performance. Bajaj and Jain (2020) in a doctoral study have recently empirically substantiated the positive link of Metacognitive Ability scores with performance as adjudged by a mentor among 1216 working professionals. Their results have supported our assertion that there is a measurable link between metacognition and Employee Performance at the workplace, which require further investigation. Metacognition has also been found to be an important predictor of firm performance at the workplace in the field of operational management and sustainable management (Cho & Linderman, 2019; Yoo et al., 2018). Therefore, It is essential to investigate this relationship to understand which individuals may have a priori ability to metacognate and then appropriately recruit employees (Walker, 2016) based on the cognitive requirement of the task.

Further, the review also highlights an essential need to identify a measure of metacognitive awareness suited to the workplace. The existing managerial metacognition measures employed in the empirical studies are adopted from measures used in the context of entrepreneurial metacognition (Haynie, 2005; Haynie et al., 2012; Haynie & Shepherd, 2009), cultural metacognition (Chua et al., 2012; Van Dyne et al., 2012) and metacognition in learning (Yoo et al., 2018). The Metacognitive Awareness Inventory (MAI) by Schraw and Dennison (1994) can be considered a reliable measure of measuring metacognitive awareness at work.

Metacognition has also been seen as critical to performance in online learning environments (Liu & Liu, 2020b; Reisoglu et al., 2020; Zumbach et al., 2020) and a predictor of performance on cognitive tasks (Woolfolk & Shaughnessy, 2004) and learning (Pintrich et al., 2000). This evidence highlights the need to explore the role of managerial metacognition in the context of virtual teams and virtual work environments. Virtual work practices are likely to stay in place post-pandemic as firms realize the cost savings from having workers connect technologically (Kniffin et al., 2020). Therefore, the implications of an empirical examination of the role of managerial metacognition in a virtual context will provide insights for HRD scholars and practitioners.

The review has also confirmed the linkage between the components of metacognition with collaboration in a group or team tasks (Kwon et al., 2013; Nonose et al., 2014). This linkage further sheds light on the expected impact of metacognition on team-level outcomes. A review of the literature and empirical investigation of the same can establish a stronger case for metacognition and team-level outcomes at the workplace.

Finally, we propose to the researchers and practitioners in the field of organisational and business management to empirically examine metacognitive ability as a distinct construct among management professionals influencing workplace cognitive competencies and performance at work.

Conflict of interest

We have no conflicts of interest to disclose.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Ackoff, R. L., & Addison, H. J. (2010). *Systems Thinking for Curious Managers*. Devon: Triarchy Press.

Aghaei Chadegani, A., Salehi, H., Md Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A comparison between two main academic literature collections: Web of science and scopus databases. *Asian Social Science*, *9*(5), 18–26. https://doi.org/10.5539/ass.v9n5p18

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Bajaj, B., & Jain, S. (2020). *A Study of Metacognitive Ability, Emotional Intelligence and Social Intelligence In Relation to Employee's Individual Performance and Team Effectiveness*. University School of Management Studies, Guru Gobind Singh Indraprastha University.

Bamel, U., Pereira, V., Del Giudice, M., & Temouri, Y. (2022). The extent and impact of intellectual capital research: a two decade analysis. *Journal of Intellectual Capital*, *23*(2), 375–400. https://doi.org/10.1108/JIC-05-2020-0142

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, *52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038

Batha, K., & Carroll, M. (2007). Metacognitive training aids decision making. *Australian Journal of Psychology*, *59*(2), 64–69. https://doi.org/10.1080/00049530601148371

Berraies, S. (2020). Effect of middle managers' cultural intelligence on firms' innovation performance: Knowledge sharing as mediator and collaborative climate as moderator. *Personnel Review*, *49*(4), 1015–1038. https://doi.org/10.1108/PR-10-2018-0426

Bonesso, S., Gerli, F., Pizzi, C., & Cortellazzo, L. (2018). Students' Entrepreneurial Intentions: The Role of Prior Learning Experiences and Emotional, Social, and Cognitive Competencies. *Journal of Small Business Management*, *56*(00), 215–242. https://doi.org/10.1111/jsbm.12399

Bonner, J. (1998). Implications of Cognitive Theory for Instructional Design: Revisited. *Educational Communication and Technology*, *36*(1), 4–14.

