INTERNATIONAL REVIEW OF SOCIAL PSYCHOLOGY

Brohmer, H., et al. (2021). Do Behavioral Observations Make People Catch the Goal? A Meta-Analysis on Goal Contagion. *International Review of Social Psychology*, 34(1): 3, 1–15. DOI: https://doi.org/10.5334/irsp.428

RESEARCH ARTICLE

Do Behavioral Observations Make People Catch the Goal? A Meta-Analysis on Goal Contagion

Hilmar Brohmer*, Lisa V. Eckerstorfer†, Robbie C. M. van Aert‡ and Katja Corcoran*,§

Goal contagion is a social-cognitive approach to understanding how other people's behavior influences one's goal pursuit: An observation of goal-directed behavior leads to an automatic inference and activation of the goal before it can be adopted and pursued thereafter by the observer. We conducted a meta-analysis focusing on experimental studies with a goal condition, depicting goal-directed behavior and a control condition. We searched four databases (PsychInfo, Web of Science, ScienceDirect, and JSTOR) and the citing literature on Google Scholar, and eventually included $e = 48$ effects from published studies, unpublished studies and registered reports based on 4751 participants. The meta-analytic summary effect was small – $g = 0.30$, *95%CI* [0.21; 0.40], $\tau^2 = 0.05$, *95%CI* [0.03, 0.13] – implying that goal contagion might occur for some people, compared to when this goal is not perceived in behavior. However, the original effect seemed to be biased through the current publication system. As shown by several publication-bias tests, the effect could rather be half the size, for example, selection model: $g = 0.15$, *95%CI* [–0.02; 0.32]. Further, we could not detect any potential moderator (such as the presentation of the manipulation and the contrast of the control condition). We suggest that future research on goal contagion makes use of open science practices to advance research in this domain.

Keywords: goal contagion; goal pursuit; automatic inference; meta-analysis; publication bias

Introduction

Goals are essential to many species' existence. Understood as representations of desired states that are attainable through action (Kruglanski & Kopetz, 2009), they determine the upcoming steps of living beings as they strive to achieve something, be it nourishment, sex, company, or a place to hide. Humans are no exception, but beyond these aforementioned basic needs (Jolly, 1976; Maslow, 1943), our complex cognitive structure allows us to incorporate many goals in our daily life, like 'catching the train to the workplace,' 'going grocery shopping,' or 'going jogging after work.' Moreover, we can plan ahead so that our daily goals serve as a means for higher-order goals like 'earning money,' 'being healthy,' or 'keeping in shape,' which usually serve self-regulatory purposes in the long run (Carver & Scheier, 1981, 2012). However, it is also due to our cognitive architecture that we often monitor other people's behavior as this can contain important information (e.g., My colleagues bring their home-cooked meals for lunch – I also want to live healthily!). Consequently, observing other people's goal-directed behavior might affect our own goals.

We can take others as a source to adjust our goals. A social-cognitive approach to this phenomenon is provided by the theory of goal contagion (henceforth, GC), which was introduced by Aarts and colleagues (Aarts, Dijksterhuis, & Dik, 2008; Aarts, Gollwitzer, & Hassin, 2004) more than a decade ago. As much research has been conducted on this topic, we intend to summarize the evidence for GC in a meta-analysis and search for potential moderators. To do so, we will first provide a clear description of GC, based on theoretical introductions and empirical studies in the literature. We need to overcome vague concepts and paradigms to formulate precise guidelines for the later extraction of studies and effects in the meta-analysis.

The process of and evidence for goal contagion

The original authors of the first GC studies based their theoretical approach on the spontaneous causal inferences framework (Hassin, Bargh, & Uleman, 2002), which posits that people make spontaneous inferences about traits. For instance, the observation of a person offering to help someone else likely leads to the inference that this person

* Institute of Psychology, University of Graz, Graz, AT

† Max Planck Institute for Intelligent Systems, Rationality Enhancement Group, Tübingen, DE

‡ Tilburg School of Social and Behavioral Sciences and Department of Methodology, Tilburg University, Tilburg, NL

§ BioTechMed – Graz, Graz, AT

Corresponding author: Hilmar Brohmer (hilmar.brohmer@uni-graz.at)

might be helpful in general. However, Aarts et al. (2004) extended this idea from traits of the observed person to the goals she/he is pursuing. Thus, an observed helpful behavior could be an indicator for both the observed person's trait of being helpful in general, but also for his/her current goal of offering help. The latter is the prerequisite for the GC process.

The process of GC can be described as follows: People observe others behaving in a certain way and infer their goals quickly, automatically, which may happen outside their conscious awareness (Aarts et al., 2004; Hassin et al., 2005). If this inferred goal has some relevance to the observers, they are inclined to adopt this goal thereafter. This way, GC is actually a two-step process, predicting a goal-directed behavior of an observer, following the automatic inference and activation of that goal through observations. Laurin (2016) even compared GC to a misattribution process, because the automatically inferred goal of the other person is mistakenly attributed to the self and thus directs the observer's thoughts and actions. Hence, this inference step is often operationalized through variables that do not refer to the goal itself, but assess the activation of the goal indirectly (e.g., a lexical decision task with goal-related words refers to the speed of participants). This is also theoretically sensible and in line with regard to GC's origin in the spontaneous causal inference framework (Hassin et al., 2002), as described above.[1]

Nonetheless, research on GC also does not entirely exclude less automatic pathways from the model: indeed GC research sometimes reports direct (sometimes also referred to as 'explicit') measures of goal inference as manipulation checks or additional dependent variables (DVs) (Dik & Aarts, 2007; Jia, Tong, & Lee, 2014) that make direct reference to the goal (e.g., asking participants 'what the person in the text tries to achieve').

Several studies have presented evidence for the existence of GC. Three aspects become apparent from the empirical literature: First, different research teams have tested a wide array of diverse goals. Therefore, the GC effect was shown for goals ranging from having casual sex (Aarts et al., 2004) to behaving prosocially (Dik & Aarts, 2007) to achieving high scores in a task (Leander & Shah, 2013) to dieting (Lee & Shapiro, 2015). Second, in most of the literature, GC manipulation is accompanied by moderators that might operate in a unique way for some goals but potentially not for others. For instance, an observed person showing extra effort in his or her behavior (as moderating condition) might be beneficial for the observer adopting a prosocial goal (Dik & Aarts, 2007), but not for self-serving goals such as earning money (Corcoran et al., 2018). Third, although goal contagion is often conceptualized as a mediation (where automatic activation mediates the effect between goal observation and goal adoption), most studies focus on either demonstrating the automatic activation of the observed goal *or* behavioral measures as goal adoption without looking at activation (for exceptions that focus on both, see Corcoran et al., 2018; Dik & Aarts, 2007; Jia et al., 2014). Recently, more labs, including our own lab, have attempted to contribute to the body of research on GC by testing both formerly used and new goals, as well as the two-step model, including moderators (Brohmer et al., 2018; Corcoran et al., 2018; Wessler & Hansen, 2016). Interestingly, these attempts often yielded effects close to zero, although sample sizes were much larger than in previous studies.

Because these studies demonstrated that the GC effect does not seem to be as robust as could be expected from the previously published literature, we reasoned that a meta-analysis on GC would be advantageous. On the one hand, we wanted to 1) summarize the evidence for GC, 2) discern the statistical evidence for the automatic activation process and behavioral measures, and 3) identify further moderating effects that might turn out to be important across goals. Hence, our motivation was to see which goals that people perceive in others are truly affecting their own goals and under which conditions. On the other hand, we also wanted to test and correct for potential publication bias to obtain more accurate effect size estimates, which has proven to be effective in other social-psychological research (Francis, 2012; Kühberger, Fritz, & Scherndl, 2014; Lane & Dunlap, 1978).

Potential moderators

In accordance with the aforementioned three points, we selected the DV that was used (henceforth: *DV category*) as a first moderator of interest, which has the advantage of being relatively objectively identifiable: it can be automatic goal activation or goal pursuit (including behavioral intention). Automatic activation as indicator of a goal inference is usually measured via variations of the lexical decision or word completion task (e.g., Dik & Aarts, 2007). Goal pursuit contains both behavioral measures and an expression of intention (see supplementary document: https://osf.io/jx7rc/).

Our second moderator of interest differentiates elicited goals to test the idea that some goals might be more contagious than others. There are many dimensions on which goals might differ. We decided to look at goals based on how many people would pursue this goal, which hints at whether a goal can be perceived as quite 'common' (henceforth: *common goal*). Crucially, this moderator is useful because the GC process is theorized to be strengthened when observers assign a high or positive value to the goal that they infer (Aarts et al., 2004; Brohmer et al., 2018; Corcoran et al., 2018). If a goal is pursued by a majority of people, this indicates a high value in the eyes of many and demonstrates a broader relevance. This in turn could make it more likely from the perspective of a specific individual (such as a study participant) that he/she might also assign a high value to this goal, which should foster the GC process. Therefore, common goals might be more contagious overall.[2]

The last two moderators have a more methodological focus and might provide insights into how best to study GC. The third moderator will be the presentation of stimulus material, depicting a goal-directed behavior (henceforth: *presentation*). This measure is relatively objective: as most stimulus material could be expected to be identifiable as texts or video clips and animations, these forms being most likely to depict behavior in a standardized

form. More vivid materials – such as videos – showed stronger effects in other intervention contexts (e.g., Soetens et al., 2014; Walthouwer et al., 2015). Therefore, we assume that video clips and animations might be more effective to elicit GC than texts.

