Received: 20 September 2021 | Revised: 18 April 2022 | Accepted: 19 April 2022

DOI: 10.1002/acp.3952

RESEARCH ARTICLE

WILEY

Meta-analytic evidence shows no relationship between task-based and self-report measures of thought control

Kyra Göbel | Lisa Hensel | Oliver C. Schultheiss | Cornelia Niessen

Institute of Psychology, Friedrich-Alexander University of Erlangen-Nürnberg, Erlangen, Germany

Correspondence
Kyra Göbel and Lisa Hensel, Institute of Psychology, Friedrich-Alexander University of Erlangen-Nürnberg, Nägelsbachstr. 49c, 91052 Erlangen, Germany.
Email: kyra.goebel@fau.de and lisa.hensel@fau.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Numbers: NI 1066/3-1 (DFG), NI 1066/4-1 (DFG)

Abstract

We meta-analytically examined correlations between task-based and self-report thought control measures as well as potential moderators. Results of a frequentist meta-analysis indicate that the two measures do not overlap, $r = .07$, 95% CI [−0.01;0.15] ($k = 20$; $N = 1194$); subgroup analyses for different self-report measures $r = .12$; 95% CI [−0.06;0.30] and $r = .03$, 95% CI [−0.05;0.11], respectively. Additionally, we found no evidence for the presence of moderators. To further test whether the identified effect size is more consistent with a null effect or the assumption of a positive association, robust Bayesian meta-analyses were conducted. We did not obtain sufficiently strong evidence to decide for or against a null effect. In sum, our findings suggest that task-based and self-report thought control measures might capture different psychological processes and thus call for more research on the different underlying processes and constructs they measure.

KEYWORDS

Bayesian meta-analysis, meta-analysis, think/no-think, thought control, thought suppression

1 | INTRODUCTION

Facing problems and coping with unwanted thoughts are a regular part of human life. Thought control ability, defined as the ability to control one's intrusive, unpleasant or unwanted thoughts (Luciano et al., 2005; Williams et al., 2010), is thus an important skill, as it helps regulate emotions (Engen & Anderson, 2018; Erdelyi, 2006; Gagnepain et al., 2017; Geraerts et al., 2010; Nørby, 2015, 2017), maintain task focus and cognitive flexibility (van Vugt & van der Velde, 2018), preserve a positive self-perception (Göbel & Niessen, 2021; Sedikides & Green, 2009), and maintain beliefs, attitudes (Waldum & Sahakyan, 2012), relationships (Freyd, 1996) and a generally positive and coherent self-image (Anderson & Hanslmayr, 2014; Fawcett & Hulbert, 2020). The ability to control one's thoughts is of great interest in both clinical and non-clinical psychological research. Researchers measure thought control with standardized laboratory tasks requiring inhibitory control, such as the think/no-think paradigm (TNT; Anderson & Green, 2001; Anderson & Weaver, 2009), and with self-report questionnaires (e.g., Thought Control Ability Questionnaire [TCAQ]; Luciano et al., 2005). Anderson and colleagues assume that memory inhibition, initiated by intentionally avoiding recall of a specific thought, is the main mechanism causing reduced memory access in the TNT (Anderson & Green, 2001; Racsmány et al., 2011). Questionnaires such as the TCAQ aim to assess individual differences in the perceived control of intrusive, unwanted thoughts (Feliu-Soler et al., 2019; Luciano et al., 2005). But how well do self-report questionnaires capture the process that the TNT is intended to reflect? Is there sufficient convergence to trust self-report questionnaires as an indicator for the process assessed by the TNT paradigm?

Kyra Göbel and Lisa Hensel contributed equally to this work and share first authorship.

Although there is clear theoretical overlap between the two measures, published empirical data on their association with one another are scarce (for exception, see Catarino et al., 2015; Küpper et al., 2014; Waldhauser et al., 2018). However, we assume that the respective data have often been collected without being explicitly reported, as self-report measures are often part of the TNT procedure, but not central to the research question. To our knowledge, no line of research has directly addressed this question, but similar debates about task-based versus self-report measures exist in the contexts of inhibition (Von Gunten et al., 2020), self-regulation (Enkavi et al., 2019), and self-control (Dreves et al., 2020). Wennerhold and Friese (2020) argue that small-to-zero correlations between self-report and behavioral measures are a common phenomenon that is often due more to substantial theoretical differences than to psychometric properties.

The main aim of the present study is to meta-analytically investigate the consistency and effect size of the relationship between task-based and self-report measures of thought control in the extant literature. Furthermore, we want to test whether inconsistencies in effect sizes are due to methodological artifacts or potential moderator variables. In addition, we aim to discuss commonalities and differences between the two assessment methods based on our findings.

1.1 | Thought control

Thought control is a cognitive control concept that refers to the suppression of unwanted, intrusive thoughts—that is, deliberate attempts to not think about specific memory contents (Luciano et al., 2005; Wegner et al., 1987). Originally, the term "thought control" emerged from research in clinical psychology, but as both clinical (e.g., depressed patients) and non-clinical samples experience unwanted thoughts with similar form and content, individual differences in dealing with these thoughts have gradually become a focus of mainstream psychological research.

To assess thought suppression and its efficacy, Wegner et al. (1987) designed the famous "white bear" experiment instructing participants to not think about a specific target, (i.e., a white bear) and to ring a bell when they think about it. Since then, researchers have developed alternative experimental thought suppression tasks based on this procedure (e.g., Grisham & Williams, 2009; Rassin, 2003), but have also designed self-report questionnaires to capture corresponding self-attributed abilities (Feliu-Soler et al., 2019).

Thought control can also be seen as a mental control strategy involving inhibitory control processes that prevent the targeted reactions or thoughts from interfering (Anderson & Huddleston, 2011; Luciano et al., 2005). In 2001, Anderson and Green introduced the TNT paradigm, which provides an alternative way to study the success of active thought suppression in a standardized, task-based setup. The paradigm assumes that thought suppression attempts elicit intentional inhibitory control processes, which then hinder the retrieval of the to-be-suppressed memory. There is also neurocognitive evidence suggesting that thought suppression activates a general executive control mechanism in the lateral prefrontal cortex, which is also involved in stopping motor responses (Anderson & Hanslmayr, 2014). Levy and Anderson (2008) further propose that individual differences in thought control can be explained by pre-existing differences in executive control ability. Although other factors, such as the type of suppression strategy, might influence the success of thought suppression (e.g., Hertel & Calcaterra, 2005), the existence of a general underlying inhibitory mechanism suggests a trait-based view of thought control ability.

1.2 | The measures

1.2.1 | Think/no-think task (TNT)

On the TNT, participants first study and practice a list of word pairs (cue–response), which are evenly split into three groups: baseline, think, and no-think. In the main phase of the experiment, think and no-think cue words are presented and participants are instructed to further maintain and strengthen or actively suppress the studied association, respectively, as typically indicated by green or red font. Baseline word pairs only serve to calculate the TNT effect and thus are not processed any further during this phase. In a final recall test, participants are instructed to again recall the response words for *all* cue words, regardless of the instructions given in the previous phase. The typical pattern of results is that individuals have a better memory for baseline compared to no-think word pairs, which provides evidence for active suppression that goes beyond passive memory loss. This suppression effect is primarily attributed to memory inhibition (Anderson & Green, 2001; Anderson & Huddleston, 2011)—that is, active avoidance of a specific thought results in its decreased mental activation. It therefore supports the idea of an effective inhibitory control mechanism and a general ability to suppress unwanted memories (Anderson & Green, 2001). As the TNT compares experimental conditions within each participant, it can also be used to calculate individual thought control ability measures by subtracting no-think recall from baseline recall (Anderson & Green, 2001) and thus assess individual differences.

As research has progressed, the TNT has been modified in various ways: There are versions differing in the number of stimuli and suppression attempts (e.g., Anderson & Green, 2001; Anderson & Huddleston, 2011; Hanslmayr et al., 2010), stimulus valence (e.g., Joormann et al., 2005; van Schie et al., 2013) and type (e.g., autobiographical memories; Noreen & MacLeod, 2013). However, as all subsequent TNT versions draw on the original TNT (Anderson & Green, 2001), rely on the same underlying mechanism (i.e., inhibitory memory control) and usually yield comparable results concerning the suppression effect (Anderson & Huddleston, 2011; Niessen, Göbel, Siebers, & Schmid, 2020), we assume that they all represent valid procedures for assessing thought control ability.

1.2.2 | Self-report thought control measures

Self-report questionnaires are practical and efficient instruments for assessing thought control. Over the years, several self-report questionnaires with somewhat diverging validity domains have emerged: In 1994, Wegner and Zanakos introduced the White Bear Suppression Inventory (WBSI) to capture dispositional thought suppression tendencies using 15 items. However, subsequent validation studies questioned its unidimensional structure and revealed at least two subscales (experience of unwanted intrusive thoughts, attempts to suppress thoughts). Moreover, based on theoretical analyses and empirical findings, they concluded that the WBSI measures suppression failure and difficulties in thought control rather than dispositional suppression tendencies (Höping & de Jong-Meyer, 2003; Rassin, 2003; Williams et al., 2010).

