More green than gray? Toward a sustainable overview of environmental spillover effects: A Bayesian meta-analysis

Geiger, S.J.; Brick, C.; Nalborczyk, L.; Bosshard, A.; Jostmann, N.B.

DOI
10.1016/j.jenvp.2021.101694

Publication date
2021

Document Version
Final published version

Published in
Journal of Environmental Psychology

Journal of Environmental Psychology 78 (2021) 101694

Contents lists available at ScienceDirect

Journal of Environmental Psychology

journal homepage: www.elsevier.com/locate/jep

More green than gray? Toward a sustainable overview of environmental spillover effects: A Bayesian meta-analysis☆

Sandra J. Geiger[a,b,*], Cameron Brick[a], Ladislas Nalborczyk[c], Anna Bosshard[a], Nils B. Jostmann[a]

[a] Department of Social Psychology, University of Amsterdam, Amsterdam, the Netherlands
[b] Environmental Psychology, Department of Cognition, Emotion, and Methods, Faculty of Psychology, University of Vienna, 1010 Vienna, Austria
[c] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

ARTICLE INFO

Handling Editor: Sander van der Linden

Keywords:
Spillover
Pro-environmental behavior
Conservation
Bayesian
Meta-analysis
Open science

ABSTRACT

In response to climate change, interventions have been implemented to encourage sustainable behavior. Such interventions may not only promote the target behavior but also increase (positive spillover) or reduce (negative spillover) non-targeted outcomes. This pre-registered meta-analysis integrated the experimental research on environmental spillover to update a previous meta-analysis (Maki et al., 2019). Database searches in several languages supplemented by searches to retrieve unpublished literature yielded 63 aggregated effect sizes from 38 studies and 29 articles (N = 26,613 unique participants). A three-level Bayesian meta-analysis provided weak support for no spillover on intentions and strong support for no spillover on behaviors. If spillover was present, it would likely be small and positive for intentions, $\delta = 0.15$, 95% CrI [-0.01, 0.31], but negligible for behaviors, $\delta = 0.01$, 95% CrI [-0.13, 0.16]. Positive spillover was most likely when interventions were autonomy-supportive (very strong evidence), provided a rationale (moderate to strong evidence), did not use financial (dis)incentives (weak to strong evidence), and addressed normative (extreme evidence) or a combination of normative and personal gain goals (strong evidence). Spillover was similar across research settings (moderate evidence) and partly across samples (weak to moderate evidence), which may suggest generalizability. To set standards for robust spillover research, we developed the Power-Reporting-Open science (PRO) guidelines. The Bayesian approach allows for robust conclusions and continuous updating with new evidence. We hope that this supports future revisions toward a sustainable overview of robust and high-powered spillover studies that independent researchers can easily update.

1. Introduction

Individuals are faced with environmental choices every day—whether taking the subway instead of the car or throwing away waste instead of recycling. When opting against the sustainable choice, individuals often contribute to climate change (Milinski et al., 2008). In response to climate change, governments worldwide have introduced interventions to encourage intentions and behaviors that protect or avoid harming the environment (i.e., sustainable intentions and behaviors; Steg & Vlek, 2009). While most intervention studies measure the target behavior (Abrahamse et al., 2005), effects on other outcomes, so-called *spillover effects*, often go unnoticed (Nilsson et al., 2017; Truelove et al., 2014). To illustrate, when free public transportation was introduced in several European cities (The Guardian, 2020), this boosted commuting but also may have promoted (positive spillover) or reduced (negative spillover) other sustainable actions, such as recycling.

Along a continuum, spillover can be positive or 'green', negative or 'gray', or absent (Dolan & Galizzi, 2015). Positive spillover may result

☆ This meta-analysis was pre-registered on the Open Science Framework (OSF; https://osf.io/u67dp). The datasets and analysis scripts are publicly available on the OSF (https://osf.io/tu7yx/). We have no conflict of interest to disclose. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We thank Quentin F. Gronau for his advice on the analysis, Janneke Staaks for her support with developing the search strategy, and Alexander Maki for his valuable insights on his meta-analysis on spillover effects.

* Corresponding author. Environmental Psychology, Department of Cognition, Emotion, and Methods, Faculty of Psychology, University of Vienna, Wächtergasse 1, 1010 Vienna, Austria.

E-mail address: sandra.geiger@univie.ac.at (S.J. Geiger).

https://doi.org/10.1016/j.jenvp.2021.101694
Received 24 February 2021; Received in revised form 18 September 2021; Accepted 19 September 2021
Available online 6 October 2021

from social and self-identity processes as well as a desire for consistency (Truelove et al., 2014). This means that positive spillover may occur when individuals identify with environmental groups (social identity; Dono et al., 2010; Fielding et al., 2008), or when they see themselves as environmentally friendly (self-identity; Van der Werff, Steg, & Keizer, 2014) which may prompt further sustainable actions to mitigate cognitive dissonance (Festinger, 1962). Negative spillover can be explained in terms of moral licensing (Truelove et al., 2014) which implies that individuals justify indulgent behaviors (e.g., taking the car) based on past behaviors considered as morally right (e.g., sustainable consumption). Several studies have, however, questioned the replicability of moral licensing (Blanken et al., 2014; Urban, Bahník, & Kohlová, 2019). Finally, spillover can be absent, which may happen when individuals are unaware of the environmentally harmful impact of the non-targeted behaviors (Reams et al., 1996), or when interventions activate both positive and negative spillover mechanisms that ultimately balance (Lacasse, 2016).

Taken together, various mechanisms can result in different spillover effects. This is also reflected in the mixed empirical evidence (Nash et al., 2017) showing positive (e.g., Lanzini & Thøgersen, 2014), negative (e.g., Thøgersen & Ölander, 2003), and null effects (e.g., Schütte & Gregory-Smith, 2015). A previous frequentist meta-analysis (Maki et al., 2019) synthesized the evidence until 2017 and included 13 articles with 20 studies and 60 spillover effects on intentions and behaviors. Spillover was positive but small for intentions ($d = 0.17$, 95% CI [0.05, 0.29], $p < .01$) and negative albeit very small for behaviors ($d = -0.03$, 95% CI [-0.06, -0.01], $p < .05$). Positive spillover most likely occurred when interventions triggered intrinsic motivation and neither used (dis)incentives nor guilt, when intentions rather than behaviors were measured, as well as when the targeted and non-targeted outcomes were similar.

1.1. The present meta-analysis

Relative to Maki et al. (2019), the present meta-analysis included 16 additional articles and 18 additional studies, totaling to 183 non-aggregated and 63 aggregated effect sizes from 38 studies and 29 articles. It primarily aimed to update the previous work by investigating whether interventions change non-targeted sustainable (a) intentions and (b) behaviors (Research Question 1). It also investigated seven moderators, including the type of non-targeted outcome (intentions vs. behaviors), intervention characteristics (autonomy support, rationale provision, financial (dis)incentive, and goal), and characteristics regarding the generalizability of spillover (research setting and sample type). In contrast to the previous meta-analysis, this work employed a Bayesian approach to meta-analysis. The Bayesian approach has three major advantages: Firstly, it allows robust conclusions about the presence or absence of spillover (Gronau et al., 2017; Wagenmakers, 2017), which is important for practical application. Secondly, it helps to accurately assess the strength of findings and prevents "the field from incorporating ambiguous findings as if these were real and reliable" (Wetzels et al., 2011, p. 296). Lastly, unlike traditional frequentist meta-analyses, Bayesian analyses can be continuously updated without inflating false-positive rates. Updating frequentist meta-analyses can entail a two- to five-fold inflation of false-positive rates, thereby outweighing the inflation caused by publication bias (Borm & Donders, 2009). The field of environmental spillover will likely grow rapidly in the next years, calling for further meta-analytical updates. As such, the Bayesian meta-analysis represents a further step toward an accurate, sustainable, and constantly evolving overview of spillover research (Simmonds et al., 2017).

1.2. Potential moderators of spillover effects

As suggested by the mixed evidence, spillover may be contingent on several factors. This meta-analysis investigated seven potential moderators: the type of non-targeted outcome (intentions vs. behaviors), autonomy support (whether interventions contained autonomy-supportive vs. controlling elements), rationale provision (whether interventions provided a rationale why individuals should engage in the target behavior), financial (dis)incentives, intervention goal (whether interventions addressed normative goals, such as environmental protection, and/or personal gain goals, such as saving money), research setting (laboratory, online, or field), and sample type (university student vs. non-student sample).

1.2.1. Type of non-targeted outcome: Intentions vs. behaviors

Spillover effects may depend on whether intentions or behaviors are assessed. Although intentions are an important precursor to actual behavior (Ajzen, 1991), interventions that change intentions do not necessarily trigger corresponding behaviors (intention-behavior gap; Webb & Sheeran, 2006). The intention-behavior gap thus implies that intentions are easier to influence than behaviors (Sheeran & Webb, 2016), which also applies to the environmental domain where intentions to recycle rarely result in recycling behavior (Echegaray & Hansstein, 2017). In line with this, the previous meta-analysis (Maki et al., 2019) found stronger spillover on intentions than behaviors. We thus expected that interventions change non-targeted intentions more than behaviors (Hypothesis 1).

1.2.2. Autonomy-supportive vs. controlling interventions

Spillover may also vary depending on whether interventions contain autonomy-supportive or controlling elements. Self-determination theory (Deci & Ryan, 2012) broadly distinguishes between two types of motivation—autonomous and controlled motivation. Autonomous motivation implies that individuals act voluntarily, while controlled motivation means that individuals act due to pressure or control. Interventions support autonomous motivation when they preserve a sense of choice (e.g., perceived choice of whether to engage in a behavior or which behavior to adopt), provide an explicit rationale (e.g., understanding about why the behavior is important, either for individuals or the environment), involve perspective-taking (e.g., acknowledgement of difficulty of the behavior), or use non-controlling language (Bartholomew et al., 2009; Bradshaw et al., 2021; Deci et al., 1994; Legate et al., 2021; Su & Reeve, 2011). By contrast, interventions likely trigger controlling motivation when pressuring individuals to act in a certain way, for example by using bans, punishments, coercive language (e.g., must or should; Bradshaw et al., 2021), blaming, shaming, injunctive norms, evaluative feedback, or (dis)incentives (for more details on the coding, see Appendix A).

Autonomous but not controlled motivation has been associated with more frequent and stable sustainable intentions (Osbaldiston & Sheldon, 2003), engagement in a wide range of sustainable behaviors (Green-Demers et al., 1997), and spillover from sustainable behaviors within to outside the workplace (Hicklenton et al., 2019). By contrast, controlled motivation has been linked to negative spillover (Thomas et al., 2016), and incentivized interventions have been associated with less positive/more negative spillover than non-incentivized interventions (Maki et al., 2019).

In line with this argumentation, we expected that interventions with autonomy-supportive elements are associated with more positive changes in non-targeted (a) intentions and (b) behaviors than interventions with controlling elements (Hypothesis 2). Additionally, we investigated rationale provision as a prominent example of an autonomy-supportive intervention and financial (dis)incentives as a prominent example of a controlling intervention. We hypothesized that interventions with compared to without a rationale are associated with more positive changes in non-targeted (a) intentions and (b) behaviors (Hypothesis 3). Interventions with as compared to without financial (dis)incentives were expected to be associated with more negative changes in non-targeted (a) intentions and (b) behaviors (Hypothesis 4). Hypothesis 2 was modified and Hypothesis 3 and 4 were added after peer review but before data extraction of these additional moderators.

1.2.3. Intervention goal

Spillover effects may also depend on which type of goals interventions address. Goal-framing theory (Lindenberg & Steg, 2007, 2013) posits that interventions can trigger overarching goals to become focal: normative goals to adopt appropriate intentions and behaviors (doing what is right, e.g., environmental protection), personal gain goals (e.g., improving one's health), or both types of goals. The focal goal can then activate intentions and behaviors that align with this goal. For example, normative goals can trigger sustainable purchases because individuals aim to protect the environment, while personal gain goals can yield the same direct effect but due to different motives, such as health benefits of sustainable food products.

Although both goals can be effective at promoting the target behavior, their spillover effects may differ. Normative goals are likely to cause more positive spillover than personal gain goals or a combination of both types. Normative goals activate pro-environmental goals and identities which facilitate non-targeted sustainable intentions and behaviors (Thøgersen & Noblet, 2012; Truelove et al., 2014; Xu et al., 2018a). In contrast, personal gain goals activate self-serving goals as well as non-targeted intentions and behaviors from which individuals personally benefit. This limits positive spillover because sustainable intentions and behaviors (e.g., buying relatively expensive sustainable products) often contradict personal gain goals (e.g., saving money; Milinski et al., 2008). Personal gain goals can also weaken positive spillover by crowding out intrinsic motivation (Xu et al., 2018a). Supporting these claims, studies have shown that normative goals are associated with more positive spillover (Evans et al., 2013; Steinhorst et al., 2015; Steinhorst & Matthies, 2016) compared to personal gain goals which can even induce negative effects (Geng et al., 2019). These described mechanisms also apply to interventions that emphasize both types of goals. Personal gain goals can undermine normative goals for the target outcome (Schwartz et al., 2015), presumably carrying over to non-targeted intentions and behaviors.

We hypothesized that interventions that appeal to normative rather than personal gain (Hypothesis 5) or both types of goals (Hypothesis 6) are associated with more positive changes in non-targeted (a) intentions and (b) behaviors.

1.2.4. Generalizability: Research setting and sample type

Spillover research has been conducted in the laboratory, online, and in the field (Galizzi & Whitmarsh, 2019). Effect sizes typically vary depending on the setting of the intervention (Cadario & Chandon, 2020; Mitchell, 2012). The magnitude of spillover may thus depend on situational factors inherent to specific settings. Spillover may be smaller in the field compared to more controlled settings, as field research may involve more costly and difficult behaviors (e.g., switching to a green energy provider) that have been shown to reduce spillover (Maki et al., 2019). Effects may, however, also be larger in the field due to factors operating only in the field, such as peer influences or establishing a green reputation (Babutsidze & Chai, 2018). Alternatively, the same mechanisms may operate in different settings, inducing generalizable spillover. As research on this is still in its infancy, we explored whether spillover on (a) intentions and (b) behaviors differed between settings to inform generalizability of the results from laboratory to online and field settings (Research Question 2).

Spillover research has used both university student and non-student samples. In recent years, concerns arose about conducting psychological research with university student samples, as they may limit the generalizability of the results. Spillover effects may be comparable across student and non-student samples, as the underlying mechanisms (e.g., social identity, self-identity, desire for consistency, moral licensing) seem to suggest general phenomena. On the other hand, a recent meta-analysis found differences between student and non-student samples in the association between self- and social identity as well as sustainable behavior (Udall et al., 2021). The previous meta-analysis (Maki et al., 2019) found that spillover on intentions was similar, whereas it was less negative for behaviors in the general population compared to student samples. We explored whether spillover differed between student and non-student samples to further inform generalizability (Research Question 3).

2. Method

2.1. Open science practices

We adhered to the PRISMA-S and -P guidelines (Preferred Reporting Items for Systematic Review and Meta-Analysis Search Extension and Protocols Statement; Moher et al., 2015; Rethlefsen et al., 2019) as well as recommendations on reproducibility in meta-analyses (Lakens et al., 2016). We pre-registered a systematic review protocol on the Open Science Framework (April 18, 2020; https://osf.io/u67dp) before conducting the searches. We added an addendum to the pre-registration (June 21, 2021) that further specified Hypotheses 2 to 4 and Research Question 3. All deviations from these protocols are reported in Appendix B.

2.2. Meta-analytic criteria

To be included in this meta-analysis, articles needed to meet several criteria based on the PICOS framework, including population, intervention, comparator, outcome, and study design (Center for Review and Center for Dissemination, 2009).

