Routledge
Taylor & Francis Group

Threatening communication: a critical re-analysis and a revised meta-analytic test of fear appeal theory

Gjalt-Jorn Ygram Peters*, Robert A.C. Ruiter and Gerjo Kok

Work & Social Psychology, Faculty of Psychology & Neuroscience, Maastricht University, Maastricht, The Netherlands

Despite decades of research, consensus regarding the dynamics of fear appeals remains elusive. A meta-analysis was conducted that was designed to resolve this controversy. Publications that were included in previous meta-analyses were re-analysed, and a number of additional publications were located. The inclusion criteria were full factorial orthogonal manipulations of threat and efficacy, and measurement of behaviour as an outcome. Fixed and random effects models were used to compute mean effect size estimates. Meta-analysis of the six studies that satisfied the inclusion criteria clearly showed a significant interaction between threat and efficacy, such that threat only had an effect under high efficacy ($d = 0.31$), and efficacy only had an effect under high threat ($d = 0.71$). Inconsistency in results regarding the effectiveness of threatening communication can likely be attributed to flawed methodology. Proper tests of fear appeal theory yielded the theoretically hypothesised interaction effect. Threatening communication should exclusively be used when pilot studies indicate that an intervention successfully enhances efficacy.

Keywords: fear appeals; behaviour change; intervention development; risk perception; perceived risk; perceived threat

Introduction

In 2004, the European Union's ASPECT consortium published a report where they made a recommendation to make health warnings 'mandatory on both sides of all tobacco products' (ASPECT Consortium, 2004, p. 23). This advice is based on results from studies such as those by Hammond, Fong, McDonald, Brown, and Cameron (2004, 2006), and the more recent study by Borland et al. (2009), which states:

> Based on the findings from this study, and the absence of any credible alternative hypothesis to explain the findings, we conclude that the stronger the warnings the greater the reactions, and thus the greater the quitting activity they evoke. That said, the total impact on smoking prevalence is too small to estimate accurately. Health policy-makers should be seeking to implement strong, large, graphic warnings to maximise their effectiveness in achieving their goal of warning people off continuing to smoke. (p. 674 in Borland et al., 2009)

At the same time, strong criticism has been voiced (Ruiter & Kok, 2005, 2006), expressing the view '[...] that the evidence so far is not supportive of scary graphic

---

warnings'. This discussion focuses on the strong, scary graphic warnings that are widely used in mass media interventions (of which over 80% lacks efficacy components; Cohen, Shumate, & Gold, 2007), designed to enhance the perceived severity of a threat (as opposed to emotionally neutral communications explaining one's susceptibility to a threat).

Despite the past six decades of research into threat communication effectiveness (see Witte & Allen, 2000, for an excellent overview), consensus remains elusive. Several reviews and meta-analyses have been conducted into the topic (e.g., de Hoog, Stroebe, & de Wit, 2007; Ruiter, Abraham, & Kok, 2001; Witte & Allen, 2000), but their outcomes were not consistently in line with current theory. Two theories currently prevail: the extended parallel process model (Witte, 1992) and the stage model of processing of fear-arousing communications (de Hoog et al., 2007). Both postulate that behaviour change is the function of a perceived threat, but only when there is sufficient perceived efficacy. A threat is a danger of harm, characterised by the degree of severity and the degree to which one is susceptible to this threat (and a threatening communication is a message conveying one or both of these elements). Efficacy is one's ability to negate the harm, a function of the effectiveness of a potential response in negating the harm (response efficacy) and one's capability to enact that response (self-efficacy). Both theories predict no behaviour change when a threat is not severe, one is not susceptible to it, there exists no effective response, or when one is incapable to execute an effective response. When threat is increased but efficacy is low, defensive reactions are predicted, such as denying the severity of or susceptibility to a threat.

Thus, these theories predict that studies into threatening or fear-arousing persuasive messages (referred to as fear appeals in social psychology) should show an effect on behaviour only if both efficacy and threat are successfully manipulated; and that this effect should manifest itself as an interaction between threat and efficacy in their effect on behaviour (see the left-most graph in Figure 1). However, meta-analyses consistently found main effects of threat and efficacy, and there was no evidence for the predicted interaction effect, implying that higher threat simply results in more behaviour change (de Hoog et al., 2007; Witte, 1994). The fact that these main
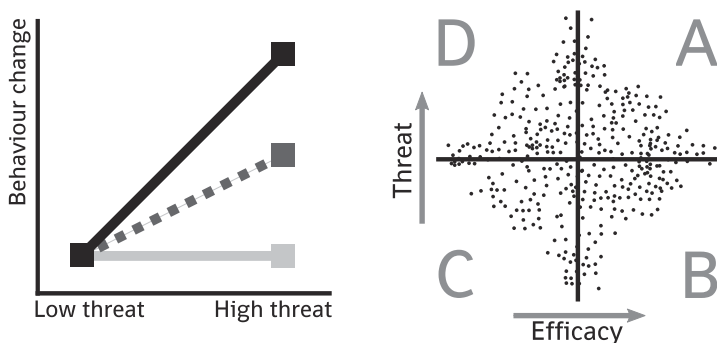

Figure 1. The left graph shows the effect fear appeals have in theory (the black line represents the effect under high efficacy, the grey line the effect under low efficacy, and the dotted line is the main effect observed when efficacy levels are disregarded); the right plot shows the distribution of threat and efficacy associated to a random sample of behaviour-population combinations.

effects of threat were found is in itself in line with theory, as the predicted simple effects average into a small main effect when the (crucial) distinction between low and high efficacy is neglected (see the dotted line in the left-most graph in Figure 1). However, the fact that this interaction was not observed reflects a problematic inconsistency between theory and the empirical evidence. There are two problems with research into threatening communication that may explain this divergence between theory and evidence: the significant outcome bias (SOB) and the intention-behaviour gap.

Fear appeal theory and significant outcome bias (SOB)

The first problem is an unfortunate interaction between the dynamics of fear appeals and the significant outcome bias manifested in publication bias (Levine, Asada, & Carpenter, 2009). Publication bias is the phenomenon that null findings rarely get published. Although it is hard to quantify the extent of publication bias, it is clear that it is widespread. In 1959, Sterling (1959) found that 97% of the hypothesis-testing studies reported significant findings, and 30 years later, this percentage had only decreased slightly to 95% (Sterling, Rosenbaum, & Weinkam, 1995). In fact, many researchers do not even submit null findings. Greenwald found that only 6.4% of researchers submit null findings versus 58.8% for significant effects (Greenwald, 1975). Thus, when a study does not reject the null hypothesis, it has a low probability of being submitted, and even if it is submitted, it has a low probability of being published. This significant outcome bias (SOB), the combination of publication bias and the resulting reluctance of researchers to submit non-significant findings, particularly impacts fear appeal research because of the dynamics of fear appeals.

Theory predicts that fear appeals exclusively work when both efficacy and threat are sufficiently high. Base levels of efficacy and threat, however, vary between individuals. For any given behaviour, individuals can be indifferent (low threat, low efficacy), avoidant (high threat, low efficacy), proactive (low threat, high efficacy) or responsive (high threat, high efficacy; see the risk perception attitude framework, Rimal & Juon, 2010; Rimal & Real, 2003). Of course, these baseline efficacy and threat values also differ per population. For any given behaviour-population combination, self-efficacy can vary from low to high; for example, self-efficacy to exercise four times a week is higher in a population of 20–30 year olds who already exercise three times a week than for a population of 60–80 year olds who have been sedentary all their lives. Similarly, threat can vary from low to high. For smokers who smoke one cigarette at special occasions, say four times a year, the threat of lung cancer is considerably lower than for smokers who smoke 30 cigarettes per day (a complication is that regular smokers have particularly low self-efficacy; Hamilton, Cross, & Resnicow, 2000). Therefore, it is possible to construct a diagram plotting threat against efficacy, for these behaviour-population combinations (see right-most plot in Figure 1).

Roughly speaking, populations in quadrant A are responsive: their members are high in both efficacy and perceived threat. The EPPM predicts that these participants will be easy to convince to conduct the desired behaviour (if perceived threat is not already at the highest level). Participants in quadrant B are proactive: they have high efficacy but low threat. When presented with threatening information, their high efficacy will prevent defensive reactions and their perceived threat will increase, bringing them to quadrant A and causing them to perform the desired

behaviour. Participants in quadrant C are indifferent: low in both threat and efficacy. Presenting them with threatening information will cause them to react defensively, dismissing the threat. Finally, participants in quadrant D are avoidant: they are already high in threat but low in efficacy. Their perceived threat cannot be increased further, and threatening communications again trigger defensive reactions. Delineating behaviour-population combinations as a function of the different baseline threat and efficacy values allows predictions as to the outcomes of experiments where threatening communications (i.e., fear appeals lacking an efficacy component) are presented to behaviour-population combinations in each quadrant.