Boyatzis, R.E. (1991). *The competent manager: A model for effective performance*. John Wiley & Sons.

Boyatzis, Richard E., Thiel, K., Rochford, K., & Black, A. (2017). Emotional and Social Intelligence Competencies of Incident Team Commanders Fighting Wildfires. *Journal of Applied Behavioral Science*, *53*(4), 498–516. https://doi.org/10.1177/0021886317731575

Brand, S., Reimer, T., & Opwis, K. (2003). Effects of metacognitive thinking and knowledge acquisition in dyads on individual problem solving and transfer performance. *Swiss Journal of Psychology*, *62*(4), 251–261. https://doi.org/10.1024/1421-0185.62.4.251

Brick, N., MacIntyre, T., & Campbell, M. (2015). Metacognitive processes in the self-regulation of performance in elite endurance runners. *Psychology of Sport and Exercise*, *19*, 1–9. https://doi.org/10.1016/j.psychsport.2015.02.003

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. *Metacognition, Motivation, and Understanding,* pp. 65–116. New York: Wiley.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In *Cognitive Development*. New York: Wiley.

Ceschi, A., Demerouti, E., Sartori, R., & Weller, J. (2017). Decision-making processes in the workplace: How exhaustion, lack of resources and job demands impair them and affect performance. *Frontiers in Psychology*, *8*(May), 1–14. https://doi.org/10.3389/fpsyg.2017.00313

Chen, X., Chen, J., Wu, D., Xie, Y., & Li, J. (2016). Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project. *Procedia Computer Science*, *91*, 547–555. https://doi.org/10.1016/j.procs.2016.07.140

Cho, Y. S., & Jung, J. Y. (2014). The Relationship between Metacognition, Entrepreneurial Orientation, and Firm Performance: An Empirical Investigation. *Academy of Entrepreneurship Journal*, *20*(2), 71–86.

Cho, Y. S., & Linderman, K. (2019). Metacognition-based process improvement practices. *International Journal of Production Economics*, *211*, 132–144. https://doi.org/10.1016/j.ijpe.2019.01.030

Chua, R. Y. J., Morris, M. W., & Mor, S. (2012). Collaborating across cultures: Cultural metacognition and affect-based trust in creative collaboration. *Organizational Behavior and Human Decision Processes*, *118*(2), 116–131. https://doi.org/10.1016/j.obhdp.2012.03.009

Clause, C. S., Delbridge, K., Schmitt, N., Chan, D., & Jennings, D. (2001). Test Preparation Activities and Employment Test Performance. *Human Performance*, *14*(2), 149–167. https://doi.org/10.1207/S15327043HUP1402_02

Cobo, M. J., Martínez, M. A., Gutiérrez-Salcedo, M., Fujita, H., & Herrera-Viedma, E. (2015). 25 years at Knowledge-Based Systems: A bibliometric analysis. *Knowledge-Based Systems*, *80*, 3–13. https://doi.org/10.1016/j.knosys.2014.12.035

Crook, A. E., & Beier, M. E. (2010). When Training With a Partner Is Inferior to Training Alone: The Importance of Dyad Type and Interaction Quality. *Journal of Experimental Psychology: Applied*, *16*(4), 335–348. https://doi.org/10.1037/a0021913

Crotty, S. K., & Brett, J. M. (2012). Fusing creativity: Cultural metacognition and teamwork in multicultural teams. *Negotiation and Conflict Management Research*, *5*(2), 210–234. https://doi.org/10.1111/j.1750-4716.2012.00097.x