Lastly, we are also interested in to what extent the control conditions for each study might be perceived as neutral or contrary to the goal (henceforth: *contrast control*). For instance, for an observed prosocial goal in an experimental condition (i.e., someone provides help to another person), a neutral condition could be a situation without a prosocial context (i.e., nobody needs help). A situation contrary to the goal could be when selfish behavior is observed (i.e., someone does not provide help, although he or she could). It is expected that control conditions that are contrary to the goal might result in a stronger GC effect as the contrary condition might inhibit the goal-directed behavior much stronger than a neutral control condition.

Method

We defined general inclusion criteria for extracting published articles from databases and specific inclusion criteria to identify relevant studies in those papers. In addition, we developed our coding scheme for the GC studies based on a preliminary coding of five original papers. This was necessary due to differences in reported statistical information, which often affect the extraction of relevant effects. Hence, we used those five studies to gain general experience for the coding and how to deal with limited information. Those three steps – general inclusion of papers, specific inclusion of studies, and the coding of relevant effects – will be described in the following paragraph. All materials, data, codes, and a PRISMA- guidelines checklist for meta-analyses can be accessed in the accompanying Open Science Framework (OSF) project folder (https://osf.io/mxepy/).

General inclusion criteria for articles and database search

General theoretical eligibility

Prior to the database search, we committed ourselves to a definition and general criteria of GC as the benchmark, which we derived from the literature (e.g., Aarts et al., 2004; Dik & Aarts, 2007) and summarized in three crucial points.

First, for GC to occur, an observer has to observe or read about a behavior by another person that implies a certain goal, but the goal is not explicitly mentioned. This is crucial for GC from a theoretical perspective, as a behavior is the originator of the proposed cognitive process of goal inference. This distinguishes GC from other routes of goal activation, such as goal priming (e.g., Bargh, et al. 2001; see also Weingarten et al., 2016), in which the goal itself is presented as a semantic concept. Second, the goal gets activated in the observer based on an inference of the observed behavior. This inference is assumed to happen automatically and outside the observer's conscious awareness (Aarts et al., 2008; see also De Houwer & Moors, 2010). Third, even though the GC process should ultimately lead to the adoption of the goal by the observer resulting in goal-directed behavior or intentions, we also accepted studies focusing on studies on goal activation. From a design-specific viewpoint, GC should be demonstrated in an experimental-psychological study. That is, there has to be a goal manipulation including an experimental condition, in which participants observe a goal- directed behavior and (some sort of) control condition, followed by a measure of automatic goal activation or goal pursuit.

The second and third point are not independent: the inference of the goal in the observer is theorized to occur quickly and automatically outside conscious awareness after the observation (Aarts et al., 2004, pp. 24–25). It has to be reiterated that non-automatic inference is usually of minor interest in GC-related research and therefore direct measures of the outcome rather serve as a manipulation check, which is why we do not include it as a central part of our definition. Only after an automatic inference should goal adoption occur, which is typically measured in participants' goal-directed behavior (i.e., goal pursuit) or intention for goal-directed behavior. Notably, inference is often described as a mediator between observation and adoption of the goal, but as the path from inference to goal pursuit or intention is rarely studied, we will focus on the relationships between goal manipulation and goal inference or goal pursuit.

The theory on GC does not set restrictions on the goals to be elicited. In accordance, the preliminary coding of the five papers revealed the use of a diverse range of goals. Therefore, we set no restriction either, and recognized all kinds of goals – be they 'academic achievements' (Wessler & Hansen, 2016), 'being helpful' (Dik & Aarts, 2007), or 'having casual sex' (Aarts et al., 2004). In the same vein, we expected that in an experimental setting the focal goal toward which the behavior is directed should be rather obvious to strengthen the manipulation (although behaviors are sometimes multifinal, see Shah, Kruglanski, & Friedman, 2003).

Systematic search

After fixing definition and criteria, we conducted a systematic search between March and April 2018 in four databases for published work, namely PsychInfo, Web of Science, ScienceDirect, and JSTOR. In all four databases, we applied a similar search logic: we looked in the title, abstract, and keywords for the term 'goal' in combination with 'contagion,' 'social learning,' 'modelling,' 'modeling,' 'role model,' 'social standard' and its alterations, 'comparison standard' and its alterations, or 'observational learning' (for specific search syntaxes, see https://osf.io/w8b9m/). The initial systematic search resulted in $k = 2821$ articles.

Screening of articles

The articles were split equally among three trained student assistants, who applied the criteria of the general theoretical eligibility to separate relevant from irrelevant papers based on titles and abstracts. This culminated in $k = 30$ articles that remained of interest. Afterwards, the same coders applied the same criteria again, by looking into the 30 articles in more detail. After this screening, 17

articles had to be excluded, as it turned out they were not eligible, and 13 relevant articles (including the five preliminary articles) and one registered report (see section on unpublished studies) were kept.

Additional searches

However, some potentially relevant GC articles did not show up during this search, which is why we decided to perform an additional browsing in citing literature on Google Scholar and in the reference lists of articles that we obtained from the systematic search. We also performed another extended search on PsychInfo for studies with adult samples, using the keywords 'observational learning,' 'role model,' and 'social learning.' For this search, we excluded the word 'goal' to make sure we would not miss studies conceptually related to GC, despite not employing the same wording.

These procedures – the additional browsing and extended search – yielded a total of k = 2908 articles and documents (including the ones from the previous paragraph). We again checked the content of the articles, which yielded an additional k = 12 articles. These seemed to be of relevance due to a fitting experimental setup, although they did not necessarily self-identify as being GC related. In total, we found k = 24 articles that were of relevance as they potentially contained studies that would fit in this meta-analysis (see **Figure 1**).

Specific inclusion criteria for studies in published articles

After coding all studies that were initially theoretically eligible (resulting in a total of e = 96 effects that measured automatic goal activation or goal pursuit; e = 127 when also including explicit inference measures), we proceeded to look into the method sections of all selected studies to see if they also fit our specific inclusion criteria.[3] These criteria encompassed points such as whether there was an *identifiable goal* the authors wanted to elicit in their experimental design, whether *participants* were adults of at least 18 years of age, and whether the specific goal the experimenters wanted to trigger was *not mentioned* in between the manipulation and the measurement of the DV. This last point is crucial for a clear distinction of GC from goal priming because GC includes a goal-inference step, which would be obsolete if the goal is mentioned before the DV is measured. Other points were that the goal-directed behavior of the observed person was *not identical* to the one shown by the participant, in order to distinguish GC from role modeling (Morgenroth, Ryan, & Peters, 2015) or mimicry (Chartrand & Lakin, 2013), and whether the control condition *differed sufficiently* from the experimental condition (i.e., by not using an attenuated version of the experimental conditions). Finally, it was important that sufficient statistical information was reported according to our preliminary set *criteria for the extraction of effects* (see below).

We applied these specific criteria study by study to identify suitable effects (as there is often more than one DV measured per study) and had to exclude 38 effects, leaving e = 58 effects. Some of these effects were taken from the same studies and had to be either combined (e.g., if two effects from the same study were based on automatic activation) or reduced to the preferred effect (i.e., only pursuit was used when there was automatic activation and pursuit measured). Finally, this left us with e = 48 effects for the confirmatory analysis. It has to be noted that some effects from self-identified GC studies had to be eliminated from the confirmatory analysis – for instance, if goal pursuit was too close to the manipulation or the goal manipulation itself was too explicit about the goal. Those individual coding decisions that did not fit the criteria are marked in the accompanying spreadsheet as excluded (see https://osf.io/w8b9m/).

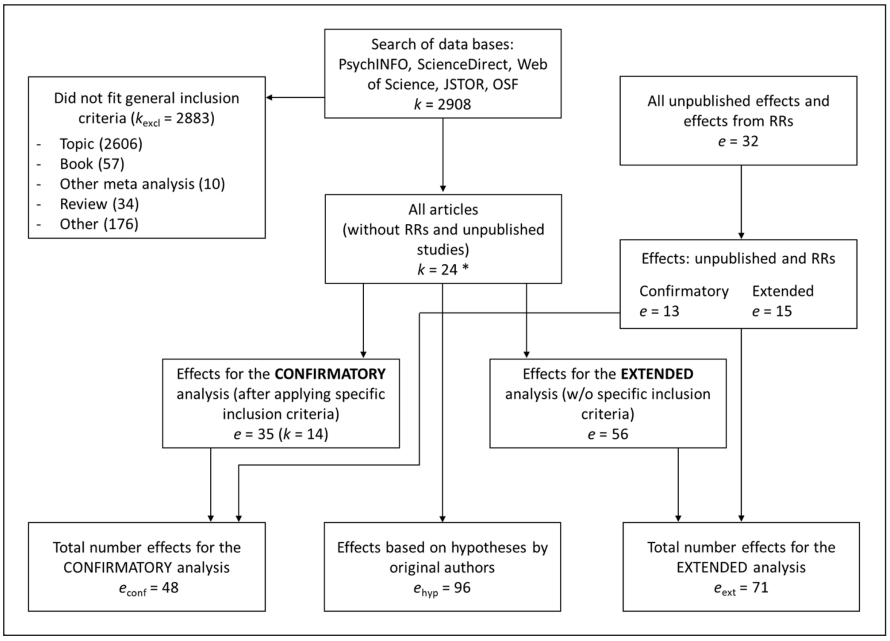

Figure 1: Decision tree for search for articles leading to the confirmatory and extended analysis; * including effects that did not fit inclusion criteria, explicit inference, and effects from five preliminarily coded articles; all e = 127 coded effects can be found in the data matrix online.