When an experiment presents a threatening communication to a behaviour-population combination in the upper parts of quadrants A or D, threat is already high, so the manipulation of threat will fail, and the SOB causes these studies to be unlikely to be published. When doing a threat communication experiment using a behaviour-population combination in quadrant C, participants will exhibit defensive reactions (see e.g., Cho & Salmon, 2006), no desirable behaviour change is achieved, and the study is not published (SOB). Thus, current fear appeal theories, in combination with the SOB, predict that most published threat communication experiments are those conducted in populations and with behaviours in quadrant B, where the low baseline threat leaves room for a successful threat manipulation, and the high baseline efficacy inhibits defensive reactions, allowing the threat manipulation to change behaviour.

This interaction of fear appeal dynamics with SOB predicts the two conclusions of meta-analyses of fear appeal experiments so far: there is evidence that higher threat corresponds to more behaviour change, and no evidence for an interaction between threat and efficacy. If this interaction-theory is correct, the conclusions of these meta-analyses exclusively hold for behaviour-population combinations where efficacy is already high (and perceived threat low). To draw conclusions that generalise to all behaviour-population combinations, a meta-analysis should include only studies that manipulate both threat and efficacy. Such an inclusion criterion would prevent the introduction of bias by excluding studies that manipulate threat in a sample that is already high in efficacy (when meta-analysed, such studies would seem to provide evidence for a main effect of threat, whereas in reality, they provide evidence for a simple effect of threat under high efficacy). In theory, studies establishing that the level of efficacy in a population is low (or moderate) could also be included. However, as no validated scales for efficacy for the different behaviours exist, claims as to the relative base level of efficacy for a sample will be problematic, and requiring manipulation avoids these problems by enabling comparison of performance of fear appeals under low and high levels of efficacy.

The association of threat to intention and behaviour

The second problem is the so-called 'intention-behaviour gap' (Godin, Conner, & Sheeran, 2005; Reuter et al., 2010; Sniehotta, Scholz, & Schwarzer, 2005), the phenomenon that intention to perform a behaviour does not consistently result in subsequent performance (intention typically predicts around one-third of behaviour; Cooke & Sheeran, 2004; and medium-to-large changes in intention typically lead to only small-to-medium changes in behaviour; Webb & Sheeran, 2006). Thus, although

intention is typically an accepted determinant of behaviour, at the same time it is clear that some situations exist where intention is not a reliable determinant.

There is reason to believe that threat-induced increases in intention (or attitude) do not engender the desired behaviour. For example, the reasoned action approach (Fishbein & Ajzen, 2010) hypothesises that behaviour is predicted by both intention and self-efficacy; specifically, changes in intention only result in behaviour change if self-efficacy is sufficiently high. This same theory postulates that intention is in turn predicted by attitude and subjective norm (and self-efficacy), and that attitude encompasses, amongst other beliefs, perceived threat. The reasoned action approach predicts that a threatening communication increases perceived threat, which increases attitude, which increases intention, which does *not* increase behaviour, unless self-efficacy is high (Fishbein & Ajzen, 2010, p. 64).

There is a second theory that predicts a decoupling of the intention–behaviour relationship in the case of threatening information: the terror management health model (Goldenberg & Arndt, 2008; Hansen, Winzeler, & Topolinski, 2010). This model postulates that activation of death-related cognitions triggers attempts to jettison these cognitions from consciousness: behaviour change to reduce the actual threat and/or defensive reactions to reduce the salience of the threatening cognitions. One possible defensive reaction is escaping self-awareness by engaging in health-defeating behaviour. Whether one engages in behaviour change or defensive reactions depends on a number of moderators, one of which is again efficacy. Another moderator is self-esteem: it has been shown that when threatening information is presented to persons deriving self-esteem from a risk behaviour, their intention to engage in the risk behaviour decreases, but they simultaneously engage in more risk behaviour (Ben-Ari, Florian, & Mikulincer, 2000).

This decoupling of intention and behaviour in the context of threatening communication is in line with the results of a review that showed that although threat components (severity, susceptibility) significantly predict intention, these same components are not associated to subsequent behaviour (Milne, Sheeran, & Orbell, 2000). Therefore, when a study shows an effect of threatening communication on intention, but fails to measure behaviour, it is possible that the fear appeal is not associated to, or even *decreases*, the desirable behaviour. As it is not possible to establish whether this is the case for studies where behaviour is not measured, only studies measuring behaviour will be included.

The rationale of the current meta-analysis

Thus, this re-analysis and extension of existing fear appeal meta-analyses is based on two hypotheses. First, interaction of fear appeal dynamics with SOB caused the outcomes of meta-analyses so far to be biased; and second, in the case of threatening information, intention is not a reliable outcome measure. If these hypotheses are incorrect, this meta-analysis will draw the same conclusions as prior meta-analyses. However, if the more stringent inclusion criteria presently employed do lead to different conclusions, this is an indication that one or both of the hypotheses may be true. Regardless of the truth of these hypotheses, by only including studies that (1) manipulated threat and efficacy independently, and (2) measured behaviour as an outcome, the current meta-analysis will provide a robust answer to the question whether fear appeal theory holds. In line with the current theoretical state of the art

we hypothesised to find evidence for a so-called simple effect of threat, in that threat is positively associated to behaviour under high efficacy (but not under low efficacy), and evidence for a simple effect of efficacy in that efficacy is positively associated to behaviour under high threat (but not under low threat).

Methods

In the description of the Methods section, we followed the QUOROM guidelines (Moher et al., 2000) as much as possible (where these could be translated to a psychological experiment context, rather than a clinical intervention context). However, this led to a very lengthy Methods section. Therefore, we have decided to make this detailed Methods section available at http://fearappeals.com/meta-analysis/, and have included the most relevant information here.

Inclusion (exclusion) criteria

The current meta-analysis used four inclusion criteria. First, threat, or a component of threat (severity or susceptibility) had to be manipulated. Second, efficacy, or a component of efficacy (response efficacy or self-efficacy) had to be manipulated. In these criteria, 'manipulated' not only means 'influenced'; there also had to exist at least two conditions with different values of these variables (e.g., 'low' and 'high' or 'absent' and 'present'), so studies increasing threat for all participants were excluded. Third, these manipulations had to be orthogonal, resulting in (at least) a 2×2 full factorial experimental design. Fourth, the outcome measure had to be behaviour. Only real behavioural measures were accepted. Requests for information or registrations for sessions, albeit technically behavioural measures, were not accepted, as such behaviours are predicted by very different determinants from target behaviours (e.g., whereas self-efficacy for requesting information about quitting smoking can be assumed to be very high, self-efficacy for actually quitting is likely lower). In addition, because the review is a test of the EPPM which deals with threat to individuals themselves, the behaviour and threat had to concern the participants themselves directly. For example, studies about problems that concern others or about distal problems were excluded (such as a study about industrial whaling by Shelton & Rogers, 1981).

Although the aim of the current meta-analysis is mainly a re-analysis of the study outcomes already examined by previous reviews, we also tried to locate additional studies. Because the most recent meta-analysis was very recent (de Hoog et al., 2007), we limited ourselves to identifying additional articles using queries in bibliographic databases (see below). We imposed no restrictions regarding year of publication or language (but did not encounter any non-English articles).

Data sources

There were two channels for identifying relevant publications. First, a list of studies cited in prior reviews made available by de Hoog et al. (2007) was used to create a spreadsheet with each article's authors, year, title and journal, which was then complemented with the abstracts (see the spreadsheet at http://fearappeals.com/meta-analysis/). In the first screening round, the screener (GJP) excluded studies on the basis of their abstract (i.e., if

it was clear that there was no orthogonal manipulation of threat and efficacy, and/or no behavioural outcome measure). The full-text was acquired for studies that could not be excluded (see Figure 2), and these full-texts were screened by GJP in the second screening round to yield 13 initial inclusions (Chu, 1966; Dabbs & Leventhal, 1966; Duval & Mulilis, 1999; Griffeth & Rogers, 1976; Leventhal, Jones, & Trembly, 1966; Leventhal, Singer, & Jones, 1965; Leventhal, Watts, & Pagano, 1967; Mulilis & Duval, 1995, 1997; Mulilis & Lippa, 1990; Witte, 1994; Wurtele, 1988; Wurtele & Maddux, 1987).