Dierdorff, E. C., & Ellington, J. K. (2012). Members matter in team training: Multilevel and longitudinal relationships between goal orientation, self-regulation, and team outcomes. *Personnel Psychology*, *65*(3), 661–703. https://doi.org/10.1111/j.1744-6570.2012.01255.x

Duman, B. (2018). The Relationship Between the Entrepreneurship Characteristics and Metacognitive Awareness Levels of Pre-service Teachers. *Journal of Education and Training Studies*, *6*(5), 152–160. https://doi.org/10.11114/jets.v6i5.3080

Elferink-Gemser, M. T., Faber, I. R., Visscher, C., Hung, T. M., De Vries, S. J., & Nijhuis-Vandersanden, M. W. G. (2018). Higher-level cognitive functions in Dutch elite and sub-elite table tennis players. *PLoS ONE*, *13*(11), 1–13. https://doi.org/10.1371/journal.pone.0206151

Fiore, S. M., Cuevas, H. M., Scielzo, S., & Salas, E. (2002). Training individuals for distributed teams: Problem solving assessment for distributed mission research. *Computers in Human Behavior*, *18*(6), 729–744. https://doi.org/10.1016/S0747-5632(02)00027-4

Flavell, J. H. (1976). Metacognitive aspects of problem solving. *The Nature of Intelligence*. NY: Academic New York.

Flavell, J. H. (1981). Cognitive monitoring. *Children's Oral Communication Skills*. NY: Academic New York.

Flavell, John H. (1979). Metacognition and Cognitive Monitoring. *American*

Psychologist, *34*(10), 906–911. https://doi.org/10.1093/nq/CLVII.dec14.424-a

Ford, J. K., Weissbein, D. A., Smith, E. M., Gully, S. M., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology*, *83*(2), 218–233. https://doi.org/10.1037/0021-9010.83.2.218

Fornes, S. L., Rocco, T. S., & Wollard, K. K. (2008). Workplace commitment: A conceptual model developed from integrative review of the research. *Human Resource Development Review*, *7*(3), 339–357. https://doi.org/10.1177/1534484308318760

Frye, L. M., & Wearing, A. J. (2016). A model of metacognition for bushfire fighters. *Cognition, Technology and Work*, *18*(3), 613–619. https://doi.org/10.1007/s10111-016-0372-4

Gavelek, J. and Raphael, T. E. (1985). *Metacognition,Instruction, and the Role of Questioning Activities*. ,in Forrest-Pressley, D. L., MacKinnon, G. E. andWaller, T. G. (Eds), Metacognition, Cognition andHuman Performance (Volume 2 – Instructional Prac-tices), Academic Press, New York, pp. 103–136.

Giampaoli, D., Ciambotti, M., & Bontis, N. (2017). Knowledge management, problem solving and performance in top Italian firms. *Journal of Knowledge Management*, *21*(2), 355–375. https://doi.org/10.1108/JKM-03-2016-0113

Greco, L. M., Charlier, S. D., & Brown, K. G. (2018). Trading off learning and performance : Exploration and exploitation at work. *Human Resource Management Review*, *March*, 0–1. https://doi.org/10.1016/j.hrmr.2018.06.001

Guan, X., & Frenkel, S. (2019). How perceptions of training impact employee performance manufacturing firms. *Personnel Review*, *48*(1), 163–183. https://doi.org/10.1108/PR-05-2017-0141

Guz, A. N., & Rushchitsky, J. J. (2009). Scopus: A system for the evaluation of scientific journals. *International Applied Mechanics*, *45*(4), 351–362. https://doi.org/10.1007/s10778-009-0189-4

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). *Metacognition in Educational Theory and Practice*. Routledge, Taylor & Francis Group.