The 30-item Thought Control Questionnaire (TCQ; Wells & Davies, 1994) assesses the extent to which individuals use various strategies to control unwanted and unpleasant thoughts. It includes the subscales of distraction (replacing the to-be-suppressed thought with other positive thoughts), worry (replacing the thought with other negative thoughts), social control (obtaining support in coping with the thought), punishment (getting angry with oneself for having the thought), and reappraisal (assessing the thought more neutrally; Valdez & Lilly, 2012; Wells & Davies, 1994). As neither of the aforementioned questionnaires distinguishes between thought control tendencies and the ability to control one's thoughts, Luciano et al. (2005) developed the unidimensional Thought Control Ability Questionnaire (TCAQ), which assesses the ability to control one's thoughts and intrusive cognitions using 25 items. Lastly, the Profiles of Everyday Thought Suppression questionnaire (PETS; Ie, 2014) measures the frequency of suppressing different unwanted thoughts in order to identify interindividual differences in everyday thought suppression. The PETS includes 60 items representing 14 factors (e.g., aggression).

In a study assessing psychometric properties of the TCAQ, such as its reliability, Williams et al. (2010) also investigated the relationships among several self-report thought control measures, including the TCAQ, the TCQ, and the WBSI. The TCAQ negatively related to the WBSI and the WBSI subscales, confirming that the WBSI indeed assesses failure to suppress thoughts instead of suppression tendency (Rassin, 2003; Williams et al., 2010). Additionally, the TCAQ negatively related to the TCQ (Williams et al., 2010). However, associations between the TCAQ and TCQ subscales (distraction, worry, social control, punishment, reappraisal) varied in terms of strength and direction. Distraction correlated positively with the TCAQ, worry and punishment correlated negatively with the TCAQ, and social control and reappraisal did not correlate with the TCAQ at all.

1.3 | The relationship between task-based and self-report measures of thought control

The TNT is usually conducted under controlled laboratory conditions and with neutral stimuli. It is therefore shielded from internal (e.g., previous experiences) and external (e.g., distractions) interferences and provides a relatively pure, non-distorted measure of individual thought control. The suppression effect has also been found with negative stimuli (Depue et al., 2006; Küpper et al., 2014) and with autobiographical memories (Noreen & MacLeod, 2013), demonstrating that it is not limited to neutral word pairs. However, it is important to note that some studies also failed to find suppression effects for emotional and especially negative stimuli, suggesting that although controlling negative memories seems possible, it might be harder compared to neutral content (e.g., Marx et al., 2008; Nørby et al., 2010). To further examine the external validity of the TNT, its correlates and consequences outside the laboratory have also been addressed (Göbel & Niessen, 2021; Niessen, Göbel, Siebers, & Schmid, 2020; Streb et al., 2016). For example, findings for the TNT show that higher thought control ability predicts lower intensities of intrusive memories after trauma (Streb et al., 2016) and protects well-being and self-esteem (Göbel & Niessen, 2021). Additionally, psychometric investigations of self-report questionnaires revealed that the TCAQ negatively relates to anxiety, stress, depression, and obsessive–compulsive symptoms; that is, markers of emotional vulnerability (Luciano et al., 2005; Williams et al., 2010). The WBSI-Intrusion scale also moderately relates to measures of anxiety and depression (Höping & de Jong-Meyer, 2003).

Some evidence suggests that people are generally capable of having insight into their thought suppression activities. Friedman and Castel (2011) showed that people were able to make valid predictions about what they would remember and forget in an experimental setup. Grisham and Williams (2009) corroborated this observation by showing that self-reported thought control ability significantly predicts the frequency of intrusions during an experimental thought suppression task, even when controlling for depression and anxiety. Moreover, the negative correlation between involuntary intrusion ratings during the TNT main phase and subsequent forgetting in the final recall test (e.g., Levy & Anderson, 2012; van Schie & Anderson, 2017) indicates that people are able to monitor their thought control performance.

According to Wenzlaff and Wegner (2000), believing in one's mental control abilities is also crucial for successful engagement in thought control activities. This, too, would suggest a positive correlation between task-based and self-report thought control measures. In the related research field of self-control, for example, some studies find a moderate correlation between self-report and behavioral measures (for an overview, see de Ridder et al., 2012).

However, standardized laboratory tasks and self-report questionnaires do not necessarily measure interchangeable psychological constructs. There can be various reasons for a lack of convergent validity, such as low reliability or low validity of one or both measures (Barkley & Fischer, 2011), or differences in the underlying constructs they actually capture (e.g., Sharma et al., 2014).

Noreen and MacLeod (2015) explored inhibitory control using several different memory and behavioral inhibition tasks. They compared the TNT with a variant of the retrieval-practice task (Storm et al., 2006), both of which assess memory inhibition abilities. Although they found standard inhibitory forgetting effects in both tasks, the results were not correlated. There was also no significant relationship between the TNT and behavioral inhibition measures such as the Stroop-Color Naming (adapted from Trenerry et al., 1989) and the Go/No-Go (Schulz et al., 2007) task. At first glance, this pattern of results raises questions about underlying common inhibition processes and therefore also about whether all tasks measure the same core inhibitory control construct.

Wennerhold and Friese (2020) contributed to a parallel debate in the self-control and inhibitory control literature by arguing that there are often fundamental differences between self-report and task-based measures that go beyond psychometric issues—that is, the two measures likely focus on different core aspects of the underlying theoretical construct. More specifically, they consider differences between typical and maximum performance (i.e., self-report measures are often designed to assess typical performance, whereas task-based measures rather assess maximum performance), single measurement versus repeated performance (i.e., self-report measures aggregate different instances of behavior, whereas task-based measures usually rely on a point assessment), and domain-specific behavioral differences (i.e., self-report measures usually capture general, cross-domain tendencies, whereas task-based measures represent narrower, domain-specific constructs) to be responsible for the lack of convergence. We assume that these mechanisms and explanations can also be applied to the assessment of thought control.

Based on these considerations and previous findings, differences between TNT data and thought control questionnaire data seem more likely than perfect agreement. Although there is evidence for a common, underlying inhibitory control mechanism (Anderson et al., 2004; Anderson & Green, 2001; Levy & Anderson, 2008) that should be reflected in both behavioral and self-report measures, differences in the assessed performance level (i.e., the TNT likely assesses maximum performance, whereas self-report questionnaires capture typical performance), differences in performance frequency (i.e., the TNT is usually conducted once, whereas questionnaire items refer to recurrent thought control attempts in everyday life) and domain-specific variations (i.e., self-report questionnaires usually focus on everyday experiences and therefore capture a wider field of inhibition than a standardized laboratory task) cannot be dismissed. Even if psychometric properties were perfect, these fluctuations would account for differences between task-based and self-report measures.

Taken together, we assume a positive association between TNT and self-report thought control measures, because the targeted constructs overlap and are assumed to tap the same underlying inhibitory control mechanism. However, given the different methods of assessment involved, we considered it reasonable to expect only a medium correlation between the two measures.

1.4 | Study aim

The primary focus of this meta-analysis is the association between two different methodological approaches to assess thought control, the experimental TNT task and self-report thought control questionnaires. Moreover, we aim to explain the observed between-study variance by investigating potential moderators such as mean age and gender composition of the samples, publication status, presentation order of the assessment methods, and type of self-report measure used.

We examine age and gender because both variables are fundamental sample characteristics that vary among studies and could therefore have an impact on the studied associations. For example, there is some evidence that cognitive performance and inhibition ability may decline with age (e.g., Anderson et al., 2011; Murman, 2015), that women report engaging in more (unsuccessful) thought suppression activities (Wegner & Zanakos, 1994), and that men report higher thought control abilities (e.g., Peterson et al., 2009; Williams et al., 2010). These findings may be a consequence of more rumination and unsuccessful thought suppression attempts in women, but also reflect gender differences in the social acceptability of experiencing and expressing emotions (Robinson & Clore, 2002). Following the recommendations of Hunter and Schmidt (2004), we also investigate publication status, as published studies often produce higher effect sizes than unpublished material (Simmons et al., 2011). Additionally, researchers have pointed out that preceding measures may impact later measures (e.g., Finley et al., 2015; Schultheiss & Pang, 2007). Thus, considering procedural aspects of the studies by investigating the order of presentation of the two thought control assessment methods appears to be appropriate. Furthermore, as there are different self-rating questionnaires that assess thought control, and some may be more valid than others, we also included the type of measure as a moderator variable.

For transparency and to facilitate future updates of this meta-analytic research, we made our data set and calculations publicly available on the Open Science Framework (https://osf.io/mqzrp).

2 | MATERIALS AND METHOD

2.1 | Inclusion criteria

To be included in the meta-analysis, studies had to meet the following criteria: (1) studies must have reported both the TNT (Anderson & Green, 2001) and one of the following self-report thought control measures: TCQ (Wells & Davies, 1994), TCAQ (Luciano et al., 2005), WBSI (Rassin, 2003), or PETS (Ie, 2014). (2) The sample size and correlation between task-based and self-report measures of thought control must be available. (3) Studies must report which TNT indicator the findings are based on (e.g., baseline minus no-think) and the mean score of the self-report thought control measure. (4) The TNT must be conducted in a laboratory setting and in the native language of participants when employing verbal stimuli.