Second, additional relevant literature was identified through a series of queries in the databases PsycINFO and MedLine (the last one at 24 January 2012). The description of this procedure is very lengthy, and as these queries yielded only three additional studies (Grala & McCauley, 1976; Muthusamy, Levine, & Weber, 2009; Ordoñana, González-Javier, Espín-López, & Gómez-Amor, 2009), the detailed description of this procedure, including the flowcharts, is not included here (but available at http://fearappeals.com/meta-analysis/).

Data extraction

For the studies that were included on the basis of the first query, data were extracted through independent coding by RR and GK, using a coding form (available at http://fearappeals.com/meta-analysis/). However, the variation in designs, used measures
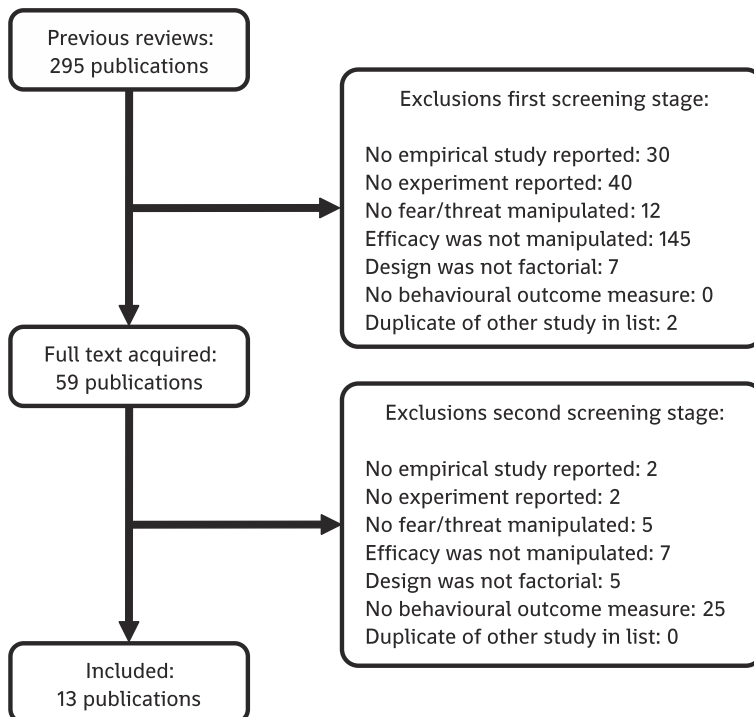

Figure 2. Flowchart showing how the final 13 included publications resulted from processing the 295 publications cited in earlier reviews.

and analyses, and the scarcity of reported information that allowed straightforward computation of effect sizes, eventually necessitated examination of all original texts by one of the authors (GJP). Therefore, for the studies included on the basis of the second and third query, as well as on the basis of the list of publications cited in previous reviews, the effect data was extracted by GJP (the manipulation quality data was rated independently by RR and GK, see below). The extracted data were entered into two spreadsheets (available at http://fearappeals.com/meta-analysis/): one summarising descriptive study data and one to compute effect sizes (and convert effect sizes to Cohen's d where necessary).

The quality of the behavioural measure was directly extracted from the paper (self-report versus an objective measure). When several follow-up measurements were reported, the last measure was taken (Grala & McCauley, 1976; Leventhal et al., 1967). Manipulation quality was coded independently by GK and RR on the basis of extracts of the relevant fragments from the methods section. This coding was based, for threat, on validity of the severity manipulation and validity of the susceptibility manipulation; for efficacy, on validity of the response efficacy manipulation and validity of the self-efficacy manipulation; and for both, on the contrast between the conditions (as difference between the groups, not just the contents of the manipulation for the experimental group, determines effect size, see also de Bruin, Viechtbauer, Hospers, Schaalma, & Kok, 2009). The two judgements were then averaged and dichotomised (see Table 1).

While extracting data, it became clear that a number of included studies did not report data that allowed effect size computation (e.g., only significant statistics from multivariate analyses of variance were reported). In these cases, the authors were contacted (Dabbs & Leventhal, 1966; Griffeth & Rogers, 1976; Leventhal et al., 1966; Mulilis & Duval, 1995, 1997; Mulilis & Lippa, 1990; Witte, 1994; Wurtele, 1988; Wurtele & Maddux, 1987). In all cases where authors replied, the data had been lost, with two exceptions. One author still had the original data and sent this (typed up as an appendix to a thesis; Wurtele & Maddux, 1987), and another author sent the appendix to a meta-analysis she conducted, which contained useable statistics (Witte, 1994). In the first case, it proved impossible to reconstruct which data belonged to which conditions (see the spreadsheet at http://fearappeals.com/meta-analysis/ for more information), so unfortunately this study could still not be included. In the second case, only the main effects of threat and efficacy were available. This was also the case for one other study (Leventhal et al., 1966). Thus, in the end, nine studies were included, seven of which provided effect sizes for the main effects of threat and efficacy, the interaction between those two, and the resulting simple effects; and two of which only provided effect sizes for the main effects (see Table 1).

Analyses

For each study, effect sizes were calculated for the main effects of threat and efficacy, their interaction and the simple effects. When a study examined additional factors, these were collapsed to yield 2×2 designs (e.g., imminency in Chu, 1966), and when a study's threat and/or efficacy manipulation had three levels, only the lowest and highest levels were used (e.g., Duval & Mulilis, 1999). When the dependent variable was dichotomous and had empty cells, 0.5 was added to all frequencies in the study to enable odds ratio calculation (cf. Deeks & Higgins, 2010). Odds ratios, the effect

Table 1. Descriptive information about the included studies.

				Population				Behaviour				Manipulations[a]		
Id	First author	Year	Country	Population	Age[b]	Female	n	Behaviour	Behavioural measure	Type & rating[c]	FU[d]	Manipulations	Threat (quality)	Efficacy (quality)
43	Chu	1966	Taiwan	High school students	?	?	471	Roundworm medication	Willingness to take medicine	W (1)	0	Presentation by experimenter	2.5 (1)	1.5 (0)
61	Duval	1999	US (ca)	Home-owners	45	58%	52	Earthquake preparation	Earthquake preparation	SR (2)	4	Text	2 (0)	2 (0)
147	Leventhal	1965	US (ct)	Students	?	?	59	Tetanus vaccination	Actual vaccination	O (3)	4	Text and photographs	2.5 (1)	2 (0)
150	Leventhal	1967	US (ct)	Students	?	?	80	Smoking	Reported nr. of cigarettes smoked	SR (2)	12	Threat: movie Efficacy: pamphlet	1.5 (0)	3 (1)
317	Muthusamy	2009	Namibia	Students	23	69%	144	Condom use	Self-reported condom use	SR (2)	2	Text and photographs	2.5 (1)	2.5 (1)
328	Ordoñana	2009	Spain	Students	20–30	83%	92	Tetanus vaccination	Self-reported vaccination	SR (2)	8	Video (pictures and voice-over)	2 (0)	2 (0)
334	Grala	1976	US (pa)	High school students	13–17	0%	32	Truancy	100% presence at school	O (3)	4	Presentation by experimenter	2 (0)	2.5 (1)
142	Leventhal	1966	US (ct)	Students	74% 18–24	63%	417	Tetanus vaccination	Actual vaccination	O (3)	?	Text and photographs	2.5 (1)	2 (0)
288	Witte	1994	US (ca)	Students	93% 17–24	50%	54	Condom use	Self-reported condom use	SR (2)	6	Text and pictures	2.5 (1)	2.5 (1)

[a]Quality rating classification in low versus high is given in parentheses;
[b]Duval reported the median;
[c]Willingness (w), self-report (SR) or objective (O). Rating provided in parentheses;
[d]FU = follow-up (in weeks).

size calculated for studies with dichotomous dependent variables, were converted to Cohen's d (Cohen, 1988) to enable synthesis (Borenstein, Hedges, Higgins, & Rothstein, 2009; or see for a simpler formula Chinn, 2000). The values of Cohen's d were corrected for small sample sizes (i.e., d was converted to Hedges' g) by multiplication with *J* (Borenstein et al., 2009).

Finally, the data were imported into R (R Development Core Team, 2012) and the *metafor* package was used to fit fixed and random effects models (Viechtbauer, 2010). Fitting a random effects model in *metafor* provides a number of estimates of heterogeneity, of which we report Q (to test for significance of heterogeneity), and I^2 (as indicator of percentage of variation attributable to between-study effects; Higgins & Thompson, 2002; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006). Funnel plots and forest plots were generated and funnel plot asymmetry was tested using the regression test (with standard error as the predictor). Note that in acknowledgement of the concerns of Simmons, Nelson, and Simonsohn (2011), who explain how limited disclosure of the decisions made in the analysis process can endanger the integrity of science, we have made everything we used for our analyses available through http://fearappeals.com/meta-analysis/.