Unpublished studies and studies from registered reports

We also looked for unpublished studies via the OSF and ProQuest using the same search terms as before (see search syntax in the OSF). Furthermore, we contacted relevant labs via email, asking whether there are more unpublished GC studies available. Neither approach yielded further results. Hence, we could only include effects from our own lab, which were not published at the time of the coding procedure, and published studies from registered reports (henceforth: RRs). As RRs employ the main peer-review procedure before the data collection and therefore before any bias can occur, they have to be treated differently from common publications (see Chambers, 2019). Data and codes for the extraction of effects from these studies, along with preregistrations when available, are also provided online.

Criteria for the extraction of relevant effects

We intended to code effects and variables that can broadly be summarized in four categories and which will be discussed in the following sections: effects relevant for GC (confirmatory effects); effects from all studies that passed the initial criteria of general theoretical eligibility (extended effects); effects hypothesized by the original authors (originally hypothesized effects); and potential moderators. The coding was done for all effects that passed the initial criteria for general theoretical eligibility.

Effects for the confirmatory analysis

We determined clear criteria for the extraction of relevant confirmatory effects. We were interested in DVs that represented automatic activation, goal pursuit or behavioral intention. However, these effects of the DV could be present as either *main effects* or *simple effects* in factorial designs (i.e., with independent variables in interaction).

We took the main effect if this was the only manipulation (i.e., a goal vs. control group) or when the original authors expected an attenuation of the GC effect through a second factor as a moderator. This latter case implied that the GC effect was present in both conditions of the second factor, although to a different degree.

There were two situations, in which we would consider a simple effect (i.e., the GC effect as observed in one specific condition of factor 2): first, when a knockout effect was expected (i.e., there is the GC effect in the first condition of factor 2, but there is no effect in the second condition of the factor 2); and second, when a crossover effect was expected (i.e., there is a reversed effect in the second condition of factor 2; see Giner-Sorolla, 2018).

Furthermore, it could be possible that there were more than two conditions present on a factor. If this was the case, whereby the additional condition was a second control condition (e.g., control 1 vs. control 2 vs. goal), we took the effect that was most neutral and least opposed in comparison to the goal condition to ensure that the GC effect was not driven by the opposed control group. If more than two groups were present on the goal factor, whereby the additional condition was a second goal condition (e.g., control vs. goal low vs. goal high), we would aggregate the goal conditions as they both manipulated the goal of interest (see coding scheme: https://osf.io/jy9m3/).

Means, standard deviations (or standard errors), and group sizes were crucial descriptive statistics for the calculation of the effect size Hedges' *g*, which is the standardized mean difference Cohen's *d* corrected for positive bias (Hedges, 1981). During the coding phase, we found that reporting standards varied strongly across papers. Therefore, when descriptive statistics were not reported in studies, we based the calculation of the effect sizes on test statistics like *t*-values, *F*-values, χ^2-values and *r*-estimates. An automated procedure is provided by Del Re's *R* package *compute.es* (Del Re, 2015).[4]

Effects for the extended analysis

Using our definition and our conservative inclusion criteria for GC might have a side effect: our attempt at higher precision could work at the expense of variance that could explain moderating effects. That is, because we will probably have to exclude several studies for the confirmatory analysis, we will also reduce our chance of identifying moderating conditions. Therefore, we will conduct an extended, not preregistered analysis using all effects that passed the initial criteria for general theoretical eligibility, $e = 96$. Again, similar effects from these studies were combined, resulting in a sample of $e = 71$ effects.

Originally hypothesized effects

The originally hypothesized effects, which are the hypothesized effects in the primary studies by the original authors, were not always equivalent to the relevant effects for this meta-analysis as the former could contain specific interactions with other variables (e.g., other manipulations) that could deactivate or even reverse the goal contagion effect. Therefore, both effects have to be clearly distinguished to avoid certain effects being falsely attributed to GC alone (for a recent example of such a case see Crede, 2019, in response to Cuddy, Schultz, & Fosse, 2018). *P*- and *Z*-values for these effects were extracted for additional publication bias tests and power estimations (Schimmack & Brunner, 2017, 2019; Simonsohn, Nelson, & Simmons, 2014, 2019). Results of these tests complement the main results and are reported in the supplementary document (see https://osf.io/jx7rc/).

Moderators

As the GC effect was expected to be heterogeneous across studies, we intended to identify potential moderators of this effect, which we described in the theory section and in the supplementary materials. They encompass *DV category* (activation vs. pursuit), *common goal* (number of people expected to pursue the goal), *presentation* of manipulation material (texts vs. video clips and animations; excluding single pictures), *contrast control* (neutrality of the goal condition), and *self- versus other-directed goal* (exploratory moderator). Other differentiations, such as whether the goals are more short-term versus more long-term oriented or approach-related versus avoidance-related implying wins and losses might also be of interest.

However, due to the relatively low number of studies (see results section), we decided to test a restricted number of moderators, thereby avoiding an inflation of the false-positive error rate.

Results of the Meta-analysis

Analyses were conducted in *R* (R Core Team, 2019), using the *metafor* package (Viechtbauer, 2010, 2017), the *robumeta* package (Fisher, Tipton, & Zhipeng, 2017), *puniform* (van Aert, Wicherts, & van Assen, 2016), *weightr* (Coburn & Vevea, 2019; Vevea & Hedges, 1995), and *ggplot2* (Wickham, 2019).

Agreement for all codings was assessed based on the rating of two trained raters, who applied the specific inclusion criteria for all studies and extracted the relevant effects. The agreement was low for all coded effects, $\kappa = .35$ and also for the hypothesized effects, $\kappa = .42$, despite our carefully developed coding scheme (https://osf.io/jy9m3/; for details, see supplementary document: https://osf.io/jx7rc/). We conducted the confirmatory and extended analysis using the second rater's codings to see whether the summary effect would be different compared to rater 1. Interestingly, this was not the case: Despite nominal differences in individual estimates across raters, which yielded low reliability scores, the conclusion remained the same (see supplementary Figure S2). In the following, we will report results only from the first rater as the large majority of discussions upon initial disagreement resulted in agreeing with his codings.

Confirmatory and extended analysis

We conducted the random-effects meta-analysis using the restricted maximum likelihood (REML) estimator for estimating the between-study variance in true effect size for $e = 48$ effects (see **Figure 1**) from the published and unpublished literature. We opted for the random-effects model instead of the equal-effect model, because we wanted to estimate the average true effect size in the population from which the effects were randomly sampled (Borenstein et al., 2010). Those effects represented either a measure of automatic activation or goal pursuit and were based on 4751 participants. The results indicated a small summary effect of GC, Hedges' $g = 0.30$, 95%*CI* [0.21, 0.40][5] (see **Figure 2A**), which was accompanied by some heterogeneity across studies, $Q(47) = 113.00$, $p < 0.001$, $\tau^2 = 0.05$, 95%*CI* [0.03, 0.13], $I^2 = 57.57\%$, 95%*CI* [40.19%, 77.67%], $H^2 = 2.36$, 95%*CI* [1.67, 4.48]. The extended analysis was based on more effects, $e = 71$, and yielded the same effect, $g = 0.30$, 95%*CI* [0.22, 0.37], and a similar amount of heterogeneity, $Q(70) = 153.38$, $p < .001$, $\tau^2 = 0.05$, 95%*CI* [0.03, 0.11], $I^2 = 53.46\%$, 95%*CI* [37.93%, 71.61%], $H^2 = 2.15$, 95%*CI* [1.61, 3.52] (see **Figure 2B**).

For the previous analyses, we combined intention and behavior/goal pursuit. As a behavioral intention can also be seen as different from actual behavior, we additionally report pure behavioral results for the confirmatory data ($e = 26$, **Figure 2C**), $g = 0.30$, 95%*CI* [0.16, 0.44], $Q(25) = 73.27$, $p < .001$, $\tau^2 = 0.08$, 95%*CI* [0.04, 0.26], $I^2 = 67.90\%$, 95%*CI* [48.93%, 87.29%], $H^2 = 3.12$, 95%*CI* [1.96, 7.87], and extended data ($e = 40$, **Figure 2D**), $g = 0.31$, 95%*CI* [0.21, 0.41], $Q(39) = 92.96$, $p < .001$, $\tau^2 = 0.05$, 95%*CI* [0.02, 0.14], $I^2 = 56.78\%$, 95%*CI* [36.18%, 77.84%], $H^2 = 2.31$, 95%*CI* [1.57, 4.51]. As one can see, these effects are no different from the larger data sets.