2.2 | Literature search

We conducted a systematic literature search for English- and German-language studies that assessed thought control with the TNT task and reported a self-report thought control measure. We identified the original article of the TNT paradigm (Anderson &

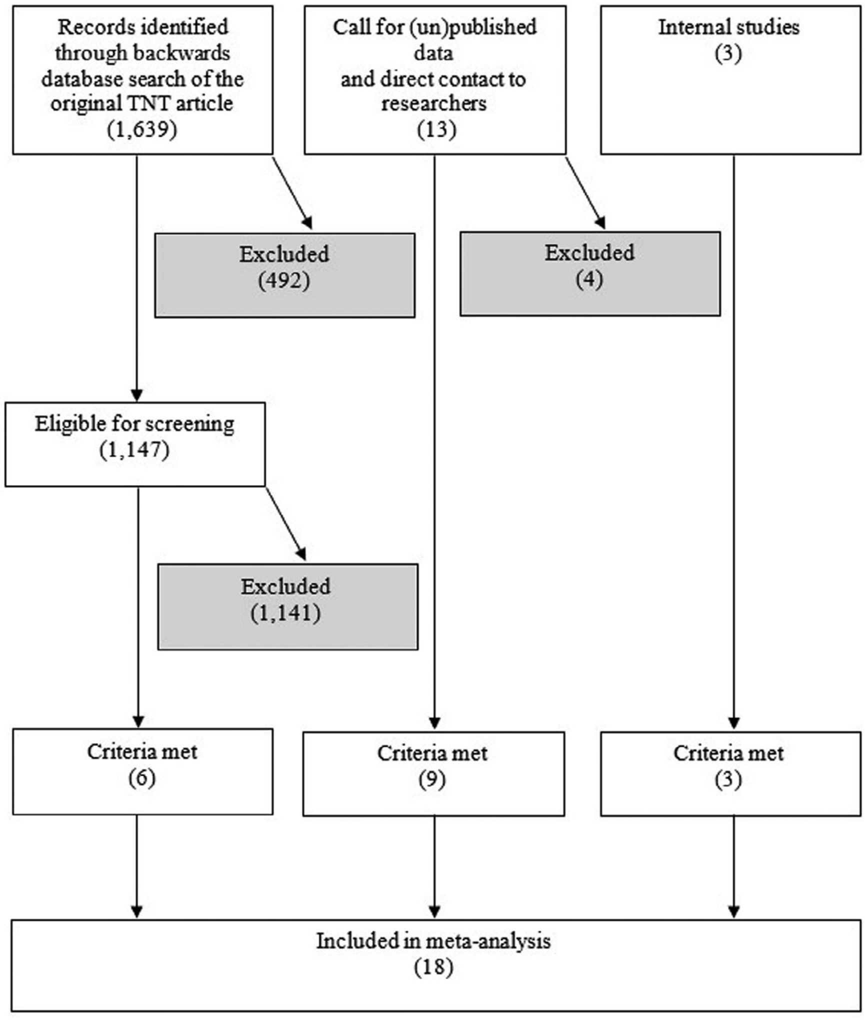

FIGURE 1 Literature search and study selection

Green, 2001) and in turn screened all subsequent articles that cited this original article in the following databases: PsycINFO, Web of Science, and Google Scholar. We then systematically searched for all of the existing self-report measures within the cited articles. Furthermore, we directly contacted researchers who work with the TNT and sent out calls for data through various mailing lists (German Psychological Society, Society for Personality and Social Psychology). The first study search was closed in April 2021 and revealed 15 studies. In a second attempt, we again contacted more than 50 researchers working with the TNT and were able to identify three more studies until January 2022. The final database resulting from these various search processes consists of 18 studies (see Figure 1).

2.3 | Coding of study characteristics

Both first authors independently coded all relevant information from the eligible studies, using a classification system including all variables of interest and a short description for each. Specifically, it included the following variables: (a) TNT indicator,[1] (b) self-report thought control measure, (c) correlation between TNT and self-report thought control measure, (d) sample size, (e) publication status, (f) presentation order of thought control measures, (g) gender composition, (h) mean age, (i) reliability of self-report thought control measure. Data was marked as missing if a variable could not be determined based on the given information and we were unable to obtain it by contacting the respective author(s).

Different self-report questionnaires capture different core aspects of thought control that vary in their content, which may lead to correlations in opposite directions. Thus, we inverted the WBSI and TCQ-Worry correlations for theoretical reasons and due to correlational evidence (see above).

For all of the coded variables, interrater reliability was satisfactory to excellent, with coefficients ranging from .89 to 1.00. All discrepancies were resolved by discussion, which led to complete agreement between the coders.

2.4 | Meta-analytic procedures

Analyses were conducted in R (R Core Team, 2017) using the *meta* package (Schwarzer, 2007). We applied the random effects model instead of the fixed effects model, as the latter posits a single true effect size among all identified studies. The random effect model assumes that the true effect sizes vary across the aggregated studies (Berkey et al., 1995; Borenstein et al., 2014). We began our analysis by investigating the overall relationship between the TNT (Anderson & Green, 2001) and self-report thought control measures with the *metacor* function (Schwarzer, 2007). Metacor automatically z-transforms all study correlations, which are then used as effect sizes. To further address possible heterogeneity issues stemming from differences between self-report questionnaires, we conducted subgroup meta-analyses based on the type of self-report measure.

In addition to estimating the overall correlation, we performed moderation analyses using the *metareg* function (Schwarzer, 2007) for continuous moderators (participant mean age, proportional gender composition) and analysis of variance (*aov* function; Chambers et al., 1992) for categorical moderators (publication status, presentation order of the assessment methods, type of self-report measure). Due to mostly small sample sizes of the included studies, all analyses were estimated using the restricted maximum likelihood method (for further details see Novianti et al., 2014; Viechtbauer, 2005).

To counteract possible bias due to the small sample sizes, we also ran trim-and-fill analysis using the *trimfill* function (Schwarzer, 2007). Trim-and-fill is a method for data augmentation which estimates the number of missing studies because of the inhibition of extreme results (Viechtbauer, 2010). The method's foundation is to first remove (trim) the smaller studies that cause funnel plot asymmetry; second, to estimate the true center of the funnel; and third, to replace the excluded studies and their missing counterparts (fill; Duval & Tweedie, 2000). Trim-and-fill provides an estimate of missing study numbers and adjusted results of the meta-analysis integrating the filled studies.

In addition, we performed outlier analysis to determine whether the pooled effect size estimate is robust—that is, that the effect size found does not depend on extreme effect sizes of particular studies (Viechtbauer & Cheung, 2010). Studies are considered outliers if their confidence interval does not overlap with the confidence interval of the pooled effect (Viechtbauer & Cheung, 2010). To conservatively evaluate the effect size, the calculations are then performed without the identified outliers.

2.5 | Correction for measurement error

Following Hunter and Schmidt (1990), we corrected for imperfect reliability of self-report thought control measures. Specifically, we controlled for measurement error, as observed correlations are usually attenuated. We corrected for measurement error at the individual level by dividing the respective correlation by the product of the square roots of the reliability coefficients of the TNT and the self-report thought control measure (for detailed information see Borenstein et al., 2014; Hunter & Schmidt, 1990). All obtained reliability measures were Cronbach's alpha coefficients. However, seven of the 18 studies did not report sufficient reliability measures. In three of these cases, we used the mean reliability level of the corresponding self-report measure across all included studies with valid information. In the remaining four cases, missing coefficients were replaced with the reliability coefficient from the measure's validation study. For the TNT, no explicit reliability estimates were reported, because the TNT was conducted only once in each of the studies. Generally, it is not reasonable to repeat the paradigm within persons as the suppression effect can only be induced once. Retest coefficients were therefore not available, and internal consistency is not typically calculated for recall item lists. Therefore, we set the TNT's reliability to 1. While this may be an unrealistically high estimate, it also limits the effect size inflation that a correction for reliability <1 would incur.

3 | RESULTS

3.1 | Descriptive statistics

We found 18 studies that satisfied our criteria. Because one of these studies was split into two substudies testing different samples and another in two substudies with different presentation orders, this resulted in $k = 20$ samples ($N = 1194$). Four of the 18 studies have not been published so far. Published studies were dated from 2003 to 2020. Unpublished studies covered a time period ranging from 2016 to 2021. The samples' average age ranged from 18.50 to 35.77 years. Of the 1194 participants 743 were women (62.23%), 450 participants were men (37.69%), and one person nonbinary (0.08%). Table 1 lists sample sizes, self-report thought control measures, and Pearson correlations, which ranged from −.248 to .559.

3.2 | Overall correlation

Across $k = 20$ samples, we found that the pooled association between the TNT and self-report thought control measures was not significant ($Z = 1.72$, $p = .085$). The correlation was positive and small, $r = .07$, 95% CI = −0.01 to 0.15 (see Figure 2). Observed correlations varied considerably across studies, with variation being greater than would be expected from sampling error alone ($Q(19) = 35.45$, $p = .012$; $\tau^2 = .01$, 95% CI: [0.002;0.085]). This implies that a substantial proportion of the observed variance is attributable to between-study differences, justifying the search for potential moderators.