Results

Descriptive information about the included studies is shown in Table 1, and extracted effect size measures and their significance levels per study are shown in Table 2. At first glance, heterogeneity appears large: studies have been conducted in four countries, with three populations (students, high school students and home owners) of different ages and gender distributions, and with a variety of behaviours (tetanus vaccination, condom use, smoking, roundworm medication, earthquake preparation and truancy). Indeed, when we ran random effects models, almost half the heterogeneity in the effect sizes of the threat main effect and the effect of threat under low efficacy were explained by heterogeneity (see Table 3). Both corresponding Qs almost achieved significance [$Q(8) = 14.66$, $p = 0.066$ and $Q(6) = 10.66$, $p = 0.099$]. Given the low power of Q when only a few studies are used (Borenstein et al., 2009; Huedo-Medina et al., 2006), we took this as an indication of heterogeneity.

Heterogeneity

To explore this heterogeneity, we ran a number of explorative moderation analyses using the quality ratings of the behaviour measure and the two manipulations as moderators. We ran five analyses for each effect (i.e., the two main effects, the interaction and the four simple effects): one with two dummy variables for the behavioural measure ('low', corresponding to rating 1, and 'high' corresponding to rating 3), one for each dummy variable separate (given the low power of these moderation analyses), and one for each dichotomised manipulation quality variable. Thus, in total, we ran 35 moderation analyses (see http://fearappeals.com/meta-analysis/ for the output). The only moderator that emerged was the quality of the behavioural measure, specifically, low quality (this was the only moderator to reach $p < 0.05$: specifically, $p = 0.0065$ in the prediction of the threat main effect, and $p = 0.0171$ in the prediction of the effect of threat under high efficacy).

S18

G.-J.Y. *Peters* et al.

Table 2. The extracted effect size measures and their significance levels in each study.

Study id	Threat (main effect)		Efficacy (main effect)		Threat × Efficacy (interaction)		Threat (low efficacy)		Threat (high efficacy)		Efficacy (low threat)		Efficacy (high threat)	
	d	p	d	p	d	p	d	p	d	p	d	p	d	p
43	0.64	0.000	0.29	0.050	−1.34	0.025	0.36	0.059	1.71	0.000	0.12	0.507	1.46	0.001
61	−0.14	0.316	−0.03	0.825	−0.44	0.003	−0.58	0.008	0.30	0.151	−0.47	0.032	0.41	0.049
147	−0.14	0.697	0.99	0.017	−0.93	0.356	−0.78	0.358	0.15	0.729	0.60	0.269	1.53	0.023
150	−0.14	0.220	0.57	0.000	−0.50	0.000	−0.63	0.000	0.36	0.033	0.08	0.625	1.07	0.000
317	0.15	0.071	0.33	0.000	−0.24	0.005	−0.09	0.468	0.39	0.002	0.09	0.468	0.56	0.000
328	0.55	0.150	0.82	0.049	0.03	0.975	0.62	0.470	0.59	0.177	0.88	0.263	0.85	0.089
334	−0.45	0.371	1.01	0.074	0.62	0.622	0.00	1.000	−0.62	0.293	1.28	0.089	0.66	0.453
142	0.41	0.079	0.51	0.037										
288	0.34	0.219	0.94	0.001										

The only study with a very weak behavioural measure was study #43, where high school students indicated their willingness to take medication for roundworm, with the understanding that willingness would mean they would indeed be taking the medication. This willingness was measured immediately after the manipulation (i.e., using a follow-up of 0 weeks). Thus, study #43 differed from the other studies in two important respects: first, the influenced behaviour was a behaviour that was extremely easy to perform, with medication simply provided to the participants (i.e., it was likely that baseline efficacy is high); and second, behaviour (i.e., willingness) was measured immediately after the manipulation. Perhaps this outcome measure resembles attitude or intention more than behaviour: the effect sizes from this study do suggest this, as this study showed the highest effects of threat, achieving an exceptionally high effect size for the effect of threat under high efficacy ($d = 1.71$; recall the stronger associations of threat to intention). In retrospect, therefore, the judgement call that this study used a valid behavioural outcome measure appears to have been erroneous, and the study should not have been included.

Therefore, we ran a second random effects model excluding study #43 (see Table 3). Heterogeneity was considerably reduced, with at most 29% of the heterogeneity explained by between-study differences (only for the main effect of efficacy). Heterogeneity did not approach significance for any effect. Therefore, we decided to run a fixed effects model (still excluding study #43). The outcomes were identical to the random effects model excluding study #43, and these are the outcomes that we will base our conclusions on.

Meta-analysis

The mean effect size measures are shown in Table 3, a graphical representation of the meta-analysis is shown in Figure 3 (a forest plot), and a graphical representation of the mean effect sizes including the confidence intervals of these mean effect size measures is shown in Figure 4. The interaction between threat and efficacy was significant ($p = 0.004$), which means that the effect size and significance levels for the main effects of threat ($d = 0.11$, $p = 0.215$) and efficacy ($d = 0.49$, $p < 0.0001$) have no meaning (i.e., there exists no effect of threat, as the effect depends on the value of efficacy, and vice versa). This interaction effect in each included study is illustrated in Figure 5. To illustrate the meta-analysed interaction effect, for each study, the outcome of every combination of threat and efficacy was ranked from 1 (worst outcome) to 4 (best outcome). A graph showing the average ranks, weighed by study size, is shown in Figure 6.

When looking at the simple effects, our hypotheses regarding the nature of this interaction were confirmed. There was no effect of threat when efficacy was low ($d = -0.31$, $p = 0.061$), but there was an effect when efficacy was high ($d = 0.31$, $p = 0.036$); and there was no effect of efficacy when threat was low ($d = 0.07$, $p = 0.689$), but there was an effect when threat was high ($d = 0.71$, $p < 0.0001$). Interestingly, the effect of threat under low efficacy is borderline significant and *negative*, suggesting that high threat may decrease behaviour when efficacy is low. These opposite effects of threat under low and high efficacy are clearly visible in Figure 6. Threat had such opposite effects in two-thirds of the included studies (see Figure 5).

Table 3. The outcomes of the meta-analytic random and fixed effects models including and excluding study #43.

| Model	Threat (main effect)		Efficacy (main effect)		Threat × Efficacy (interaction)		Threat (at low efficacy)		Threat (at high efficacy)		Efficacy (at low threat)		Efficacy (at high threat)	
	$d/Q/I^2$	p	$d/Q/I^2$	p	$d/Q/I^2$	p	$d/Q/I^2$	p	$d/Q/I^2$	p	$d/Q/I^2$	p	$d/Q/I^2$	p
Random, incl. #43	0.20	0.0925	0.46	<0.0001	−0.38	0.0011	−0.14	0.4993	0.40	0.0053	0.07	0.6114	0.77	<0.0001
Q [8 or 6 Df]	14.66	0.0662	10.72	0.5902	4.64	0.5902	10.66	0.0994	8.61	0.1968	4.12	0.6604	4.64	0.5902
Heterogeneity (%)	45%		21%		0%		48%		0%		0%		0%	
Random, excl. #43	0.11	0.2154	0.52	<0.0001	−0.34	0.0039	−0.32	0.0742	0.31	0.0356	0.07	0.6886	0.71	<0.0001
Q [7 or 5 Df]	7.24	0.4045	9.473	0.2205	1.98	0.8515	3.73	0.5889	2.92	0.7118	4.10	0.5347	3.09	0.6859
Heterogeneity (%)	0%		29%		0%		7%		0%		0%		0%	
Fixed, excl. #43	0.11	0.2154	0.49	<0.0001	−0.34	0.0039	−0.31	0.0609	0.31	0.0356	0.07	0.6886	0.71	<0.0001
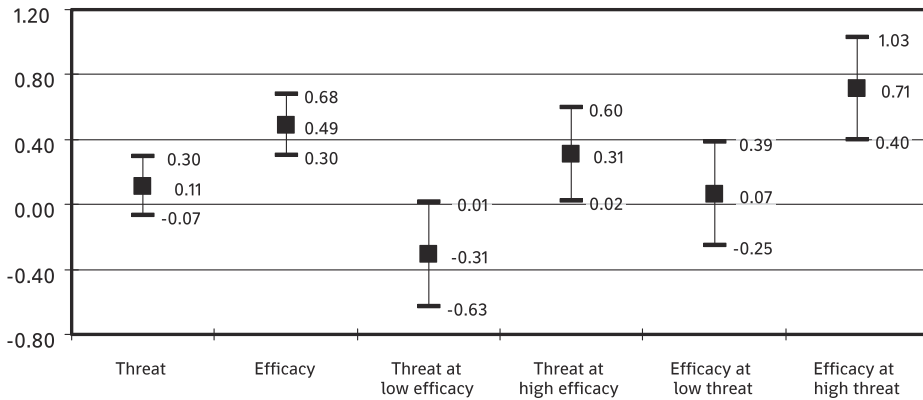
Figure 3. Visual representation of the meta-analysis, including 95% confidence intervals for the main and simple effects.