As displayed in Table 1, articles needed to include human participants and adopt an experimental, quasi-experimental, or natural-experimental design with a control (no intervention) or comparison (alternative intervention) condition, as such robust designs allow valid conclusions about spillover effects (Galizzi & Whitmarsh, 2019). Articles were required to have reported results from an intervention (e.g., policy or experimental manipulation) that significantly increased the targeted sustainable intention or behavior (Ghesla et al., 2019) or a manipulation to prime participants with reminders of past behavior. Regarding outcomes, articles were only included if they assessed spillover on at least one sustainable intention or self-reported or objective behavior. Both published and unpublished empirical articles, doctoral dissertations, and master theses were eligible. Inclusion was further restricted to articles published in English, German, French, and Dutch. Articles were only considered for inclusion when their full text, relevant statistics, and information to assess eligibility were available online or after request to the authors.

Consequently, articles were excluded based on nine hierarchically structured, pre-registered criteria: other language, other publication type, non-environmental or other outcome, other study design, other intervention, other sample, other (e.g., contextual) or no spillover measured, unobtainable full text, and insufficient reporting. The tenth exclusion criterion of duplicated data was added after the pre-registration.

2.3. Literature search

We devised a comprehensive search strategy in consultation with librarians. We searched PsycINFO (Ovid), Business Source Premier (EBSCOhost), GreenFile (EBSCOhost), and the Social Science Citation Index (Web of Science) on April 20, 2020, and for an update on October 29, 2020. These searches were limited to title, abstract, keywords, and human participants (where possible) without date restrictions. The pilot-tested search terms for all databases are in Tables C1, C2, and C3. Additionally, the first 300 results of a full-text search in Google Scholar were screened for both the initial and updated search (Haddaway et al., 2015), and reference lists of included articles and past related reviews were hand-searched for eligible articles. The OSF preprint repository

Table 1
Inclusion and exclusion criteria based on PICOS and additional criteria.

PICOS Criterion	Inclusion	Exclusion
Population	Human sample	Other sample (6)
Intervention	Intervention (e.g., information provision, policy, charge, nudge, or experimental manipulation) that significantly increased at least one sustainable intention or behavior, or effectively primed participants with reminders of past behavior	Other intervention (5)
Comparator	Control (no intervention) or comparison (alternative intervention) condition	Other study design (4)
Outcome	Spillover on non-targeted sustainable intention or self-reported or objective sustainable behavior	Non-environmental or other outcome (e.g., attitude) (3)
	Statistical outcome: Cohen's *d*	Other or no spillover (7)
Study design	Experimental, quasi-experimental, or natural- experimental design	Other study design (4)
Additional Criterion	**Inclusion**	**Exclusion**
Language	English, German, French, or Dutch	Other language (1)
Publication type and status	Published or unpublished empirical articles, dissertations, or master theses	Other publication type (2)
Full text	Full text available online or one month after request to the authors	Full text unobtainable (8)
Sufficient reporting	Statistical information available online or one month after request to the authors	Insufficient reporting (9)
Duplicated data [a]	If two records (e.g., dissertation and journal article) were based on the same dataset, only the one with the more detailed reporting or the earlier record was included.	Duplicated data (10)

Note. [a] Not pre-registered. PICOS = population (P), intervention (I), comparator (C), outcome (O), and study design (S; Center for Review and Center for Dissemination, 2009). Numbers in parentheses indicate the hierarchy of exclusion criteria.

was scoped to retrieve unpublished literature (Table C4).

2.4. Screening process and data extraction

The retrieved articles were de-duplicated first in *Zotero* and then manually before importing them to *Rayyan* (Ouzzani et al., 2016). The screening comprised two stages—an initial screening based on title and abstract and a full-text eligibility selection including the data extraction. Both stages were conducted by a trained rater who was provided with guidelines on the meta-analytic criteria and data extraction (Appendix A). For the initial screening, the rater categorized the articles based on title and abstract as 'include' (i.e., fulfills inclusion criteria), 'exclude' (i.e., certainly fulfills at least one exclusion criterion), or 'unclear'. For the subsequent eligibility selection, the rater applied the same system to the full text of articles previously labeled as 'include' or 'unclear' and recorded the first reason for exclusion in the hierarchy. The rater then extracted the pre-specified study characteristics and statistical information from the included articles. To minimize subjective judgments, the moderators were coded by an additional independent rater. Discrepancies were resolved through discussion. Interrater reliability in terms of Cohen's kappa was significantly different from zero (all $ps < .001$) and was almost perfect for non-targeted outcome ($\kappa = 0.94$), autonomy support ($\kappa = 0.84$), rationale provision ($\kappa = 0.83$), financial (dis) incentives ($\kappa = 1.00$), intervention goal ($\kappa = 0.93$), and sample type ($\kappa = 0.96$), but only substantial for setting ($\kappa = 0.62$).

2.4.1. Study characteristics

General information (authors, year, title, journal) and extrinsic characteristics (country, language, publication status, peer-reviewed) were documented alongside the PICOS elements. Age, gender distribution, sample type, total sample size, and sample size per condition were extracted as sample characteristics. Characteristics of the intervention were coded, including effectiveness regarding the target outcome, type and a narrative description, autonomy support (autonomy-supportive, controlling, or both kinds of elements), rationale provision, financial (dis)incentives, and intervention goal (normative, personal gain, or both types of goals). Moreover, characteristics of the design were distinguished, including study design (laboratory, online, field, or combination), sample type (university student or non-student sample), and comparator (control or comparison). For the targeted and non-targeted outcomes, their type was coded (intention, self-reported behavior, or objective behavior); for spillover, their direction was coded (positive, negative, or zero).

2.4.2. Computation and transformation of effect sizes

To compute Cohen's *d*, means and standard deviations (*SDs*) of the non-targeted outcome, sample size per condition, effect sizes, test statistics, and *p*-values were extracted. If any information was only graphically available, *WebPlotDigitizer* (Rohatgi, 2019) was used for data extraction. The available information was extracted in the following preference order: interaction effect of the intervention and targeted outcome on the non-targeted outcome, effect of the intervention on the non-targeted outcome, or means and *SDs* of the non-targeted outcome for the control and intervention condition. If spillover was reported for several measurement points, the statistical information was extracted from the first post-intervention measurement point. If the information was reported in a way other than Cohen's *d*, this information was transformed using the *esc* package (Version 0.5.1, Lüdecke, 2019), the *compute.es* package (Version 0.2–2, Del Re, 2013), or the Campbell Collaboration effect size calculator (Wilson, 2001).

3. Results

3.1. Search results and study characteristics

The literature search identified 4,465 articles through electronic database searches and an additional 1,227 articles through other sources, totaling 5,692 articles (Fig. 1). After deduplication, 4,557 articles remained for the screening based on title and abstract. Of those, 4,168 articles did not meet the inclusion criteria, while the remaining 389 articles were assessed for eligibility based on their full text. Thereby, 360 articles were excluded based on ten criteria: language ($n = 1$), publication type ($n = 54$), non-environmental ($n = 81$), study design ($n = 63$), intervention ($n = 49$), sample ($n = 0$), other (i.e., contextual) or no spillover measured ($n = 100$), unobtainable full text ($n = 5$), insufficient reporting ($n = 5$), duplicated data ($n = 2$). This yielded 183 extracted and 63 aggregated effect sizes from 55 interventions, 38 studies, and 29 articles.

The 29 articles were written between 2011 and 2021, and most were published and peer-reviewed ($n = 23$). All studies employed experimental ($n = 27$), quasi-experimental ($n = 10$), or natural experimental ($n = 1$) designs. Overall, they included 26,613 unique participants with a mean age of 44.1 years and 54% female participants (where reported). A summary and more details are in Table 2.

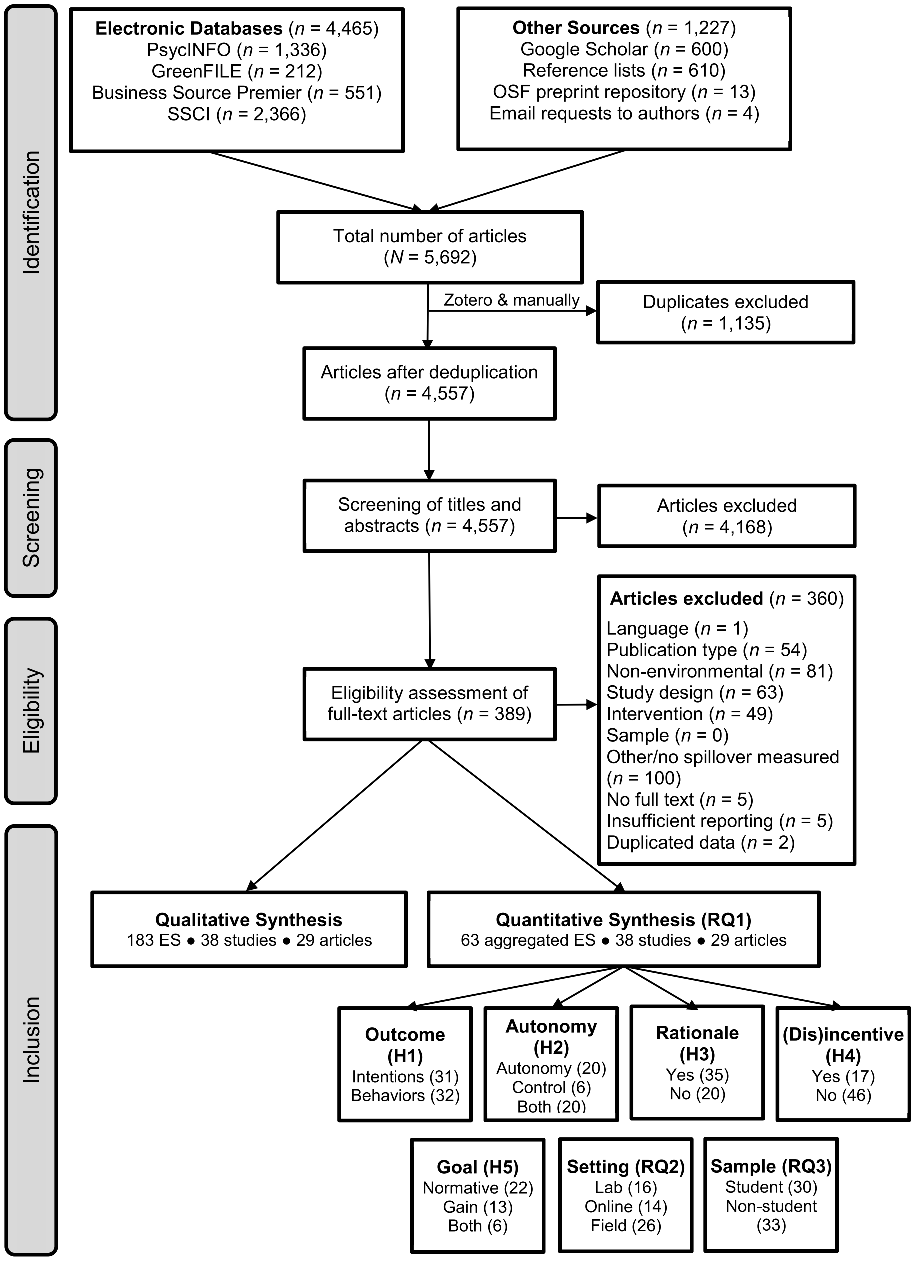

Fig. 1. PRISMA Flow Diagram.
Note. ES = Effect sizes. SSCI = Social Science Citation Index. The flow diagram was based on the template and recommendations provided by Vu-Ngoc et al. (2018) and Moreau and Gamble (2020).

3.2. Overview of included articles

The morphological box in Fig. 2 summarizes the most prevalent characteristics of the final set of articles, including three categories and 13 dimensions: intervention (type, autonomy, rationale provision, (dis) incentive, and goal), spillover effect (measurement, non-targeted outcome, direction, and magnitude), and method (comparator, setting, sample, and country). The intervention and method dimensions were evaluated at the level of interventions (N = 55), while the spillover dimension was evaluated at the level of non-aggregated effect sizes (N = 183).

The most frequent interventions were combined interventions (e.g., incentives and information; n = 15), framing (n = 9), others (e.g., shopping lists and imagining engaging in a pro-environmental activity; n = 9), changes in the structural environment (e.g., online store with sustainable products; n = 6), and (dis)incentives (n = 5). Most interventions had autonomy-supportive and controlling elements (n = 18), followed by only autonomy-supportive (n = 16) and only controlling (n = 6) elements. Most interventions provided a rationale (n = 29) and did not use financial (dis)incentives (n = 39). They either addressed a normative goal (n = 18), a personal gain goal (n = 12), a combination of goals (n = 6), or no/an unclear goal (n = 19).

Spillover was most frequently measured as self-reported behaviors (n = 91), followed by intentions (n = 73) and objective behaviors (n = 19). The most prominent non-targeted outcomes were energy saving (n = 48), recycling (n = 31), water consumption (n = 25), and dieting (n = 20). Remarkably, most reported spillover effects were non-significant (n = 140) and approximately equally many were negative (n = 18) or positive (n = 17). In line with this, 165 effect sizes were small, with few medium (n = 12) or large (n = 6) effect sizes.

Regarding methods, interventions were almost equally often compared to a control (n = 29) or alternative intervention condition (n

Table 2
Articles in the meta-analysis.

Article Information		Effect			Moderators						
Authors	Intervention	ES	*d*	*N*	Type	Autonomy	Rationale	(Dis) Incentive	Goal	Setting	Sample
Baca-Motes et al. (2013), S1	Commitment	1	0.41	567	B	Yes	Yes	No	Normative	Field	Non-student
Bergquist et al. (2019), S1	Combined intervention with norms, posters, incentive, and feedback	1	−0.62	48	B	Both	Yes	Yes	Both	Field	Student
Carlsson et al. (2020), S1	Social information campaign	1	0.01	768	B	Both	Yes	No	NA	Field	Non-student
Carrico et al. (2018), S1*	Combined intervention with information provision, diary, behavior change goal, reminders, and feedback	1	−0.25	272	B	Yes	Yes	No	Normative	Online	Non-student
Carrico et al. (2018), S1*	Information about dietary change to improve health	1	−0.48	262	B	Yes	Yes	No	Personal gain	Online	Non-student
Clot et al. (2013), S1	Incentive for imagined pro-environmental activity	1	−0.22	195	B	No	NA	Yes	Personal gain	Lab	Student
Clot et al. (2013), S1	No incentive for imagined pro-environmental activity	1	−0.30	192	B	Yes	NA	No	NA	Lab	Student
Elf et al. (2019), S1	Combined intervention with education, modeling, incentive, training, and persuasion	1	0.79	232	B	Both	Yes	Yes	Both	Field	Non-student
Geng et al. (2016), S1*	Shopping list with green products	1	−0.67	40	B	NA	No	No	NA	Lab	Student
Geng et al. (2016), S2*	Shopping list with green products	1	−0.68	40	I	NA	No	No	NA	Lab	Student
Geng et al. (2016), S3*	Goal progress and shopping list with green products	1	−0.99	83	I	NA	No	No	NA	Lab	Student
Geng et al. (2016), S3*	Goal commitment and shopping list with green products	1	0.36	84	I	NA	No	No	NA	Lab	Student
Geng et al. (2016), S4*	Shopping list with green products and attribution task	1	0.21	80	I	NA	No	No	NA	Lab	Student
Geng et al. (2016), S5*	Shopping list with green products and recalling commitment	1	0.58	82	I	NA	No	No	NA	Lab	Student
Geng et al. (2019), S1	Framing many past behavior and environmental frame	1	0.52	74	I	Yes	Yes	No	Normative	NA	Student
Geng et al. (2019), S1	Framing many past behavior and monetary frame	1	−0.52	63	I	Both	Yes	Yes	Personal gain	NA	Student
Geng et al. (2019), S2	Framing many past behaviors (non-environmental label) and environmental frame	1	1.83	82	I	Yes	Yes	No	Normative	NA	Student
Geng et al. (2019), S2	Framing many past behaviors (environmental label) and monetary frame	1	2.76	86	I	Both	Yes	Yes	Both	NA	Student
Lacasse (2019), S1*	Combined intervention with behavior adoption, calendar tracking, and information	12	−0.04	84	B	Yes	Yes	No	Normative	Field	Non-student
Lacasse (2019), S1*	Combined intervention with behavior adoption, calendar tracking, and information	4	0.07	84	I	Yes	Yes	No	Normative	Field	Non-student
Lacasse (2019), S1*	Combined intervention with behavior adoption and calendar tracking	12	0.02	80	B	Yes	Yes	No	Normative	Field	Non-student
Lacasse (2019), S1*	Combined intervention with behavior adoption and calendar tracking	4	0.10	80	I	Yes	Yes	No	Normative	Field	Non-student
Lacasse (n.d.), S1*	Many past behaviors and environmentalist label	1	−0.26	71	I	NA	NA	No	Normative	Online	Non-student
Lacasse (n.d.), S1*	Few past behaviors	1	−0.44	57	I	NA	NA	No	Normative	Online	Non-student
Lacasse (n.d.), S2*	Many past behaviors and environmentalist label	1	−0.20	34	I	NA	NA	No	Normative	Online	Student
Lacasse (n.d.), S2*	Few past behaviors	1	−0.17	28	I	NA	NA	No	Normative	Online	Student
Lanzini (2013), S1*	Framing monetary incentives	4	−0.04	80	I	Both	No	Yes	Personal gain	Field	Student
Lin and Chang (2017), S1	Green shopping	2	−1.01	160	B	NA	No	No	NA	Lab	Student
Liu et al. (2021), S1	Self-set goals and feedback	1	0.30	75	B	Both	No	No	NA	Field	Non-student
Liu et al. (2021), S1	Assigned goals and feedback	1	0.19	75	B	Both	No	No	NA	Field	Non-student
Liu et al. (2021), S1	Assigned goals, incentives, and feedback	1	0.14	74	B	Both	Yes	Yes	Personal gain	Field	Non-student
Maki et al. (2015), S1*	Message intervention	2	0.61	65	I	Both	Yes	No	Normative	Combined	Student
Maki et al. (2015), S1*	Modeling intervention	2	0.27	69	I	Yes	No	No	NA	Combined	Student
Maki et al. (2015), S1*	Modeling intervention	2	−0.06	69	B	Yes	No	No	NA	Combined	Student
Margetts and Kashima (2017), S1*	Green shopping	2	0.04	158	I	NA	No	No	NA	Lab	Student