Egger's test indicated no funnel plot asymmetry, $t(18) = 2.09$, $p = .051$, intercept = −0.21, 95% CI = [0.34;−0.08]. Funnel plot asymmetry can originate from publication bias (Ioannidis & Trikalinos, 2007). However, the absence of an effect should be interpreted with caution, as our results suffer from a low power due to the small sample size. The funnel plot did not indicate an absence of correlations close to zero (see Figure 3).

The trim-and-fill analysis imputed three more studies to our database. Subsequent results show a slightly smaller effect of $r = .02$ ($Z = 0.40$, $p = .692$). Together, the observed and filled correlations varied considerably across studies, with more variation than would be

TABLE 1 Overview of studies in descending order of correlation, including sample size, self-report thought control measure and correlation between TNT and self-report thought control measure

Number	Study	*N*	Self-report thought control measure	*r*
1	Waldhauser et al., 2018	24	WBSI	.559[c]
2	Catarino et al., 2015	36	TCAQ	.500
3	Küpper et al., 2014	24	TCAQ	.380[d]
4	Streb et al., 2016	24	TCQ-Worry	.230[c]
5	Hensel et al., 2021	55	TCAQ	.188
6	Legrand et al., 2020	23	TCAQ	.183
7	Peters, 2012 I	89	WBSI-Intrusion	.158[c]
8	Göbel & Niessen, 2019[a] I	66	TCAQ	.105
9	Jäckel & Sievers, 2016	59	WBSI-Intrusion	.050[c]
10	Peters, 2012 II	136	WBSI-Intrusion	.044[c]
11	Wessel et al., 2020	36	WBSI-Intrusion	.038[c]
12	Peters, 2012 III	75	WBSI-Intrusion	.020[c]
13	Marx et al., 2008	48	WBSI-Intrusion	.013[c]
14	Niessen, Göbel, Lang, & Schmid, 2020	158	TCAQ	−.019
15	Peters, 2012 IV	87	WBSI-Intrusion	−.030[c]
16	Göbel & Niessen, 2019[a] II	49	TCAQ	−.063
17	Hertel & Gerstle, 2003	61	WBSI	−.079[c]
18	Eckstein, 2017[b] I	41	WBSI-Intrusion	−.080[c]
19	Eckstein, 2017[b] II	42	WBSI-Intrusion	−.100[c]
20	James, 2020	61	TCAQ short	−.248

Abbreviations: TCAQ, Thought Control Ability Questionnaire; TCQ, Thought Control Questionnaire; WBSI, White Bear Suppression Inventory.
[a]Study was split due to different presentation order.
[b]Study was split due to differences in sample.
[c]Correlations inverted due to the content composition of the questionnaire.
[d]Correlation inverted due differences in the calculation of the TNT indicator.

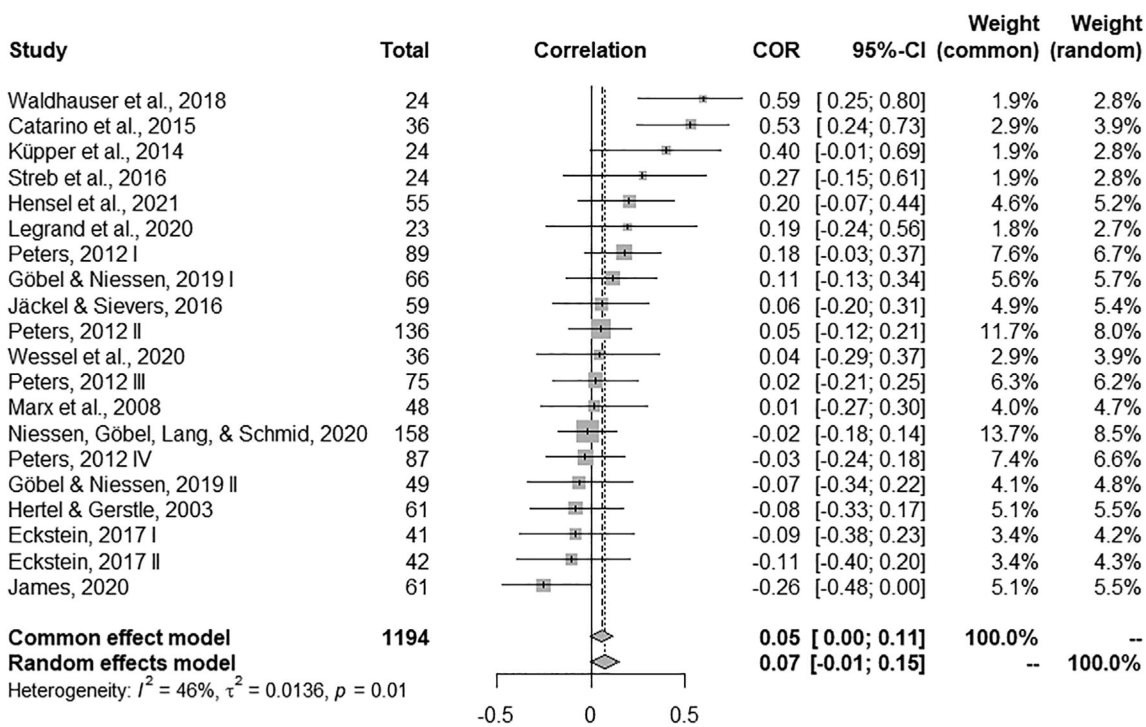

FIGURE 2 Forest plot for the conducted meta-analysis of the correlation between TNT and self-report thought control measure

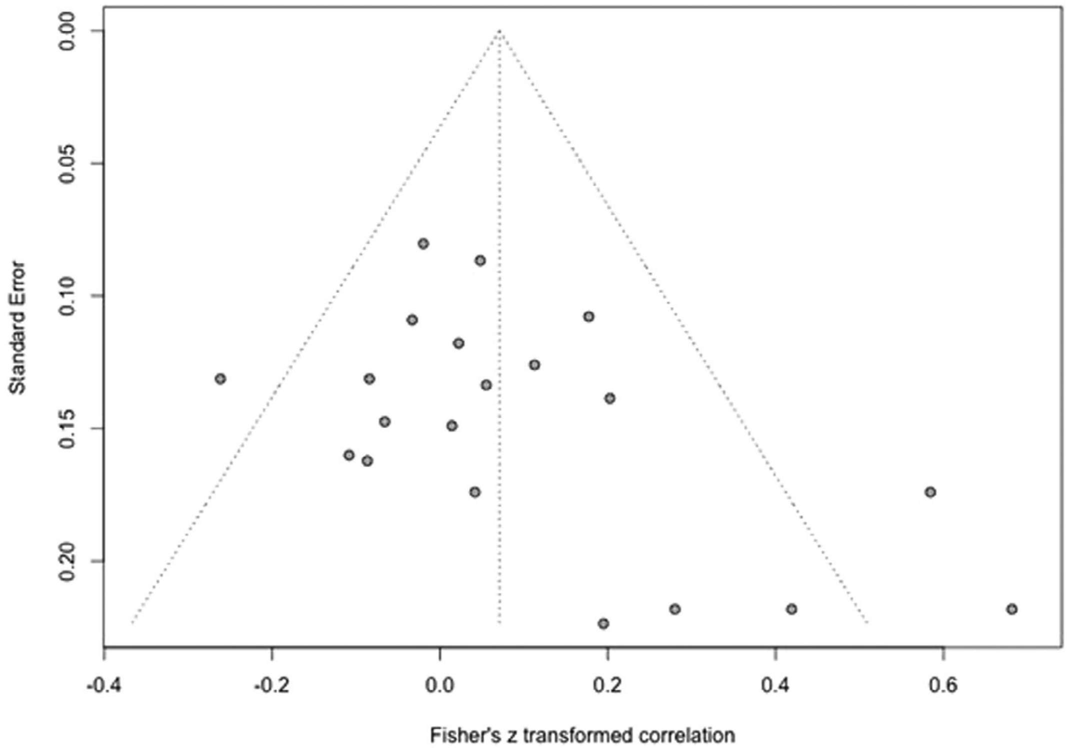

FIGURE 3 Funnel plot of the correlation between TNT and self-report thought control measure

expected from sampling error alone ($Q(22) = 60.12$, $p < .001$; $\tau^2 = .04$, 95% CI: [0.021;0.156]).

Using outlier analysis, we identified two studies that heavily impact the pooled effect size. We removed them and reran the overall analysis. Across $k = 18$ samples, we again found a slightly smaller correlation of $r = .03$ (CI: [−.03;.09]; $Z = 0.86$, $p = .391$; $Q(17) = 17.05$, $p = .451$; $\tau^2 = <.001$, 95% CI: [0.000;0.029]).

3.3 | Moderator analyses

Age, gender, publication status, presentation order, and type of self-report measure did not moderate the association between TNT and self-report thought control measures, suggesting that moderators could not explain a substantial amount of variance (Gender: $Z = -1.34$, $SE = 0.23$, $p = .181$; age: $Z = -0.57$, $SE = 0.01$, $p = .567$; publication status: $f(1) = 0.90$, $p = .358$; presentation order: $f(1) = 0.54$, $p = .471$; type of self-report measure: $f(1) = 0.67$, $p = .426$). However, again, the absence of moderator effects could also be attributed to the low power.