Publication bias and correction

To assess potential publication bias, we first correlated the sample size per study with the size of the effect. The negative and significant correlation (see **Figure 3A**) hinted at potential publication bias as studies with a smaller sample size yielded larger effects (Kühberger et al., 2014). We split the data into published and unpublished relevant effects, where unpublished effects also included effects from RRs, as most publication bias corrections require traditional publications (without preregistration) only. Effects from both subgroups differed considerably: The summary effect of the published studies was larger, $g = 0.42$, 95%*CI* [0.34; 0.50], $Q(34) = 41.12$, $p = 0.187$, $\tau^2 = 0.003$, 95%*CI* [0.00, 0.08], $I^2 = 5.83\%$, 95%*CI* [0.00%, 61.19%], $H^2 = 1.06$, 95%*CI* [1.00, 2.58] than that of the unpublished studies, $g = -0.01$, 95%*CI* [−0.10; 0.08], $Q(12) = 15.76$, $p = 0.900$, $\tau^2 = 0.004$, 95%*CI* [0.00, 0.10], $I^2 = 15.24\%$, 95%*CI* [0.00%, 80.62%], $H^2 = 1.18$, 95%*CI* [1.00, 5.16], which in turn yielded an effect very close to zero (i.e., no effect). Interestingly, heterogeneity was much smaller and nonsignificant in the subgroups, indicating limited potential for moderating effects.

We proceeded by funnel-plotting the effect sizes of the published effects against the standard errors of the effects (see **Figure 3B**). Egger's regression (Egger et al., 1997), which is depicted as the diagonal dashed line in **Figure 3a**, $b = 2.45$, $t = 4.86$, $p < 0.001$, also suggested that small-study effects were present and one possible cause of small-study effects is publication bias. The estimates for the other three models yielded similar results and can be found in Supplementary Table S1.

Next, we applied several older and more recent correction methods for the effect size estimate (descriptions are provided in **Table 1**).[6] For both the confirmatory and extended model, all corrections brought the estimate closer to zero. Assuming that the true effect is around $g = 0.15$, this corresponds to 17 to 22 participants that have to be exposed to a goal contagion manipulation in order to find one person who is actually influenced by the observation compared to a control group.[7]

Results for the moderators

We conducted meta-regressions for all preregistered (i.e., presentation, DV category, common goal, and contrast control) and exploratory moderators (prosocial/cooperative goal vs. self-serving goal) individually and controlling for the other variables. However, results were always similar: There was virtually no evidence that any moderator showed an expected effect (see **Figure 4**). Only the zero-order effect of contrast control had a slope coefficient different from zero, which was also tiny in size. However, when we accounted for whether the effect came from the published or unpublished literature, this effect vanished as well (for

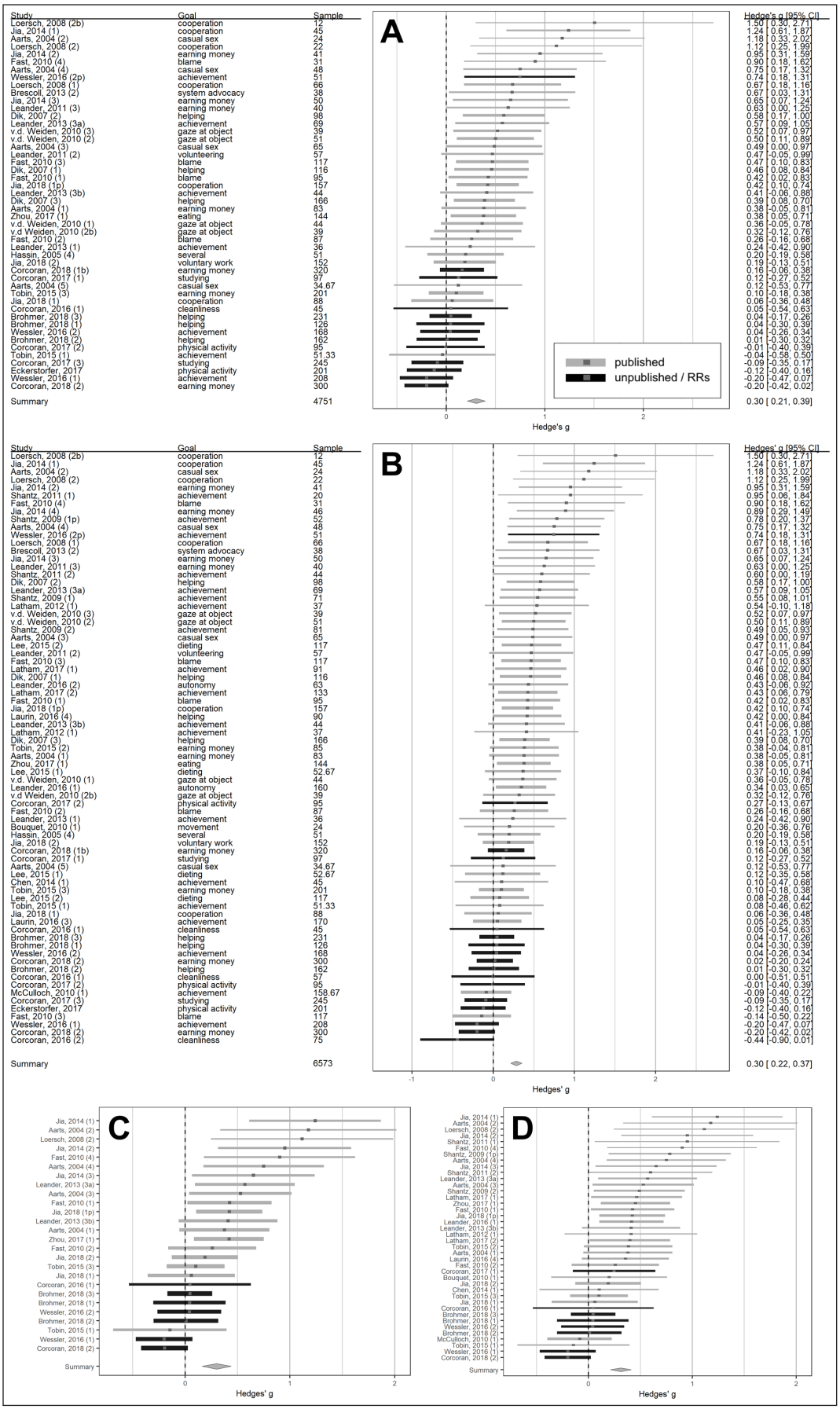

Figure 2: Forest plots for the confirmatory analysis **(A)**, extended analysis **(B)**, behavior only in the confirmatory data **(C)**, and extended data **(D)**. Note that sample sizes sometimes include decimals as insufficient information on subsamples were provided in the original articles and we assumed equal-sized cells; thickness of CIs indicates higher precision.

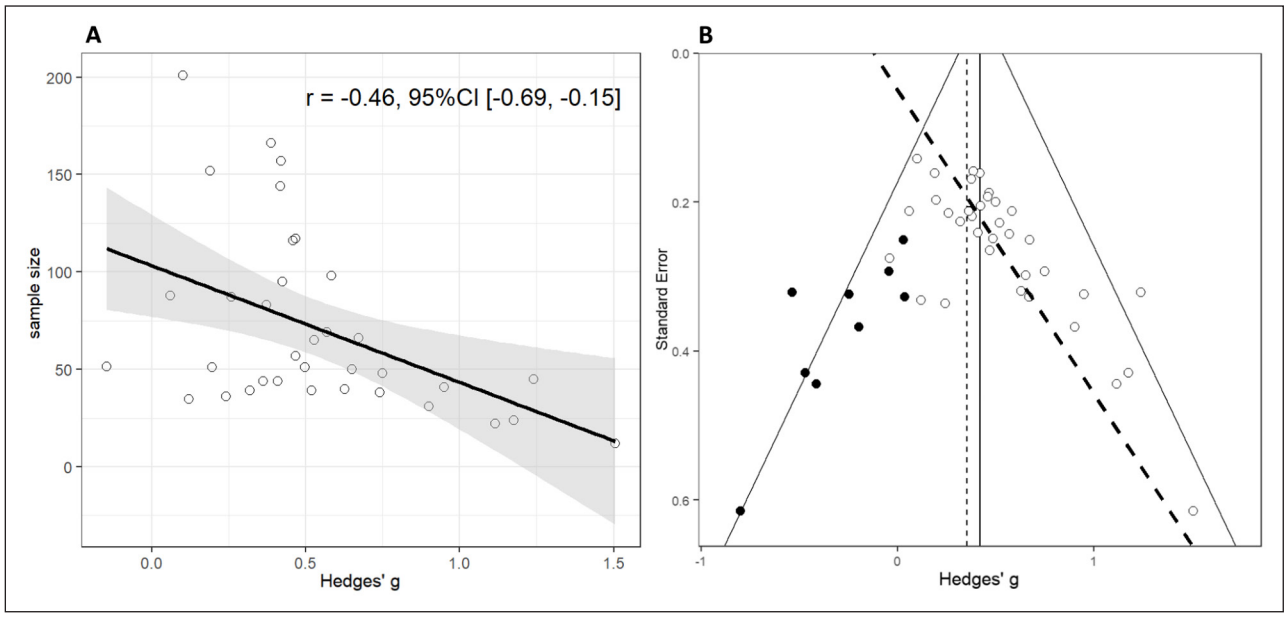

Figure 3: Publication bias methods. **(A)** association between effect size (Hedges' *g*) and sample size; shades are the 95% CI; **(B)** association between effect size and standard error in a funnel plot; vertical full line is the raw summary effect; black full dots are imputed effects for trim-and-fill; vertical dashed line is the trim-and-fill corrected effect; the diagonal dashed line is Egger's regression.