(*continued on next page*)

Table 2 (*continued*)

Article Information		Effect			Moderators						
Authors	Intervention	ES	*d*	*N*	Type	Autonomy	Rationale	(Dis) Incentive	Goal	Setting	Sample
Margetts and Kashima (2017), S2*	Green shopping	2	0.28	43	I	NA	No	No	NA	Lab	Student
Margetts and Kashima (2017), S3*	Green shopping	1	0.89	172	I	NA	No	No	NA	Online	Non-student
Margetts and Kashima (2017), S3*	Green shopping	1	−0.15	172	B	NA	No	No	NA	Online	Non-student
Parag et al. (2011), S1*	Personal carbon allowance	1	0.35	730	I	Both	Yes	Yes	Both	Online	Non-student
Poortinga et al. (2013), S1*	Charge on single-use plastic bags vs. no charge (England)	2	−0.01	500	B	No	No	Yes	Personal gain	Field	Non-student
Poortinga et al. (2013), S1*	Charge on single-use plastic bags vs. no charge (Wales)	2	0.05	500	B	No	No	Yes	Personal gain	Field	Non-student
Schwartz et al. (2015), S1*	Using Social Marketing to Spur Residential Adoption of ENERGY STAR®-Certified LED Lighting	7	0.01	80	B	Both	Yes	Yes	Both	Field	Non-student
Sintov et al. (2019), S1	Organic carts and descriptive norms	10	0.10	284	B	Yes	Yes	No	Normative	Field	Non-student
Steinhorst et al. (2015), S1*	Environmental framing	2	0.24	448	I	Yes	Yes	No	Normative	Online	Non-student
Steinhorst et al. (2015), S1*	Monetary framing	2	0.04	433	I	Both	Yes	Yes	Personal gain	Online	Non-student
Suffolk (2016), S1	Energy efficiency improvement	5	0.03	89	B	Both	NA	Yes	NA	Field	Non-student
Swim and Bloodhart (2013), S1	Feedback (admonishment or praise)	1	0.31	173	B	Both	Yes	No	Normative	Lab	Student
Swim and Bloodhart (2013), S2	Feedback (admonishment or praise)	1	0.39	176	I	Both	Yes	No	Normative	Lab	Student
Swim and Bloodhart (2013), S2	Feedback (admonishment or praise)	1	0.27	176	B	Both	Yes	No	Normative	Lab	Student
Thomas et al. (2016), S1*	Charge on single-use plastic bags	10	0.03	17,636	B	No	No	Yes	Personal gain	Field	Non-student
Tiefenbeck et al. (2013), S1*	Water consumption campaign with feedback and social norms	1	−0.12	907	B	Both	Yes	No	Normative	Field	Non-student
Touhey (2019), S1	Ban of single-use plastic bags	3	0.01	100	B	No	No	No	NA	Field	Non-student
Wolstenholme et al. (2020), S1	Message on positive impacts of eating less meat on health	10	0.03	115	I	Yes	Yes	No	Personal gain	Online	Student
Wolstenholme et al. (2020), S1	Message on positive impacts of eating less meat on environment	10	−0.02	124	I	Yes	Yes	No	Normative	Online	Student
Wolstenholme et al. (2020), S1	Message on positive impacts of eating less meat on health and environment	10	0.11	126	I	Yes	Yes	No	Both	Online	Student
Xu et al. (2018a), S1	Environmental framing of waste separation	1	0.11	159	I	Yes	Yes	No	Normative	Field	Non-student
Xu et al. (2018a), S1	Environmental framing of waste separation	12	0.01	159	B	Yes	Yes	No	Normative	Field	Non-student
Xu et al. (2018a), S1	Monetary framing of waste separation and incentives	1	−0.16	115	I	Both	Yes	Yes	Personal gain	Field	Non-student
Xu et al. (2018a), S1	Monetary framing of waste separation and incentives	12	0.00	115	B	Both	Yes	Yes	Personal gain	Field	Non-student
Xu et al. (2018b), S1	Monetary inducement	1	−0.74	64	B	No	NA	Yes	Personal gain	Field	Non-student
Xu et al. (2018b), S1	Education	1	0.56	57	B	Yes	Yes	NA	NA	Field	Non-student
Zawadzki (2015), S1	Sorting recyclable materials	1	0.48	178	B	NA	Yes	No	NA	Lab	Student
Zawadzki (2015), S1	Sorting recyclable materials	1	−0.36	173	I	NA	Yes	No	NA	Lab	Student

Note. Studies marked with an asterisk (*) were included in the previous meta-analysis by Maki et al. (2019).
ES = Number of non-aggregated effect sizes, I = intention, B = behavior.

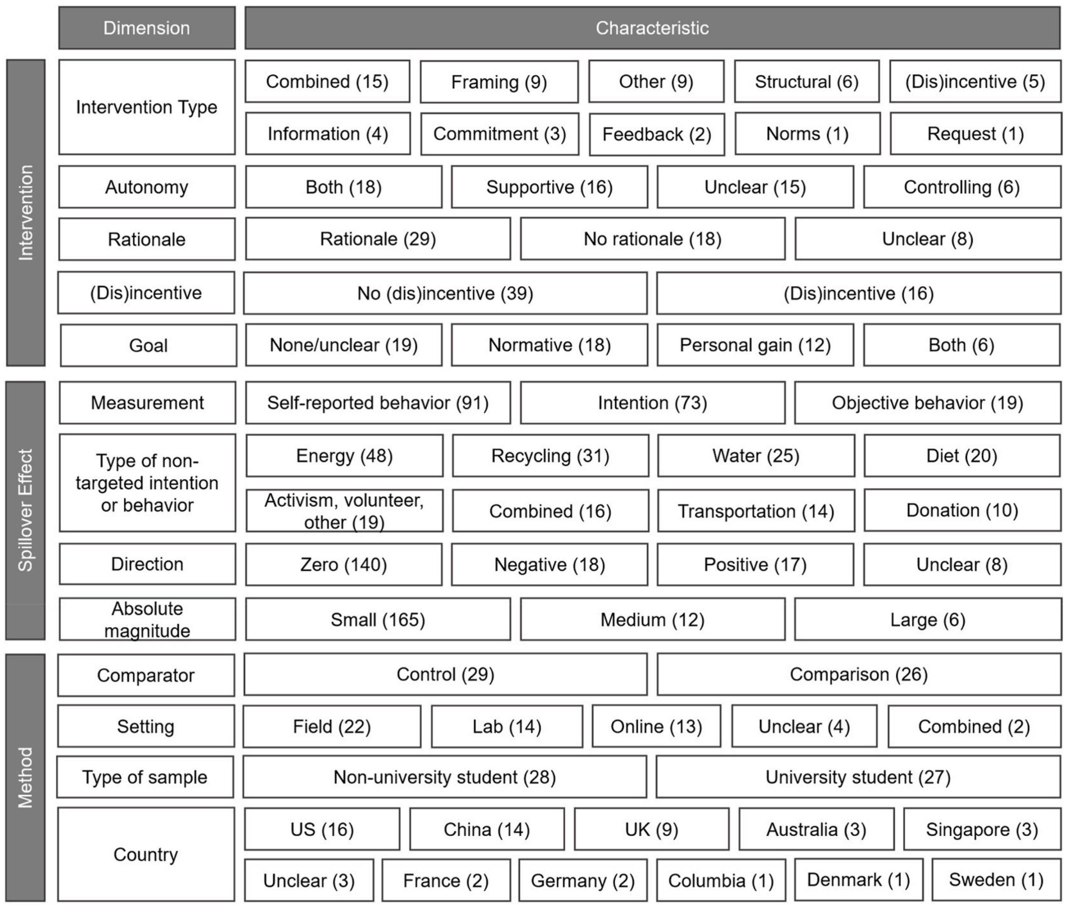

Fig. 2. Morphological Box of Included Articles.
Note. Small: $0 \leq d \leq 0.49$, medium: $0.50 \leq d \leq 0.79$, large: $d \geq 0.80$. Intervention and method dimensions were evaluated at the level of interventions ($N = 55$), while the spillover dimension was evaluated at the level of non-aggregated effect sizes ($N = 183$).

= 26). Field interventions ($n = 22$) were more common than laboratory ($n = 14$) and online interventions ($n = 13$). Approximately the same number of interventions targeted non-university student ($n = 28$) and university student ($n = 27$) samples. The samples came from ten countries, with the United States ($n = 16$), China ($n = 14$), and the United Kingdom ($n = 9$) being most prominent.

3.3. Planned analysis

3.3.1. Meta-analytic model

We conducted a multi-level Bayesian meta-analysis to test the overall spillover effect and the seven moderators. The multi-level model accounted for dependencies among the data. Dependencies can arise when articles include more than one study that compared several interventions to the same control/comparison condition, and when these intervention-control comparisons yielded more than one non-targeted outcome (Cheung, 2014). The meta-analytic model in Fig. 3 includes three levels with effect sizes at level 1, studies/comparisons at level 2, and articles at level 3. Specifying three levels allows estimating the overall effect size δ, the between-article heterogeneity τ_a, and the between-study/comparison heterogeneity τ_s. Effect sizes were aggregated to avoid including more levels and potentially overweighting comparisons yielding many outcomes. This means that effect sizes and standard errors (*SE*) obtained from one intervention were weighted by sample size, such that each intervention-control comparison yielded at most one effect size for intentions and behaviors, respectively.

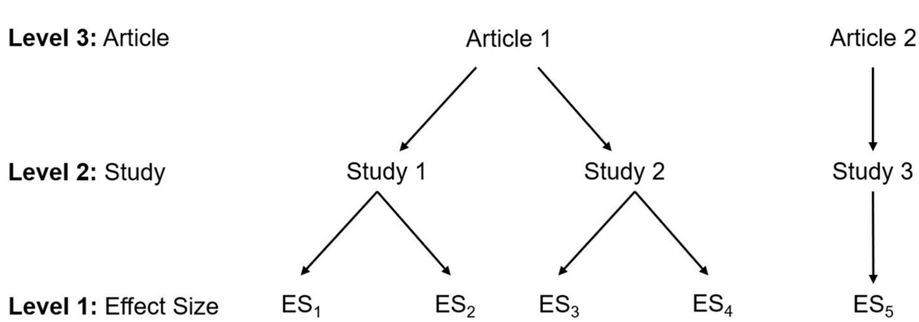

Fig. 3. The Meta-Analytic Model.
Note. Based on Molto et al. (2020). ES = Effect size. Three-level meta-analytic model to estimate the overall effect size, accounting for sampling variance (Level 1) as well as between-study/comparison (Level 2) and between-article (Level 3) heterogeneity.

3.3.2. The Bayesian framework

The Bayesian analyses were performed in *R* (Version 4.0.2, R Core Team, 2020) with *brms* (Version 2.13.5, Bürkner, 2018) and *RStan* (Version 2.19.3, Stan Development Team, 2018). We deviated from the pre-registration in some respects (e.g., priors and moderator analyses) but transparently reported all deviations in Appendix B. The prior distribution for the overall effect size and the contrast-coded moderators were set to a weakly informative normal distribution, δ, $\widehat{\beta} \sim$ Normal(0, 1), as effect sizes in psychology typically vary between 0 and 1.5 (Szucs & Ioannidis, 2017). For the between-article and between-study heterogeneity, mildly informative priors were specified, τ_a, $\tau_s \sim$ HalfNormal(0, 0.1). Both priors are depicted in Fig. 4. During the analyses, these prior distributions were updated with new information, the data, to obtain the posterior distributions (Figs. 7–13).

Models were estimated by specifying four Markov Chain Monte Carlo analyses with 20,000 iterations and a warm-up of 5,000 iterations. Posterior convergence was evaluated based on the potential scale reduction factor $\widehat{R}$ (Gelman & Rubin, 1992) and trace plots. The four chains seemed to have converged well for all analyses, as all $\widehat{R}$s were close to 1. The chains were also well-mixed, as the corresponding trace plots spread randomly around the mean (Nalborczyk et al., 2019).

The two-sided primary research question (Research Question 1) was tested by comparing the null hypothesis that spillover effects are zero, H_0: $\delta = 0$, against the alternative hypothesis that the effects differ from zero, H_1: $\delta \neq 0$. Support for a hypothesis was based on the posterior distribution and the Bayes factor (BF; Wagenmakers et al., 2018). The BF quantifies the empirical evidence in favor of a hypothesis over a competing hypothesis; specifically, it describes the marginal likelihood of the data under one hypothesis compared to another (Stefan et al., 2019) based on the standard inference criteria in Table 3. The 95% credible interval (CrI) indicates a range of values in which the population value falls with 95% probability, given the data and the priors.

3.3.3. Moderator analytic plans

As pre-registered, the moderator analyses were conducted separately for intentions and behaviors if at least four effect sizes existed per subgroup (intention vs. behavior) and subcategory of the moderator (e.g., rationale vs. no rationale; Fu et al., 2011). This was not the case for the moderators autonomy support, goal, and setting, for which we combined spillover effects on intentions and behaviors. Effect sizes categorized as 'unclear' (Appendix A) on one of the moderators were excluded for the respective analysis. Separate meta-regression models with contrasts for dichotomous (Hypotheses 1 to 4 and Research Question 3) and categorical moderators (Hypothesis 5 and 6 and Research Question 2) were fit using the aforementioned priors. Two-sided moderator hypotheses (Research Questions 2 and 3) were tested by comparing the null hypothesis that spillover effects do not differ between the categories of a moderator, H_0: $\widehat{\beta} = 0$, against a two-sided alternative hypothesis, that the effects differ, H_1: $\widehat{\beta} \neq 0$. For one-sided moderator hypotheses (Hypotheses 1 to 6), the hypothesis that the effects are larger for one category (A) than another (B), H_1: $\widehat{\beta} > 0$, was compared against the alternative hypothesis that the effects are smaller for this category (A) than the other (B), H_0: $\widehat{\beta} < 0$.