3.4 | Subgroup frequentist meta-analyses for TCAQ and WBSI-intrusion measures

We ran subgroup frequentist meta-analyses based on $k = 8$ samples ($N = 472$) for the TCAQ and $k = 9$ samples ($N = 613$) for the WBSI-Intrusion scale. Results show that the correlation between TCAQ and TNT measures is slightly larger ($r = .12$; 95% CI [−0.06;0.30]; $\tau^2 = .05$) than the correlation between WBSI-Intrusion and TNT ($r = .03$, 95% CI [−0.05; 0.11]; $\tau^2 = .00$). Yet, both are positive, very small, and statistically insignificant. Additionally, there is no statistically significant subgroup effect ($Q(1) = 0.73$, $p = .392$), which indicates that the correlations between the measures do not reliably differ. However, it needs to be mentioned that the respective databases are comparatively small.[2]

3.5 | Advanced analysis: Bayesian meta-analysis

We also performed a robust Bayesian model-averaged meta-analysis using JASP version 0.14.1.0 (JASP Team, 2020). Compared to frequentist meta-analysis, this approach allowed us to robustly estimate the effect size by averaging across varying assumptions regarding effect size, heterogeneity, and publication bias from prior distributions. It also allowed us to test whether the observed effect size is more consistent with a null effect or the assumption of a positive association between TNT and self-report thought control measures. Finally, this approach also makes it easier to update our findings in the future as more evidence regarding the association between the two types of measures accumulates (Gronau et al., 2021).

Following recommendations by Gronau et al. (2021), we used default settings for the heterogeneity prior distribution (inverse gamma distribution [1, 0.15]) and the publication bias weight function (two-sided, cut-point at .05 and parameters $\alpha = [1, 1]$). Based on Cohen (1988) and findings resulting from similar research questions addressing the correlations between task-based and self-report measures (e.g., Hofmann et al., 2005; Yovel et al., 2022), we settled for $r = .30$ (medium effect size) for the alternative hypothesis. Thus, we set the effect size prior distribution to Normal (0.30, 0.15) for the alternative hypothesis of a positive association and to Normal (0.00, 0.15) for the null hypothesis of no association between TNT and self-reported thought control. Results indicated that for the effect size neither the null nor the alternative hypothesis can be ruled out, $BF_{10} = 0.574$, $P(\text{Model}|\text{data}) = .365$. Likewise, there was no conclusive evidence for or against the assumption of heterogeneity, $BF_{rf} = 1.000$, $P(\text{Model}|\text{data}) = .500$, and publication bias, $BF_{\omega\varpi} = 1.000$, $P(\text{Model}|\text{data}) = .333$. The model-averaged effect size was close to our previous estimate, $r = .05$, 95% CI [−0.01;0.13]. The heterogeneity estimate was $\tau = .15$, 95% CI [0.04;0.36]. MCMC diagnostics were sufficient, with all *R*-hat values below 1.02 and all ESS above 691.

The Bayes factor is a relative measure, and in this case, the resulting BF_{10} expresses how much more likely the proposed effect (alternative hypothesis) is compared to the null effect. Unlike frequentist approaches, the Bayes factor does not report a dichotomous decision and instead quantifies the amount of support for each hypothesis under consideration (Hoijtink et al., 2019). Generally, a BF_{10} of about 1 indicates that there is no preference for either the null or the alternative hypothesis, whereas a BF_{10} greater than 1 points toward the alternative model and a BF_{10} smaller than 1 points toward the null model. The BF_{10} of 0.574 we calculated for our data implies that the null model is 1.74 times more likely compared to the alternative model (1/0.574).

3.6 | Subgroup Bayesian meta-analyses for TCAQ and WBSI-intrusion measures

To again address possible heterogeneity issues, we also performed separate Bayesian meta-analyses for the TCAQ ($k = 8$) and the WBSI-Intrusion scale ($k = 9$). Using the same prior setup as described above, results for the TCAQ indicated no evidence for or against the null effect ($BF_{10} = 1.077$, $P(\text{Model}|\text{data}) = .519$). Again, there was no evidence for or against the assumption of heterogeneity $BF_{rf} = 1.000$, $P(\text{Model}|\text{data}) = .500$, and publication bias, $BF_{\omega\varpi} = 1.000$, $P(\text{Model}|\text{data}) = .333$. The model-averaged effect size was slightly smaller than the original estimate, $r = .08$, 95% CI [−0.04;0.21]. The heterogeneity estimate was $\tau = .29$, 95% CI [0.05;0.70].

For the WBSI-Intrusion scale, Bayesian meta-analytic results pointed more toward the null than to the alternative medium-sized effect, $BF_{10} = 0.373$, $P(\text{Model}|\text{data}) = .272$). Again, there was no evidence for or against the assumption of heterogeneity $BF_{rf} = 1.000$, $P(\text{Model}|\text{data}) = .500$, and publication bias, $BF_{\omega\varpi} = 1.00$, $P(\text{Model}|\text{data}) = .333$. The model-averaged effect size was identical to the original estimate, $r = .03$, 95% CI [−0.05;0.11]. The heterogeneity estimate was $\tau = .10$, 95% CI [0.03;0.26].

Whereas the Bayesian meta-analytic results do not allow to draw any conclusion with regard to the relation between TCAQ and TNT, the relation between WBSI-Intrusion and TNT tends to be more consistent with the null than with a medium effect size by a factor of 2.68 (1/0.373).

4 | DISCUSSION

Based on 20 samples, frequentist and robust Bayesian meta-analyses showed that TNT and self-report thought control measures do not substantially correlate with each other. Although results of the overall robust Bayesian meta-analysis point toward a null correlation ($BF_{10} < 1$), the Bayes factor does not provide clear evidence for or against a null correlation. Results of the subgroup analyses indicate that the average correlation between TNT and TCAQ is slightly higher than between TNT and WBSI-Intrusion scale, but effect size estimates are still low and non-significant. Robust Bayesian meta-analyses confirm this pattern of results: They do not provide evidence for either a medium-size or a null effect for the TNT–TCAQ relation, but rather tend toward a null effect for the TNT–WBSI-Intrusion relation. In addition, results of the trim-and-fill and outlier analyses fit into our overall findings which indicate a lack of an overlap between task-based and self-report measures of thought control ability: Both corrected analyses shifted the medium correlation even further toward zero.

Overall, our findings show that there is rather low convergence between task-based and self-report thought control measures. The literature suggests several explanations for this small correlation. A possible explanation is that limited introspective insight into thought control processes prevents persons from appropriately assessing their own abilities. However, a few empirical findings call this account into question by showing that persons do have some insight into their ability to control their thoughts (Friedman & Castel, 2011; Grisham & Williams, 2009).

We rather argue that there are systematic differences between task-based and self-report measures of thought control. The first of these differences refers to the thought control demands imposed: When completing the TNT task, participants are repeatedly instructed and motivated to concentrate, put in effort, and strive for the best outcome. In addition, the TNT is usually conducted in the laboratory, where participants are shielded from external interferences and do not experience distractions or competing tasks. Thus, we suggest that the TNT measures maximum performance. In contrast, when completing questionnaire measures, participants are asked to rate experiences that reflect their general thought control abilities, for example, “I am usually successful when I decide not to think about something” (TCAQ; Luciano et al., 2005). Given the questionnaires' general wording and the diversity of everyday experiences that they cover, this type of measure reflects average performance across frequently unspecified contexts (Wennerhold & Friese, 2020). Therefore, self-report questionnaires assess typical performance in a decontextualized fashion, whereas the TNT captures maximum performance on a specific laboratory task. However, this might not be the only explanation for the comparatively low intercorrelation that we found, as existing meta-analytic findings on the association between typical and maximum performance in other contexts tend to yield moderate rather than low effects (e.g., $r = .42$; Beus & Whitman, 2012).

Second, the TNT measures thought control at one particular point in time and therefore also depends on fluctuating mental states: For example, thought control becomes more difficult with sleep deprivation (Harrington et al., 2021), exhaustion (van Schie & Anderson, 2017), or a state of high anxiety (Marzi et al., 2014). Due to these and other factors, individuals may not be able to fully apply their thought control skills when performing the TNT, which can reduce the correlation with a person's self-report assessment of their general ability to control their thoughts in daily life. Furthermore, thought control might also naturally fluctuate within persons. Parallel debates regarding state and trait measures are held in the field of self-control (e.g., de Ridder et al., 2018; Fleeson, 2004; Wennerhold & Friese, 2020), for example.

Third, the two assessment methods differ in the specific thought control domains and strategies they primarily measure. The TNT is usually conducted under controlled laboratory conditions, often with explicit suppression instructions, and thus assesses a very specific and small range of thought control (i.e., inhibition of word associations). In contrast, questionnaire data are not limited to a single suppression strategy and usually cover more than one domain (e.g., thoughts, emotions). Overall, we argue that the TNT measures a narrower and the questionnaires a wider array of thought control activities and abilities.