Hardy, J. H., Day, E. A., & Steele, L. M. (2019). Interrelationships Among Self-Regulated Learning Processes: Toward a Dynamic Process-Based Model of Self-Regulated Learning. *Journal of Management*, *45*(8), 3146–3177. https://doi.org/10.1177/0149206318780440

Harrison, G. M., & Vallin, L. M. (2018). Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition and*

Learning, *13*(1), 15–38. https://doi.org/10.1007/s11409-017-9176-z

Hart, J. T. (1965). Memory and the Feeling of Knowing Experience. *Journal of Educational Psychology*, *56*(4), 208–216. http://arxiv.org/abs/1012.0819

Haynie, J. Michael, Shepherd, D., Mosakowski, E., & Earley, P. C. (2010). A situated metacognitive model of the entrepreneurial mindset. *Journal of Business Venturing*, *25*(2), 217–229. https://doi.org/10.1016/j.jbusvent.2008.10.001

Haynie, James M. (2005). The Role of Metacognition and Feedback in Entrepreneurial. *Pro Quest, Dissertation*.

Haynie, M. J., & Shepherd, D. A. (2009). A measure of adaptive cognition for entrepreneurship research. *Entrepreneurship: Theory and Practice*, *33*(3), 695–714.

Haynie, M. J., Shepherd, D. A., & Patzelt, H. (2012). Cognitive Adaptability and an Entrepreneurial Task: The Role of Metacognitive Ability and Feedback. *Entrepreneurship: Theory and Practice*, *36*(2), 237–265. https://doi.org/10.1111/j.1540-6520.2010.00410.x

Hidayat, R., Zulnaidi, H., & Zamri, S. N. A. S. (2018). Roles of metacognition and achievement goals in mathematical modeling competency: A structural equation modeling analysis. *PLoS ONE*, *13*(11), 1–25. https://doi.org/10.1371/journal.pone.0206211

Jia, X., Li, W., & Cao, L. (2019). The role of metacognitive components in creative thinking. *Frontiers in Psychology*, *10*(2404), 1–11. https://doi.org/10.3389/fpsyg.2019.02404

Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology*, *90*(4), 677–691. https://doi.org/10.1037/0021-9010.90.4.677

Kim, J. H. (2018). The effect of metacognitive monitoring feedback on performance in a computer-based training simulation. *Applied Ergonomics*, *67*(October 2017), 193–202. https://doi.org/10.1016/j.apergo.2017.10.006

Kiso, H., & Hershey, D. A. (2017). Working adults' metacognitions regarding financial planning for retirement. *Work, Aging and Retirement*, *3*(1), 77–88. https://doi.org/10.1093/workar/waw021

Kniffin, K. M., Narayanan, J., Anseel, F., Antonakis, J., Ashford, S. P., Bakker, A. B., Bamberger, P., Bapuji, H., Bhave, D. P., Choi, V. K., Creary, S. J., Demerouti, E., Flynn, F. J., Gelfand, M. J., Greer, L. L., Johns, G., Kesebir, S., Klein, P. G., Lee, S.

Y., ... van Vugt, M. (2020). COVID-19 and the Workplace: Implications, Issues, and Insights for Future Research and Action. *American Psychologist*, *76*(1), 63–77. https://doi.org/10.1037/amp0000716

Knox, B. J., Lugo, R. G., Josok, O., Helkala, K., & Sütterlin, S. (2017). Towards a cognitive agility index: The role of metacognition in human computer interaction. *Communications in Computer and Information Science*, *713*, 330–338. https://doi.org/10.1007/978-3-319-58750-9_46

Kuhn, D., & Dean, D. J. (2004). Metacognition : A Bridge Between. *Theory Into Practice*, *43*(4), 268–273. https://doi.org/10.1207/s15430421tip4304

Kwon, K., Hong, R. Y., & Laffey, J. M. (2013). The educational impact of metacognitive group coordination in computer-supported collaborative learning. *Computers in Human Behavior*, *29*(4), 1271–1281. https://doi.org/10.1016/j.chb.2013.01.003

Liu, S., & Liu, M. (2020a). Data on player activity and characteristics in a Serious Game Environment. *Data in Brief*, *28*, 104965. https://doi.org/10.1016/j.dib.2019.104965

Liu, S., & Liu, M. (2020b). The impact of learner metacognition and goal orientation on problem-solving in a serious game environment. *Computers in Human Behavior*, *102*, 151–165. https://doi.org/10.1016/j.chb.2019.08.021

Livingstone, J. A. (1997). *Metacognition: An Overview.* http. www, gse. buffalo, edu/fas/shuell/CEP564/Metaeog. htm.