3.3.4. Additional planned analyses

Sensitivity analyses were conducted with different priors to check the robustness of the overall estimate. A leave-one-out approach identified outliers, and contour-enhanced funnel plots assessed publication bias of the overall effect as well as statistical power of the included studies.

3.4. Analysis

3.4.1. Overall spillover effects (Research Question 1)

This meta-analysis investigated whether effective interventions change non-targeted sustainable intentions ($n = 31$) and behaviors ($n = 32$). It showed moderate evidence against overall spillover on intentions and behaviors combined, $BF_{01} = 8.27$. This implies that the data are 8.27 times more likely under the null hypothesis that the effect is zero compared to the alternative hypothesis that the effect is different from zero. If spillover effects were present, they would likely be small, $\delta = 0.07$, $SE = 0.07$, 95% CrI [−0.06, 0.20], $\tau_a = 0.10$, $SE = 0.06$, 95% CrI [0.01, 0.24], $\tau_s = 0.34$, $SE = 0.04$, 95% CrI [0.26, 0.43]. The posterior effect size estimates of all articles are displayed in the forest plot in Fig. 5.

Estimating the effect separately for intentions and behaviors showed weak evidence against spillover on intentions, $BF_{01} = 2.32$. This means that the data are 2.32 times more likely under the null hypothesis that spillover is zero compared to the alternative hypothesis that it is different from zero. To get a feeling of how little evidence this is, we can inspect the probability wheel in Fig. 6 (Panel A). Imagine that the probability plot represents a pizza with green pepperoni (green area) and mozzarella (white area). You blindly poke your finger onto this pizza. How surprised are you if your finger is covered with the non-dominant topping, in this case, pepperoni? The more surprised you are, the stronger the evidence that spillover on intentions is absent (Wagenmakers, 2018). In this case, you would not be overly surprised since pepperoni still covers 30% of the pizza. This means that there is

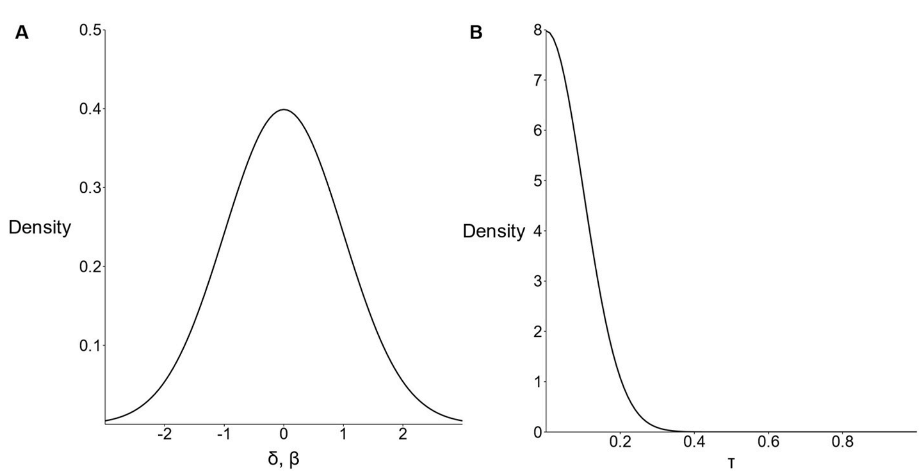

Fig. 4. Prior Distributions for the Overall Effect Size δ, Moderators $\widehat{\beta}$ (A), and Heterogeneity τ (B).

Note. Panel A displays the prior distribution for the overall effect size δ and the $\widehat{\beta}$-coefficients of the moderators, δ, $\widehat{\beta} \sim$ Normal(0, 1). Panel B depicts the prior distribution for the between-article heterogeneity τ_a and between-study heterogeneity τ_s which was truncated at zero, $\tau \sim$ HalfNormal(0, 0.1).

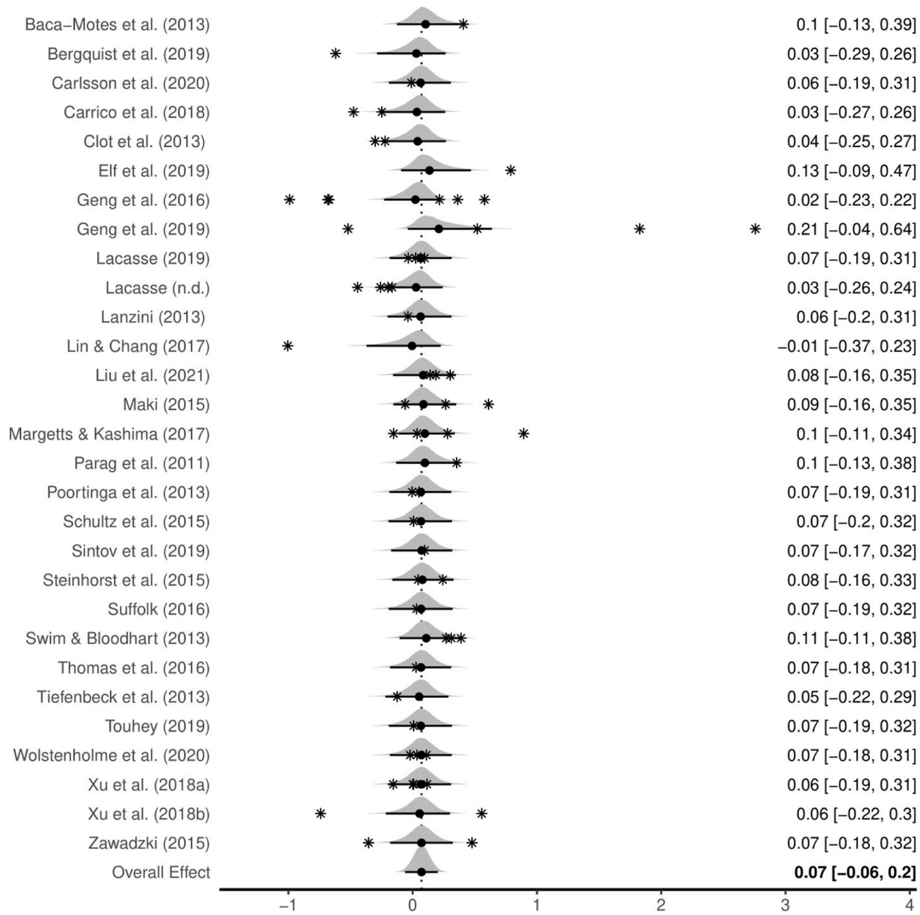

Fig. 5. Forest Plot of Spillover Effects on Intentions and Behaviors Combined.
Note. The forest plot displays the estimated overall spillover effect for intentions and behaviors combined. For each article, the posterior distribution (gray area), the estimated mean effect size (black dot), the 95% CrI (error bars), and the average effect size of each study (asterisk) are shown.

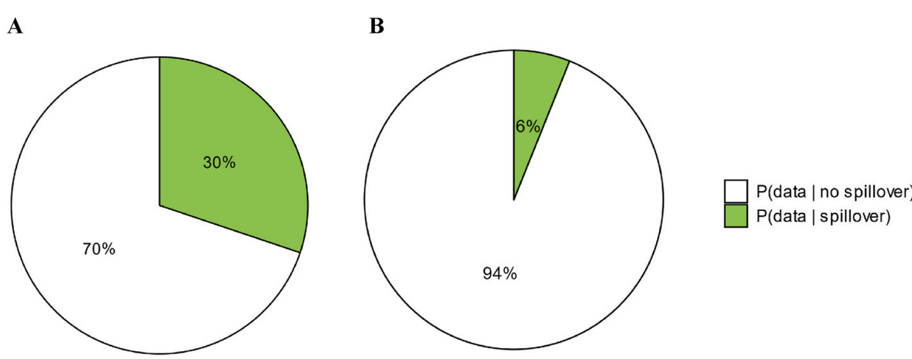

Fig. 6. Probability Wheels Visualizing the Bayes Factor for Intentions (A) and Behaviors (B).
Note. The green area depicts the probability of the data under the alternative hypothesis that overall spillover effects exist (H_1), whereas the white area visualizes the probability of the data under the null hypothesis that overall spillover does not exist (H_0). Both hypotheses were assigned a probability of 50% before seeing the data. This means that after seeing the data, the probability of H_0 for no overall spillover on intentions increased to 70% but there is still a 30% chance for overall spillover. The probability for no overall spillover on behaviors increased to 94%, leaving only a small chance of 6% for the alternative that overall spillover exists. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

only weak evidence that overall spillover on intentions is absent. Even if the effect was present, it would likely be small, $\delta = 0.15$, $SE = 0.08$, 95% CrI [−0.01, 0.31], and uncertain, as the posterior distribution in Fig. 7 spans a wide range of positive and negative values.

Regarding behaviors, there was strong evidence against an overall spillover effect, $BF_{01} = 15.44$. Even if the effect existed, it would likely be negligible, $\delta = 0.01$, $SE = 0.07$, 95% CrI [−0.13, 0.16]. We can repeat the thought experiment about poking a pizza for the probability wheel in Panel B. This time, you would likely be surprised if your finger is covered with the non-dominant topping, pepperoni, as it only represents 6% of the pizza. Your surprise would imply strong evidence against overall spillover on behaviors.

The analysis also showed infinite evidence for heterogeneity at the level of articles, $\tau_a = 0.10$, $SE = 0.06$, 95% CrI [0.01, 0.23], and studies, $\tau_s = 0.34$, $SE = 0.04$, 95% CrI [0.25, 0.43], indicating that further moderators most likely exist.

3.4.2. *Non-targeted outcome (Hypothesis 1)*

In line with Hypothesis 1, we found strong support that interventions caused larger spillover on intentions than behaviors, $BF_{10} = 28.11$, $\widehat{\beta} = 0.14$, $SE = 0.08$, 95% CrI [0.01, 0.26].

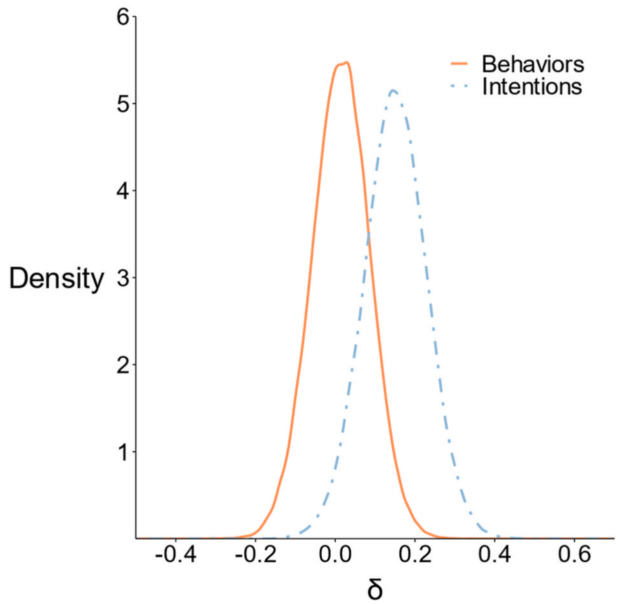

Fig. 7. Posterior Distribution of Overall Spillover Effects on Intentions and Behaviors.
Note. Intentions ($n = 31$): $\delta = 0.15$, 95% CrI [-0.01, 0.31], behaviors ($n = 32$): $\delta = 0.01$, 95% CrI [-0.13, 0.16].

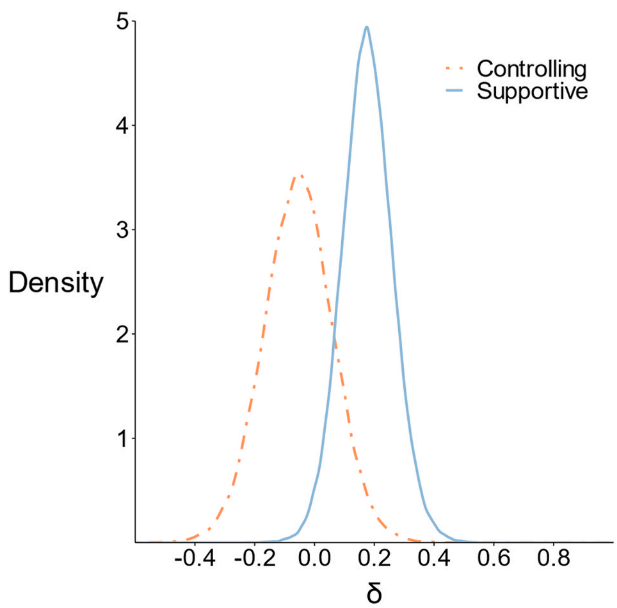

Fig. 8. Posterior Distribution of the Spillover Effect Depending on Autonomy.
Note. Autonomy-supportive ($n = 20$): $\delta = 0.18$, 95% CrI [0.01, 0.35], controlling ($n = 6$): $\delta = -0.05$, 95% CrI [-0.29, 0.17].

3.4.3. Autonomy-supportive vs. controlling interventions (Hypothesis 2)

Supporting Hypothesis 2 with very strong evidence, spillover effects were more positive when interventions included autonomy-supportive ($n = 20$) rather than controlling ($n = 6$) elements, $BF_{10} = 31.41$, $\widehat{\beta} = 0.23$, $SE = 0.13$, 95% CrI [0.03, 0.44]. However, this should be interpreted with caution due to the small number of effect sizes for controlling interventions, which is also reflected in the wide posterior distribution (Fig. 8).

3.4.4. Rationale provision (Hypothesis 3)

Interventions with ($n = 17$) compared to without ($n = 10$) a rationale were associated with more positive spillover on intentions, as indicated by the moderate evidence, $BF_{10} = 6.03$, $\widehat{\beta} = 0.19$, $SE = 0.18$, 95% CrI [−0.11, 0.50]. Similarly, there was strong evidence that spillover on behaviors was more positive for interventions with ($n = 18$) rather than without ($n = 10$) a rationale, $BF_{10} = 23.00$, $\widehat{\beta} = 0.21$, $SE = 0.12$, 95% CrI [0.01, 0.41]. But once again, the posterior distributions in Fig. 9 span a wide range of values, especially for intentions (Panel A).

3.4.5. Financial (dis)incentive (Hypothesis 4)

The analysis showed strong evidence for more positive spillover on intentions when interventions did not include a financial (dis)incentive ($n = 25$) vs. when they did ($n = 6$), $BF_{10} = 20.06$, $\widehat{\beta} = 0.17$, $SE = 0.10$, 95% CrI [0.00, 0.33]. However, the evidence was not compelling for behaviors, $BF_{10} = 1.80$, $\widehat{\beta} = 0.04$, $SE = 0.10$, 95% CrI [−0.12, 0.20], based on 11 interventions with and 21 without financial (dis)incentives. This also becomes apparent from the overlapping posterior distributions in Fig. 10 (Panel B).

3.4.6. Intervention goal (Hypothesis 5 and 6)

There was extreme support for Hypothesis 5 that positive spillover on intentions and behaviors combined was more likely when interventions addressed normative ($n = 22$) rather than personal gain ($n = 13$) goals, $BF_{10} = 409.96$, $\widehat{\beta} = 0.23$, $SE = 0.08$, 95% CrI [0.10, 0.37]. Contrary to Hypothesis 6, there was strong evidence that interventions with both types of goals ($n = 6$) were associated with more positive spillover than those addressing normative goals, $BF_{01} = 16.67$, $\widehat{\beta} = -0.24$, $SE = 0.15$, 95% CrI [−0.48, 0.00]. The latter finding should, however, be interpreted with caution, as it was based on only six effect sizes from interventions addressing both types of goals. This uncertainty is also reflected in the posterior distribution of a combination of goals in Fig. 11.

3.4.7. Generalizability: Setting and sample type (Research Question 2 and 3)

Exploratory analyses revealed moderate to strong evidence (Table 4) against differences in spillover on intentions and behaviors combined across settings (laboratory: $n = 16$, online: $n = 14$, field: $n = 26$). If differences were present, they would likely be small, whereby effect sizes obtained in online and field studies tended to be larger and more positive than those in the laboratory. These findings correspond to the overlapping posterior distributions in Fig. 12 and imply that based on the currently available data, findings seem to generalize across settings.