In addition to our main results, the moderator analysis revealed that none of the identified moderator variables significantly predicted the observed between-study variance. Thus, neither age, gender, publication status, presentation order, nor type of self-report measure explained a substantial amount of variance. It is notable that effect size differences were at least somewhat independent of basic sample characteristics (mean age, gender composition), highlighting the robustness of our findings. However, as our samples' average age range is relatively narrow, it might be that age differences within this range are too small to impact the overall correlation.

Several studies report individual differences in self-reported thought suppression: Women exhibit significantly higher (unsuccessful) thought suppression scores than men in the WBSI (Wegner & Zanakos, 1994), and men usually indicate having higher thought control abilities in the TCAQ (e.g., Peterson et al., 2009). However, to our knowledge there is no evidence for gender differences in the TNT. Differences in the questionnaire measures might therefore be of minor importance for explaining the variance between task-based and self-report measures of thought control.

Additionally, it is not surprising that we did not find publication status to moderate the overall correlation. First, almost one-third of the underlying studies comprising this meta-analysis remained unpublished; second, none of the underlying studies focused primarily on the association between task-based and self-report thought control measures.

We found that the presentation order of task-based and self-report thought control measure did not affect correlations between them. Although there is evidence that preceding measures affect later ones (e.g., Finley et al., 2015), such that if self-report measures are administered after the TNT, participants might reflect on their ability to control their thoughts and thus assess their ability more appropriately, our meta-analytic data show that these procedural aspects had no effect on the overall correlation. Nevertheless, it might be interesting to further explore how (task-based) thought control experiences can affect more global self-evaluations. Incorporating intrusion rating scales directly after suppression attempts (e.g., Levy & Anderson, 2012; van Schie & Anderson, 2017) might be a promising approach to connect task-based and self-reported thought control success.

Lastly, we found that the type of the questionnaire did not moderate between-measure associations. To some extent, this finding is unexpected, as the underlying questionnaires focus on somewhat different content domains (e.g., dispositional tendencies regarding suppression attempts, use of strategies to control unwanted thoughts). However, it could also be explained by the rather low number of studies as the analyses might not have had enough statistical power to detect systematic differences (cf. Hempel et al., 2013).

4.1 | Implications

One possible implication of our findings concerns the validity of various thought control measures. The validity of the TNT is rooted in the nature of the task itself, which documents how well individuals succeed in suppressing an association once they intend to do so. But what evidence is there for the validity of self-report thought control measures? What are the processes that make a person attribute more (or less) thought control capacity to herself or himself? This is a largely unresolved question, because validation of such measures typically follows a nomological network approach centered on convergent, discriminant, and criterion validity relative to other self-report measures—an approach whose usefulness for establishing a measure's validity has been questioned (see Borsboom et al., 2009). It typically does not attempt to examine the effects of variation in actual thought control processes on scores of self-report measures (cf. Bornstein, 2011; Borsboom et al., 2004) or entail a substantive analysis of which types of memory a person draws on to arrive at a response to items such as "It is very easy for me to stop having certain thoughts" (Luciano et al., 2005). In the absence of such substantial evidence for the validity of self-report thought control measures, perhaps higher correlations with procedural measures like the one we meta-analytically examined should not be expected.

Overall, our findings highlight the importance of theoretical considerations regarding the measurement of thought control.

4.2 | Strengths and limitations

Our research has several strengths. By analyzing and discussing correlations between the behavioral TNT task and self-report questionnaire data, we provide a data-based and theoretically grounded overview of their relations. We combined frequentist with Bayesian meta-analysis to distinguish absence of evidence from evidence of absence. Additionally, we identified a substantial amount of unpublished literature and corrected for imperfect reliability in self-report measures.

Despite the aforementioned strengths, our study also has limitations. Although we identified several unpublished studies, 20 samples represent a rather small dataset. We assume that in many studies employing the TNT, self-reported thought control data were collected as a control variable, but not always reported in the resulting publications. Although we supplemented missing data through individual queries, the presented study situation might be incomplete. Moreover, analyses of heterogeneity, publication bias analyses and moderator analyses must be interpreted with caution, as possible effects are difficult to detect with small data bases like ours. However, we argue that our meta-analysis still provides a valuable contribution to research, as it approaches the topic under study from a top-down perspective and systematically summarizes current empirical findings. Thus, the present meta-analysis represents a snapshot of current nascent research, not a final evaluation of a mature field (see Lakens et al., 2016). To future-proof our analyses, we have uploaded our database and scripts to the Open Science Framework. This allows our data to be viewed, extended, and used for future publications.

Furthermore, there is variance among the reported TNT tasks, as they differ in aspects such as stimulus type (e.g., word-based vs. picture-based) or suppression instructions (e.g., specific vs. unspecific). There is meta-analytical evidence that the type of suppression strategy employed might impact thought control success (Stramaccia et al., 2021). Hence, we cannot rule out effects of variations of TNT task design in the meta-analyzed studies.

Likewise, we included different self-report thought control questionnaires in our meta-analysis that differ in their interpretation. As described above, Williams et al. (2010) found that the TCQ, TCAQ, WBSI and their respective subscales show complex associations with each other. Additionally, findings by Streb et al. (2016) indicate that the subscales of the TCQ differentially relate to the TNT. This makes it particularly difficult to assess whether all of the included self-report measures capture the same construct and should be treated equally. We tried to limit this problem by considering their focus, by inverting several correlations, and by conducting subgroup analyses. Still, there might be a substantial amount of variance remaining, at least in the overall calculations. Additionally, one study deviated from typical TNT scoring and reported only the no-think value, which may have further reduced comparability across studies, although we inverted the respective correlation.

5 | CONCLUSION

In sum, the present findings show that task-based and self-report thought control measures do not substantially overlap, which is probably due to differences in the respective underlying psychological processes. Whereas the laboratory task aims at measuring memory inhibition as a maximum performance outcome under specific controlled conditions, self-report questionnaires try to cover a wider range of inhibition strategies in everyday life. Our findings encourage future research on differences between these measures, with a focus on what each type actually assesses and when and why they overlap or fail to overlap.

CONFLICT OF INTEREST

All authors confirm that they do not have any kind of competing interests.

DATA AVAILABILITY STATEMENT

Data are available under the following link: https://osf.io/mqzrp/.

ORCID

Kyra Göbel https://orcid.org/0000-0003-3850-9761
Lisa Hensel https://orcid.org/0000-0001-6341-265X
Oliver C. Schultheiss https://orcid.org/0000-0002-7892-5228
Cornelia Niessen https://orcid.org/0000-0003-1994-3677

ENDNOTES

[1] Individual thought control ability is usually measured by subtracting no-think recall from baseline recall, with higher values reflecting higher thought control ability. However, for one study we only obtained no-think recall data with higher values indicating lower thought control ability. We therefore inverted the correlation obtained in this study before conducting our analyses.

[2] Detailed analyses and plots for the two subgroup analyses can be found in the Data S1.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, *410*(6826), 366–369. https://doi.org/10.1038/35066572

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, *18*(6), 279–292. https://doi.org/10.1016/j.tics.2014.03.002

Anderson, M. C., & Huddleston, E. (2011). Towards a cognitive and neurobiological model of motivated forgetting. In R. F. Belli (Ed.), *True and false recovered memories: Toward a reconciliation of the debate* (pp. 53–120). Springer. https://doi.org/10.1007/978-1-4614-1195-6_3

Anderson, M. C., & Weaver, C. (2009). Inhibitory control over action and memory. *The New Encyclopedia of Neuroscience*, *5*, 153–163. https://doi.org/10.1016/b978-008045046-9.00421-6

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. E. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*(5655), 232–235. https://doi.org/10.1126/science.1089504

Anderson, M. C., Reinholz, J., Kuhl, B. A., & Mayr, U. (2011). Intentional suppression of unwanted memories grows more difficult as we age. *Psychology and Aging*, *26*(2), 397–405. https://doi.org/10.1037/a0022505

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests.

Developmental Neuropsychology, *36*(2), 137–161. https://doi.org/10.1080/87565641.2010.549877

Berkey, C. S., Hoaglin, D. C., Mosteller, F., & Colditz, G. A. (1995). A random-effects regression model for meta-analysis. *Statistics in Medicine*, *14*(4), 395–411. https://doi.org/10.1002/sim.4780140406

Beus, J. M., & Whitman, D. S. (2012). The relationship between typical and maximum performance: A meta-analytic examination. *Human Performance*, *25*(5), 355–376. https://doi.org/10.1080/08959285.2012.721831

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2014). *Comprehensive meta-analysis, version 3*. Biostat.