Publication bias

To assess publication bias, we ordered two sets of funnel plots for every effect: one set based on the random effects meta-analysis including all studies, and one based on the fixed effects meta-analysis excluding study #43 (see http://fearappeals.com/meta-analysis/). We also tested the null hypothesis that each funnel plot was symmetrical. Of these 10 significance tests, only the test for the main effect of efficacy was significant (z $= 2.07$, $p = 0.039$), and only for the random effects model including study #43. The funnel plots suggest that the shift to non-significance (due to omitting study #43) is likely to reflect decreased power, as the plot shows only two effect sizes below the mean effect size estimate (and six studies at or above the mean effect size estimate). Thus, there is some evidence of publication bias regarding the main effect of efficacy, which may mean that this estimate is biased (i.e., the true effect size is likely lower).

In addition to inspection of the funnel plots, correlations between effect size and sample size and methodological quality measures were calculated (Levine et al., 2009). We examined the distribution of positive and negative correlations, as an abundance of strong negative correlations indicates publication bias. This distribution appeared symmetrical, providing no evidence of publication bias (for the exact correlations, see the spreadsheet available at http://fearappeals.com/meta-analysis/).

Discussion

The current paper reports the results of a meta-analysis of fear appeal dynamics, specifically, a critical re-analysis and extension of previous meta-analyses. These previous meta-analyses suffered two problems that were currently addressed: problematic interaction of the significant outcome bias, SOB, with theoretical fear appeal dynamics; and the inclusion of attitude and/or intention as outcome measures. The results are in line with the theoretical predictions. Threat and efficacy interact in their effects on behaviour, such that threat only has an effect if efficacy is high, and efficacy only has an effect is threat is high.
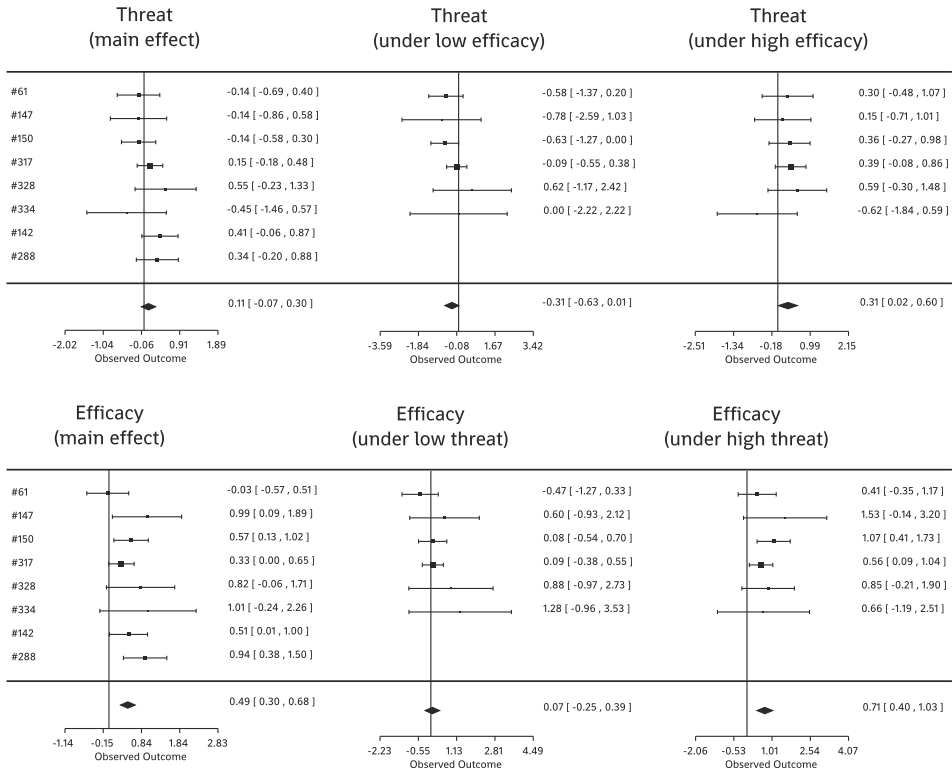
Figure 4. Forest plots with 95% confidence intervals of the meta-analyses of the main and simple effects of threat and efficacy.

Although the hypotheses were confirmed, the current meta-analysis does not allow conclusions as to the hypothesised problems of previous meta-analyses (the interaction of SOB with fear appeal dynamics and the inclusion of non-behavioural
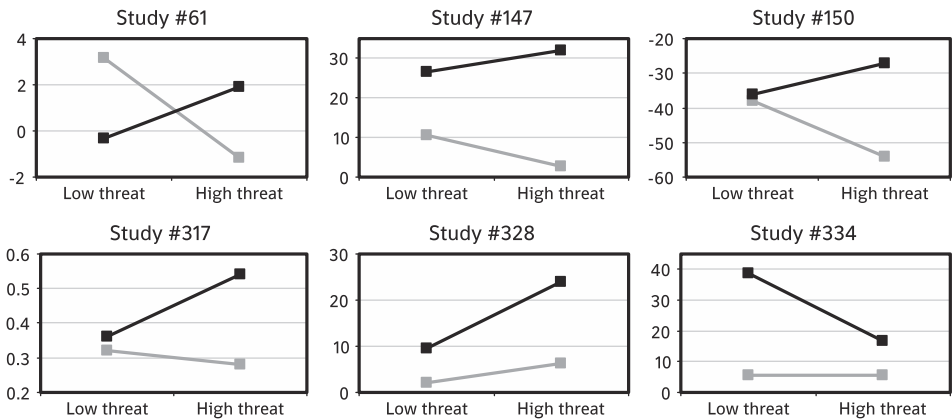

Figure 5. Illustration of the interaction effect between threat and efficacy on behaviour, in each included study. The grey lines reflect the effects of threat under low efficacy and the black lines reflect the effects of threat under high efficacy.
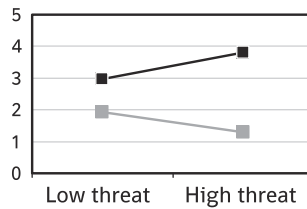
Figure 6. Illustration of the interaction effect between threat and efficacy on behaviour, based on weighed ranked outcomes. The grey line reflects the effect of threat under low efficacy and the black line reflects the effect of threat under high efficacy.

outcome measures). Such conclusions would require dedicated tests (although the results of study #43 support the relevance of true behavioural outcome measures). To examine the interaction of SOB with fear appeal dynamics, uniform threat and efficacy measures need to be developed, and baseline threat and efficacy should be established for different behaviour-population combinations. To examine the differential effects of fear appeals on attitude, intention and behaviour, full factorial experiments should include all three as outcome measures. However, regardless of whether these hypotheses were correct, the current meta-analysis does resolve the controversy regarding the outcomes of previous fear appeal meta-analyses.

The evidence clearly supports theories that hold that efficacy needs to be high before there can be an effect of threatening information [e.g., the EPPM, Witte, 1992; the terror management health model (Goldenberg & Arndt, 2008); and, to a lower degree, the stage model of processing of fear-arousing communications (de Hoog et al., 2007), as this model focuses mainly on the relative roles of susceptibility and severity]. These results are in line with the outcomes of two other meta-analyses. One meta-analysis compared different behaviour change methods (or techniques) and found that 'no threat-inducing argument had any positive behavioural effect whatsoever' (Albarracín et al., 2005, p. 882). Similarly, as another meta-analysis into condom use, where long-term effects of fear appeals were studied, '[...] clearly shows, inducing fear is not an effective way to promote HIV-relevant learning or condom use either immediately following the intervention or later on' (Earl & Albarracín, 2007, p. 504). These outcomes are also in line with recent social neurocognitive evidence (Kessels, Ruiter, & Jansma, 2010) and eye-tracking evidence (Nielsen & Shapiro, 2009) that shows that attention is automatically diverted from threatening information in a high-risk population.