The analysis also showed moderate evidence against differences in spillover on intentions between university student ($n = 21$) and non-university student ($n = 10$) samples, $BF_{01} = 4.96$, $\widehat{\beta} = 0.01$, $SE = 0.20$, 95% CrI [−0.39, 0.41]. For spillover on behaviors, it is still unclear whether effects generalize between sample types (student: $n = 9$, non-student: $n = 23$), $BF_{01} = 2.62$, $\widehat{\beta} = 0.20$, $SE = 0.14$, 95% CrI [−0.07, 0.48]. As shown in Fig. 13 (Panel B), spillover on behaviors seemed more positive in non-student compared to student samples.

3.5. Additional analyses: Sensitivity, outliers, publication bias, and power

To test the findings' robustness, we conducted sensitivity analyses for the overall spillover effect, with the default prior distribution in psychology (Gronau et al., 2017), δ, $\widehat{\beta} \sim \text{Cauchy}(0, \frac{1}{\sqrt{2}})$, and a weakly informative prior for the heterogeneity, $\tau_a, \tau_s \sim \text{HalfCauchy}(0, 0.5)$. To test whether deviations from the pre-registration influenced the results, we followed the original plan and used bridge sampling in *brms* to compare the model with an intercept to the same model with the intercept fixed to zero. The sensitivity analysis showed moderate evidence against spillover on intentions and behaviors combined, $BF_{01} = 7.97$. If the effect was present, it would likely be small, $\delta = 0.06$, $SE = 0.08$, 95% CrI [−0.10, 0.23], $\tau_a = 0.14$, $SE = 0.09$, 95% CrI [0.01, 0.34],

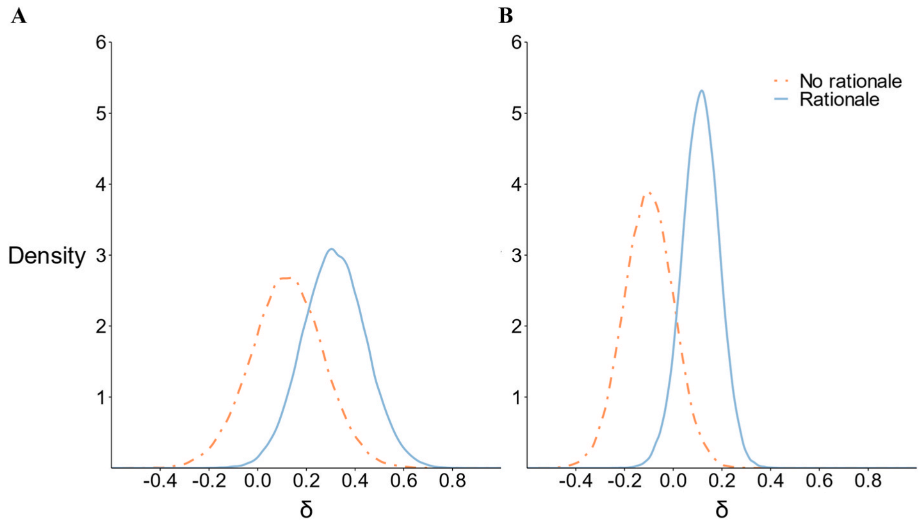

Fig. 9. Posterior Distribution of the Spillover Effect on Intentions (A) and Behaviors (B) Depending on Rationale Provision.
Note. Panel A (intentions): no rationale ($n = 10$): δ = 0.12, 95% CrI [-0.18, 0.41], rationale ($n = 17$): δ = 0.31, 95% CrI [0.06, 0.57]. Panel B (behaviors): no rationale ($n = 10$): δ = −0.10, 95% CrI [-0.30, 0.10], rationale ($n = 18$): δ = 0.11, 95% CrI [-0.04, 0.26].

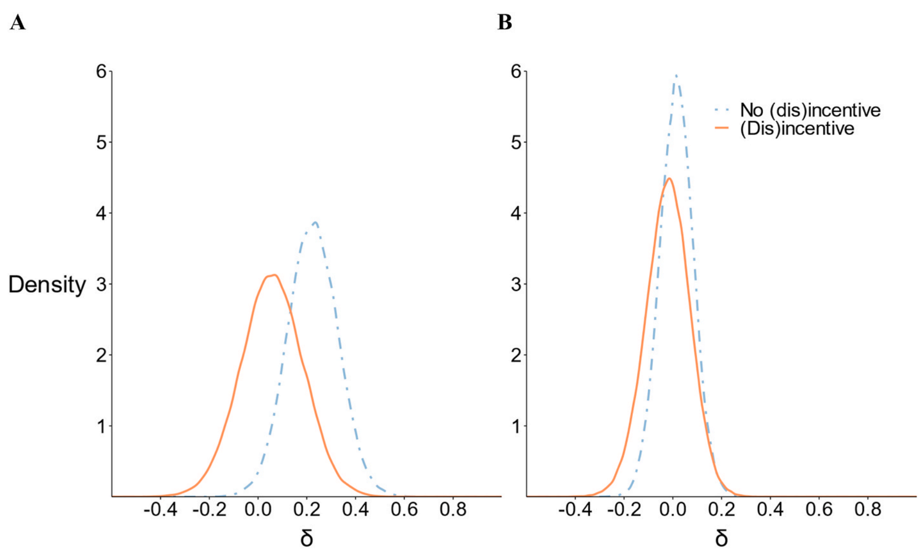

Fig. 10. Posterior Distribution of the Spillover Effect on Intentions (A) and Behaviors (B) Depending on Financial (Dis)incentives.
Note. Panel A (intentions): no (dis)incentives ($n = 25$): δ = 0.22, 95% CrI [0.02, 0.43], (dis)incentives ($n = 6$): δ = 0.06, 95% CrI [-0.20, 0.31]. Panel B (behaviors): no (dis)incentives ($n = 21$): δ = 0.01, 95% CrI [-0.12, 0.15], (dis)incentives ($n = 11$): δ = −0.02, 95% CrI [-0.20, 0.15].

$\tau_s = 0.43$, $SE = 0.07$, 95% CrI [0.31, 0.58]. As this conclusion aligns with the previous analysis, the results seem insensitive to the choice of priors.

A leave-one-out analysis repeatedly estimated the model but dropped one effect size at a time. The posterior estimate for the overall spillover on intentions and behaviors combined varied between 0.05 and 0.10, indicating no outliers.

To assess publication bias of the overall effect, we used a contour-enhanced funnel plot (Peters et al., 2008) based on a frequentist mixed-effects model with restricted Maximum-Likelihood estimation, as Bayesian tools are limited to single-level models. The contour-enhanced funnel plot in Fig. 14 is slightly asymmetrical. However, this asymmetry likely stems from factors other than publication bias, such as heterogeneity, as almost two-thirds of the effect sizes ($n = 40$) are non-significant and lie inside the inner white area ($p > .05$), while significant results are missing in the outer white area ($p < .01$). With publication bias, fewer effect sizes would lie within the area of non-significance, whereas more effect sizes would lie within the areas of significance. Thus, there is no evidence for publication bias.

We additionally examine the power of primary studies to detect the non-aggregated effect size, using power-enhanced funnel plots (Kossmeier et al., 2020). We used δ = 0.15 for intentions and δ = 0.01 for behaviors, as best guesses of the 'true' (population) fixed-effect estimate. As displayed in Fig. 15, all studies were highly underpowered. Most effect sizes lie in the dark red and dark orange areas indicating less than 20% power, and none of them reached more than 40% power. In line with this, the median power to detect spillover on intentions was 12.7% ($n = 73$ effect sizes), while it was 5.0% for behaviors ($n = 110$ effect sizes).

4. Discussion

This meta-analysis investigated spillover effects of effective interventions on sustainable intentions and behaviors. As potential moderators, it examined the type of non-targeted outcome (intentions vs. behaviors), several intervention characteristics (autonomy support, rationale provision, financial (dis)incentive, and goal), and

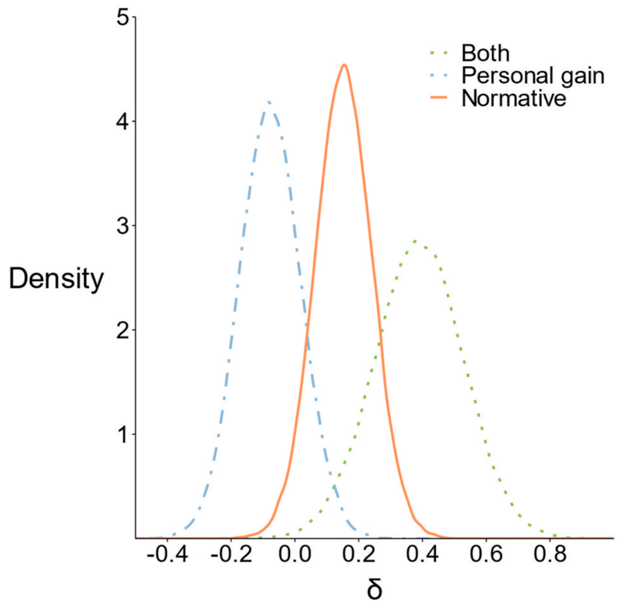

Fig. 11. Posterior Distribution of the Spillover Effect on Intentions and Behaviors Combined Depending on Intervention Goal.
Note. Normative goals (n = 22): δ = 0.15, 95% CrI [-0.02, 0.33], personal gain goals (n = 13): δ = −0.08, 95% CrI [-0.27, 0.11], both types of goals (n = 6): δ = 0.39, 95% CrI [0.11, 0.66].

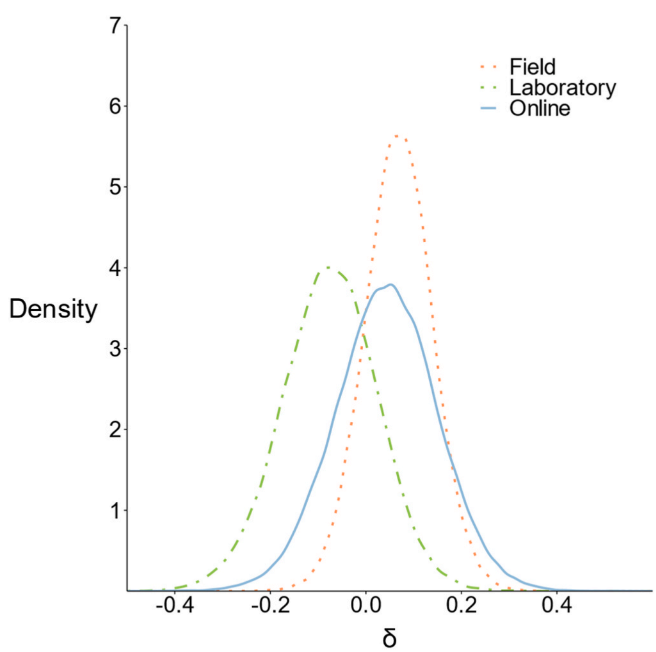

Fig. 12. Posterior Distribution of the Spillover Effect Depending on the Research Setting.
Note. Laboratory (n = 16): δ = −0.08, 95% CrI [-0.28, 0.12], online (n = 14): δ = 0.04, 95% CrI [-0.17, 0.25], field (n = 26): δ = 0.07, 95% CrI [-0.07, 0.21].

characteristics regarding the generalizability of spillover (research setting and sample type). The three-level Bayesian analysis included 63 aggregated effect sizes from 38 studies and 29 articles between 2011 and 2021, based on more than 26,000 unique participants. It showed weak evidence against spillover on intentions. It is thus still uncertain whether overall spillover on intentions exists; if so, it would likely be small, δ = 0.15, 95% CrI [−0.01, 0.31]. A Cohen's d of 0.15 means that the control and intervention condition overlap to 94.0% (Fig. 16). Returning to the introductory example, if free public transport increased usage in 100 individuals, four to five more individuals would show increased recycling intentions in the intervention (e.g., city with free public transport) compared to the control condition (e.g., city with paid public transport; Magnusson, 2021), assuming that spillover existed. Regarding behaviors, the analysis showed strong support against spillover, and even if spillover was present, it would likely be negligible, δ = 0.01, 95% CrI [−0.13, 0.16]. This means that if free public transport increased the target outcome in 100 individuals, there is strong evidence that non-targeted intentions, such as recycling, would remain unaffected.

These results partly align with the previous frequentist meta-analysis (Maki et al., 2019) which showed small positive spillover on intentions (d = 0.17, $p < .01$) and very small negative spillover on behaviors (d = −0.03, $p < .05$). The two meta-analyses agree that spillover is small for intentions (assuming it exited) and larger for intentions than behaviors. However, they come to different conclusions about the presence of spillover and the direction of behavioral spillover. The previous meta-analysis concluded that effective interventions slightly increase non-targeted intentions but very slightly decrease non-targeted behaviors. This meta-analysis found weak evidence that effective interventions do not influence non-targeted intentions and strong evidence that they do not affect non-targeted behaviors. We believe that these updated results are important to consider. On the dark side, the evidence for spillover on intentions seems much weaker than previously assumed; on the bright side, there seems to be no spillover on behaviors, rather than negative spillover as previously concluded.

These different conclusions may partly result from the kind of analysis. P-values from .005 to .05, as found in the previous meta-analysis (Maki et al., 2019), "deserve skepticism, curiosity, and modest optimism" (Wagenmakers, 2017, para. 2) and may translate into weak evidence in the Bayesian framework. This was the case in the present meta-analysis, probably also because many newer studies showed non-significant effects (e.g., Carlsson et al., 2020; Sintov et al., 2019; Wolstenholme et al., 2020). A definite comparison of the two meta-analyses is, however, not possible because of other differences (e.g., multi-level modeling and aggregation methods).

4.1. Intervention characteristics

Despite the non-existent overall effects, different interventions can still show positive or negative spillover. The present meta-analysis identified ingredients of interventions that may help to promote positive and avoid negative spillover, such as autonomy support, rationale provision, financial (dis)incentives, and goals. Although these factors moderated the effect, it is important to consider that spillover was never larger than δ = 0.31 for intentions (providing a rationale), δ = 0.11 for behaviors (providing a rationale), and δ = 0.39 for intentions and behaviors combined (normative and personal gain goals combined). A summary of all findings is in Table 5.

4.1.1. Autonomy, rationale, (dis)incentives, and goals

Autonomy-supportive (e.g., information) rather than controlling (e.g., bans) interventions were associated with more positive spillover on intentions and behaviors combined. Despite the evidence being very strong, it is still limited, as only a few interventions were purely controlling. This finding is related to earlier results that intrinsic rather than extrinsic motivation is more likely to result in positive spillover (Maki et al., 2019). This makes sense as individuals are usually intrinsically motivated when they experience a sense of autonomy (Deci & Ryan, 2008). However, they experience extrinsic or amotivation when their sense of autonomy is threatened. Autonomy support thus goes hand in hand with different kinds of motivations. What should, however, be considered is that this meta-analysis distinguished between purely autonomous, purely controlling, and both kinds of interventions, as compared to the previous meta-analysis which focused on intrinsic and extrinsic motivation but did not distinguish mixed-motivation

Journal of Environmental Psychology 78 (2021) 101694

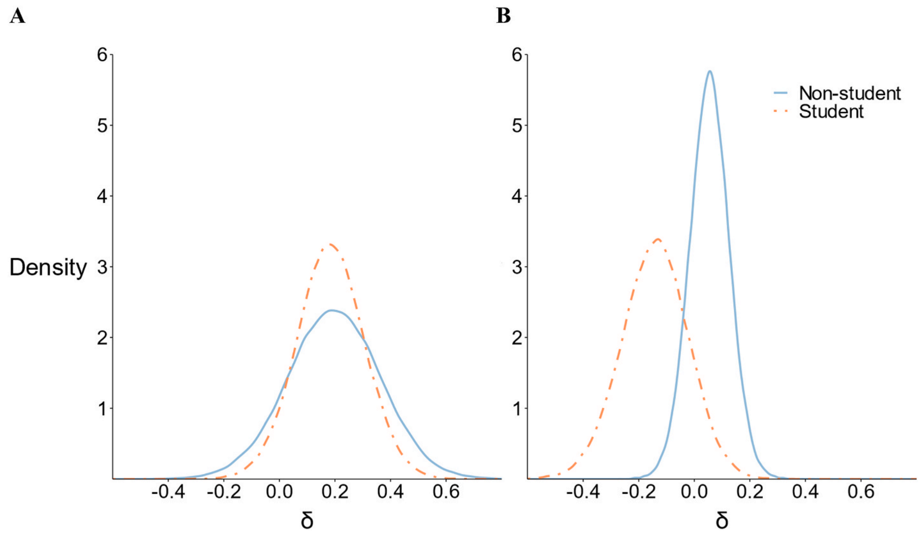

Fig. 13. Posterior Distribution of the Spillover Effect on Intentions (A) and Behaviors (B) Depending on the Sample Type.
Note. Panel A (intentions): student samples (n = 21): δ = 0.18, 95% CrI [-0.06, 0.42], non-student samples (n = 10): δ = 0.20, 95% CrI [-0.14, 0.52]. Panel B (behaviors): student samples (n = 9): δ = −0.14, 95% CrI [-0.39, 0.09], non-student samples (n = 23): δ = 0.05, 95% CrI [-0.09, 0.19].