Lyons, P. R., & Bandura, R. P. (2019). Exploring linkages of performance with metacognition. *Journal of Management Development*, *38*(3), 195–207. https://doi.org/10.1108/JMD-07-2018-0192

Martinez, M. E. (2006). What Is Metacognition? *Phi Delta Kappan*, *87*(9), 696–699.

Matsaggouras, E. G. (1994). Theory and Practice of Instruction. *Teaching Strategies: From Information to Critical Thinking*, *2*(2). Taylor & Francis.

Mattingly, E. S., Kushev, T. N., Ahuja, M. K., & Ma, D. (2016). Switch or persevere? The effects of experience and metacognition on persistence decisions. *International Entrepreneurship and Management Journal*, *12*(4), 1233–1263. https://doi.org/10.1007/s11365-016-0391-x

Miller, T. M., & Geraci, L. (2014). Improving metacognitive accuracy: How failing to retrieve practice items reduces overconfidence. *Consciousness and Cognition*, *29*, 131–140. https://doi.org/10.1016/j.concog.2014.08.008

Mitchell, R. J., Shepherd, D. A., & Sharfman, M. P. (2011). Erratic Strategic Decisions: When and Why Managers are Inconsistent in Strategic Decision Making.

Strategic Management Journal, *32*, 683–704. https://doi.org/10.1002/smj

Mor, S., Morris, M. W., & Joh, J. (2013). Identifying and Training Adaptive Cross-Cultural Management Skills : The Crucial Role of. *Academy of Management Learning & Education*, *12*(3), 453–475.

Najmaei, A., & Sadeghinejad, Z. (2019). Metacognition, Entrepreneurial Orientation, and Firm Performance: An Upper Echelons View. In *The Anatomy ofEntrepreneurial Decisions, Contributions to Management Science* (pp. 79–114). https://doi.org/10.1007/978-3-030-19685-1_5

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *69*(2 2), 1–15. https://doi.org/10.1103/PhysRevE.69.026113

Ng, K. H., Lee, C. B., & Teo, T. (2011). The role of self-questioning: Problem solving in a security organization. *Systems Research and Behavioral Science*, *28*(1), 91–104. https://doi.org/10.1002/sres.1062

Nietfeld, J. (2003). An Examination of Metacognitive Strategy Use and Monitoring Skills by Competitive Middle Distance Runners. *Journal of Applied Sport Psychology*, *15*(4), 307–320. https://doi.org/10.1080/714044199

Nonose, K., Kanno, T., & Furuta, K. (2014). Effects of metacognition in cooperation on team behaviors. *Cognition, Technology and Work*, *16*(3), 349–358. https://doi.org/10.1007/s10111-013-0265-8

Norman, E., & Furnes, B. (2016). The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media? *Computers in Human Behavior*, *54*, 301–309. https://doi.org/10.1016/j.chb.2015.07.043

Oh, H. (2016). Impact of Metacognition on Simulation Effectiveness in Problembased Learning using Simulation. *International Journal of Computer Science and Information Technology for Education*, *1*(1), 59–64. https://doi.org/10.21742/ijcsite.2016.1.10

Pan, W., & Sun, L. Y. (2018). A Self-Regulation Model of Zhong Yong Thinking and Employee Adaptive Performance. *Management and Organization Review*, *14*(1), 135–159. https://doi.org/10.1017/mor.2017.33

Paris, S. G., & Winograd, P. (1990). Promoting Metacognition and Motivation of Exceptional Children. *Remedial and Special Education*, *11*(6), 7–15. https://doi.org/10.1177/074193259001100604

Piaget, J. (1950). *The psychology of intelligence*. London, UK: Routledge and Kegan

Paul.