Bornstein, R. F. (2011). Toward a process-focused model of test score validity: Improving psychological assessment in science and practice. *Psychological Assessment*, *23*(2), 532–544. https://doi.org/10.1037/a0022402

Borsboom, D., Cramer, A. O. J., Kievit, R. A., Scholten, A. Z., & Franić, S. (2009). The end of construct validity. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 135–170). IAP Information Age Publishing. https://psycnet.apa.org/record/2009-23060-007

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

*Catarino, A., Küpper, C. S., Werner-Seidler, A., Dalgleish, T., & Anderson, M. C. (2015). Failing to forget: Inhibitory-control deficits compromise memory suppression in posttraumatic stress disorder. *Psychological Science*, *26*(5), 604–616. https://doi.org/10.1177/0956797615569889

Chambers, J. M., Freeny, A., & Heiberger, R. M. (1992). Chapter 5: Analysis of variance; designed experiments. In J. M. Chambers & T. J. Hastie (Eds.), *Statistical models in S*. Wadsworth & Brooks/Cole.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

de Ridder, D., Kroese, F., & Gillebaart, M. (2018). Whatever happened to self-control? A proposal for integrating notions from trait self-control studies into state self-control research. *Motivation Science*, *4*(1), 39–49. https://doi.org/10.1037/mot0000062

de Ridder, D. T., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors. *Personality and Social Psychology Review*, *16*(1), 76–99. https://doi.org/10.1177/1088868311418749

Depue, B. E., Banich, M. T., & Curran, T. (2006). Suppression of emotional and nonemotional content in memory: Effects of repetition on cognitive control. *Psychological Science*, *17*(5), 441–447. https://doi.org/10.1111/j.1467-9280.2006.01725.x

Dreves, P. A., Blackhart, G. C., & McBee, M. T. (2020). Do behavioral measures of self-control assess construct-level variance? *Journal of Research in Personality*, *88*, 104000. https://doi.org/10.1016/j.jrp.2020.104000

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

*Eckstein, M. (2017). *Intentional forgetting: How the level of difficulty in the association establishment process of word pairs influences the think/no-think effect* [Unpublished thesis]. University of Bamberg.

Engen, H. G., & Anderson, M. C. (2018). Memory control: A fundamental mechanism of emotion regulation. *Trends in Cognitive Sciences*, *22*(11), 982–995. https://doi.org/10.1016/j.tics.2018.07.015

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, *116*(12), 5472–5477. https://doi.org/10.1073/pnas.1818430116

Erdelyi, M. H. (2006). The unified theory of repression. *Behavioral and Brain Sciences*, *29*(5), 499–551. https://doi.org/10.1017/s0140525x06009113

Fawcett, J. M., & Hulbert, J. C. (2020). The many faces of forgetting: Toward a constructive view of forgetting in everyday life. *Journal of Applied Research in Memory and Cognition*, *9*(1), 1–18. https://doi.org/10.1016/j.jarmac.2019.11.002

Feliu-Soler, A., Pérez-Aranda, A., Montero-Marín, J., Herrera-Mercadal, P., Andrés-Rodríguez, L., Angarita-Osorio, N., Williams, A. D., & Luciano, J. V. (2019). Fifteen years controlling unwanted thoughts: A systematic review of the Thought Control Ability Questionnaire (TCAQ). *Frontiers in Psychology*, *10*, 1446. https://doi.org/10.3389/fpsyg.2019.01446

Finley, A. J., Tang, D., & Schmeichel, B. J. (2015). Revisiting the relationship between individual differences in analytic thinking and religious belief: Evidence that measurement order moderates their inverse correlation. *PLoS One*, *10*(9), e0138922. https://doi.org/10.1371/journal.pone.0138922

Fleeson, W. (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science*, *13*(2), 83–87. https://doi.org/10.1111/j.0963-7214.2004.00280.x

Freyd, J. J. (1996). *Betrayal trauma: The logic of forgetting childhood abuse*. Harvard University Press.

Friedman, M. C., & Castel, A. D. (2011). Are we aware of our ability to forget? Metacognitive predictions of directed forgetting. *Memory & Cognition*, *39*(8), 1448–1456. https://doi.org/10.3758/s13421-011-0115-y

Gagnepain, P., Hulbert, J., & Anderson, M. C. (2017). Parallel regulation of memory and emotion supports the suppression of intrusive memories. *The Journal of Neuroscience*, *37*(27), 6423–6441. https://doi.org/10.1523/jneurosci.2732-16.2017

Geraerts, E., Hauer, B. J. A., & Wessel, I. (2010). Effects of suppressing negative memories on intrusions and autobiographical memory specificity. *Applied Cognitive Psychology*, *24*(3), 387–398. https://doi.org/10.1002/acp.1684

*Göbel, K., & Niessen, C. (2019). *Experimental dataset and personnel survey* [Unpublished raw data]. Friedrich-Alexander-Universität Erlangen-Nürnberg.

Göbel, K., & Niessen, C. (2021). Thought control in daily working life: How the ability to stop thoughts protects self-esteem. *Applied Cognitive Psychology*, *35*(4), 1011–1022. https://doi.org/10.1002/acp.3830

Grisham, J. R., & Williams, A. D. (2009). Cognitive control of obsessional thoughts. *Behaviour Research and Therapy*, *47*(5), 395–402. https://doi.org/10.1016/j.brat.2009.01.014

Gronau, Q. F., Heck, D. W., Berkhout, S. W., Haaf, J. M., & Wagenmakers, E.-J. (2021). A primer on Bayesian model-averaged meta-analysis. *Advances in Methods and Practices in Psychological Science*. *4*(3):1–19. https://doi.org/10.1177/25152459211031256

Hanslmayr, S., Leipold, P., & Bäuml, K.-H. (2010). Anticipation boosts forgetting of voluntarily suppressed memories. *Memory*, *18*(3), 252–257. https://doi.org/10.1080/09658210903476548

Harrington, M. O., Ashton, J. E., Sankarasubramanian, S., Anderson, M. C., & Cairney, S. A. (2021). Losing control: Sleep deprivation impairs the suppression of unwanted thoughts. *Clinical Psychological Science*, *9*(1), 97–113. https://doi.org/10.1177/2167702620951511

Hempel, S., Miles, J. N., Booth, M. J., Wang, Z., Morton, S. C., & Shekelle, P. G. (2013). Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews*, *2*, 107. https://doi.org/10.1186/2046-4053-2-107

*Hensel, L., Rohleder, N., & Niessen, C. (2021). *Effects of psychosocial stress on prosociality: The moderating role of current life stress and thought control* [Manuscript submitted for publication]. Friedrich-Alexander-Universität Erlangen-Nürnberg.

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, *12*(3), 484–489. https://doi.org/10.3758/bf03193792

*Hertel, P. T., & Gerstle, M. (2003). Depressive deficits in forgetting. *Psychological Science*, *14*(6), 573–578. https://doi.org/10.1046/j.0956-7976.2003.psci_1467.x

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality & Social Psychology Bulletin*, *31*(10), 1369–1385. https://doi.org/10.1177/0146167205275613

Hoijtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*, *24*(5), 539–556. https://doi.org/10.1037/met0000201

Höping, W., & de Jong-Meyer, R. (2003). Differentiating unwanted intrusive thoughts from thought suppression: What does the white bear suppression inventory measure? *Personality and Individual Differences*, *34*(6), 1049–1055. https://doi.org/10.1016/S0191-8869(02)00089-2

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publications, Inc.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publications, Inc.

Ie, A. Y. L. (2014). *Profiles of everyday thought suppression* (Doctoral dissertation). Harvard University.

Ioannidis, J. P., & Trikalinos, T. A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *Canadian Medical Association*, *176*(8), 1091–1096. https://doi.org/10.1503/cmaj.060410

*Jäckel, E., & Sievers, K. (2016). *The influence of the direction of word pair association on the think/no-think effect* [Unpublished thesis]. University of Bamberg.

*James, R.-M. B. M. (2020). *The relationship between memory control, emotion dysregulation and subclinical borderline symptoms* [Master's thesis, Lund University]. https://lup.lub.lu.se/student-papers/search/publication/9030462

JASP Team. (2020). JASP (Version 0.14.1) [Computer software]. Retrieved from https://jasp-stats.org/pronamic_faq_category/other/

Joormann, J., Hertel, P. T., Brozovich, F., & Gotlib, I. H. (2005). Remembering the good, forgetting the bad: Intentional forgetting of emotional material in depression. *Journal of Abnormal Psychology*, *114*(4), 640–648. https://doi.org/10.1037/0021-843x.114.4.640

*Küpper, C. S., Benoit, R. G., Dalgleish, T., & Anderson, M. C. (2014). Direct suppression as a mechanism for controlling unpleasant memories in daily life. *Journal of Experimental Psychology: General*, *143*(4), 1443–1449. https://doi.org/10.1037/a0036518

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC. Psychology*, *4*, 24. https://doi.org/10.1186/s40359-016-01

*Legrand, N., Etard, O., Vandevelde, A., Pierre, M., Viader, F., Clochon, P., Doidy, F., Peschanski, D., Eustache, F., & Gagnepain, P. (2020). Long-term modulation of cardiac activity induced by inhibitory control over emotional memories. *Scientific Reports*, *10*(1), 1–19. https://doi.org/10.1038/s41598-020-71858-2

Levy, B. J., & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica*, *127*(3), 623–635. https://doi.org/10.1016/j.actpsy.2007.12.004

Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *Journal of Neuroscience*, *32*(47), 16785–16794. https://doi.org/10.1523/jneurosci.2640-12.2012

Luciano, J. V., Algarabel, S., Tomás, J. M., & Martínez, J. L. (2005). Development and validation of the thought control ability questionnaire. *Personality and Individual Differences*, *38*(5), 997–1008. https://doi.org/10.1016/j.paid.2004.06.020