Table 3
Interpretation of the Bayes factor.

BF_{10} Evidence for H_1	Interpretation as proposed by Jeffreys (1961)	BF_{01} Evidence for H_0
>100	Extreme evidence	>100
30–100	Very strong evidence	30–100
10–30	Strong evidence	10–30
3–10	Moderate evidence	3–10
1–3	Weak evidence	1–3
1	No evidence	1

Note. Based on Lee and Wagenmakers (2014).

Table 4
Comparison of spillover on intentions and behaviors combined in different research settings.

Comparison	Estimate $\hat{\beta}$ and 95% CrI	BF_{01}
Laboratory – Online	−0.12 [-0.39, 0.16]	5.10
Laboratory – Field	−0.14 [-0.39, 0.09]	4.15
Online – Field	−0.03 [-0.28, 0.23]	10.88

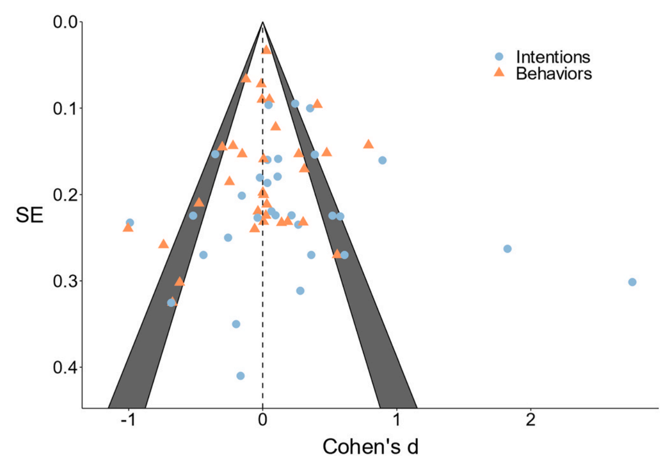

Fig. 14. Contour-Enhanced Funnel Plot of the Overall Spillover Effect on Intentions and Behaviors.
Note. The contour-enhanced funnel plot displays the overall spillover effect on intentions and behaviors in terms of Cohen's d (N = 63) against the standard error (SE). Unlike traditional funnel plots, this plot is centered on zero with the inner white area representing $p > .05$, the gray area $.01 \leq p < .05$, and the outer white area $p < .01$.

interventions.

Investigating a prominent autonomy-supportive element, rationale provision, showed moderate to strong evidence that interventions with rather than without a rationale were associated with more positive spillover, especially for behaviors. Financial (dis)incentives as a typically controlling element were associated with more negative spillover, as suggested by the strong evidence for intentions but only weak evidence for behaviors. Similarly, normative goals (e.g., environmental protection) were associated with more positive spillover than personal gain goals (e.g., monetary or health benefits), with extreme support. Contrary to the expectations, a combination of goals was associated with more positive spillover than normative goals, with strong evidence. One potential explanation for this unexpected finding may be rooted in the outcomes that the two goals change. Normative goals may primarily spill over to intentions and behaviors with environmental benefits (e.g., recycling). A combination of goals may not only spill over to intentions and behaviors with environmental benefits but also to those with self-interested benefits (e.g., saving water; Steinhorst et al., 2015), thereby being more effective. An alternative explanation may be that individuals differ in their core values. Normative goals may be most strongly activated in individuals with biospheric values (i.e., judgments based on benefits and costs for the environment; Steg et al., 2014), instigating positive spillover to non-targeted outcomes, while they leave individuals with egoistic values (i.e., judgments based on benefits and costs for own resources) unaffected. A combination of goals may target both, individuals with biospheric and individuals with egoistic values, thereby reaching more individuals and thus driving more positive spillover.

4.2. Generalizability: Research setting and sample type

We investigated whether spillover effects generalize across settings and sample types. Moderate evidence suggested that spillover was similar across laboratory, online, and field studies. Although exploratory, this finding carefully proposes that laboratory research produces externally valid results and reliably predicts findings online and in the field. In short, "psychological laboratories [may] reveal truths rather than trivialities" (Mitchell, 2012, p. 110), meaning that researchers may

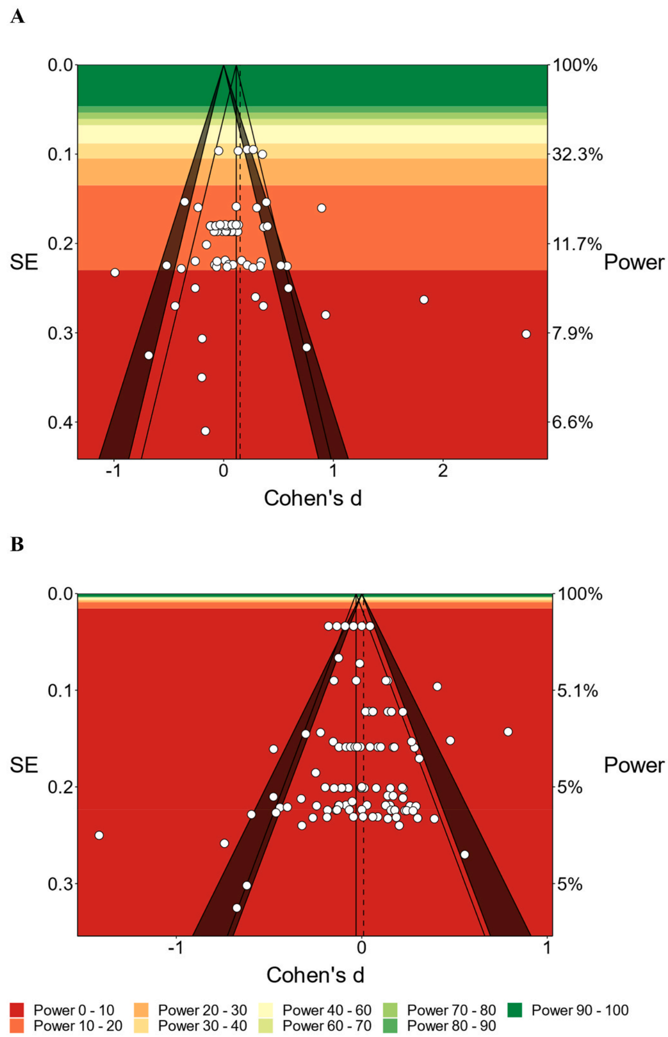

Fig. 15. Power-enhanced funnel plot for intentions (A) and behaviors (B).

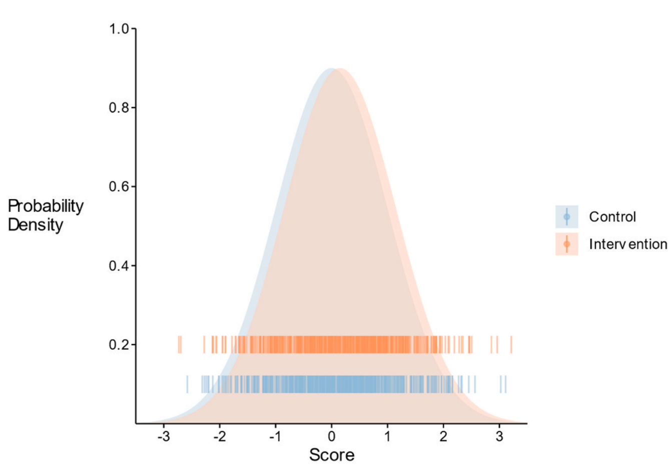

Fig. 16. Simulation of Distributions for the Intervention and Control Condition Based on Overall Spillover on Intentions (δ = 0.15).
Note. The simulations were based on 500 observations per condition. Intervention: Score ~ Normal(0.15, 1) and Control: Score ~ Normal(0, 1).

investigate valid spillover in the laboratory. Regarding sample type, there was moderate evidence that spillover on intentions generalized between student and non-student samples, in contrast to the weak evidence for the generalizability of spillover on behaviors. This means that spillover on intentions may be accurately measured in student samples, whereas this may be difficult for spillover on behaviors.

4.3. Limitations

This meta-analysis was limited by underpowered primary research and the available evidence. As most primary studies were highly underpowered, this meta-analysis may under- or overestimate the true spillover effect (Stanley & Spence, 2014). Similarly, the available evidence is still limited, which is reflected in the number of included articles ($n = 29$) and the sometimes weak to moderate, not strong or extreme, evidence in favor of a hypothesis. This limited evidence prevents definite conclusions. Although both shortcomings restrict the interpretation of the current findings, the Bayesian approach helps to overcome these limitations by allowing us to continuously update the analysis once more, hopefully high-powered studies become available. We encourage researchers and practitioners to think "like Bayesians, updating [their] knowledge [and this meta-analysis] as new information comes in" (Levitin, 2016, p. 202).

4.4. Recommendations for policymakers

Considering the current evidence, we recommend policymakers choosing sustainable interventions based on their expected direct effects (e.g., effect of free public transportation on sustainable commuting) and consider potential spillover (e.g., recycling) to a lesser degree unless they explicitly aim for widespread change. For policymakers who aim for positive spillover, we would like to give first tentative recommendations on how to design more effective, evidence-based interventions. We suggest interventions that support individuals' autonomy, such as information campaigns without demanding language, non-evaluative feedback, and interventions that provide a rationale (e.g., why individuals should engage in the target behavior). We recommend refraining from interventions with financial (dis)incentives, such as rewards for engaging in the target behavior, if spillover on intentions is the primary consideration. Lastly, we recommend policymakers emphasize normative (e.g., environmental protection) instead of personal gain (e.g., monetary or health benefits) goals. Examples are WWF's paper dispenser campaign "Save paper, save the planet" (Saatchi & Saatchi Copenhagen, 2007) or descriptive norms that others behave in a sustainable way (Steg et al., 2014). Although emphasizing both types of goals seemed to be most effective in this meta-analysis, we cannot give practical recommendations, as this finding was unexpected and the sample size of interventions that addressed both goals is still relatively small. Future research should thus investigate interventions that jointly address normative and personal gain goals, particularly because many real-world campaigns employ this strategy.

4.5. Recommendations for researchers and future directions

Based on the current evidence, we provide three recommendations for spillover researchers, including generalizability, the PRO guidelines, and regular meta-analytical updates.

4.5.1. Generalizability

Spillover findings seem to generalize from the laboratory to online settings and the more naturalistic field. This finding implies that researchers may use laboratory studies to investigate valid spillover effects, which could save them time and money, as field research is often more time and resource-intense than laboratory research. This also seems to apply to spillover of intentions—but not necessarily behaviors—in student and non-student samples. However, the evidence for

Table 5
Summary of the meta-analysis.

Research question/ Hypothesis	Estimate [95% CrI]	Bayes Factor	Evidence strength	Hypothesis supported?	Summary
RQ1a: spillover on intentions ≠ 0	δ = 0.15 [-0.01, 0.31]	BF_{01} = 2.32	Weak ●○○○○		Weak evidence against spillover on intentions
RQ1b: spillover on behaviors ≠ 0	δ = 0.01 [-0.13, 0.16]	BF_{01} = 15.44	Strong ●●●○○		Strong evidence against spillover on behaviors
H1: spillover on intentions > behaviors	$\hat{\beta}$ = 0.14 [0.01, 0.26]	BF_{10} = 28.11	Strong ●●●○○	Yes	Strong evidence that spillover on intentions is more positive than on behaviors
H2: spillover of autonomy-supportive interventions > controlling interventions	$\hat{\beta}$ = 0.23 [0.03, 0.44]	BF_{10} = 31.41	Very strong ●●●●○	Yes	Very strong evidence that spillover on intentions and behaviors is more positive for autonomy-supportive than controlling interventions
H3a: spillover on intentions of interventions with rationale > without rationale	$\hat{\beta}$ = 0.19 [-0.11, 0.50]	BF_{10} = 6.03	Moderate ●●○○○	Yes	Moderate evidence that spillover on intentions is more positive for interventions with than without a rationale
H3b: spillover on behaviors of interventions with rationale > without rationale	$\hat{\beta}$ = 0.21 [0.01, 0.41]	BF_{10} = 23.00	Strong ●●●○○	Yes	Strong evidence that spillover on behaviors is more positive for interventions with than without a rationale
H4a: spillover on intentions of financial (dis)incentive < no (dis)incentive	$\hat{\beta}$ = 0.17 [0.00, 0.33]	BF_{10} = 20.06	Strong ●●●○○	Yes	Strong evidence that spillover on intentions is more negative for interventions with than without (dis)incentives
H4b: spillover on behaviors of financial (dis)incentive < no (dis)incentive	$\hat{\beta}$ = 0.04 [-0.12, 0.20]	BF_{10} = 1.80	Weak ●○○○○	No	Weak evidence that spillover on behaviors is more negative for interventions with than without (dis)incentives
H5: spillover of normative goals > personal gain goals	$\hat{\beta}$ = 0.23 [0.10, 0.37]	BF_{10} = 409.96	Extreme ●●●●●	Yes	Extreme evidence that spillover on intentions and behaviors combined is more positive for normative than gain goals
H6: spillover of normative goals > both goals	$\hat{\beta}$ = -0.24 [-0.48, 0.00]	BF_{01} = 16.67	Strong ●●●○○	No	Strong evidence that spillover on intentions and behaviors combined is more positive for both types of goals than normative goals
RQ2: spillover in lab = online	$\hat{\beta}$ = -0.12 [-0.39, 0.16]	BF_{01} = 5.10	Moderate ●●○○○		Moderate evidence that spillover on intentions and behaviors generalizes between lab and online and lab and field settings
lab = field	$\hat{\beta}$ = -0.14 [-0.37, 0.13]	BF_{01} = 4.15	Moderate ●●○○○		
RQ3a: spillover on intentions in students = non-students	$\hat{\beta}$ = 0.01 [-0.39, 0.41]	BF_{01} = 4.96	Moderate ●●○○○		Moderate evidence that spillover on intentions generalizes across samples
RQ3b: spillover on behaviors in students = non-students	$\hat{\beta}$ = 0.20 [-0.07, 0.48]	BF_{01} = 2.62	Weak ●○○○○		Weak evidence that spillover on behaviors generalizes across samples

Note. Gray indicates that a research question was tested. Green indicates evidence in favor of the tested hypothesis, while orange indicates evidence against the tested hypothesis.

none of these findings is extremely strong. Additionally, the results regarding the generalizability of spillover across settings may depend on whether spillover on intentions or behaviors is measured, as measures of intentions may be similar across settings, while measures of behaviors may differ (e.g., number of pages printed or napkins used in the laboratory vs. diet and electricity use; Lacasse, 2019; Lin & Chang, 2017; Zawadzki, 2015; Tiefenbeck et al., 2013). Whether the generalizability of spillover across settings depends on intentions and behaviors could, however, not be investigated in this meta-analysis due to small sample sizes.