Table 1: Different effect size correction methods and Hedges' *g* with 95% CI. See supplementary document for details.

Correction Approach	Description	Source	Confirmatory Model (*e* = 48)	Pursuit Confirm (*e* = 26)	Extended Model (*e* = 71)	Pursuit Extend (*e* = 40)
Trim & Fill	Liberal correction based on mirrored studies	Duval & Tweedie, 2000	0.33 [0.22; 0.43]	0.16 [–0.003; 0.33]	0.19 [0.11; 0.26]	0.19 [0.08; 0.30]
PET method	Conservative correction based on the intercept of Egger's regression (PEESE with asterisk)	Stanley, 2008	–0.12 [–0.35; 0.11]	–0.11* [–0.24; 0.02]	–0.03* [–0.11; 0.06]	–0.05* [–0.17; 0.06]
Selection model	Assigns different weights to significant and nonsignificant effects	Vevea & Hedges, 1995	0.15 [–0.02; 0.32]	0.13 [–0.07; 0.34]	0.06 [0.01; 0.11]	0.04 [–0.002; 0.08]
P-uniform	Assumes a uniform distribution of p-values ($p < .05$) conditioned on the true underlying effect size.	van Assen et al., 2015	0.21 [–0.05; 0.41]	0.27 [–0.07; 0.54]	0.13 [–0.10; 0.32]	0.16 [–0.14; 0.39]
P-uniform*	Also contains information from nonsignificant effects ($p \geq .05$)	van Aert et al., 2016; van Aert & van Assen, 2019	0.17 [0.05; 0.29]	0.20 [0.03; 0.38]	0.16 [0.07; 0.25]	0.15 [0.03; 0.28]
Hybrid method	Assumes a bias in published effects, but not in unpublished effects	van Aert & van Assen, 2018	0.11 [0.001; 0.23]	0.10 [–0.04; 0.29]	0.07 [–0.03; 0.18]	0.07 [–0.07; 0.22]

more detailed descriptions, see supplementary document). Consistent with the results from the subgroups, publication status as predictor had the strongest effect and was superior to all other moderators, $b = 0.43$, 95%*CI* [0.28; 0.58].

General Discussion

Observing someone pursuing his or her goals can have a profound effect on ourselves – maybe even to the extent that we adjust our own goals. The theory on GC (Aarts et al., 2004) allows a social-cognitive approach to this phenomenon, according to which an observation of someone's goal-directed behavior will lead to an automatic inference and activation in the observer and potentially to a behavior towards a similar goal. Here, we set out to summarize the evidence for GC in a meta-analysis and to identify moderators of this process, which is based on 48 effects for the confirmatory analysis and 76 effects for the extended analysis.

First, we initially found an overall summary effect of Hedges' $g = 0.30$. But this effect can be described as

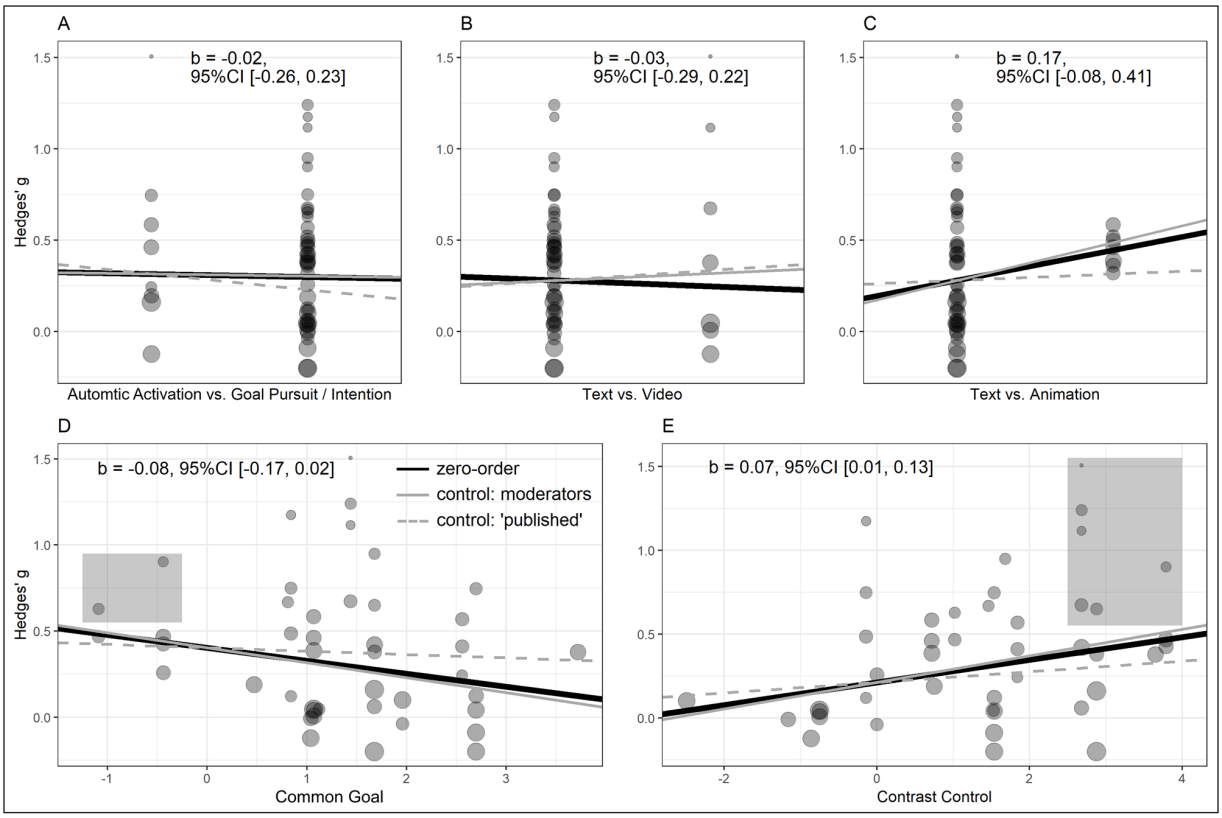

Figure 4: Meta-regression to identify moderators. Bubble size indicates sample size of the study; estimates in the panels are zero-order and for the black lines; continuous moderators are based on a pre-study and are centered on the scale-midpoint; rectangles contain potentially interesting effects (details, see supplementary document); additional estimates: **(A)**: gray line: $b = –0.01$, 95%CI [–0.30, 0.28], dashed line: $b = –0.09$, 95%CI [–0.25, 0.06]; **(B)**: gray line: $b = 0.04$, 95%CI [–0.23, 0.30], dashed line: $b = 0.06$, 95%CI [–0.11, 0.22]; **(C)**: gray line: $b = 0.21$, 95%CI [–0.18, 0.59], dashed line: $b = 0.04$, 95%CI [–0.15, 0.22]; **(D)**: gray line: $b = –0.09$, 95%CI [–0.18, 0.01], dashed line: $b = –0.02$, 95%CI [–0.09, 0.05]; **(E)**: gray line: $b = 0.08$, 95%CI [0.02, 0.14], dashed line: $b = 0.03$, 95%CI [–0.01, 0.07].

small, and it seems to be biased through the current state of the publication system. On average, published studies reported larger effects, $g = 0.42$, than unpublished studies and RRs, $g = –0.01$. Publication bias correction methods estimated the effect size to be around half the size of the uncorrected summary effect. The suggestion of a true effect around $g = 0.15$ was further supported by the extended analyses after correction. All in all, GC appears to have a rather soft effect, as one needs around 20 people to find one person being affected by the observation. Hence, with regard to the GC effect, observing others is not something that can be expected to distract people all the time from their daily activities. Rather, the GC effect, if it exists, might be limited to particular instances in contexts that are hard to pinpoint, which also becomes clear by looking at the moderators.

Second, when looking at the DV category, there was no noteworthy difference between behavioral and automatic-activation outcome measures. This is somewhat surprising as one could expect stronger effects of activation, regarded as prerequisite for a goal to be adopted. Furthermore, focusing on behavior only goal pursuit (without intention for behavior) yielded similar effects as the overall analysis. Taken together, we found no evidence indicating that the GC is stronger or more easily detectable depending on the deployed DV (automatic activation, intention for behavior, or behavioral goal pursuit).