Thus, it seems that the controversy that has long surrounded fear appeal research can at least partly be attributed to methodology. It appears that conclusions were based on results from studies with designs that in fact precluded conclusions as to threat communication effectiveness. Studies that at first glance appeared to support the effectiveness of threatening communications may, at closer inspection, reveal designs (not manipulating efficacy), populations (high in baseline efficacy), or outcome measures (other than behaviour) that preclude conclusions concerning the effectiveness of threatening communications. For example, Latour, Snipes, and Bliss (1996) report the results on an experiment where higher threat was positively associated to behaviour. However, this behaviour was the purchase of a stun gun, which participants were informed was for sale in the area for $60–$80: an excellent example of a behaviour for which most people will have a high efficacy. Evidence

from such a study means that threatening communications are effective *when efficacy is high*: not in general. Another example is a recent paper addressing disgust and fear, concluding that 'health campaign producers would be well advised to continue using message strategies that focus on presenting either tobacco-related health threats or presenting disgusting images' whenever message goals included enhancing awareness and/or learning (Leshner, Bolls, & Thomas, 2009, p. 457). However, this study did not contain any behavioural outcome measures; this conclusion was based on the observation that recognition was lower for low fear (and disgust). The current meta-analysis clearly indicates, however, this although such communications may increase recognition, they either enhance health-defeating behaviour or have no effect.

Similar methodological flaws can explain studies that inspired such strong recommendations in favour of threatening information as cited in the Introduction (e.g., those by Hammond et al., 2004, and Borland et al., 2009). These studies, for example, had no experimental designs, and therefore, allow no causal conclusions (Wilkinson et al., 1999). The inconsistency of these studies' outcomes with evidence from studies that did have experimental designs suggests that the associations that were observed in the non-experimental studies had extraneous causes (e.g., bias in self-reports).

Perhaps the persistence of this controversy can partly be explained by the intuitive appeal of threatening communication. For example, lay target group members often suggest the use of confronting, threatening communications (Goodall & Appiah, 2008; Gorn, Lavack, Pollack, & Weinberg, 1996). In addition, there is the equally intuitive assumption that a target populations' judgement of a communications effectiveness indicates true effectiveness (Vardavas, Connolly, Karamanolis, & Kafatos, 2009). The current meta-analysis clearly suggests that this assumption is wrong: target group members' ideas about effective behaviour change methods do not translate into actual effectiveness, at least not when it concerns threatening communication. Note that this does not mean that the target group should not be involved in intervention development: there are a number of moments in the intervention development process where their input is crucial (see e.g., pp. 174–176 of Bartholomew, Parcel, Kok, Gottlieb, & Fernández, 2011).

A somewhat worrying finding of this meta-analysis was that under low efficacy, the effect of threat was negative and almost significant. This is in line with the terror management health model (Goldenberg & Arndt, 2008), which suggests that threatening information can cause people to engage in health-defeating behaviour. If future research confirms this association, that would render the use of threatening information very risky. If an intervention developer is not very certain that either the population is high in response *and* self-efficacy, or that a given relevant intervention will manage to considerably increase response *and* self-efficacy, threatening communications should be avoided (in fact, the significant main effect of efficacy and the strong effect of efficacy under high threat that were found in the current meta-analysis suggest that a focus on efficacy is a better bet in any case).

If fear appeals are the wrong instrument to achieve behaviour change, which behaviour change method (or technique) should health promoters use? In a meta-analysis of Internet interventions, the five most powerful behaviour change methods were stress management, general communication skills training, modelling or demonstration of the behaviour, relapse prevention/coping planning and facilitation of social comparison (Webb, Joseph, Yardley, & Michie, 2010). The meta-analysis

into condom use interventions cited above concluded that in the case of condom use promotion, active interventions outperform passive interventions and within such active interventions, it is desirable to present information and behavioural skills in combination with self-management training or HIV testing and counselling (but see the decision trees; Albarracín et al., 2005). A number of recent reviews into other specific behaviours and populations is also available (e.g., Dombrowski et al., 2012; Taylor et al., 2012; Thoolen, de Ridder, & Lensvelt-Mulders, 2012). Ultimately, however, the behaviour change method (or technique) to use should ideally be decided specifically for every situation on the basis of relevant determinants and additional information compiled through the use of protocols such as Intervention Mapping (Bartholomew et al., 2011).

The main strength of the current meta-analysis is formed by the stringent inclusion criteria. Unfortunately, these inclusion criteria also caused the weakness of this meta-analysis: only eight studies were included in the meta-analysis of the main effects, and only six studies in the meta-analysis of the interaction effect and the simple effects (excluding study #43). Therefore, it is possible that this meta-analysis suffers from a power problem. If more studies had been conducted, outcomes could have been different in one of two ways. First, we may have made Type 1 errors: it is possible that additional studies would have different outcomes, nullifying our current conclusions. In this case, the currently included studies would show mainly error-variance, reflected by little consistency in the effect size patterns of the individual studies. Second, it is possible that we have made Type 2 errors: additional studies may show similar outcomes, increasing the significance of the currently identified effect sizes (e.g., the negative effect of threat under low efficacy would achieve significance). In this case, the currently included studies would show mainly true effect sizes, reflected by consistency in the effect size patterns of the individual studies. These patterns (shown in Figure 5) seem to support the second scenario. Note that regardless of which scenario is true, the conclusion remains the same: threatening communication is better avoided. After all, if we made Type 1 errors, threats are ineffective; and if we made Type 2 errors, threats are potentially harmful. Still, this paucity of research remains a shortcoming of the current meta-analysis, also because it is very difficult to assess the extent of publication bias with only six studies.

In any case, it is important to realise that these six studies together constitute *all the evidence* that exists on the effectiveness of threatening communications. Therefore, the conclusion that this evidence base is so small in the first place is in fact an important contribution of this study, given the prevalence of recommendations to widely apply threatening communications (e.g., Borland et al., 2009; Hammond, Fong, McDonald, Brown, & Cameron, 2006; Latour et al., 1996; Leshner et al., 2009), in combination with the fact that health campaign planners follow these recommendations (e.g., see this recent CDC campaign; Stobbe, 2012). The current meta-analysis strongly signals that these recommendations should cease as soon as possible, because (1) there are very few studies that could theoretically have supported the use of threatening communications; and (2) those studies that do exist do not support the wide application of threatening communications. Instead, they indicate that using threatening communication is at best ineffective, and at worst causes health-defeating behaviour, unless the intervention contains an element that effectively enhances response efficacy and (especially) self-efficacy.

There are three practical implications of this meta-analysis. One concerns intervention developers; one concerns researchers; and one concerns journal editors and managers.

First, intervention developers should employ utmost caution when using severity-based threatening communication (note that emotionally neutral, susceptibility-based information appears less dangerous and equally effective; de Hoog et al., 2007). Threatening communication only works when either the baseline efficacy is high in the relevant behaviour-population combination, or when the intervention also includes a potent efficacy-enhancing element. An overview of effective methods to enhance efficacy is provided on pages 342–344 of Bartholomew et al. (2011). Note that to enhance efficacy, a simple recommendation (e.g., the recommendation to quit smoking, to call a phone number, to moderate drinking) is insufficient. The efficacy manipulations reported in the studies included in this meta-analysis entailed at least a paragraph of text, and in fact, using active rather than passive methods is recommended (Albarracín et al., 2005). Mere recommendations at best enhance response efficacy, but without high self-efficacy, threatening communication will have no effect, or worse, backfire. Given this danger of backfiring, developing a behaviour change intervention that uses threatening elements is ineffective and unethical unless pilot tests indicate that the intervention also reliably enhances response and self-efficacy. In general, the evidence indicates that intervention developers should look to the theory for guidance, as theory-based interventions have been found to be more effective than non-theory-based interventions (Peters, Kok, Ten Dam, Buijs, & Paulussen, 2009; Webb et al., 2010). When restricted to mass media, it will probably be wisest to resort to a behaviour change method that does not involve emphasising negative consequences of a behaviour, and if that cannot be avoided, at least make sure the communication is not threatening, emotional or confronting (see also de Hoog et al., 2007).

As for researchers in behaviour change research: proper tests of the EPPM should receive higher priority. Although the current meta-analysis sheds light on the controversy that dominated the fear appeal literature for the past decades, these conclusions are only based on six studies. The fact that in all these years, only six real tests of the EPPM have been conducted, or at least published, is very worrisome given the popularity of this method among intervention developers. Such wide application should ideally be based on clear evidence: several empirical tests, full factorial $2 \times 2 \times 2 \times 2$ experiments where severity, susceptibility, response efficacy and self-efficacy are manipulated and where the target behaviour is measured at follow-up (see Wurtele & Maddux, 1987). Note that 2×2 manipulations (only 'threat' and 'efficacy') are insufficiently specific; for example, there is evidence that susceptibility has a different (much stronger) effect than severity (de Hoog et al., 2007).