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In *Issues in the measurement of metacognition* (pp. 43–97). https://digitalcommons.unl.edu/burosmetacognition/3

Plumlee, R. D., Rixom, B. A., & Rosman, A. J. (2015). Training auditors to perform analytical procedures using metacognitive skills. *Accounting Review*, *90*(1), 351–369. https://doi.org/10.2308/accr-50856

Pohl, S., Zobel, J., & Moffat, A. (2010). Extended Boolean retrieval for systematic biomedical reviews. *Conferences in Research and Practice in Information Technology Series*, *102*(ACSC), 117–125.

Reisoglu, I., Eryılmaz Toksoy, S., & Erenler, S. (2020). An analysis of the online information searching strategies and metacognitive skills exhibited by university students during argumentation activities. *Library and Information Science Research*, *42*(3), 101019. https://doi.org/10.1016/j.lisr.2020.101019

Rhodes, J., Cheng, V., Sadeghinejad, Z., & Lok, P. (2018). The relationship between management team (TMT) metacognition, entrepreneurial orientations and small and medium enterprises (SMEs) firm performance. *International Journal of Management Practice*, *11*(2), 111–140. https://doi.org/10.1504/IJMP.2018.090830

Rozhkov, M., Cheung, B. C. F., & Tsui, E. (2017). Workplace context and its effect on individual competencies and performance in work teams. *International Journal of Business Performance Management*, *18*(1), 49–81. https://doi.org/10.1504/ijbpm.2017.10001261

Ryan, G., Emmerling, R. J., & Spencer, L. M. (2009). Distinguishing high-performing European executives: The role of emotional, social and cognitive competencies. *Journal of Management Development*, *28*(9), 859–875. https://doi.org/10.1108/02621710910987692

Schmidt, A. M., & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology*, *56*(2), 405–429. https://doi.org/10.1111/j.1744-6570.2003.tb00156.x

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, *36*(1–2), 111–139. https://doi.org/10.1007/s11165-005-3917-8

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness.

Contemporary Educational Psychology, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Shahbaz, W., & Parker, J. (2021). Workplace mindfulness: An integrative review of antecedents, mediators, and moderators. *Human Resource Management Review*, 100849. https://doi.org/10.1016/j.hrmr.2021.100849

Shepherd, D. A., Patzelt, H., & Haynie, J. M. (2010). Entrepreneurial spirals: Deviation-amplifying loops of an entrepreneurial mindset and organizational culture. *Entrepreneurship: Theory and Practice*, *34*(1), 59–82. https://doi.org/10.1111/j.1540-6520.2009.00313.x

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*(March), 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Spencer, & Spencer, L. M. (1993). *Competence at work*. NY: John Wiley & Sons, Inc.

Tempelaar, D. T. (2006). The role of metacognition in business education. *Industry and Higher Education*, 291–297.

Torraco, R. J. (2002). Cognitive Demands of New Technologies and the Implications for Learning Theory. *Human Resource Development Review*, *1*(4), 439–467. https://doi.org/10.1177/1534484302238436

Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, *4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Urban, B., & Wood, E. (2017). The innovating firm as corporate entrepreneurship. *European Journal of Innovation Management*, *20*(4), 534–556. https://doi.org/10.1108/EJIM-10-2016-0100

Van Dyne, L., Ang, S., Ng, K. Y., Rockstuhl, T., Tan, M. L., & Koh, C. (2012). Sub-Dimensions of the Four Factor Model of Cultural Intelligence: Expanding the Conceptualization and Measurement of Cultural Intelligence. *Social and Personality Psychology Compass*, *6*(4), 295–313. https://doi.org/10.1111/j.1751-9004.2012.00429.x

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*(1), 3–14. https://doi.org/10.1007/s11409-006-6893-0

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: M.I.T. Press.