*Marx, B. P., Marshall, P. J., & Castro, F. (2008). The moderating effects of stimulus valence and arousal on memory suppression. *Emotion*, *8*(2), 199–207. https://doi.org/10.1037/1528-3542.8.2.199

Marzi, T., Regina, A., & Righi, S. (2014). Emotions shape memory suppression in trait anxiety. *Frontiers in Psychology*, *4*, 1001. https://doi.org/10.3389/fpsyg.2013.01001

Murman, D. L. (2015). The impact of age on cognition. *Seminars in Hearing*, *36*(3), 111–121. https://doi.org/10.1055/s-0035-1555115

*Niessen, C., Göbel, K., Lang, J., & Schmid, U. (2020). Stop thinking: An experience sampling study on suppressing distractive thoughts at work. *Frontiers in Psychology*, *11*, Article 1616. https://doi.org/10.3389/fpsyg.2020.01616

Niessen, C., Göbel, K., Siebers, M., & Schmid, U. (2020). Time to forget: Intentional forgetting in the digital world of work. *Zeitschrift für Arbeits- und Organisationspsychologie A&O*, *64*(1), 30–45. https://doi.org/10.1026/0932-4089/a000308

Nørby, S. (2015). Why forget? On the adaptive value of memory loss. *Perspectives on Psychological Science*, *10*(5), 551–578. https://doi.org/10.1177/1745691615596787

Nørby, S. (2017). Forgetting and emotion regulation in mental health, anxiety and depression. *Memory*, *26*(3), 342–363. https://doi.org/10.1080/09658211.2017.1346130

Nørby, S., Lange, M., & Larsen, A. (2010). Forgetting to forget: On the duration of voluntary suppression of neutral and emotional memories. *Acta Psychologica*, *133*(1), 73–80. https://doi.org/10.1016/j.actpsy.2009.10.002

Noreen, S., & MacLeod, M. D. (2013). It's all in the detail: Intentional forgetting of autobiographical memories using the autobiographical think/no-think task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 375–393. https://doi.org/10.1037/a0028888

Noreen, S., & MacLeod, M. D. (2015). What do we really know about cognitive inhibition? Task demands and inhibitory effects across a range of memory and behavioural tasks. *PLoS One*, *10*(8), e0134951. https://doi.org/10.1371/journal.pone.0134951

Novianti, P. W., Roes, K. C., & van der Tweel, I. (2014). Estimation of between-trial variance in sequential meta-analyses: A simulation study. *Contemporary Clinical Trials*, *37*(1), 129–138. https://doi.org/10.1016/j.cct.2013.11.012

*Peters, J. H. (2012). *Angstbewältigung und Erinnerung: Eine funktionale Sicht des Gedächtnisses [Coping with anxiety and memory: A functional view of memory]*. Springer. https://doi.org/10.1007/978-3-531-19526-1_5

Peterson, R. D., Klein, J., Donnelly, R., & Renk, K. (2009). Predicting psychological symptoms: The role of perceived thought control ability. *Cognitive Behaviour Therapy*, *38*(1), 16–28. https://doi.org/10.1080/16506070802561215

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Racsmány, M., Conway, M. A., Keresztes, A., & Krajcsi, A. (2011). Inhibition and interference in the think/no-think task. *Memory & Cognition*, *40*(2), 168–176. https://doi.org/10.3758/s13421-011-0144-6

Rassin, E. (2003). The White Bear Suppression Inventory (WBSI) focuses on failing suppression attempts. *European Journal of Personality*, *17*(4), 285–298. 10.1002%2Fper.478

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 322–344). Guilford Press.

Schulz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., & Halperin, J. M. (2007). Does the emotional go/no-go task really measure behavioral inhibition? Convergence with measures on a non-emotional analog.

Archives of Clinical Neuropsychology, 22(2), 151–160. https://doi.org/10.1016/j.acn.2006.12.001

Schwarzer, G. (2007). meta: An R package for meta-analysis. *R News*, *7*(3), 40–45.

Sedikides, C., & Green, J. D. (2009). Memory as a self-protective mechanism. *Social and Personality Psychology Compass*, *3*(6), 1055–1068. https://doi.org/10.1111/j.1751-9004.2009.00220.x

Sharma, L., Markon, K. E., & Clark, L. A. (2014). Toward a theory of distinct types of "impulsive" behaviors: A meta-analysis of self-report and behavioral measures. *Psychological Bulletin*, *140*(2), 374–408. https://doi.org/10.1037/a0034418

Simmons, J., Nelson, L., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review*, *13*(6), 1023–1027. https://doi.org/10.3758/BF03213919

Stramaccia, D. F., Meyer, A. K., Rischer, K. M., Fawcett, J. M., & Benoit, R. G. (2021). Memory suppression and its deficiency in psychological disorders: A focused meta-analysis. *Journal of Experimental Psychology: General*, *150*(5), 828–850. https://doi.org/10.1037/xge0000971

*Streb, M., Mecklinger, A., Anderson, M. C., Lass-Hennemann, J., & Michael, T. (2016). Memory control ability modulates intrusive memories after analogue trauma. *Journal of Affective Disorders*, *192*, 134–142. https://doi.org/10.1016/j.jad.2015.12.032

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *Stroop neuropsychological screening test*. Psychological Assessment Resources.

Valdez, C. E., & Lilly, M. M. (2012). Thought control: Is it ability, strategies, or both that predicts posttraumatic symptomatology in victims of interpersonal trauma? *Journal of Psychopathology and Behavioral Assessment*, *34*(4), 531–541. https://doi.org/10.1007/s10862-012-9300-y

Van Schie, K., & Anderson, M. C. (2017). Successfully controlling intrusive memories is harder when control must be sustained. *Memory*, *25*(9), 1201–1216. https://doi.org/10.1080/09658211.2017.1282518

Van Schie, K., Geraerts, E., & Anderson, M. C. (2013). Emotional and nonemotional memories are suppressible under direct suppression instructions. *Cognition & Emotion*, *27*, 1122–1131. https://doi.org/10.1080/02699931.2013.765387

van Vugt, M. K., & van der Velde, M. (2018). How does rumination impact cognition? A first mechanistic model. *Topics in Cognitive Science*, *10*(1), 175–191. https://doi.org/10.1111/tops.12318

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3). 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Von Gunten, C. D., Bartholow, B. D., & Martins, J. S. (2020). Inhibition tasks are not associated with a variety of behaviours in college students. *European Journal of Personality*, *34*(3), 412–430. https://doi.org/10.1002/per.2250

*Waldhauser, G. T., Dahl, M. J., Ruf-Leuschner, M., Müller-Bamouh, V., Schauer, M., Axmacher, N., Elbert, T., & Hanslmayr, S. (2018). The neural dynamics of deficient memory control in heavily traumatized refugees. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-31400-x

Waldum, E. R., & Sahakyan, L. (2012). Putting congeniality effects into context: Investigating the role of context in attitude memory using multiple paradigms. *Journal of Memory and Language*, *66*(4), 717–730. https://doi.org/10.1016/j.jml.2011.12.013

Wegner, D. M., & Zanakos, S. (1994). Chronic thought suppression. *Journal of Personality*, *62*(4), 615–640. https://doi.org/10.1111/j.1467-6494.1994.tb00311.x

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*(1), 5–13. https://doi.org/10.1037/0022-3514.53.1.5

Wells, A., & Davies, M. I. (1994). The thought control questionnaire: A measure of individual differences in the control of unwanted thoughts. *Behaviour Research and Therapy*, *32*(8), 871–878. https://doi.org/10.1016/0005-7967(94)90168-6

Wennerhold, L., & Friese, M. (2020). Why self-report measures of self-control and inhibition tasks do not substantially correlate. *Collabra: Psychology*, *6*(1), 9. https://doi.org/10.1525/collabra.276

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology*, *51*(1), 59–91. https://doi.org/10.1146/annurev.psych.51.1.59

*Wessel, I., Albers, C. J., Zandstra, A. R. E., & Heininga, V. E. (2020). A multiverse analysis of early attempts to replicate memory suppression with the think/no-think task. *Memory*, *28*(7), 870–887. https://doi.org/10.1080/09658211.2020.1797095

Williams, A. D., Moulds, M. L., Grisham, J. R., Gay, P., Lang, T., Kandris, E., Werner-Seidler, A., & Yap, C. (2010). A psychometric evaluation of the thought control ability questionnaire (TCAQ) and the prediction of cognitive control. *Journal of Psychopathology and Behavioral Assessment*, *32*(3), 397–405. https://doi.org/10.1007/s10862-009-9171-z

Yovel, I., Aviram, G., Kahana, N., & Katz, B. A. (2022). Testing a new indirect measure of general self-worth: The self-esteem questionnaire-based implicit association test. *The British Journal of Social Psychology*, *61*(1), 100–120. https://doi.org/10.1111/bjso.12472

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Göbel, K., Hensel, L., Schultheiss, O. C., & Niessen, C. (2022). Meta-analytic evidence shows no relationship between task-based and self-report measures of thought control. *Applied Cognitive Psychology*, 1–14. https://doi.org/10.1002/acp.3952