To enable future meta-analysis of such studies, researchers should pay attention to two things. First, means, standard deviations and sample sizes should always be reported for all cells in the design. Second, the manipulations should be described extensively. Ideally, such descriptions use definitions consistent with existing behaviour change taxonomies. A recent example is the taxonomy developed by Abraham and Michie (2008), which has the advantage that it is accompanied by a coding scheme that is being tested for reliability. This taxonomy was preceded by Intervention Mapping, which also contains a taxonomy (IM; Bartholomew, Parcel,

& Kok, 1998; Bartholomew et al., 2011). IM has the advantage that it has a further developed vocabulary, distinguishing theoretical methods for behaviour change from parameters for effectiveness and practical applications (Schaalma & Kok, 2009). Ideally, therefore, researchers use the more extensive IM vocabulary to describe the theoretical methods and practical applications used in their manipulations, at the same time making sure to make their descriptions fit with the Abraham and Michie (2008) taxonomy.

Finally, part of the responsibility rests with editors and managers of journals. Researchers can hardly be blamed for not conducting replications when replications are as a rule rarely published (Hubbard & Armstrong, 1994; Sterling et al., 1995). Similarly, the significant outcome bias (SOB) has already caused most researchers to no longer even submit non-significant findings (Greenwald, 1975). As long as journals do not eliminate their SOB, it will remain virtually impossible for an evidence base to accumulate that enables the development of a toolbox of effective behaviour change methods (or the development of a cumulative science of behaviour change; Michie & Johnston, 2012). Decades ago, there may have existed a justification for a SOB, as journals were forced to be selective in their publications as the state of technology limited the number of papers they could publish each issue. Now that paper versions of journals have been rendered obsolete in favour of online publications, the only effect of a SOB that remains is to hamper scientific progress, as shown by the very limited number of proper tests of one of the most widely applied behaviour change methods.

In conclusion, warning labels on packs of cigarettes seem ill-advised. They may in fact increase smoking among smokers who derive self-esteem from their identity as a smoker. More health benefits would be achieved if the areas currently reserved for warning labels would be used for a message to enhance efficacy or influence other determinants that have been found to play a role in ceasing smoking (such as subjective norm; note that attitude, the construct encompassing perceived threat, has been found to have only a weak influence; Topa & Moriano, 2010). Given the minimal, or even negative, effects we can expect from threatening communication, the potential of evidence- and theory-based communications on cigarette pack labels is promising.

References

ASPECT Consortium. (2004). *Tobacco or health in the European Union: Past, present and future. The ASPECT consortium*. Luxembourg: Office for Official Publications of the European Communities. Retrieved from http://ec.europa.eu/health/archive/ph_determinants/life_style/tobacco/documents/tobacco_fr_en.pdf

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology*, 27, 379–387.

Albarracín, D., Gillette, J.C., Earl, A.N., Glasman, L.R., Durantini, M.R., & Ho, M.H. (2005). A test of major assumptions about behavior change: A comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychological Bulletin*, 131(6), 856–897. doi:10.1037/0033-2909.131.6.856

Bartholomew, L.K., Parcel, G.S., & Kok, G. (1998). Intervention Mapping: A process for developing theory- and evidence-based health education programs. *Health Education and Behavior*, 25(5), 545–563.

Bartholomew, L.K., Parcel, G.S., Kok, G., Gottlieb, N.H., & Fernández, M.E. (2011). *Planning health promotion programs: An intervention mapping approach* (3rd ed). San Francisco, CA: Jossey-Bass.

Ben-Ari, O.T., Florian, V., & Mikulincer, M. (2000). Does a threat appeal moderate reckless driving? A terror management theory perspective. *Accident Analysis and Prevention*, *32*(1), 1–10. doi:10.1016/S0001-4575(99)00042-1

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons.

Borland, R., Yong, H.-H., Wilson, N., Fong, G.T., Hammond, D., Cummings, K.M., … McNeill, A. (2009). How reactions to cigarette packet health warnings influence quitting: Findings from the ITC Four-Country survey. *Addiction*, *104*(4), 669–675. doi:10.1111/j.1360-0443.2009.02508.x

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*(22), 3127–3131. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11113947

Cho, H., & Salmon, C.T. (2006). Fear appeals for individuals in different stages of change: Intended and unintended effects and implications on public health campaigns. *Health Education*, (781062704). doi:10.1207/s15327027hc2001

Chu, G.C. (1966). Fear arousal, efficacy, and imminency. *Journal of Personality and Social Psychology*, *4*(5), 517–524.

Cohen, E.L., Shumate, M.D., & Gold, A. (2007). Anti-smoking media campaign messages: Theory and practice. *Health Communication*, *22*(2), 91–102.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooke, R., & Sheeran, P. (2004). Moderation of cognition-intention and cognition-behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *The British Journal of Social Psychology*, *43*(2), 159–186. doi:10.1348/0144666041501688

Dabbs, J.M., & Leventhal, H. (1966). Effects of varying the recommendations in a fear-arousing communication. *Journal of Personality and Social Psychology*, *4*(5), 525–531. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5972084

Deeks, J.J., & Higgins, J.P.T. (2010). Statistical algorithms in Review Manager 5. Retrieved from http://ims.cochrane.org/revman/documentation/Statistical-methods-in-RevMan-5.pdf

de Bruin, M., Viechtbauer, W., Hospers, H.J., Schaalma, H.P., & Kok, G. (2009). Standard care quality determines treatment outcomes in control groups of HAART-adherence intervention studies: Implications for the interpretation and comparison of intervention effects. *Health Psychology*, *28*(6), 668–674. doi:10.1037/a0015989

de Hoog, N., Stroebe, W., & de Wit, J.B.F. (2007). The impact of vulnerability to and severity of a health risk on processing and acceptance of fear-arousing communications: A meta-analysis. *Review of General Psychology*, *11*(3), 258–285. doi:10.1037/1089-2680.11.3.258

Dombrowski, S.U., Sniehotta, F.F., Avenell, A., Johnston, M., MacLennan, G., & Araujo-Soares, V. (2012). Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co- morbidities or additional risk factors for co-morbidities: A systematic review. *Health Psychology Review*, *6*, 7–32.

Duval, T.S., & Mulilis, J.-P. (1999). A person-relative-to-event (PrE) approach to negative threat appeals and earthquake preparedness: A field study. *Journal of Applied Social Psychology*, *29*(3), 495–516.

Earl, A.N., & Albarracín, D. (2007). Nature, decay, and spiraling of the effects of fear-inducing arguments and HIV counseling and testing: A meta-analysis of the short- and long-term outcomes of HIV-prevention interventions. *Health Psychology*, *26*(4), 496–506. doi:10.1037/0278-6133.26.4.496

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press.

Godin, G., Conner, M., & Sheeran, P. (2005). Bridging the intention-behaviour "gap": The role of moral norm. *The British Journal of Social Psychology*, *44*(Pt. 4), 497–512. doi:10.1348/014466604X17452

Goldenberg, J.L., & Arndt, J. (2008). The implications of death for health: A terror management health model for behavioral health promotion. *Psychological Review*, *115*(4), 1032–1053. doi:10.1037/a0013326

Goodall, C., & Appiah, O. (2008). Adolescents' perceptions of Canadian cigarette package warning labels: Investigating the effects of message framing. *Health Communication*, *23*(2), 117–127. doi:10.1080/10410230801967825

Gorn, G.J., Lavack, M., Pollack, C.R., & Weinberg, C.B. (1996). An experiment in designing effective warning labels. *Health Marketing Quarterly*, *14*(2), 43–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10164447

Grala, C., & McCauley, C. (1976). Counseling truants back to school: Motivation combined with a program for action. *Journal of Counseling Psychology*, *23*(2), 166–169. doi:10.1037//0022-0167.23.2.166

Greenwald, A.G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, *82*(1), 1–20. doi:10.1037/h0076157

Griffeth, R.W., & Rogers, R.W. (1976). Effects of fear-arousing components of driver education on students' safety attitudes and simulator performance. *Journal of Educational Psychology*, *68*(4), 501–506. doi:10.1037//0022-0663.68.4.501

Hamilton, G., Cross, D., & Resnicow, K. (2000). Occasional cigarette smokers: Cue for harm reduction smoking education. *Addiction Research*, *8*(5), 419–437.

Hammond, D., Fong, G.T., McDonald, P.W., Brown, K.S., & Cameron, R. (2004). Graphic Canadian cigarette warning labels and adverse outcomes: Evidence from Canadian smokers. *American Journal of Public Health*, *94*(8), 1442–1445.

Hammond, D., Fong, G.T., McDonald, P.W., Brown, K.S., & Cameron, R. (2006). Showing leads to doing: Graphic cigarette warning labels are an effective public health policy. *European Journal of Public Health*, *16*(2), 223–224. doi:10.1093/eurpub/ckl037

Hansen, J., Winzeler, S., & Topolinski, S. (2010). When the death makes you smoke: A terror management perspective on the effectiveness of cigarette on-pack warnings. *Journal of Experimental Social Psychology*, *46*(1), 226–228. Elsevier Inc. doi:10.1016/j.jesp.2009.09.007

Higgins, J.P.T., & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. doi:10.1002/sim.1186

Hubbard, R., & Armstrong, J.S. (1994). Replications and extensions in marketing: Rarely published but quite contrary. *International Journal of Research in Marketing*, *11*(3), 233–248. doi:10.1016/0167-8116(94)90003-5

Huedo-Medina, T.B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, *11*(2), 193–206. doi:10.1037/1082-989X.11.2.193

Kessels, L.T.E., Ruiter, R.A.C., & Jansma, B.M. (2010). Increased attention but more efficient disengagement: Neuroscientific evidence for defensive processing of threatening health information. *Health Psychology*, *29*(4), 346–354. doi:10.1037/a0019372

Latour, M.S., Snipes, R.L., & Bliss, S.J. (1996). Don't be afraid to use fear appeals; an experimental study. *Journal of Advertising Research*, *36*(2), 59–67.

Leshner, G., Bolls, P., & Thomas, E. (2009). Scare' em or disgust 'em: The effects of graphic health promotion messages. *Health Communication*, *24*(5), 447–458. doi:10.1080/10410230903023493

Leventhal, H., Jones, S., & Trembly, G. (1966). Sex Differences in attitude and behavior change under conditions of fear and specific instructions. *Journal of Experimental Social Psychology*, *2*, 387–399.

Leventhal, H., Singer, R., & Jones, S. (1965). Effects of fear and specificity of recommendation upon attitudes and behavior. *Journal of Personality and Social Psychology*, *34*(1), 20–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14313839

Leventhal, H., Watts, J.C., & Pagano, F. (1967). Effects of fear and instructions on how to cope with danger. *Journal of Personality and Social Psychology*, *6*(3), 313–321. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6075213

Levine, T., Asada, K., & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analyses: Evidence and implications of a publication bias against nonsignificant findings. *Communication Monographs*, *76*(3), 286–302. doi:10.1080/03637750903074685

Michie, S., & Johnston, M. (2012). Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. *Health Psychology Review*, *6*, 1–6.

Milne, S., Sheeran, P., & Orbell, S. (2000). Prediction and intervention in health-related behavior: A meta-analytic review of protection motivation theory. *Journal of Applied Social Psychology*, *30*(1), 106–143.

Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D.F. (2000). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *Onkologie*, *23*(6), 597–602. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11441269

Mulilis, J.-P., & Duval, T.S. (1995). Negative threat appeals and earthquake preparedness: A person-relative-to-event (PrE) model of coping with threat. *Journal of Applied Social Psychology*, *25*(15), 1319–1339. doi:10.1111/j.1559-1816.1995.tb02620.x

Mulilis, J.-P., & Duval, T.S. (1997). The PrE model of coping and tornado preparedness: Moderating effects of responsibility. *Journal of Applied Social Psychology*, *27*(19), 1750–1766. doi:10.1111/j.1559-1816.1997.tb01623.x

Mulilis, J.-P., & Lippa, R. (1990). Behavioral change in earthquake preparedness due to negative threat appeals; A test of protection motivation theory. *Journal of Applied Social Psychology*, *20*(8), 619–638.

Muthusamy, N., Levine, T.R., & Weber, R. (2009). Scaring the already scared: Some problems with HIV/AIDS fear appeals in Namibia. *Journal of Communication*, *59*(2), 317–344. doi:10.1111/j.1460-2466.2009.01418.x

Nielsen, J., & Shapiro, S. (2009). Coping with fear through suppression and avoidance of threatening information. *Journal of Experimental Psychology. Applied*, *15*(3), 258–274. doi:10.1037/a0017018

Ordoñana, J.R., González-Javier, F., Espín-López, L., & Gómez-Amor, J. (2009). Self-report and psychophysiological responses to fear appeals. *Human Communication Research*, *35*(2), 195–220. doi:10.1111/j.1468-2958.2009.01344.x

Peters, L.W.H., Kok, G., Ten Dam, G.T.M., Buijs, G.J., & Paulussen, T.G.W.M. (2009). Effective elements of school health promotion across behavioral domains: A systematic review of reviews. *BMC Public Health*, *9*, 182. doi:10.1186/1471-2458-9-182

R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Reuter, T., Ziegelmann, J.P., Wiedemann, A.U., Lippke, S., Schuz, B., & Aiken, L.S. (2010). Planning bridges the intention-behaviour gap: Age makes a difference and strategy use explains why. *Psychology & Health*, *25*, 873–87. doi:10.1080/08870440902939857

Rimal, R.N., & Juon, H.-S. (2010). Use of the risk perception attitude framework for promoting breast cancer prevention. *Journal of Applied Social Psychology*, *40*(2), 287–310.

Rimal, R.N., & Real, K. (2003). Perceived risk and efficacy beliefs as motivators of change. *Human Communication Research*, *29*(3), 370–399.

Ruiter, R.A.C., Abraham, C., & Kok, G. (2001). Scary warnings and rational precautions: A review of the psychology of fear appeals. *Psychology & Health*, *16*, 613–630. doi:10.1080/08870440108405863

Ruiter, R.A.C., & Kok, G. (2005). Saying is not (always) doing: Cigarette warning labels are useless. *European Journal of Public Health*, *15*(3), 329. doi:10.1093/eurpub/cki095

Ruiter, R.A.C., & Kok, G. (2006). Response to Hammond et al. Showing leads to doing, but doing what? The need for experimental pilot-testing. *European Journal of Public Health*, *16*(2), 225. doi:10.1093/eurpub/ckl014

Schaalma, H., & Kok, G. (2009). Decoding health education interventions: The times are a-changin'. *Psychology & Health*, *24*(1), 5–9. doi:10.1080/08870440801995802

Shelton, M.L., & Rogers, R.W. (1981). Fear-arousing and empathy-arousing appeals to help; the pathos of persuasion. *Journal of Applied Social Psychology*, *11*(4), 366–378.

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

Sniehotta, F., Scholz, U., & Schwarzer, R. (2005). Bridging the intention-behaviour gap: Planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology & Health*, *20*(2), 143–160. doi:10.1080/08870440512331317670

Sterling, T.D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance--or vice versa. *Journal of the American Statistical Association*, *54*(285), 30–34.

Sterling, T.D., Rosenbaum, W.L., & Weinkam, J.J. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *The American Statistician*, *49*(1), 108–112.

Stobbe, M. (2012, April 15). CDC Anti-Smoking Ad Campaign Set To Launch. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/2012/03/15/cdc-anti-smoking-ad-campaign_n_1347174.html

Taylor, N., Conner, M., & Lawton, R. (2012). The impact of theory on the effectiveness of worksite physical activity interventions: A meta-analysis and meta-regression. *Health Psychology Review*, *6*(1), 33–73.

Thoolen, B., de Ridder, D., & Lensvelt-Mulders, G.V. (2012). Patient-oriented interventions to improve antibiotic prescribing practices in respiratory tract infections: A meta-analysis. *Health Psychology Review*, *6*(1), 92–112.

Topa, G., & Moriano, J.A. (2010). Theory of planned behavior and smoking: Meta-analysis and SEM model. *Substance Abuse and Rehabilitation*, *1*, 23–33. doi:10.2147/SAR.S15168

Vardavas, C.I., Connolly, G., Karamanolis, K., & Kafatos, A. (2009). Adolescents perceived effectiveness of the proposed European graphic tobacco warning labels. *European Journal of Public Health*, *19*(2), 212–7. doi:10.1093/eurpub/ckp015

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Webb, T.L., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research*, *12*(1), e4. doi:10.2196/jmir.1376

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. doi:10.1037/0033-2909.132.2.249

Wilkinson, L., Aiken, L., Appelbaum, M., Boodoo, G., Kenny, D.A., Kraemer, H., ... Cronbach, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594–604.

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs*, *59*(4), 329–349. doi:10.1080/03637759209376276

Witte, K. (1994). Fear control and danger control; a test of the extended parallel process model (EPPM). *Communication Monographs*, *61*, 113–134.

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education and Behavior*, *27*(5), 591–615. doi:10.1177/109019810002700506

Wurtele, S.K. (1988). Increasing women's calcium intake: The role of health beliefs, intentions, and health value. *Journal of Applied Social Psychology*, *18*(8), 627–639.

Wurtele, S.K., & Maddux, J.E. (1987). Relative contributions of protection motivation theory components in predicting exercise intentions and behavior. *Health Psychology*, *6*(5), 453–466.

All Supplementary Material is available alongside this article on www.tandfonline.com - go to http://dx.doi.org/10.1080/17437199.2012.703527