Walker, E. T. (2016). *The Application of Metacognition to Business Decision Making*.

Master thesis.

Wollard, K. K., & Shuck, B. (2011). Antecedents to employee engagement: A structured review of the literature. *Advances in Developing Human Resources*, *13*(4), 429–446. https://doi.org/10.1177/1523422311431220

Woolfolk, A., & Shaughnessy, M. F. (2004). An Interview With Anita Woolfolk : The Educational Psychology of Teacher Efficacy. *Educational Psychology Review*, *16*(2), 153–176.

Yoo, W. J., Choo, H. H., & Lee, S. J. (2018). A study on the sustainable growth of SMEs: The mediating role of organizational metacognition. *Sustainability (Switzerland)*, *10*(8), 1–17. https://doi.org/10.3390/su10082829

Zhao, L., & Ye, C. (2020). Time and Performance in Online Learning: Applying the Theoretical Perspective of Metacognition. *Decision Sciences Journal of Innovative Education*, *18*(3), 435–455. https://doi.org/10.1111/dsji.12216

Zumbach, J., Rammerstorfer, L., & Deibl, I. (2020). Cognitive and metacognitive support in learning with a serious game about demographic change. *Computers in Human Behavior*, *103*(August 2019), 120–129. https://doi.org/10.1016/j.chb.2019.09.026

Appendix A

Sub-Components of Metacognitive Knowledge and Metacognitive Regulation

Dimensions	Definition
Sub-Components of Metacognitive Knowledge (or awareness)	
Declarative Knowledge	Declarative knowledge is a learner's knowledge of own resources and capabilities (Kuhn & Dean, 2004; Schraw et al., 2006).
Procedural Knowledge	Procedural knowledge is a learner's knowledge of the purpose of a task, the processes used to solve problems and his capacity to self-regulate tasks (Nelson & Narens, 1994; Schraw et al., 2006).
Conditional Knowledge	Conditional knowledge is the learner knowing the different conditions in which his declarative and procedural knowledge of a task can be applied (Schraw et al., 2006; Schraw & Dennison, 1994).
Sub-Components of Metacognitive Regulation (or Skills or activity or strategy)	
Metacognitive Monitoring	Monitoring involves making self-aware judgments about one's learning. It includes an awareness of task complexity (Efklides, 2006) and self-enhancement motivation (Jiang & Kleitman, 2015; Schraw et al., 2006).
Metacognitive Planning	Planning refers to the evaluation and employment of the most efficient resources and strategies (Li et al., 2015; Schraw et al., 2006).
Metacognitive Evaluation	Evaluation refers to the ability to make metacognitive judgments and interpret the outcome of the monitoring process (Schraw et al., 2006).

Appendix B

Key features of studies selected for integrative review (Study 15-39)

No.	Author (Year)	Subject	Field/Industry	Sample Size	Research Design	Study Design			
						Independent Variable	Dependent Variable	Moderating Variable	Mediating Variable
15	Norman and Furnes (2016)	Students	Technology	100 & 50	Experimental Study	Metacognitive Experiences	Learning for digital text# Learning for non-digital text#		
16	Frye and Wearing (2016)	Professionals	Bushfire Fighters	3 scenarios cases	Survey Study	Metacognition	Decision-Making #*	Stressful high cognitive load conditions	
17	Mattingly et al. (2016)	Entrepreneurs	Entrepreneurs from diverse industries	124	Conjoint Experimental Study	Entrepreneurial experience	Decision-Making#	Metacognitive experience Metacognitive knowledge*	
18	Bannert et al. (2015)	Students	Technology	35 & 35	Experimental Study	Self-directed metacognitive prompts	Behaviour and learning performance#*		
19	Brick et al. (2015)	Athletes	Runners	10	Interviews	Metacognition	Attentional Focus# Cognitive Control (Strategy)		
20	Plumlee et al. (2015)	Professionals	Auditors (Accounting)	108	Experimental Study	Metacognitive processes (Creative problem solving: Divergent and	Performance#***		