4.5.2. Power-Reporting-Open science (PRO) guidelines

Researchers should aim to follow the Power-Reporting-Open science (PRO) guidelines, three practical recommendations to raise awareness for more robust spillover research (Fig. 17). Firstly, PRO recommends conducting more high-powered studies. This meta-analysis included studies with sample sizes ranging from 28 to 17,636 participants. None of these studies reached more than 40% power to detect spillover effects. The median power to detect intentional and behavioral spillover was 12.7% and 5.0%, respectively. Small spillover (given the effect exists) combined with small samples results in low statistical power and wasteful research, resulting in serious problems, such as undetected small effects, overestimating effects, and low replicability (Button et al., 2013). We recommend increasing sample sizes to advance research on spillover. Table 6 provides an example for between-subjects designs with two conditions to detect average spillover of $\delta = 0.15$ on intentions and $\delta = 0.01$ on behaviors. Alternatively, researchers could base their power analyses on the smallest effect size of interest (Lakens, 2014). This would most likely require very large samples as well, as even small effect sizes (e.g., $\delta = 0.10$) could be theoretically interesting in spillover research.

We acknowledge that such sample sizes are often out of reach, but sequential analyses, large-scale collaborations, and larger control conditions can help to solve this problem. Researchers can perform sequential analyses during data collection and stop when the results are convincing, or the effect is too unlikely to be observed (Schönbrodt et al., 2017; for a tutorial, see Breffara Bret et al., 2021). This strategy is typical in medical trials and can lower the required sample size by at least 30% (Lakens, 2014). Another option is to collaborate and pool available resources for a single study. Several big team science

Table 6
Example recommendations for the sample size per condition in between-subjects designs to detect average spillover effects.

		Statistical Power			
	Cohen's *d*	80%	85%	90%	95%
Intentions	0.15[1]	550	640	762	963
Behaviors	0.01[1]	63,088	73,361	87,387	110,431

Note. [1]Assuming the effect was present. The power analyses were based on a one-sided two-sample *t*-test with $\alpha = 0.05$.

initiatives in psychology (for an overview, see Forscher et al., 2020) have shown that sample sizes, especially for intentions, can be reached when collaborating within and across institutions, countries, and continents, including the Psychological Science Accelerator (e.g., $N = 25,718$ participants in 89 countries, Legate et al., 2021), Many Labs (e.g., $N = 6,344$ participants from 36 different sources, Klein et al., 2014), the Junior Researcher Programme (e.g., $N = 10,207$ participants in 26 countries, Ruggeri et al., 2021), as well as other collaborations of researchers (e.g., $N = 53,524$ participants in 26 countries within one week, Kowal et al., 2020). The last—although less preferred—option to higher power is to increase the number of participants in the control but not the intervention condition, which is often done in genome-wide association studies (Li et al., 2019; Spencer et al., 2009) and has been applied in spillover research, for example by Thomas et al. (2016; $N_{\text{intervention}} = 931$, $N_{\text{control}} = 16,705$).

Secondly, PRO also sets reporting standards for spillover articles. Many articles in this meta-analysis failed to disclose essential information. Specifically, PRO recommends describing samples in more detail, including age, gender, country, education, and socio-economic status as indicators of generalizability. We also advise reporting the type of intervention and control/comparison condition (i.e., example materials, time frame, and frequency), the intervention's effectiveness, and the type of targeted and non-targeted outcomes (for an extensive checklist, see Galizzi & Whitmarsh, 2019). In terms of statistics, we advocate reporting descriptive statistics and effect sizes, alongside inferential statistics (i.e., test statistics, degrees of freedom, and exact *p*-values), as well as non-significant findings. For continuous outcomes, means, *SDs*, sample sizes per condition, and effect sizes (e.g., Cohen's *d*) should be reported; for dichotomous outcomes, the number of individuals per

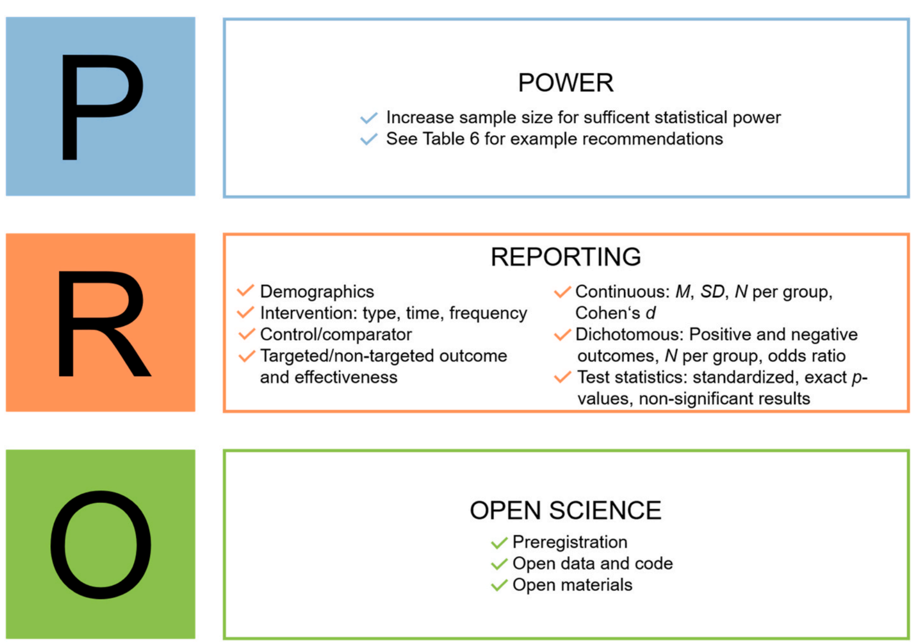

Fig. 17. The PRO guidelines for robust experimental spillover research.

outcome and condition, as well as odds ratio, should be reported.

Lastly, open science practices are not yet widely applied in experimental spillover research. We strongly encourage pre-registering studies and providing open materials, data, and code to increase transparency, robustness, and reusability of research.

4.5.3. Meta-analytical updates

Summarizing research on environmental spillover is challenging because the number of studies is increasing rapidly across various fields, including psychology, economics, business, and ecology. The current findings will eventually become outdated, calling for further meta-analytical updates. We thus encourage researchers to either send us their unpublished and published data if not already included or regularly update this meta-analysis, not only for the research community but also for policymakers. We support this endeavor by providing open data, materials, and code (https://osf.io/tu7yx/). The Bayesian approach allows for relatively easy updating, as "today's posterior is tomorrow's prior" (Lindley, 1972, p. 2), meaning that the posteriors of this meta-analysis can serve as the priors for the next meta-analysis, which will then be updated with the new evidence. We hope that updated versions of this meta-analysis will include robust studies with high statistical power, detailed reporting, and open science practices, as outlined in the PRO guidelines.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contribution CRediT

Sandra J. Geiger: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing-Original draft preparation, Visualization, Project administration. **Cameron Brick:** Resources, Writing-Review & Editing. **Ladislas Nalborczyk:** Validation, Writing-Review & Editing. **Anna Bosshard:** Moderator coding, Writing-Review & Editing. **Nils B. Jostmann:** Supervision, Writing-Review & Editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvp.2021.101694.

References

References included in the present meta-analysis are marked with an asterisk.

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology, 25*(3), 273–291. https://doi.org/10.1016/j.jenvp.2005.08.002

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1080/08870446.2011.613995

Babutsidze, Z., & Chai, A. (2018). Look at me saving the planet! the imitation of visible green behavior and its impact on the climate value-action gap. *Ecological Economics, 146*, 290–303. https://doi.org/10.1016/j.ecolecon.2017.10.017

* Baca-Motes, K., Brown, A., Gneezy, A., Keenan, E. A., & Nelson, L. D. (2013). Commitment and behavior change: Evidence from the field. *Journal of Consumer Research, 39*(5), 1070–1084. https://doi.org/10.1086/667226.

Bartholomew, K. J., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2009). A review of controlling motivational strategies from a self-determination theory perspective: Implications for sports coaches. *International Review of Sport and Exercise Psychology, 2*(2), 215–233. https://doi.org/10.1080/17509840903235330

* Bergquist, M., Nilsson, A., & Ejelöv, E. (2019). Contest-based and norm-based interventions: (How) do they differ in attitudes, norms, and behaviors?. *Sustainability, 11*(2). https://doi.org/10.3390/su11020425. e425.

Blanken, I., van de Ven, N., Zeelenberg, M., & Meijers, M. H. (2014). Three attempts to replicate the moral licensing effect. *Social Psychology, 45*(3), 232–238. https://doi.org/10.1027/1864-9335/a000189

Borm, G. F., & Donders, A. R. T. (2009). Updating meta-analyses leads to larger type I errors than publication bias. *Journal of Clinical Epidemiology, 62*(8), 825–830. https://doi.org/10.1016/j.jclinepi.2008.08.010

Bradshaw, E. L., Ryan, R. M., Noetel, M., Saeri, A. K., Slattery, P., Grundy, E., & Calvo, R. (2021). Information safety assurances increase intentions to use COVID-19 contact tracing applications, regardless of autonomy-supportive or controlling message framing. *Frontiers in Psychology, 11*, Article e3772. https://doi.org/10.3389/fpsyg.2020.591638

Breffara Bret, B., Beffara Bret, A., & Nalborczyk, L. (2021). A fully automated, transparent, reproducible, and blind protocol for sequential analyses. *Meta-Psychology, 5*, 1–15. https://doi.org/10.15626/MP.2018.869

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cadario, R., & Chandon, P. (2020). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science, 39*(3), 459–665. https://doi.org/10.1287/mksc.2018.1128

* Carlsson, F., Jaime, M., & Villegas, C. (2020). Behavioral spillover effects from a social information campaign. *Journal of Environmental Economics and Management.* , Article e102325. https://doi.org/10.1016/j.jeem.2020.102325.

* Carrico, A. R., Raimi, K. T., Truelove, H. B., & Eby, B. (2018). Putting your money where your mouth is: An experimental test of pro-environmental spillover from reducing meat consumption to monetary donations. *Environment and Behavior, 50*(7), 723–748. https://doi.org/10.1177/0013916517713067.

Center for Dissemination. (2009). *Systematic reviews CRD's guidance for undertaking reviews in health care*. CRD. https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19*(2), 211–229. https://doi.org/10.1037/a0032968

* Clot, S., Grolleau, G., & Ibanez, L. (2013). Self-licensing and financial rewards: Is morality for sale?. *Economics Bulletin, 33*, 2298–2306 http://www.accessecon.com/Pubs/EB/2013/Volume33/EB-13-V33-I3-P215.pdf.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*(1), 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology/Psychologie Canadienne, 49*(1), 14–23. https://doi.org/10.1037/0708-5591.49.1.14

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. V. Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology*. Sage Publications Ltd. https://doi.org/10.4135/9781446249215.n21.

Del Re, A. C. (2013). *Compute.es: Compute effect sizes (Version 0.2-2) [Computer software]*. https://cran.r-project.org/package=compute.es.

Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology, 47*, 1–16. https://doi.org/10.1016/j.joep.2014.12.003

Dono, J., Webb, J., & Richardson, B. (2010). The relationship between environmental activism, pro-environmental behaviour and social identity. *Journal of Environmental Psychology, 30*(2), 178–186. https://doi.org/10.1016/j.jenvp.2009.11.006

Echegaray, F., & Hansstein, F. V. (2017). Assessing the intention-behavior gap in electronic waste recycling: The case of Brazil. *Journal of Cleaner Production, 142*, 180–190. https://doi.org/10.1016/j.jclepro.2016.05.064

* Elf, P., Gatersleben, B., & Christie, I. (2019). Facilitating positive spillover effects: New insights from a mixed-methods approach exploring factors enabling people to live more sustainable lifestyles. *Frontiers in Psychology, 9*, Article e2699. https://doi.org/10.3389/fpsyg.2018.02699.

Evans, L., Maio, G. R., Corner, A., Hodgetts, C. J., Ahmed, S., & Hahn, U. (2013). Self-interest and pro-environmental behaviour. *Nature Climate Change, 3*(2), 122–125. https://doi.org/10.1038/nclimate1662

Festinger, L. (1962). *A theory of cognitive dissonance* (2nd ed.). Stanford University Press.

Fielding, K. S., McDonald, R., & Louis, W. R. (2008). Theory of planned behaviour, identity and intentions to engage in environmental activism. *Journal of Environmental Psychology, 28*(4), 318–326. https://doi.org/10.1016/j.jenvp.2008.03.003

Forscher, P. S., Wagenmakers, E., Coles, N. A., Silan, M. A., Dutra, N. B., Basnight-Brown, D., & IJzerman, H. (2020 May 20). The benefits, barriers, and risks of big team science. *PsyArXiv*. https://doi.org/10.31234/osf.io/2mdxh

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., Griffith, L., Oremus, M., Raina, P., Ismaila, A., Santaguida, P., Lau, J., & Trikalinos, T. A. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the effective health care program. *Journal of Clinical Epidemiology, 64*(11), 1187–1197. https://doi.org/10.1016/j.jclinepi.2010.08.010

Galizzi, M. M., & Whitmarsh, L. (2019). How to measure behavioral spillovers: A methodological review and checklist. *Frontiers in Psychology, 10*, Article e342. https://doi.org/10.3389/fpsyg.2019.00342

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

* Geng, L., Cheng, X., Tang, Z., Zhou, K., & Ye, L. (2016). Can previous pro-environmental behaviours influence subsequent environmental behaviours? The licensing effect of pro-environmental behaviours. *Journal of Pacific Rim Psychology, 10*, Article e9. https://doi.org/10.1017/prp.2016.6.

* Geng, L., Chen, Y., Ye, L., Zhou, K., & Chen, Y. (2019). How to predict future pro-environmental intention? The spillover effect of electricity-saving behavior under environmental and monetary framing. *Journal of Cleaner Production, 233*, 1029–1037. https://doi.org/10.1016/j.jclepro.2019.06.088.

Ghesla, C., Grieder, M., & Schmitz, J. (2019). Nudge for good? Choice defaults and spillover effects. *Frontiers in Psychology, 10*, Article e178. https://doi.org/10.3389/fpsyg.2019.00178

Green-Demers, I., Pelletier, L. G., & Ménard, S. (1997). The impact of behavioural difficulty on the saliency of the association between self-determined motivation and environmental behaviours. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 29*(3), 157–166. https://doi.org/10.1037/0008-400X.29.3.157

Gronau, Q. F., van Erp, S., Heck, D. W., Cesario, J., Jonas, K. J., & Wagenmakers, E.-J. (2017). A Bayesian model-averaged meta-analysis of the power pose effect with informed and default priors: The case of felt power. *Comprehensive Results in Social Psychology, 2*(1), 123–138. https://doi.org/10.1080/23743603.2017.1326760

Guardian, T. (2020 February 28). *Luxembourg is first country to make all public transport free*. The Guardian. https://www.theguardian.com/world/2020/feb/28/luxembourg-public-transport-free-nationwide-congestion.

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLoS One, 10*(9), 1–21. https://doi.org/10.1371/journal.pone.0138237

Hicklenton, C., Hine, D. W., & Loi, N. M. (2019). Can work climate foster pro-environmental behavior inside and outside of the workplace? *PLoS One, 14*(10), Article e0223774. https://doi.org/10.1371/journal.pone.0223774

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., & Nosek, B. A. (2014). Investigating variation in replicability. *Social Psychology, 45*, 142–152. https://doi.org/10.1027/1864-9335/a000178

Kossmeier, M., Tran, U. S., & Voracek, M. (2020). Power-enhanced funnel plots for meta-analysis. *Zeitschrift für Psychologie, 228*(1), 43–49. https://doi.org/10.1027/2151-2604/a000392

Kowal, M., Coll-Martín, T., Ikizer, G., Rasmussen, J., Eichel, K., Studzińska, A., & Ahmed, O. (2020). Who is the most stressed during the COVID-19 pandemic? Data from 26 countries and areas. *Applied Psychology: Health and Well-Being, 12*(4), 946–966. https://doi.org/10.1111/aphw.12234

*Lacasse, K. (n.d.). Unpublished data.