Third, we did not find evidence for moderating effects, which we discuss in more detail in the supplementary materials (see https://osf.io/jx7rc/). Only the effect between the GC condition and the control condition became slightly more pronounced the more contrary the goals in both conditions were. But given that several studies used control conditions contrary to the goal condition (e.g., Dik & Aarts, 2007; Laurin et al., 2016), it is surprising that this effect is not much more pronounced. One reason why contrary-goal control conditions are not as effective as one would think could be that they often indirectly imply the actual goal (Moskowitz & Gesundheit, 2009). For instance, reading about participants doing voluntary work (as a goal contrary to earning money; see Aarts et al., 2004) could still activate the goal of earning money in some participants (e.g., Corcoran et al., 2018) and hence reduce the GC effect.

Fourth, the unpublished studies used larger samples than the published studies, which contributed to their higher precision (i.e., smaller confidence intervals). However, it has to be pointed out that the unpublished studies differed on aspects other than larger samples. Foremost, they partially used different goals, such as

physical activity, which were not used in any of the published studies. Hence, it is possible that aspects regarding the goal selection contributed to the considerably smaller effects.

As with many meta-analyses in psychology, this one, too, suffers from deficiencies. An obvious limitation is probably the partially low interrater agreement scores. We intended to ensure that our preregistered coding scheme would produce similar results independent of the raters as this is currently not the standard in quantitative research (Maassen et al., 2019). This turned out to be difficult as both the relevant effects and originally hypothesized effects were often not clear enough to the coders, despite extensive training with five preliminarily coded articles as pilot. Hence, the low scores might partly depend on the experience level of the coders and partly on the large variation of the experimental designs. The latter point should not be underestimated: Different designs result in different effect sizes, which have to be transformed into common effect sizes, which in turn can introduce bias. Concerning the first point, it is important to note that discussions solved most of the disagreements. Moreover, when we conducted meta-analyses for the different raters, the overall effects did not differ, despite differences in individual effects. This result, however, has to be treated with caution as studies were not randomly assigned to these raters.

An issue related to varying designs that lowered the interrater agreement was the ambiguity of the goals and their respective operationalization as manipulation and DV. This ambiguity is problematic from a theoretical perspective because it might imply that the theory is not specific enough. This will be illustrated by looking at achievement goals here, but could really be demonstrated with other goals as well. Achievement goals were used several times to demonstrate the GC effect, but the manipulations and outcome variables were treated differently by different authors (i.e., concepts instead of studies were replicated; see Chambers, 2017), which makes it difficult to conclude that achievement goals generally show or do not show a GC effect. For instance, Leander and Shah (2013, Study 3a) had participants read about a student who had either an immediate or distant deadline for a semester paper (manipulation), which led subjects to work with higher or lower persistence on anagram tasks, respectively (DV): the closer deadline exerted a larger GC effect. Tobin, Greenaway, McCulloch, and Crittall (2015, Study 1) used a similar manipulation but had participants write an essay and its quality was assessed by different raters as DV. In both examples, the original authors argued that the behavioral outcome measure was indicative of an achievement goal being activated. In other examples, the manipulation materials also differed (e.g., Dik & Aarts, 2007; Brohmer et al., 2018). It has to be emphasized that a high degree of ambiguity of conceptually similar studies was the rule rather than the exception. And this ambiguity was fueled even further by differences of control groups and additional interacting variables.

Given this variety, it is even more interesting that only moderate heterogeneity across studies was found. In light of this, two interpretations are possible: either the corrected summary GC effect across studies that we extracted is so clear that it truly represents an existing effect, or the summary GC effect is no more than an artifact of consistently selective reporting in the literature, independent of the designs. The strong evidence for publication bias for both the extracted effects and the originally hypothesized effects (see supplementary document) along with the drop in estimated between-study variance after the publication-bias correction rather indicate the latter.

Certainly, we do not intend to dismiss individual studies or even the theory on goal contagion as a whole. In fact, goal contagion remains an elegant approach to explain how people can become inspired by their peers, which needs to be explored further. However, similar to other topics of social psychology (e.g., Friese et al., 2017; O'Donnell et al., 2018; Simmons & Simonsohn, 2017), goal contagion also suffers from many early studies that showed unreasonably large effects with low sample sizes. Future studies should avoid these pitfalls of low power, which increase both false-negative and false-positive findings.

Furthermore, the theory in its current state seems to be underspecified, despite 15 years of research. There are many examples for this underspecification: for instance, it is not clear whether measuring the accessibility of a goal concept already suffices to say that a goal inference took place, whether goal inference *could* or *needs to* occur quickly and automatically or whether a successful GC process should manifest itself in *goal pursuit only* or in *behavioral intentions* likewise. One reason for this underspecification lies potentially in a body of research without close replications and where new studies almost always introduced changes in the research designs. Changes in research designs to identify the boundaries of a theory are, of course, important for the development of a theory. But they become problematic when the empirical basis for extended designs is uncertain and thin (see Chambers, 2017, Chapter 3). Moreover, some changes in the research designs of published studies did not correspond well to the original theory, which does not allow for more nuanced conclusions.[8]

As a starting point for future research on goal contagion, we think that the theoretical processes underlying GC must be better specified (e.g., by clearly identifying causal effects, or by applying computational modeling techniques, see e.g., Rohrer, 2018; Smaldino, 2019; Guest & Martin, 2020). Also, one should assume very small GC effects of standardized mean differences of around 0.15, based on the publication bias methods for the confirmatory and extended analyses. This, of course, corresponds to sample sizes of at least 1102 participants (one-sided *t*-test, $1\text{-}\beta = 0.80$, α error probability = 0.05), which, in order to be achieved, potentially require collaborations between labs. Anything below this number does not seem to be reasonable, as the literature does not contain enough reliable information on designs that allow for much smaller sample sizes. Additionally, applying open science practices, such as data and material sharing and preregistration, will become necessary so that researchers can learn from each other more efficiently.

Conclusion

GC is a social-cognitive approach to understand how observing others can affect our own goal-directed behavior. However, there are indications of publication bias within the published literature and most recent studies yielded effects clustering around zero. Potential moderators that could advance the theory on GC could not be identified in this meta-analysis, either. We strongly suggest applying open science practices and determining the required sample sizes based on a power analysis in future research to bring goal contagion back on track.

Notes

[1] It has to be noted that the original GC articles employed the term 'implicit' for the automatic activation and inference processes. As the conceptualization of 'implicit' is vague – it could mean 'automatic', 'associative', or 'indirect' (see Corneille & Hütter, 2020) – we will stick to the term 'automatic' in this article.

[2] Note that in the preregistration, we refer to the 'common goal' moderator as 'basic goal' moderator, see https://osf.io/zgqub/.

[3] Please note that our specific inclusion criteria were extended during the coding of articles. Any changes made to the preregistered coding scheme are documented online by date (see https://osf.io/w8b9m/).

[4] Note that correlation coefficients *r* were included during the coding procedure as it turned out that some studies reported them, rather than regression coefficients with accompanying *t*-values. Additional exploratory coding that was not considered a priori is described online and corresponding exploratory analyses are reported in the supplementary document.

[5] A reviewer suggested Empirical Bayes model, which yielded a similar effect, $g = 0.31$, 95% cred. int. [0.22, 0.41].

[6] Note that Trim and Fill has been criticized by methodologists (Terrin et al., 2003; Simmonsohn et al., 2014). We also preregistered Orwin's Fail-Safe-N (Orwin, 1983), which has also been criticized and is recommended not to be used anymore (Becker, 2005; Simonsohn, Simmons, & Nelson, 2014).

[7] We would like to thank the anonymous reviewer who hinted at using the NNT (number needed to treat) as an intuitively interpretable effect size. We used Control Event Rates of 0.2 and 0.5 to reach 17 and 22 participants, respectively, see Magnussen (2020).

[8] For example, several studies (Loersh et al., 2008; Fast & Tiedens, 2010) included emotional and affective manipulations, but did not explicate how these correspond to the theory on GC. See more examples at https://osf.io/dkxsr/.

Acknowledgements

We would like to thank Larissa Titze, Edda Pavalec and Marlies "Maestra" Brunnhofer for their assistance and Malte Friese for his comments on an earlier version of this meta-analysis.

Funding Information

Katja Corcoran was supported by the Austrian Science Fund (FWF P-28393). Robbie C. M. van Aert was supported by the H2020 European Research Council (726361 IMPROVE). The views, opinions, and/or findings contained in this paper are those of the authors and shall not be construed as an official position, policy, or decision of the either the Austrian Science Fund or the European Research Council.

Competing Interests

The authors have no competing interests to declare.

References

Note that asterisks "*" indicate articles that were coded for the meta-analysis and can be found here: https://osf.io/w8b9m/.

Aarts, H., Dijksterhuis, A., & **Dik, G.** (2008). Goal contagion: Inferring goals from other's actions – And what it leads to. In J. Y. Shah & W. Gardner (Eds.), *Handbook of motivation science* (pp. 265–280). New York: Guilford.