						convergent thinking)			
21	Nonose et al. (2014)	Students	Aviation Game	13 Teams	Experimental Study	Metacognition	Team Cooperation#*		
22	Cho and Jung (2014)	MBA students and employees	Entrepreneurs	190	Survey Study	Entrepreneur's metacognition	Firm Performance#*		Entrepreneurial Orientation**
23	Mor et al. (2013)	Students	MBA Students	200	Survey Study	Cultural Metacognition	Inter-cultural cooperation#* Cultural Performance#*		
24	Kwon et al. (2013)	Students	Web design research University Students	59	Experimental Study	Group Metacognitive Activity	Positive Interdependence* Group Performance#		
25	Chua et al. (2012)	Managers	Executive MBA managers	43,60,236	Survey Study	Cultural metacognition	Intercultural creative collaborations#*		
26	Dierdorff and Ellington (2012)	Students	Management Students	338/64	Experimental Study	Goal-orientation	Team Strategic Decision Making#** Team level learning# Cooperation#	Individual Level Metacognition and Self-efficacy** Team level Metacognition and Self-efficacy**	
27	Crotty and Brett (2012)	Managers	Diverse fields	246/37	Survey Study	Cultural Metacognition	Creativity#**	Fusion Team**	Fusion Teamwork**
28	Robert Mitchell et al. (2011)	Professionals	Technology	127	Conjoint Experimental Study	Metacognitive Experience Business Environment	Strategic Decision-Making#**		
29	Ng et al. (2011)	Professionals	Border Security Services	45	Questionnaire Study	Effects of self-questioning techniques	Problem-solving Performance#* Metacognition		

30	Crook and Beier (2010)	Students (Dyads)	Technology	64	Experimental Study	Metacognitive Activity	Learner Performance #*			
31	Batha and Carroll (2007)	Students	Undergraduate Psychology students	98	Experimental Study	Metacognitive Ability and strategy Metacognitive Strategy Instruction	Decision-Making#**			
32	Tempelaar (2006)	Under-graduate students	Students	729	Experimental Study	Metacognition	Academic Performance #*			
33	Keith and Frese (2005)	Students	Primary and Secondary Education	55	Experimental Study	Training Condition (error management vs error avoidant)	Adaptive Transfer of Learning#			Metacognitive Activity** Emotional Regulation *
34	Schmidt and Ford (2003)	Students	Technology (Web-based course)	70	Experimental Study	Metacognitive Intervention	Learning Performance #**	Mastery orientation Performance Orientation		Metacognitive Activity
35	Nietfeld (2003)	Athletes	Sports (Runners)	45	Experimental Study	Racing Metacognition	Race Performance #*	General Ability Gender		
36	Brand et al. (2003)	Students (Dyads)	Business administration university students	107	Experimental Study	Metacognition	Problem-solving#*			
37	Fiore et al. (2002)	Students	Technology (Computer based training)	25	Experimental Study	Mental maps	Learning Performance #*			
						Metacognitive Predisposition	Knowledge Acquisition# *			
						Metacognitive Accuracy Dynamic Problem Solving	Task Performance #*			
38	Clause et al. (2001)	Job applicants	Law agency	493	Experimental and Questionnaire study	Self-efficacy Motivation	Test Performance #*			Metacognition Depth* Effort*
39		Students	Undergraduate Psychology students	93		Mastery Orientation*	Training Transfer	Knowledge		Self-Efficacy

Ford et al. (1998)	Experimental Study	Performance Orientation** *	and Performance#	Metacognition

For the studies with significant results *p<0.05, **p<0.01, ***p<0.001

For the variables included in the conceptual framework.