Lacasse, K. (2016). Don't be satisfied, identify! Strengthening positive spillover by connecting pro-environmental behaviors to an "environmentalist" label. *Journal of Environmental Psychology, 48*, 149–158. https://doi.org/10.1016/j.jenvp.2016.09.006

* Lacasse, K. (2019). Can't hurt, might help: Examining the spillover effects from purposefully adopting a new pro-environmental behavior. *Environment and Behavior, 51*(3), 259–287. https://doi.org/10.1177/0013916517748164.

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology, 44*(7), 701–710. https://doi.org/10.1002/ejsp.2023

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC Psychology, 4*(1), Article e24. https://doi.org/10.1186/s40359-016-0126-3

* Lanzini, P. (2013). *Pro-environmental spillover in consumer behavior: Existence and drivers [Doctoral dissertation*. Università Ca'Foscari Venezia] https://pdfs.semanticscholar.org/5b09/3354df221639ab27db3657da95673dd01b2a.pdf.

Lanzini, P., & Thøgersen, J. (2014). Behavioural spillover in the environmental domain: An intervention study. *Journal of Environmental Psychology, 40*, 381–390. https://doi.org/10.1016/j.jenvp.2014.09.006

Lee, M., & Wagenmakers, E. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Legate, N., Nguyen, T. T., Weinstein, N., Moller, A. C., Legault, L., Maniaci, M. R., & Primbs, M. (2021 May 30). A global experiment on motivating social distancing during the COVID-19 pandemic. *PsyArXiv*. https://doi.org/10.31234/osf.io/n3dyf

Levitin, D. J. (2016). *A field guide to lies: Critical thinking in the information age*. Penguin.

Li, Y., Levran, O., Kim, J., Zhang, T., Chen, X., & Suo, C. (2019). Extreme sampling design in genetic association mapping of quantitative trait loci using balanced and unbalanced case-control samples. *Scientific Reports, 9*(1), 1–9. https://doi.org/10.1038/s41598-019-51790-w

* Lin, Y. C., & Chang, C. C. A. (2017). Exploring wasteful consumption. *Journal of Environmental Psychology, 49*, 106–111. https://doi.org/10.1016/j.jenvp.2017.01.001.

Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of Social Issues, 63*(1), 117–137. https://doi.org/10.1111/j.1540-4560.2007.00499.x

Lindenberg, S., & Steg, L. (2013). Goal-framing theory and norm-guided environmental behavior. In H. C. M. van Trijp (Ed.), *Encouraging sustainable behavior: Psychology and the environment* (pp. 37–54). Psychology Press.

Lindley, D. V. (1972). *Bayesian statistics: A review*. Society for Industrial and Applied Mathematics.

Lüdecke, D. (2019). *Esc: Effect size computation for meta analysis (Version 0.5.1) [Computer software]*. https://doi.org/10.5281/zenodo.1249218

Magnusson, K. (2021). *Interpreting Cohen's d effect size: An interactive visualization (Version 2.5.0) [Web App]. R Psychologist*. https://rpsychologist.com/cohend/.

* Maki. (2015). The spread of behavior: When, how, and for whom do proenvironmental behaviors spread to other people and other behaviors?. *Doctoral dissertation, University of Minnesota*.

Maki, A., Carrico, A. R., Raimi, K. T., Truelove, H. B., Araujo, B., & Yeung, K. L. (2019). Meta-analysis of pro-environmental behaviour spillover. *Nature Sustainability, 2*(4), 307–315. https://doi.org/10.1038/s41893-019-0263-9

* Margetts, E. A., & Kashima, Y. (2017). Spillover between pro-environmental behaviours: The role of resources and perceived similarity. *Journal of Environmental Psychology, 49*, 30–42. https://doi.org/10.1016/j.jenvp.2016.07.005.

Milinski, M., Sommerfeld, R. D., Krambeck, H.-J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences, 105*(7), 2291–2294. https://doi.org/10.1073/pnas.0709546105

Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science, 7*(2), 109–117. https://doi.org/10.1177/1745691611432343

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1). e1 http://www.systematicreviewsjournal.com/content/4/1/1.

Molto, L., Nalborczyk, L., Palluel-Germain, R., & Morgado, N. (2020). Action effects on visual perception of distances: A multilevel bayesian meta-analysis. *Psychological Science, 31*(5), 488–504. https://doi.org/10.1177/0956797619900336

Moreau, D., & Gamble, B. (2020). Conducting a meta-analysis in the age of Open Science: Tools, tips, and practical recommendations. *PsyArXiv*. https://doi.org/10.31234/osf.io/t5dwg

Nalborczyk, L., Batailler, C., Lœvenbruck, H., Vilain, A., & Bürkner, P. C. (2019). An introduction to bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language, and Hearing Research, 62*(5), 1225–1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006

Nash, N., Whitmarsh, L., Capstick, S., Hargreaves, T., Poortinga, W., Thomas, G., Sautkina, E., & Xenias, D. (2017). Climate-relevant behavioral spillover and the potential contribution of social practice theory. *Wiley Interdisciplinary Reviews: Climate Change, 8*(6), Article e481. https://doi.org/10.1002/wcc.481

Nilsson, A., Bergquist, M., & Schultz, W. P. (2017). Spillover effects in environmental behaviors, across time and context: A review and research agenda. *Environmental Education Research, 23*(4), 573–589. https://doi.org/10.1080/13504622.2016.1250148

Osbaldiston, R., & Sheldon, K. M. (2003). Promoting internalized motivation for environmentally responsible behavior: A prospective study of environmental goals. *Journal of Environmental Psychology, 23*(4), 349–357. https://doi.org/10.1016/S0272-4944(03)00035-5

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews, 5*(1), Article e210. https://doi.org/10.1186/s13643-016-0384-4

* Parag, Y., Capstick, S., & Poortinga, W. (2011). Policy attribute framing: A comparison between three policy instruments for personal emissions reduction. *Journal of Policy Analysis and Management, 30*(4), 889–905. https://doi.org/10.1002/pam.20610.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*(10), 991–996. https://doi.org/10.1016/j.jclinepi.2007.11.010

* Poortinga, W., Whitmarsh, L., & Suffolk, C. (2013). The introduction of a single-use carrier bag charge in Wales: Attitude change and behavioural spillover effects. *Journal of Environmental Psychology, 36*, 240–247. https://doi.org/10.1016/j.jenvp.2013.09.001.

R Core Team. (2020). *R: A language and environment for statistical computing (version 4.0.2) [computer software]. R foundation for statistical computing*. https://www.R-project.org/.

Reams, M. A., Geaghan, J. P., & Gendron, R. C. (1996). The link between recycling and litter: A field study. *Environment and Behavior, 28*(1), 92–110. https://doi.org/10.1177/0013916596281005

Rethlefsen, M., Koffel, J., Kirtley, S., Waffenschmidt, S., & Ayala, A. P. (2019). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *OSF Preprint Repository*. https://doi.org/10.31219/osf.io/sfc38

Rohatgi, A. (2019). *WebPlotDigitizer (version 4.2) [online app]*. https://automeris.io/WebPlotDigitizer.

Ruggeri, K., Većkalov, B., Bojanić, L., Andersen, T. L., Ashcroft-Jones, S., Ayacaxli, N., & Folke, T. (2021). The general fault in our fault lines. *Nature Human Behaviour*, 1–11. https://doi.org/10.1038/s41562-021-01092-x

Saatchi & Saatchi Copenhagen. (2007). *WWF paper dispenser [WWF print advertisement]*. http://www.adsoftheworld.com/media/ambient/paper_dispenser.

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods, 22*(2), 322–339. https://doi.org/10.1037/met0000061

Schütte, L., & Gregory-Smith, D. (2015). Neutralisation and mental accounting in ethical consumption: The case of sustainable holidays. *Sustainability, 7*(6), 7959–7972. https://doi.org/10.3390/su7067959

Schwartz, D., Bruine de Bruin, W., Fischhoff, B., & Lave, L. (2015). Advertising energy saving programs: The potential environmental cost of emphasizing monetary savings. *Journal of Experimental Psychology: Applied, 21*(2), 158–166. https://doi.org/10.1037/xap0000042

Sheeran, P., & Webb, T. L. (2016). The intention–behavior gap. *Social and Personality Psychology Compass, 10*(9), 503–518. https://doi.org/10.1111/spc3.12265

Simmonds, M., Salanti, G., McKenzie, J., Elliott, J., Agoritsas, T., Hilton, J., Perron, C., Akl, E., Hodder, R., Pestridge, C., Albrecht, L., Horsley, T., Platt, J., Armstrong, R., Nguyen, P. H., Plovnick, R., Arno, A., Ivers, N., Quinn, G., & Pearson, L. (2017).

Living systematic reviews: 3. Statistical methods for updating meta-analyses. *Journal of Clinical Epidemiology, 91*, 38–46. https://doi.org/10.1016/j.jclinepi.2017.08.008
* Sintov, N., Geislar, S., & White, L. V. (2019). Cognitive accessibility as a new factor in proenvironmental spillover: Results from a field study of household food waste management. *Environment and Behavior, 51*(1), 50–80. https://doi.org/10.1177/0013916517735638.
Spencer, C. C., Su, Z., Donnelly, P., & Marchini, J. (2009). Designing genome-wide association studies: Sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics, 5*(5), Article e1000477. https://doi.org/10.1371/journal.pgen.1000477
Stan Development Team. (2018). *The Stan core library [Computer software]*. http://mc-stan.org/.
Stanley, D. J., & Spence, J. R. (2014). Expectations for replications: Are yours realistic? *Perspectives on Psychological Science, 9*(3), 305–318. https://doi.org/10.1177/1745691614528518
Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E. J. (2019). A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior Research Methods, 51*(3), 1042–1058. https://doi.org/10.3758/s13428-018-01189-8
Steg, L., Bolderdijk, J. W., Keizer, K., & Perlaviciute, G. (2014). An integrated framework for encouraging pro-environmental behaviour: The role of values, situational factors and goals. *Journal of Environmental Psychology, 38*, 104–115. https://doi.org/10.1016/j.jenvp.2014.01.002
Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology, 29*(3), 309–317. https://doi.org/10.1016/j.jenvp.2008.10.004
* Steinhorst, J., Klöckner, C. A., & Matthies, E. (2015). Saving electricity–for the money or the environment? Risks of limiting pro-environmental spillover when using monetary framing. *Journal of Environmental Psychology, 43*, 125–135. https://doi.org/10.1016/j.jenvp.2015.05.012.
Steinhorst, J., & Matthies, E. (2016). Monetary or environmental appeals for saving electricity? Potentials for spillover on low carbon policy acceptability. *Energy Policy, 93*, 335–344. https://doi.org/10.1016/j.enpol.2016.03.020
* Suffolk, C. (2016). *Rebound and spillover effects: Occupant behaviour after energy efficiency improvements are carried out (Publication No. 92990) [Doctoral dissertation, Cardiff University]. ORCA Online Research @ Cardiff* http://orca.cf.ac.uk/id/eprint/92990.
Su, Y. L., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review, 23*(1), 159–188. https://doi.org/10.1007/s10648-010-9142-7
* Swim, J. K., & Bloodhart, B. (2013). Admonishment and praise: Interpersonal mechanisms for promoting proenvironmental behavior. *Ecopsychology, 5*(1), 24–35. https://doi.org/10.1089/eco.2012.0065.
Szucs, D., & Ioannidis, J. P. A. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biology, 15*(3), Article e2000797. https://doi.org/10.1371/journal.pbio.2000797
Thøgersen, J., & Noblet, C. (2012). Does green consumerism increase the acceptance of wind power? *Energy Policy, 51*, 854–862. https://doi.org/10.1016/j.enpol.2012.09.044
Thøgersen, J., & Ölander, F. (2003). Spillover of environment-friendly consumer behaviour. *Journal of Environmental Psychology, 23*(3), 225–236. https://doi.org/10.1016/S0272-4944(03)00018-5
* Thomas, G. O., Poortinga, W., & Sautkina, E. (2016). The Welsh single-use carrier bag charge and behavioural spillover. *Journal of Environmental Psychology, 47*, 126–135. https://doi.org/10.1016/j.jenvp.2016.05.008.
* Tiefenbeck, V., Staake, T., Roth, K., & Sachs, O. (2013). For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy, 57*, 160–171. https://doi.org/10.1016/j.enpol.2013.01.021.
* Touhey, E. (2019). *The influence of plastic bag bans on pro-environmental behaviors in Rhode Island coastal communities (Publication No. 1468) [Master Thesis, University of Rhode Island]. Digital Commons@URI.*
Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., & van den Bergh, M. P. (2014). Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change, 29*, 127–138. https://doi.org/10.1016/j.gloenvcha.2014.09.004
Udall, A. M., de Groot, J. I., De Jong, S. B., & Shankar, A. (2021). How I see me—a meta-analysis investigating the association between identities and pro-environmental behaviour. *Frontiers in Psychology, 12*, Article e582421. https://doi.org/10.3389/fpsyg.2021.582421.
Urban, J., Bahník, Š., & Kohlová, M. B. (2019). Green consumption does not make people cheat: Three attempts to replicate moral licensing effect due to pro-environmental behavior. *Journal of Environmental Psychology, 63*, 139–147. https://doi.org/10.1016/j.jenvp.2019.01.011
Van der Werff, E., Steg, L., & Keizer, K. (2014). Follow the signal: When past pro-environmental actions signal who you are. *Journal of Environmental Psychology, 40*, 273–282. https://doi.org/10.1016/j.jenvp.2014.07.004
Vu-Ngoc, H., Elawady, S. S., Mehyar, G. M., Abdelhamid, A. H., Mattar, O. M., Halhouli, O., Vuong, N. L., Ali, C., Hassan, U. H., Kien, N. D., Hirayama, K., & Huy, N. T. (2018). Quality of flow diagram in systematic review and/or meta-analysis. *PLoS One, 13*(6), Article e0195955. https://doi.org/10.1371/journal.pone.0195955
Wagenmakers, E.-J. (2017 August 03). *Redefine statistical significance part I: Sleep trolls & red herrings. BayesianSpectacles*. https://www.bayesianspectacles.org/redefine-statistical-significance-part-i-sleep-trolls-red-herrings/.
Wagenmakers, E.-J. (2018 July 05). *Let's poke a pizza: A new cartoon to explain the strength of evidence in a Bayes factor*. BayesianSpectacles https://www.bayesianspectacles.org/lets-poke-a-pizza-a-new-cartoon-to-explain-the-strength-of-evidence-in-a-bayes-factor/.
Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3
Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin, 132*(2), 249–268. https://doi.org/10.1037/0033-2909.132.2.249
Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science, 6*(3), 291–298. https://doi.org/10.1177/1745691611406923
Wilson, D. B. (2001). *Practical meta-analysis effect size calculator [Online calculator]*. https://campbellcollaboration.org/research-resources/effect-size-calculator.html.
* Wolstenholme, E., Poortinga, W., & Whitmarsh, L. (2020). Two birds, one stone: The effectiveness of health and environmental messages to reduce meat consumption and encourage pro-environmental behavioral spillover. *Frontiers in Psychology, 11*, Article e2596. https://doi.org/10.3389/fpsyg.2020.577111.
* Xu, L., Zhang, X., & Ling, M. (2018a). Pro-environmental spillover under environmental appeals and monetary incentives: Evidence from an intervention study on household waste separation. *Journal of Environmental Psychology, 60*, 27–33. https://doi.org/10.1016/j.jenvp.2018.10.003.
* Xu, L., Zhang, X., & Ling, M. (2018b). Spillover effects of household waste separation policy on electricity consumption: Evidence from hangzhou, China. *Resources, Conservation and Recycling, 129*, 219–231. https://doi.org/10.1016/j.resconrec.2017.10.028.
* Zawadzki, S. J. (2015). *The interdependent nature of environmental behaviors: Testing a conceptual framework for behavioral spillover (publication No. 26452) [master thesis, penn state university]. Electronic theses and dissertations for graduate school.*