***Aarts, H., Gollwitzer, P. M.,** & **Hassin, R. R.** (2004). Goal contagion: Perceiving is for pursuing. *Journal of Personality and Social Psychology*, *87*(1), 23–37. DOI: https://doi.org/10.1037/0022-3514.87.1.23

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & **Trötschel, R.** (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*(6), 1014–1027. DOI: https://doi.org/10.1037/0022-3514.81.6.1014

Becker, B. J. (2005). Failsafe N or file-drawer number. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 111–125). Chichester, UK: Wiley. DOI: https://doi.org/10.1002/0470870168.ch7

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & **Rothstein, H. R.** (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. DOI: https://doi.org/10.1002/jrsm.12

***Bouquet, C. A., Shipley, T. F., Capa, R. L.,** & **Marshall, P. J.** (2011). Motor contagion goal-directed actions are more contagious than non-goal-directed actions. *Experimental Psychology*, *58*(1), 71–78. DOI: https://doi.org/10.1027/1618-3169/a000069

***Brescoll, V. L., Uhlmann, E. L.,** & **Newman, G. E.** (2013). The effects of system-justifying motives on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology*, *105,* 891–908. DOI: https://doi.org/10.1037/a0034701

***Brohmer, H., Corcoran, K., Kedia, G., Eckerstorfer, L. V., Fauler, A.,** & **Floto, C.** (2018). PREPRINT Inspired to lend a hand? *Attempts to elicit prosocial behavior through goal contagion*. DOI: https://doi.org/10.31219/osf.io/85wvp

Carver, C. S., & **Scheier, M. F.** (1981). Attention and self-regulation: A control-theory approach to human behavior. *SSSP Springer Series in Social Psychology.* New York, NY: Springer New York. DOI: https://doi.org/10.1007/978-1-4612-5887-2

Carver, C. S., & **Scheier, M. F.** (2012). Cybernetic control processes and the self-regulation of behavior. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 28–42). Oxford University Press, USA. DOI: https://doi.org/10.1093/oxfordhb/9780195399820.013.0003

Chambers, C. (2017). The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice. Princeton, Oxford: Princeton University Press. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1431831. DOI: https://doi.org/10.1515/9781400884940

Chambers, C. (2019). What's next for Registered Reports? *Nature*, *573*(7773), 187. DOI: https://doi.org/10.1038/d41586-019-02674-6

Chartrand, T. L., & **Lakin, J. L.** (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology*, *64*, 285–308. DOI: https://doi.org/10.1146/annurev-psych-113011-143754

***Chen, X.,** & **Latham, G. P.** (2014). The effect of priming learning vs. performance goals on a complex task. *Organizational Behavior and Human Decision Processes*, *125*(2), 88–97. DOI: https://doi.org/10.1016/j.obhdp.2014.06.004

Coburn, K. M., & **Vevea, J. L.** (2019). Package 'weightr'. Retrieved from https://cran.r-project.org/web/packages/weightr/weightr.pdf

***Corcoran, K., Brohmer, H., Eckerstorfer, L. V.,** & **Macher, S.** (2018). When your goals inspire my goals: The role of effort, personal value, and inference in goal contagion. Stage 1 Registered Report. Retrieved from https://osf.io/8qtfk/

Crede, M. (2019). A negative effect of a contractive pose is not evidence for the positive effect of an expansive pose: Comment on Cuddy, Schultz, and Fosse (2018). *Meta- Psychology*, *3*, 1–5. DOI: https://doi.org/10.15626/MP.2019.1723

Cuddy, A. J., Schultz, S. J., & **Fosse, N. E.** (2018). P-curving a more comprehensive body of research on postural feedback reveals clear evidential value for power-posing effects: Reply to Simmons and Simonsohn (2017). *Psychological Science*, *29*(4), 656–666. DOI: https://doi.org/10.1177/0956797617746749

De Houwer, J., & **Moors, A.** (2010). Implicit measures: Similarities and differences. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 176–193). New York, London: The Guilford Press.

Del Re, A. C. (2015). Package 'compute.es'. Retrieved from https://cran.r-project.org/web/packages/compute.es/compute.es.pdf

***Dik, G.,** & **Aarts, H.** (2007). Behavioral cues to others' motivation and goal pursuits: The perception of effort facilitates goal inference and contagion. *Journal of Experimental Social Psychology*, *43*(5), 727–737. DOI: https://doi.org/10.1016/j.jesp.2006.09.002

Duval, S., & **Tweedie, R.** (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. DOI: https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & **Minder, C.** (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. DOI: https://doi.org/10.1136/bmj.315.7109.629

***Fast, N. J.,** & **Tiedens, L. Z.** (2010). Blame contagion: The automatic transmission of self-serving attributions. *Journal of Experimental Social Psychology*, *46*, 97–106. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=psyc7&AN=2009-22616-005. DOI: https://doi.org/10.1016/j.jesp.2009.10.007

Fisher, Z., Tipton, E., & **Zhipeng, H.** (2017). Package 'robumeta'. Retrieved from https://cran.r-project.org/web/packages/robumeta/robumeta.pdf

Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, *19*(6), 975–991. DOI: https://doi.org/10.3758/s13423-012-0322-y

Friese, M., Frankenbach, J., Job, V., & **Loschelder, D. D.** (2017). Does self-control training improve self-control? A meta-analysis. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *12*(6), 1077–1099. DOI: https://doi.org/10.1177/1745691617697076

Giner-Sorolla, R. Powering Your Interaction. Retrieved from https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/* (Original work published 2018).

Guest, O., & **Martin, A. E.** (2020). How computational modeling can force theory building in psychological science. Retrieved from PsyArXiv DOI: https://doi.org/10.31234/osf.io/rybh9

***Hassin, R. R., Aarts, H.,** & **Ferguson, M. J.** (2005). Automatic goal inferences. *Journal of Experimental Social Psychology*, *41*(2), 129–140. DOI: https://doi.org/10.1016/j.jesp.2004.06.008

Hassin, R. R., Bargh, J. A., & **Uleman, J. S.** (2002). Spontaneous causal inferences. *Journal of Experimental Social Psychology*, *38*(5), 515–522. DOI: https://doi.org/10.1016/S0022-1031(02)00016-1

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. DOI: https://doi.org/10.2307/1164588

***Jia, L., Koh, A. H. Q.,** & **Tan, F. M.'e.** (2018). Asymmetric goal contagion: Social power attenuates goal contagion among strangers. *European Journal of Social Psychology*, *48*(5), 673–686. DOI: https://doi.org/10.1002/ejsp.2360

***Jia, L., Tong, E. M. W.,** & **Lee, L. N.** (2014). Psychological "gel" to bind individuals' goal pursuit: Gratitude facilitates goal contagion. *Emotion*, *14*, 748–760.

Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=psyc11&AN=2014-14284-001. DOI: https://doi.org/10.1037/a0036407

Jolly, R. (1976). The World Employment Conference: The Enthronement of Basic Needs. *Development Policy Review*, *A9*(2), 31–44. DOI: https://doi.org/10.1111/j.1467-7679.1976.tb00338.x

Kruglanski, A., & **Kopetz, C.** (2009). What is so special (and nonspecial) about goals? A view from the cognitive perspective. In G. B. Moskowitz (Eds.), *The psychology of goals* (pp. 27–55). New York, NY, US: Guilford Press.

Kühberger, A., Fritz, A., & **Scherndl, T.** (2014). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS ONE*, *9*(9), e105825. DOI: https://doi.org/10.1371/journal.pone.0105825

Lane, D. M., & **Dunlap, W. P.** (1978). Estimating effect size: Bias resulting from the significance criterion in editorial decisions. *British Journal of Mathematical & Statistical Psychology*, *31*, 107–112. DOI: https://doi.org/10.1111/j.2044-8317.1978.tb00578.x

***Latham, G. P., Brcic, J.,** & **Steinhauer, A.** (2017). Toward an integration of goal setting theory and the automaticity model. *Applied Psychology*, *66*(1), 25–48. DOI: https://doi.org/10.1111/apps.12087

***Latham, G. P.,** & **Piccolo, R. F.** (2012). The effect of context-specific versus nonspecific subconscious goals on employee performance. *Human Resource Management*, *51*(4), 511–523. DOI: https://doi.org/10.1002/hrm.21486

Laurin, K. (2016). Interpersonal influences on goals: Current and future directions for goal contagion research. DOI: https://doi.org/10.1111/spc3.12289

***Laurin, K., Fitzsimons, G. M., Finkel, E. J., Carswell, K. L., van Dellen, M. R., Hofmann, W., ... Brown, P. C.** (2016). Power and the pursuit of a partner's goals. *Journal of Personality and Social Psychology*, *110*, 840–868. DOI: https://doi.org/10.1037/pspi0000048

***Leander, N. P.,** & **Shah, J. Y.** (2013). For whom the goals loom: Context-driven goal contagion. *Social Cognition*, *31,* 187–200. DOI: https://doi.org/10.1521/soco.2013.31.2.187

***Leander, N. P., Shah, J. Y.,** & **Chartrand, T. L.** (2011). The object of my protection: Shielding fundamental motives from the implicit motivational influence of others. *Journal of Experimental Social Psychology*, *47*(6), 1078–1087. DOI: https://doi.org/10.1016/j.jesp.2011.04.016

***Leander, N. P., Van Dellen, M. R., Rachl-Willberger, J., Shah, J. Y., Fitzsimons, G. J.,** & **Chartrand, T. L.** (2016). Is freedom contagious? A self-regulatory model of reactance and sensitivity to deviant peers. *Motivation Science*, *2*(4), 256–267. DOI: https://doi.org/10.1037/mot0000042

***Lee, T. K.,** & **Shapiro, M. A.** (2015). Effects of a story character's goal achievement: Modeling a story character's diet behaviors and activating/deactivating a character's diet goal. *Communication Research*. Advance online publication. DOI: https://doi.org/10.1177/0093650215608236

***Loersch, C., Aarts, H., Keith Payne, B.,** & **Jefferis, V. E.** (2008). The influence of social groups on goal contagion. *Journal of Experimental Social Psychology*, *44*(6), 1555–1558. DOI: https://doi.org/10.1016/j.jesp.2008.07.009

Maassen, E., Olsson-Collentine, A. O., Wicherts, J., van Assen, M. A. L. M., & **Nuijten, M. B.** (2019, March). Investigating the Reproducibility of Psychological Meta-Analyses. *International Convention of Psychological Science*, Paris, F. Retrieved from osf.io/7nsmd/

Magnussen, K. (2020). Interpreting Cohen's d Effect Size. Teaching tool retrieved from https://rpsychologist.com/d3/cohend/

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. DOI: https://doi.org/10.1037/h0054346

***McCulloch, K. C., Fitzsimons, G. M., Chua, S. N.,** & **Albarracín, D.** (2011). Vicarious goal satiation. *Journal of Experimental Social Psychology*, *47*(3), 685–688. DOI: https://doi.org/10.1016/j.jesp.2010.12.019

Morgenroth, T., Ryan, M. K., & **Peters, K.** (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology*, *19*(4), 465–483. DOI: https://doi.org/10.1037/gpr0000059

Moskowitz, G. B., & **Gesundheit, Y.** (2009). Goal Priming. In G. B. Moskowitz & H. Grant (Eds.), *The Psychology of Goals* (pp. 203–233). New York, London: Guilford Press.

O'Donnell, M., Nelson, L. D., Ackermann, E., Aczel, B., Akhtar, A., Aldrovandi, S., ... Zrubka, M. (2018). Registered Replication Report: Dijksterhuis and van Knippenberg (1998). *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *13*(2), 268–294. DOI: https://doi.org/10.1177/1745691618755704

Orwin, R. G. (1983). A fail-safe for effect size in meta-analysis. *Journal of Educational Statistics*, *8*(2), 157–159. DOI: https://doi.org/10.3102/1076998600800 2157

***Palomares, N. A.** (2013). When and how goals are contagious in social interaction. *Human Communication Research*, *39*(1), 74–100. DOI: https://doi.org/10.1111/j.1468-2958.2012.01439.x

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. DOI: https://doi.org/10.1177/2515245917745629

Schimmack, U., & **Brunner, J.** (2017). Z-curve. *A method for the estimating replicability based on test statistics in original studies*. Preprint retrieved from the OSF. DOI: https://doi.org/10.31219/osf.io/wr93f

Schimmack, U., & **Brunner, J.** (2019). Estimating Replicability with Z-Curve. Retrieved from https://zcurve.shinyapps.io/zcurve19/

Shah, J. Y., Kruglanski, A. W., & **Friedman, R.** (2003). Goal systems theory: Integrating the cognitive and motivational aspects of self-regulation. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.). *Motivated social perception: The Ontario symposium, 9*, 247–275. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

***Shantz, A.,** & **Latham, G. P.** (2009). An exploratory field experiment of the effect of subconscious and conscious goals on employee performance. *Organizational Behavior and Human Decision Processes*, *109*(1), 9–17. DOI: https://doi.org/10.1016/j.obhdp.2009.01.001

***Shantz, A.,** & **Latham, G.** (2011). The effect of primed goals on employee performance: Implications for human resource management. *Human Resource Management*, *50*(2), 289–299. DOI: https://doi.org/10.1002/hrm.20418

Simmons, J. P., & **Simonsohn, U.** (2017). Power posing: P-curving the evidence. *Psychological Science*, *28*(5), 687–693. DOI: https://doi.org/10.1177/0956797616658563

Simonsohn, U., Nelson, L. D., & **Simmons, J. P.** (2014). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *9*(6), 666–681. DOI: https://doi.org/10.1177/1745691614553988

Simonsohn, U., Nelson, L. D., & **Simmons, J. P.** (2019). P-curve app. Retrieved from http://www.p-curve.com/app4/. DOI: https://doi.org/10.1038/s41562-020-00974-w

Simonsohn, U., Simmons, J., & **Nelson, L.** (2014). Trim-and-fill is full of it (bias). Retrived from https://datacolada.org/30

Smaldino, P. (2019). Better methods can't make up for mediocre theory. *Nature*, *575*(7781), 9. DOI: https://doi.org/10.1038/d41586-019-03350-5

Soetens, K. C., Vandelanotte, C., de Vries, H., & **Mummery, K. W.** (2014). Using online computer tailoring to promote physical activity: A randomized trial of text, video, and combined intervention delivery modes. *Journal of Health Communication*, *19*(12), 1377–1392. DOI: https://doi.org/10.1080/10810730.2014.894597

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, *70*(1), 103–127. DOI: https://doi.org/10.1111/j.1468-0084.2007.00487.x

Terrin, N., Schmid, C. H., Lau, J., & **Olkin, I.** (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, *22*(13), 2113–2126. DOI: https://doi.org/10.1002/sim.1461

***Tobin, S. J., Greenaway, K. H., McCulloch, K. C.,** & **Crittall, M. E.** (2015). The role of motivation for rewards in vicarious goal satiation. *Journal of Experimental Social Psychology*, *60*, 137–143. DOI: https://doi.org/10.1016/j.jesp.2015.05.010

Van Aert, R. C. M., & **van Assen, M. A. L. M.** (2018). Examining reproducibility in psychology: A hybrid method for combining a statistically significant original study and a replication. *Behavior Research Methods*, *50*(4), 1515–1539. DOI: https://doi.org/10.3758/s13428-017-0967-6

Van Aert, R. C. M., & **van Assen, M. A. L. M.** (2019). Correcting for publication bias in a meta-analysis with the p-uniform* method. Manuscript submitted for publication. Retrieved from https://osf.io/preprints/bitss/zqjr9

Van Aert, R. C. M., Wicherts, J. M., & **van Assen, M. A. L. M.** (2016). Conducting meta- analyses based on p values: Reservations and recommendations for applying p-uniform and p-curve. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *11*(5), 713–729. DOI: https://doi.org/10.1177/1745691616650874

Van Assen, M. A. L. M., van Aert, R. C. M., & **Wicherts, J. M.** (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, *20*(3), 293–309. DOI: https://doi.org/10.1037/met0000025

***Van der Weiden, A., Veling, H.,** & **Aarts, H.** (2010). When observing gaze shifts of others enhances object desirability. *Emotion*, *10*(6), 939–943. DOI: https://doi.org/10.1037/a0020501

Vevea, J. L., & **Hedges, L. V.** (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, *60*(3), 419–435. DOI: https://doi.org/10.1007/BF02294384

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

Viechtbauer, W. (2017). Package 'metafor'. Retrieved from https://cran.r-project.org/web/packages/metafor/metafor.pdf

Walthouwer, M. J. L., Oenema, A., Lechner, L., & **de Vries, H.** (2015). Comparing a video and text version of a web-based computer-tailored intervention for obesity prevention: A randomized controlled trial. *Journal of Medical Internet Research*, *17*(10), e236. DOI: https://doi.org/10.2196/jmir.4083

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & **Albarracín, D.** (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, *142*(5), 472–497. DOI: https://doi.org/10.1037/bul0000030

***Wessler, J.,** & **Hansen, J.** (2016). The effect of psychological distance on automatic goal contagion. *Comprehensive Results in Social Psychology*, *1*(1–3),

51–85. DOI: https://doi.org/10.1080/23743603.2017.1288877

Wickham, H. (2019). Package 'ggplot2'. Retrived from https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

***Zhou, S., Shapiro, M. A.,** & **Wansink, B.** (2017). The audience eats more if a movie character keeps eating: An unconscious mechanism for media influence on eating behaviors. *Appetite*, *108*, 407–415. DOI: https://doi.org/10.1016/j.appet.2016.10.028

How to cite this article: Brohmer, H., Eckerstorfer, L. V., van Aert, R. C. M., & Corcoran, K. (2021). Do Behavioral Observations Make People Catch the Goal? A Meta-Analysis on Goal Contagion. *International Review of Social Psychology*, 34(1): 3, 1–15. DOI: https://doi.org/10.5334/irsp.428

Submitted: 04 April 2020 **Accepted**: 24 November 2020 **Published:** 22 January 2021

]u[*International Review of Social Psychology* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS