

Functional brain imaging predicts public health campaign success

Emily B. Falk,^{1,2} Matthew Brook O'Donnell,¹ Steven Thompson,³ Richard Gonzalez,^{2,3} Sonya Dal Cin,^{2,4} Victor Strecher,⁵ Kenneth Michael Cummings,⁶ and Lawrence An^{5,7}

¹Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104, USA, ²Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, MI 48106, USA,

³Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA, ⁴Department of Communication Studies and ⁵Center for Health Communications Research, University of Michigan, Ann Arbor, MI 48109, USA, ⁶Department of Psychiatry & Behavioral Sciences, Medical University of South Carolina, Charleston, SC 29425, USA and ⁷Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA

Correspondence should be addressed to Emily B. Falk, Annenberg School for Communication, University of Pennsylvania, 3620 Walnut St., Philadelphia, PA 19104, USA. E-mail: falk@asc.upenn.edu.

Abstract

Mass media can powerfully affect health decision-making. Pre-testing through focus groups or surveys is a standard, though inconsistent, predictor of effectiveness. Converging evidence demonstrates that activity within brain systems associated with self-related processing can predict individual behavior in response to health messages. Preliminary evidence also suggests that neural activity in small groups can forecast population-level campaign outcomes. Less is known about the psychological processes that link neural activity and population-level outcomes, or how these predictions are affected by message content. We exposed 50 smokers to antismoking messages and used their aggregated neural activity within a 'self-localizer' defined region of medial prefrontal cortex to predict the success of the same campaign messages at the population level ($n = 400\,000$ emails). Results demonstrate that: (i) independently localized neural activity during health message exposure complements existing self-report data in predicting population-level campaign responses (model combined R^2 up to 0.65) and (ii) this relationship depends on message content—self-related neural processing predicts outcomes in response to strong negative arguments against smoking and not in response to compositionally similar neutral images. These data advance understanding of the psychological link between brain and large-scale behavior and may aid the construction of more effective media health campaigns.

Key words: fMRI; self; MPFC; smoking; media effects; health communication

Much of the worldwide burden of morbidity and mortality can be attributed to personal decisions made by individuals (Ezzati et al., 2002; Mokdad et al., 2004; Keeney, 2008). Previous studies have shown that mass media messages can influence health-related decision making (NHS, 2004; Noar, 2006; Wakefield et al., 2010); however, our understanding of the mechanisms prompting such influence is incomplete. Therefore, new ways of

understanding the mechanisms that produce effective mass-mediated health campaigns are essential.

Neural responses recorded using functional magnetic resonance imaging (fMRI) may provide insight into the mechanisms leading to behavior change, and by extension, the design of effective messages. Broadly, research in social neuroscience and neuroeconomics have advanced understanding of key concepts

Received: 23 January 2015; Revised: 20 July 2015; Accepted: 31 August 2015

© The Author (2015). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

relevant to messaging such as self-related processing, valuation and choice behavior (for reviews, see Northoff and Bermpohl, 2004; Northoff et al., 2006; Ariely and Berns, 2010; Liu et al., 2011; Denny et al., 2012; Levy and Glimcher, 2012; Bartra et al., 2013). Converging evidence demonstrates that neural activity within the brain's medial prefrontal cortex (MPFC) in response to health messages can predict individual behavior change (Falk et al., 2010, 2011; Chua et al., 2011; Wang et al., 2013), as well as out-of-sample population-level outcomes (Falk et al., 2012), above and beyond what is explained by participants' self-reports of their attitudes toward the behavior in question, their intentions to change their behavior (Falk et al., 2010), confidence in their ability to change and ability to relate to the ads (Falk et al., 2011) and risk beliefs (Cooper et al., 2015). A small number of studies have also found relationships between neural activity in the valuation system and mediofrontal regions more broadly, and large-scale behavioral outcomes in other domains (Berns and Moore, 2012; Dmochowski et al., 2014; Boksem and Smidts, in press). More broadly, fMRI offers a non-invasive way of monitoring neurocognitive responses to messages throughout the brain. This provides several advantages, including the ability to assess responses in real time, as persuasion occurs, without the need to rely on retrospective self-reports. As such, fMRI is ideally positioned to provide information that may not be accessible through other methods; for more extensive discussion see Berkman and Falk (2013) and Lieberman (2010).

MPFC is the region most commonly observed in studies of self-related processing (Lieberman, 2010), including computing the value of stimuli to oneself (Bartra et al., 2013). Furthermore, meta-analyses of self-related processing show robust activity in MPFC associated with self-related processing (Yarkoni et al., 2011; Denny et al., 2012). Although self-related processing, like all cognitive constructs (Poldrack, 2006), is not limited to activity in MPFC, we focus on MPFC given its role at the intersection of research predicting behavior change (Falk et al., 2010, 2011, 2012; Cooper et al., 2015) and research on self-related processing (Northoff and Bermpohl, 2004; Northoff et al., 2006; Denny et al., 2012). More specifically, in this study, we tested the extent to which activation in MPFC in a small sample of smokers could predict population-level responses to a mass media quit-smoking campaign. We hypothesized that messages eliciting more activity in self-related processing regions across individuals in a neuroimaging context would also be more successful at eliciting the target behavior at the population level. Given that activity in MPFC is associated with multiple cognitive functions, however, we use a validated cognitive task relying on self-knowledge as an independent 'localizer' to identify a subregion of MPFC most robustly activated by this form of self-related processing (Schmitz and Johnson, 2006; Chua et al., 2011).

We also examined whether neural activity in the subregion of MPFC identified by our self-localizer would be more predictive of population-level responses for messages that were more directly smoking-relevant. In particular, we focused on anti-smoking messages inspired by the Food and Drug Administration's (FDA's) proposed graphic warning labels for cigarette packaging delivered through a mass email campaign directing traffic to a quit-smoking website. We compared strong, negative messages, of the type often used and demonstrated to have at least short-term effectiveness, in antismoking campaigns (Hammond et al., 2003; Hammond et al., 2004; Hammond et al., 2006; Fathelrahman et al., 2010; Thrasher et al., 2012) to compositionally similar neutral images. This manipulation allowed us to test whether our hypothesized neural activity predicts outcomes selectively in response to strong, behavior-

relevant arguments or whether this neural activity in turn predicts behavioral outcomes, regardless of message content. In addition, this design allowed for a high degree of control over the stimuli, balanced with a high degree of external validity; The Centers for Disease Control and Prevention Best Practices for Comprehensive Tobacco Control Programs lists 'Promotion of available services, including . . . quitting Web sites and social media pages' (p.34) as one best practice. Visits to a quit-smoking website have been associated with longitudinal abstinence from smoking (Richardson et al., 2013), and meta-analytic evidence suggesting that web-based cessation interventions can have significant impact under some circumstances (Myung et al., 2009).

Materials and methods

This study tests whether self-related processing in the brains of one group of smokers ($n = 50$) in response to antitobacco messaging can predict the success of larger-scale media effects in response to the same messages as measured in a population-level campaign ($n = 400\,000$ emails) (Figure 1). We: (i) established population level success of the messages using a large-scale email campaign in the state of New York (Tables 1 and 2); (ii) verified the image properties using smokers identified on Amazon Mechanical Turk (MTurk) (Tables 1 and 2) and (iii) collected neural responses to the messages in a separate sample of smokers in Michigan, using fMRI.

We built on prior research demonstrating the value of neural signals implicated in self-related processing for prediction of individual behavior change (Falk et al., 2010, 2011; Chua et al., 2011; Wang et al., 2013). More specifically, within our fMRI sample, we localized a subregion of MPFC most strongly associated with self-related processing using a well-validated fMRI task that allowed us to independently identify neural activity associated with self-related processing [a 'self-localizer task' (Chua et al., 2011); Figure 1c]. We then used the fMRI data within this self-related processing region of interest (ROI) during message exposure to predict the population-level effectiveness of the messages.

Population-level email campaign

Participants. Likely smokers, ages 18+ years within the State of New York were identified by a third party email messaging service (National Data Group; NDG; Table 1 for NDG estimated demographics of the email sample). These participants had opted in to receive emails from NDG. NDG uses credit card data as one primary source of their direct email marketing lists and uses information such as purchases of cigarettes and disclosure of cigarette use in online marketing surveys or promotions as indicators of smoker status; NDG did not link to data from tobacco companies.

We partnered with the New York State Smokers' Quit Line (NYSSQL) to launch an email campaign in which the 40 images served as the basis for ads promoting internet-based quit-smoking resources (www.nysmokefree.com). Each target smoker of the email campaign received one of the 40 images paired with text encouraging smokers to quit smoking (Figures 1 and 2). The images consisted of 20 negative antismoking images, modeled off of FDA's proposed graphic warning labels and 20 compositionally matched neutral images. The emails also contained links to online quit-smoking resources managed by the NYSSQL. The text within the body of the email consisted of either a statement ($n = 400\,000$): 'Stop Smoking. Start Living.'

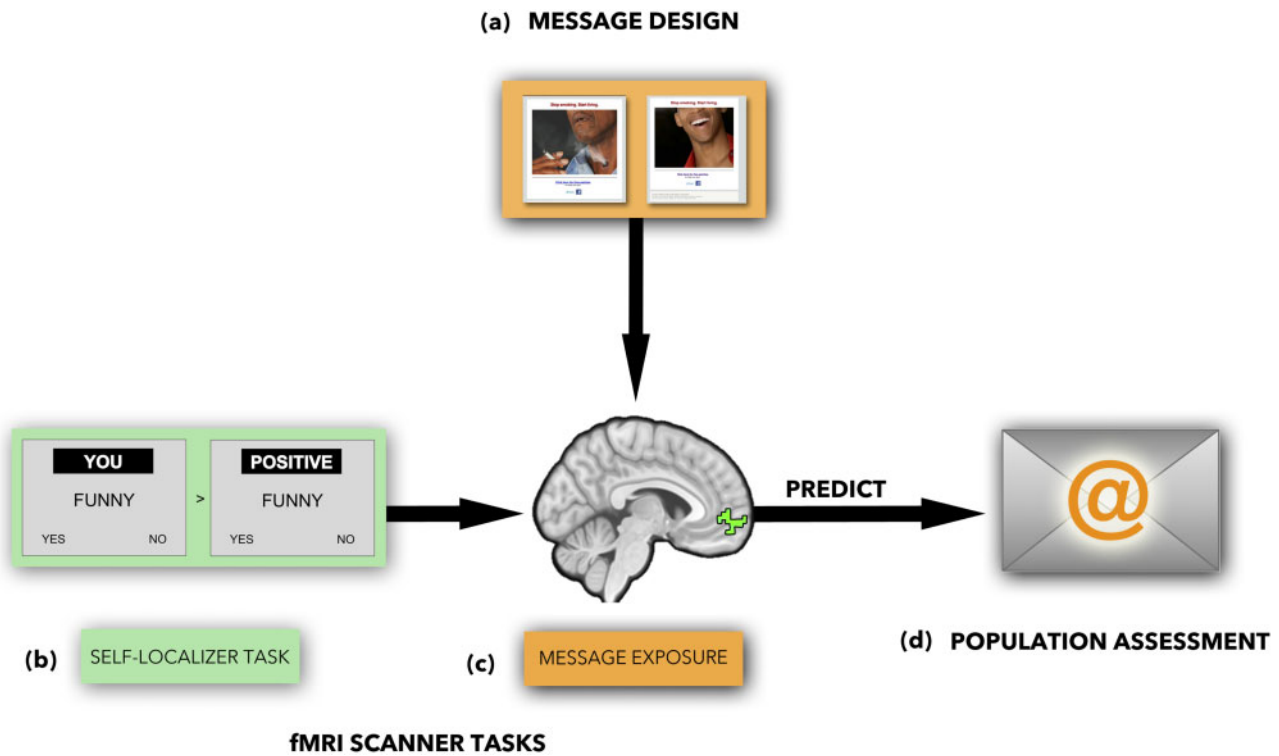


Fig. 1. Overview of methods: (a) message design; (b) self-localizer task; (c) ROI identified using the self-localizer task; (d) population-level assessment of messages.

Table 1. Participant demographics

	Email campaign recipients (n = 400 000)	fMRI sample (n = 47)	MTurk Sample 1 (n = 63)	MTurk Sample 2 (n = 19)
Gender (%M)	51.5%	59.6% (28 M, 19 F)	52.2% (33 M, 30 F)	52.6% (10 M, 9 F)
Age (years)	18–30 = 32%, 31–50 = 40%, 51+ = 28%	18–30: N = 30 (63.8%) 31–50: N = 10 (21.3%), 51+ : N = 7 (14.9%) (M = 31.89, s.d. = 12.53, range = 19–64)	18–30: N = 22 (34.9%) 31–50: N = 33 (52.4%) 51+ = 8 (12.7%) (M = 36.7, s.d. = 11.6, range = 20–66)	18–30: N = 10 (52.6%) 31–50: N = 5 (26.3%) 51+ = 4 (21.1%) (M = 34.8, s.d. = 14.8, range = 21–66)
Education	50% High school degree 34% College degree; 13% Graduate degree; 2% Vocational/Technical	High school degree or less: N = 22 (46.8%) Associate degree from a 2- year college: N = 3 (6.4%) Currently in college: N = 12 (25.5%) Bachelor's degree or post- graduate degree: N = 10 (21.3%) M = 13.1, s.d. = 6.76, range = 3–30 cigarettes per day Participants had smoked for an average of 14 years (M = 14.8, s.d. = 12.2)	High school degree or less: N = 27 (42.86%) Vocational training other than college: N = 2 (1.58%) Associate degree: N = 5 (7.94%) Bachelor's degree: N = 13 (20.631%) Unknown: N = 16 (25.4%) M = 12.1, s.d. = 8.9, range = 1–40 cigarettes per day Participants had smoked for an average of 20 years (M = 20.3, s.d. = 12.3)	High school degree or less: N = 6 (31.6%) Associate degree: N = 2 (10.5%) Bachelor's degree: N = 3 (15.8%) Unknown: N = 8 (42.1%) M = 14.1, s.d. = 1.3, range = 1–40 cigarettes per day Participants had smoked for an average of 18 years (M = 18.5, s.d. = 16.3)
Smoking status	Likely smokers (duration and frequency unknown)			
Ethnicity	Unknown	White/European-American: N = 31 (66%) African-American: N = 5 (10.64%) Hispanic/Latino: N = 5 (10.64%) selected the 'mixed' ethni- city category: N = 6 (12.77%)	White/European-American: N = 39 (61.90%) Asian-American: N = 6 (9.52%) Hispanic/Latino: N = 1 (1.58%) Unknown: N = 17 (26.98%)	White/European-American: N = 12 (63.2%) Asian-American: N = 1 (5.3%) Hispanic/Latino: N = 4 (21.1%) Unknown: N = 2 (10.5%)

Table 2. Responses to antismoking email messages by image type

Measure	Negative images mean (s.d.)	Neutral images mean (s.d.)	Difference stats
MTurk negative emotions	2.92 (1.02)	1.16 (0.07)	$t(38) = 7.7165, P < 0.001$
MTurk positive emotions	1.66 (0.20)	2.01 (0.37)	$t(38) = 3.76, P < 0.001$
MTurk image strength	3.53 (0.68)	2.54 (0.33)	$t(38) = 5.81, P < 0.001$
fMRI participants QUIT	3.28 (0.72)	2.42 (0.39)	$t(38) = 4.66, P < 0.001$
MPFC % signal change	-0.006 (0.041)	0.009 (0.017)	$t(38) = 1.58, P = 0.12$
Email CTR	0.17 (0.03)	0.14 (0.02)	$t(38) = 4.94, P < 0.001$

Note: QUIT, The degree to which a message made the participant want to quit; CTR, click through rate.

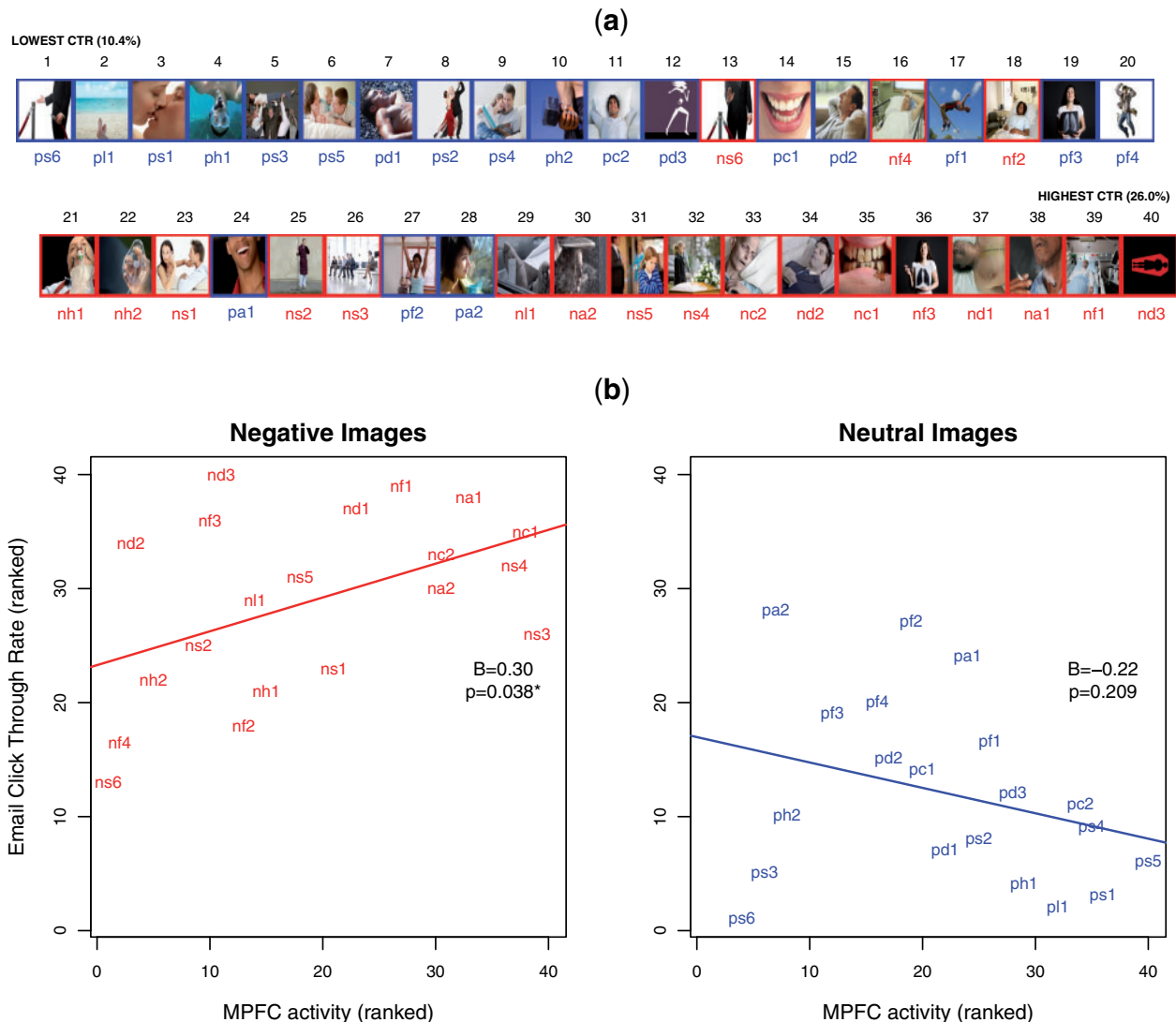


Fig. 2. Neural activity in response to negative images predicts population-level email click through rates. (a) RANK success in the email campaign. Red = negative images; Blue = neutral images; (b) Relationship between MPFC activity and population-level CTR.

(paired with each of the 40 negative and neutral images), or a question ($n = 400\,000$): ‘What are the good things that would happen if you quit smoking?’ (paired with neutral images)/ ‘What are the bad things that would happen if you don’t stop smoking?’ (paired with negative images). Given that no differences were observed between statement and question text emails, all data reported focus on the statement version of text

(Stop Smoking. Start Living.) that was used in the fMRI study ($n = 400\,000$ emails sent containing this tagline).

Prior to the launch of the email campaign, email subject lines were piloted to select an email subject line that was likely to elicit the greatest opening rate. Following the pilot, all emails were sent with the same subject line (‘NY smokers. It’s time to quit. Get free patches here.’), thus varying only message content

(image + statement/question within the email body). Email click through rates (CTRs) in the email campaign (visits to the quit-website generated by the email campaign) were treated as the primary outcome of interest. The CTR for each image in the email campaign was calculated as the number of emails opened for an image divided by the total number of clicks of links within the email (clicks/opened). Consistent with past work predicting population-level responses using neural activity (Falk et al., 2012), the population-level email data were rank ordered from lowest (CTR = 1) to highest (CTR = 40).

Image characteristics rated by smokers on mechanical Turk

Participants. Sixty-three smokers (MTurk Sample 1; Table 1 for demographics) rated the 40 images used in the email campaign using five point Likert scales according to emotion (negative: depressing, disgusting, frightening, unpleasant; positive: encouraging, hopeful, inspiring, meaningful) and image strength relative to the behavior in question [Image strength: 'This message would make smokers want to quit'; 'Is the reason the message gave for quitting a strong or a weak reason'; 'This message put thoughts in my mind about wanting to quit'; 'This message put thoughts in my mind about not wanting to quit' (Reverse coded)] (Zhao et al., 2011). Summary indices were computed as average ratings per image for negative emotions, positive emotions and image argument strength. Finally, paralleling a metric collected in the fMRI scanner (QUIT; details later), a group of 19 smokers (MTurk Sample 2; Table 1 for demographics) rated the images according to the extent to which each image 'makes me want to quit' (referred to subsequently as mturk.QUIT).

fMRI study

Participants. Fifty smokers residing in Michigan (outside of the area targeted by the email campaign) were recruited to the fMRI portion of the study; however, three participants were excluded due to excessive head motion (two participants) or data transfer difficulties (one participant). The remaining 47 subjects included 28 males and 19 females with a mean age of 31.89 (s.d. = 12.53, range of 19–64 years old; Table 1 for demographics).

The subjects were recruited using convenience sampling from the general population using Craigslist and UMClinicalStudies.org (an online registry of individuals who sign up expressing an interest in participating in research) and completed a phone screening, an appointment where they filled out initial questionnaires, an fMRI appointment and a phone call 1 month later (Figure 3). To participate in the study, during the preliminary screening subjects had to report smoking at least five cigarettes per day for the past month, have been a smoker for at least 12 months, and be between the ages of 18 and 65. Subjects also had to meet standard fMRI eligibility criteria, including no metal in their body, no history of psychiatric or neurological disorders, weight under 350 pounds, and currently not taking any psychiatric or illicit drugs. On the day of the scan, subjects reported smoking an average of 13 cigarettes per day ($M = 13.11$, $s.d. = 6.76$, range = 3–30 cigarettes per day: a small number of participants reduced their smoking between screening and fMRI appointment). Participants had smoked for an average of 14 years ($M = 14.78$, $s.d. = 12.15$). Participants were paid \$100 for completion of the study.

For the self-localizer task, an additional two subjects were excluded from analyses due to missing data (one subject) and excessive head movement specific to this task (one subject) for a total of 45 subjects included in analyzing the self-localizer task data. For the images task, one subject was excluded due to

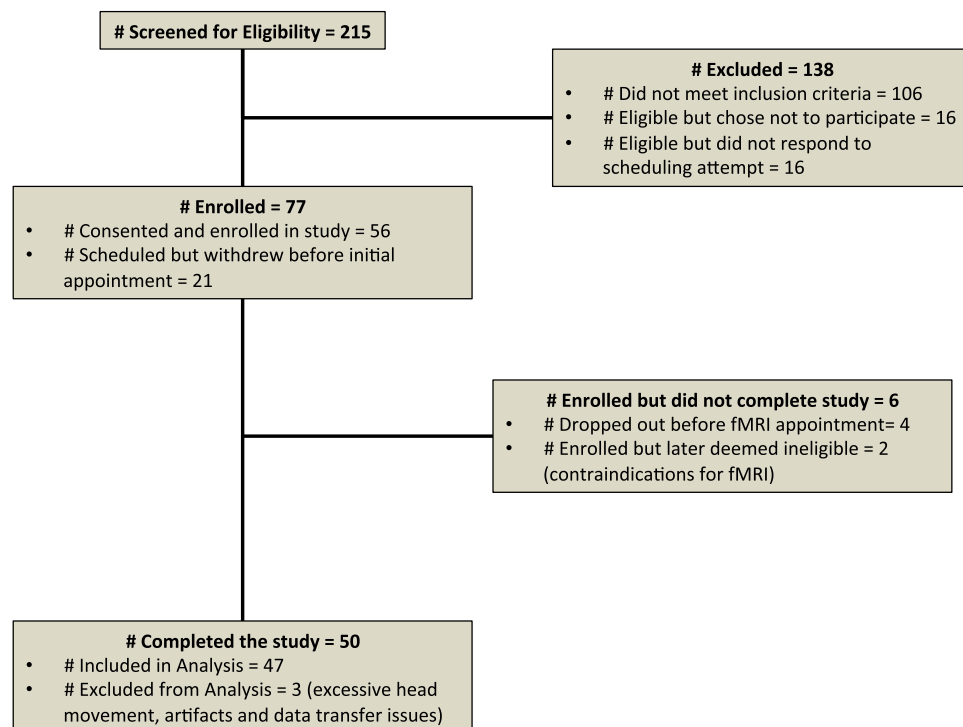


Fig. 3. Recruitment diagram. Participants were initially screened via telephone. Of the 77 deemed eligible, 21 were unable to be scheduled. An additional two participants were deemed ineligible by the study team after more detailed screening at appointment 1 and 4 participants failed to attend their scheduled fMRI appointment.

missing data for that task for a total of 46 subjects included in analyzing the images data.

fMRI data acquisition. Neuroimaging data were acquired using a 3 Tesla GE Signa MRI scanner. Two functional runs for the self-localizer task (288 volumes total), followed by two runs for the email campaign images task (300 volumes total) were acquired for each participant. Participants completed the self-localizer prior to the email campaign task. Functional images were recorded using a reverse spiral sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°, 43 axial slices, FOV = 220 mm, slice thickness = 3 mm; voxel size = $3.44 \times 3.44 \times 3.0$ mm). We also acquired in-plane T1-weighted images (43 slices; slice thickness = 3 mm; voxel size = $0.86 \times 0.86 \times 3.0$ mm) and high-resolution T1-weighted images (SPGR; 124 slices; slice thickness = $1.02 \times 1.02 \times 1.2$ mm) for use in coregistration and normalization.

Preprocessing. Functional data were preprocessed and analyzed using Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK). To allow for the stabilization of the BOLD signal, the first five volumes (10 s) of each run were discarded prior to analysis. Functional images were despiked using the 3dDespike program as implemented in the AFNI toolbox. Next, data were corrected for differences in the time of slice acquisition using sinc interpolation; the first slice served as the reference slice. Data were then spatially realigned to the first functional image. We then coregistered the functional and structural images using a two-stage procedure. First, in-plane T1 images were registered to the mean functional image. Next, high-resolution T1 images were registered to the in-plane image. After coregistration, high-resolution structural images were segmented to produce a grey matter mask, and then normalized to the skull-stripped MNI template provided by FSL ('MNI152_T1_1mm_brain.nii'). Finally, functional images were smoothed using a Gaussian kernel (8 mm FWHM).

First-level analysis. The fMRI data were modeled for each subject for each task using fixed effects models within the general linear model as implemented in SPM8 and second-level random effects models for each task were also implemented in SPM8 (see details for each task later). The six rigid-body translation and rotation parameters derived from spatial realignment were included as nuisance regressors in all first-level models. Data were high-pass filtered with a cutoff of 128 s in all models.

Self-localizer task. All participants first completed two runs of an adapted¹ version of a well-validated self-related processing task (Schmitz and Johnson, 2006; Chua et al., 2011). The task involves judging the self-relevance or valence of trait adjectives from the Anderson trait word list (Anderson, 1968). The task includes blocks of six trials with three positive and three negative words in each block. Each block was preceded with a 3 s orientation

screen and blocks were separated by 2 s of fixation. The same 36 (18 negative and 18 positive) words were shown in each condition and were arranged so that an equal number of positive and negative words occur first in each of the conditions. The task was modeled using a single boxcar function for each 18 s block. Fixation and condition preparation periods were not modeled and included with baseline rest. Neural activity during conditions in which participants made trait judgments about themselves were compared with conditions in which participants judged word valence. The resulting contrast images were combined by averaging across subjects using a random effects model in SPM8 and the resulting image map (cluster corrected FWE, $P < 0.05$) was used to identify a subregion of MPFC that was most robustly associated with self-related processing across subjects. This cluster was converted to a functionally defined ROI (fROI) using MarsBaR and served as the primary fROI in the subsequent messages task (Figure 1c).

fMRI email campaign messages task. We examined neural activity as participants were exposed to each of 40 email-campaign images that were presented along with a tag-line from the body text of the email campaign: 'Stop Smoking. Start Living.' (Figure 1b). The timing for each trial consisted of 4 s of image presentation followed immediately with a 3 s response screen with the statement 'This makes me want to quit' and a five-point rating scale (1 = definitely does not, 2 = does not, 3 = neutral, 4 = does, 5 = definitely does; referred to as 'QUIT'), followed by fixation with jitter ITI (3–7.5 s, mean = 4.10, median = 3.32, s.d. = 1.01, total = 164.0). The task also included 10 additional negative images, 10 additional neutral images, 10 personal (Facebook) or control (NimStim) images, for which no population-level email campaign data are available, and which are not the focus of the current investigation that were interspersed with the primary trials of interest.

Data for the email campaign ads tasks were modeled at the first level using the general linear model as implemented in SPM8. Consistent with past methods used to aggregate responses to cultural products (Berns and Moore, 2012), we modeled exposure to each image with the campaign separately. Our models focused on the 4 s exposure to the campaign message (before making image ratings). We separately modeled the 3 s response periods following each image exposure as a condition of no interest. Fixation rest-periods constituted an implicit baseline.

At the second level, in SPM8, we combined first-level maps averaging across subjects to form statistical parametric maps of activation relevant to each ad, relative to the implicit baseline. In other words, we conducted the analysis such that the ads were the units of analysis, as opposed to the study participants, to facilitate combination of data across levels of analysis. Activation estimates were extracted for each image for each subject using MarsBaR (Brett et al., 2002) in units of percent signal change from baseline. Parameter estimates were standardized (z-scores) within subject and outliers > 2 s.d. from mean were excluded to reduce bias from motion and other fMRI artifacts (note: this did not change the substantive ordering of the ad rankings).

Consistent with past research (Falk et al., 2012), the parameter estimates for the MPFC self-related processing neural ROI for each image were ranked first within subjects, scores were averaged across participants and a final average rank across subjects was computed for each ad. For the email campaign this resulted in an ordering from lowest (image rank = 1) to highest (image rank = 40).

1 The task contained five conditions (each condition was repeated in 6 blocks, each containing 6 trials, for a total of 36 trials per condition): you_you (from your own perspective, judging yourself), you_friend (from your perspective, judging a friend), friend_you (from a friend's perspective, judging you), friend_friend (from a friend's perspective, judging a friend) and valence (is the word positive or negative). The primary contrast of interest examined conditions in which the participant was the target of judgment (you_you & friend_you) vs conditions in which the judgment was word valence.

Analysis combining neural and population-level data

We used a brain-as-predictor framework (Berkman and Falk, 2013) to examine the relationship between individual level neural data and population-level campaign outcomes. We constructed a combined model in which ranks of neural and self-report variables were simultaneously entered to predict email CTR; rank scores of fMRI participants' ratings of QUIT, MTurk ratings of mturk.QUIT, MTurk ratings of Image Strength, and percent signal change in the MPFC self-related processing neural ROI were separately entered in regressions with image type (negative vs neutral) and the interaction between each type of rating and image type, to predict email CTR rankings implemented in R (R Development Core Team, 2014). Overall, ranks were used because they provide a non-parametric and robust way to harmonize variables across multiple data sets. The Pearson correlation between ranks is equivalent to the well-known Spearman correlation, and the extension of multiple regression on ranks has been studied by Iman and Conover (1979). In addition, from a public health context, knowing which ads are likely to be the best, independent of scale used, can aid in the selection of optimal messaging.

After conducting our a priori planned ROI analyses, we also conducted whole-brain searches to explore which regions (in addition to MPFC) are associated with message success (i.e. email CTR). More specifically, first-level models were computed for each participant in the neuroimaging data set, modeling exposure to negative (smoking-relevant) images, neutral images and parametric modulators capturing the ranked population-level CTR for each image. As with the main image-wise ROI analysis, the models focused on the main exposure period (first 4 s of exposure to each image, prior to image rating, with nuisance regressors specified for the rating period). Images were thresholded at $P < 0.005$ with a cluster extent threshold of 36 voxels, corresponding to $P < 0.05$ corrected, as determined by a monte carlo simulation conducted with AlphaSim (Ward, 2000).

Results

Email campaign data

On average, participants opened 14% of emails received (57 395 opened of 400 000 sent). Of those emails that were opened, on average, participants clicked through ~16% of the time (over the 40 ads: mean CTR = 15.6%, median CTR = 15.9%, range = 10–26%). CTRs in the email campaign were significantly higher for negative images than neutral images (Table 2, Figure 2).

Ratings of images

Ratings from the MTurk sample confirmed that the negative images elicited more negative emotions and less positive emotions than the neutral comparison images (Table 2). Both MTurk and fMRI participants rated the negative images as prompting more desire to quit (QUIT, mturk.QUIT: Table 2) as compared with the neutral images. MTurk participants confirmed that the negative images were significantly more relevant to quitting smoking/made stronger arguments for quitting (Image Strength: Table 2). There was a high degree of convergence across items in the fMRI and MTurk samples (e.g. QUIT and mturk.QUIT, $r = 0.87$, $P < 0.001$).

Prediction of email campaign CTR from neural activity

We used neural activity, image valence and their interaction to predict the success of messages within the email campaign (model $R^2 = 0.60$). Activity within an independently localized self-related MPFC subregion during health message exposure predicted population-level campaign responses [$B = 0.30$, $t(36) = 2.126$, $P = 0.04$].

Next, we examined the importance of message content in the ability to predict population-level behavior. We found that the relationship between MPFC activity and population outcomes depended on message content [interaction between MPFC and image type: $B = 0.52$, $t(36) = -2.424$, $P = 0.02$; Figure 2]—self-related neural processing predicted outcomes in response to graphic warning labels, but not in response to compositionally similar neutral images. Thus, the current data support the hypothesis that self-related processing of health messages predicts message-consistent outcomes but also suggest an important boundary condition on the relationship between neural activity and prediction of campaign success.

Prediction of email campaign CTR by self-report measures

MTurk rankings of each message's argument strength were positively associated with population-level email CTR [$B = 0.43$, $t(36) = 2.45$, $P = 0.02$, $R^2 = 0.37$]. MTurk rankings of QUIT (mturk.QUIT) were positively associated with population-level email CTR [$B = 0.37$, $t(36) = 2.13$, $P = 0.04$, $R^2 = 0.38$]. Our fMRI participants' ranked ratings of how much each message made them want to quit (labeled QUIT) were positively, but not significantly, associated with population-level CTRs in the email campaign (CTR) [$B = 0.26$, $t(36) = 1.59$, $P = 0.12$]. The relationship between self-report measures and population-level email CTR was not significantly moderated by image type ($P = 0.14$, 0.19 and 0.41, respectively).

Prediction of email campaign CTR combining neural and self-report measures

Neural activity was entered into a regression with each self-report measure to predict population-level CTR. Each of the self-report measures were entered as covariates in separate regressions given the limited degrees of freedom and collinearity between covariates. The interaction between MPFC activity and image type remained significantly related to CTR controlling image strength: $B = 0.57$, $t(35) = -2.83$, $P = 0.008$, model $R^2 = 0.65$; controlling QUIT: $B = 0.56$, $t(35) = -2.64$, $P = 0.01$, model $R^2 = 0.63$ and controlling mturk.QUIT: $B = 0.58$, $t(35) = -2.78$, $P = 0.009$, model $R^2 = 0.65$. Thus, accounting for neural responses in combination with message features improved predictive power, complementing our best self-report predictors. Together, our brain measures and self-report measures accounted for up to 65% of the variance in the success of the email campaign, suggesting that the combination of neural data with survey results may ultimately aid in more efficient prediction of the success of mass media health campaigns.

Exploratory whole-brain searches for additional regions associated with CTR

After conducting our planned ROI analyses, we conducted a series of whole-brain searches for regions associated with CTR in addition to our main MPFC ROI (Table 3; Figure 4). Reinforcing the hypothesized relationship between self-related processing

Table 3. Whole-brain search for regions associated with CTR in the larger-email campaign

Region	x, y, z	size	t-stat
a. Association with CTR for negative images			
MPFC	-2 67 10	125	3.39
VMPFC	-6 46 -8		3.16
DMPFC	15 36 43	59	3.07
DMPFC	-6 46 31		2.97
Lingual gyrus	22 -78 -11	1013	7.95
Inferior temporal lobe	-44 -57 -8	311	6.12
Inferior frontal gyrus	-44 36 10	135	5.16
Hippocampus/medial temporal lobe	-20 -26 -8	440	5.1
Hippocampus/medial temporal lobe	29 -26 -11	747	5.09
Middle cingulate gyrus	1 -2 34	100	5.08
Cerebellum	-2 -60 -38	36	3.58
b. Positive association with CTR for neutral images			
Inferior frontal gyrus	42 39 13	308	4.92
Posterior Insula	42 -16 7	52	3.61
c. Negative association with CTR for neutral images			
Middle temporal gyrus/posterior superior temporal sulcus	-54 -71 7	304	4.74
d. Association with CTR for neutral > negative Images			
Middle frontal gyrus	39 46 13	61	3.72
e. Association with CTR for negative > neutral Images			
VMPFC	-2 46 -11	38	4.1
MPFC/DMPFC	-2 67 10	354	3.83
DMPFC	1 53 31		3.82
MPFC	-2 56 4		3.64
Posterior cingulate cortex	4 -40 1	44	4.01
Amygdala	18 -2 -14	40	3.76
	29 15 -17		3.05
Fusiform/lingual gyrus	25 -71 -8	745	6.33
Inferior frontal gyrus	-37 29 -17	156	4.43
Angular gyrus/TPJ	-44 -54 25	48	4.19
Superior occipital cortex	-20 -91 37	41	3.77

Note: Threshold = $P < 0.005$, $k = 36$, corresponding to $P < 0.05$, corrected. No regions were negatively associated with CTR during exposure to the negative images. TPJ: temporoparietal junction; DMPFC: dorsomedial prefrontal cortex; VMPFC: ventromedial prefrontal cortex; PSA: public service announcement; ERP: event related potential; MNI: montreal neurological institute; fwhm: full width, half maximum; ITI: intertrial interval; BOLD: blood oxygen level dependent.

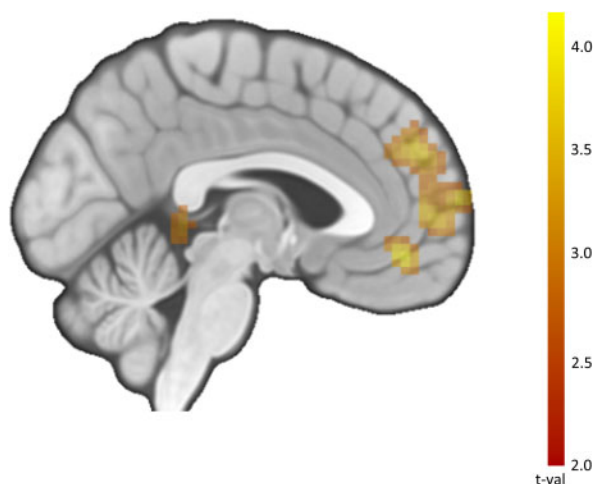


Fig. 4. Neural activity associated with CTR for negative > neutral images. Note: These results correspond to Table 3e. Image thresholded at $P < 0.005$, $k = 36$, corresponding to $P < 0.05$, corrected.

and CTR for smoking-relevant images, we observed interactions between image type and CTR in MPFC and posterior cingulate cortex. Among other regions, we also observed this relationship in regions implicated in affective salience, such as amygdala.

Discussion

In sum, we demonstrate that neural activity as measured by fMRI in the MPFC observed in relatively small groups of smokers was predictive of population-level responses to an email campaign featuring different antismoking messages. These results extend past findings by: (i) bolstering confidence that neural signals can be combined with self-report measures to aid in the design and selection of optimal public-health media messages; (ii) providing evidence for a neurocognitive mechanism linking neural and behavioral data (self-related processing) and (iii) suggesting a condition under which this method is likely to be most applicable (message content interacts with self-related neural processing to predict outcomes of interest). Furthermore, both self-report and neural variables explained significant

variance in the population-level outcome, suggesting the value of synergy across methods in understanding mass media effects in the public health domain.

This research is consistent with prior research suggesting that neural activity within MPFC can predict larger-scale population-level outcomes (Falk et al., 2012) as well as other findings linking neural activity in mediofrontal regions to larger, out-of-sample outcomes (Berns and Moore, 2012; Dmochowski et al., 2014; Boksem and Smidts, in press). We hypothesized that this effect might stem from a message's ability to elicit self-related processing across a wide range of individuals. Thus we focused on MPFC as the region most commonly predictive of behavior change in past studies, and refined our target ROI using a well-validated localizer task (Schmitz and Johnson, 2006; Chua et al., 2011). Although activity within MPFC has been linked with multiple functions, our strong, *a priori* theory and the use of a localizer strengthens our confidence in this theoretical link by refining MPFC to a more specific subregion. Furthermore, a whole-brain search for regions associated with greater CTR in smoking-relevant > neutral images showed activity in posterior cingulate—another region commonly associated with self-related processing and autobiographical memory; for reviews, see Northoff and Bermpohl (2004) and Northoff et al. (2006).

We also tested whether neural activity in self-related processing regions would predict large-scale outcomes regardless of message content, or whether predictive capacity would be stronger for antismoking messages that contain strong behavior-relevant arguments (in this case, strong arguments to quit smoking). Well-supported theories of health behavior change (e.g., Rosenstock, 1974; Leventhal et al., 1983; Strecher and Rosenstock, 1997) suggest that individuals are most likely to act on health messages when they perceive specific health risks as self-relevant combined with self-relevant pathways to address the risk. Our negative messages were easily linked to the consequences of smoking and were rated by an independent group as providing strong arguments to quit (Table 2). In contrast, the compositionally similar neutral images were not linked to smoking behavior in the absence of the 'Stop Smoking. Start Living.' tagline. Within the email campaign, the neutral messages we tested had a lower CTR than the negative messages, perhaps because they failed to elicit emotional responses and/or call to mind the relevant behavior change. In addition to behavioral relevance, the valence of the images itself may have also influenced smokers' motivations to quit. Negative emotions, such as fear, have been commonly used in public health messaging to increase motivation (Witte and Allen, 2000; Peters et al., 2013), *c.f.* (Earl and Albarracín, 2007), and can be motivating if accompanied by high levels of self-efficacy or other factors that offset threat (Peters et al., 2013; Sherman, 2013; Cohen and Sherman, 2014).

In parallel, we found that neural activity in subregions of MPFC (identified using the self-localizer task) predicted population-level CTR for the strong, negative antismoking messages, but not for the neutral images. This finding is consistent with the idea that neural responses in MPFC may only be predictive of large-scale outcomes when message content makes an argument relevant to both the behavior in question and the person in question. A neural index of self-related processing may not be meaningful in the context of neutral images that are not as directly relevant to the target behavior. In contrast, self-related processing may be more relevant to behavior when the message content highlights a need for change. In sum, the current data add confidence to prior results indicating that self-related processing of health messages may predict message-consistent

outcomes but also suggest an important boundary condition on the relationship between neural activity and prediction of campaign success.

Paralleling behavioral findings, affective salience of the graphic warning label inspired images likely also influenced message processing. One study of cigarette graphic warning labels found increased activity in amygdala and insula, brain regions associated with affective salience, accompanied by decreased craving in smokers (Do and Galvan, 2015). Likewise, our whole-brain searches suggested that amygdala activity, in addition to MPFC, is associated with higher message success for the negative > neutral messages. A second study of graphic warning labels reported ERP data consistent with diminished affective response to cigarette cues (Wang et al., 2015), suggesting that such negative images may shift the motivational value of smoking cues. Taken together, although valence and smoking relevance are confounded in our study, both likely contribute to the specific neural effects observed. It is unlikely that neural activity in MPFC predicts behavior change only for negative images [since past research has found similar relationships for humorous PSAs; (Falk et al., 2012)] but more likely that MPFC integrates behavioral relevance and affective salience in computing the value of the message to the self.

More broadly, although a growing body of studies link neural data with out-of-sample, large-scale outcomes (Berns and Moore, 2012; Falk et al., 2012; Dmochowski et al., 2014; Boksem and Smidts, in press), additional research is needed to specify further boundary conditions on type of medium, content area and psychological targets of the intervention, and to determine the generalizability of these findings across content areas and media. Furthermore, results of trials using web-based resources have been mixed (Myung et al., 2009), and as with any mass media campaign, measuring long-term behavioral outcomes is challenging; therefore additional research is also needed to explore the extent to which the present results generalize to other outcomes of interest beyond CTR (e.g. population-level smoking rates).

Finally, we found that neural variables complemented the best self-report measures pre-tested to predict campaign success. Neural activity, image characteristics and self-report predictors together accounted for up to 65% of the variance in the success of the email campaign, suggesting that the combination of neural data with survey results may ultimately aid in more efficient prediction of the success of mass media health campaigns.

Moving forward, it is possible that this knowledge may be used in the construction of more successful, evidence-based, large-scale media campaigns to improve health and ultimately reduce health care costs for preventable medical conditions. Such technology may also benefit future discussions regarding broader questions relevant to the effects of health messaging (e.g. FDA proposed graphic warning labels, cigarette and e-cigarette advertising and health-relevant messages more broadly) and help to more efficiently engage large groups of people across domains.

Acknowledgements

We acknowledge important contributions to the development of this study from Matthias Kirch, Michele Demers, Sarah Lillie, Darla Williams, Monte Montgomery at the University of Michigan and Anthony Brown, Seshadri Srinivasa, and Paula Celestino from the New York State Smokers Quit Line. We thank Francis Tinney Jr., Kristin

Shumaker, Li Chen, Nicolette Gregor, Becky Lau, Larissa S Svintsitski, Cole Schaffer for assistance with data collection, Lynda Lin and Kyle Cassidy for assistance with manuscript preparation and Robert Hornik for helpful discussions. E.F., M.O., S.T., S.D., V.S., M.C., and L.A. designed the study. E.F., M.O., S.T. collected the data. E.F., M.O., S.T., and R.G. conducted the data analysis. E.F. drafted the manuscript and all authors contributed to writing and revisions.

Funding

This work was funded by The Michigan Center of Excellence in Cancer Communication Research/ NIH-P50 CA101451 (PI Strecher), and the NIH New Innovator Award: NIH 1DP2DA03515601 (PI Falk).

Conflict of interest. Dr. Cummings has received funding from Pfizer and serves as a paid expert witness in litigation against the cigarette industry. The other authors declare no financial conflicts of interest related to this work.

References

- Anderson, N.H. (1968). Likeableness rating of 555 personality-trait words. *Journal of Personality and Social Psychology*, 9, 272–9.
- Ariely, D., Berns, G.S. (2010). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, 11(4), 284–92.
- Bartra, O., McGuire, J.T., Kable, J.W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, 76, 412–27.
- Berkman, E.T., Falk, E.B. (2013). Beyond brain mapping: using the brain to predict real-world outcomes. *Current Directions in Psychological Science*, 22(1), 45–55.
- Berns, G.S., Moore, S.E. (2012). A neural predictor of cultural popularity. *Journal of Consumer Psychology*, 22, 154–60.
- Boksem, M.A.S., Smidts, A. (2015). Brain responses to movie trailers predict individual preferences for movies and their population-wide commercial success. *Journal of Marketing Research*, 52(4), 482–92.
- Brett, M., Anton, J., Valabregue, R., Poline, J. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the The 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.
- Chua, H.F., Ho, S.S., Jasinska, A.J., et al. (2011). Self-related neural response to tailored smoking-cessation messages predicts quitting. *Nature Neuroscience*, 14(4), 426–7.
- Cohen, G.L., Sherman, D.K. (2014). The psychology of change: self-affirmation and social psychological intervention. *Annual Review of Psychology*, 65, 333–71.
- Cooper, N., Tompson, S., O'Donnell, M.B., Falk, E.B. (2015). Brain activity in self- and value-related regions in response to online antismoking messages predicts behavior change. *Journal of Media Psychology*, 27(3), 93–108.
- Denny, B.T., Kober, H., Wager, T.D., Ochsner, K.N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, 24(8), 1742–52.
- Dmochowski, J.P., Bezdek, M.A., Abelson, B.P., Johnson, J.S., Schumacher, E.H., Parra, L.C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, 5, 4567.
- Do, K.T., Galvan, A. (2015). FDA cigarette warning labels lower craving and elicit frontoinsula activation in adolescent smokers. *Social Cognitive and Affective Neuroscience*, 10(11), 1484–96.
- Earl, A., Albarracin, 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.
- Ezzati, M., Lopez, A.D., Rodgers, A., Vander Hoorn, S., Murray, C.J. (2002). Selected major risk factors and global and regional burden of disease. *Lancet*, 360(9343), 1347–60.
- Falk, E.B., Berkman, E.T., Lieberman, M.D. (2012). From neural responses to population behavior neural focus group predicts population-level media effects. *Psychological Science*, 23(5), 439–45.
- Falk, E.B., Berkman, E.T., Mann, T., Harrison, B., Lieberman, M.D. (2010). Predicting persuasion-induced behavior change from the brain. *Journal of Neuroscience*, 30(25), 8421–4.
- Falk, E.B., Berkman, E.T., Whalen, D., Lieberman, M.D. (2011). Neural activity during health messaging predicts reductions in smoking above and beyond self-report. *Health Psychology*, 30(2), 177–85.
- Fathelrahman, A.I., Omar, M., Awang, R., Cummings, K.M., Borland, R., Bin Mohd Samin, A.S. (2010). Impact of the new Malaysian cigarette pack warnings on smokers' awareness of health risks and interest in quitting smoking. *International Journal of Environmental Research Public Health*, 7(11), 4089–99.
- Hammond, D., Fong, G.T., McDonald, P.W., Cameron, R., Brown, K.S. (2003). Impact of the graphic Canadian warning labels on adult smoking behaviour. *Tobacco Control*, 12(4), 391–5.
- Hammond, D., Fong, G.T., McNeill, A., Borland, R., Cummings, K.M. (2006). Effectiveness of cigarette warning labels in informing smokers about the risks of smoking: findings from the international tobacco control (ITC) four country survey. *Tobacco Control*, 15(Suppl 3), iii19–25.
- Hammond, D., McDonald, P.W., Fong, G.T., Brown, K.S., Cameron, R. (2004). The impact of cigarette warning labels and smoke-free bylaws on smoking cessation: evidence from former smokers. *Canadian Journal of Public Health*, 95(3), 201–4.
- Iman, R.L., Conover, W.J. (1979). The use of the rank transform in regression. *Technometrics*, 21, 499–509.
- Keeney, R.L. (2008). Personal decisions are the leading cause of death. *Operations Research*, 56(6), 1335–47.
- Leventhal, H., Safer, M.A., Panagis, D.M. (1983). The impact of communications on the self-regulation of health beliefs, decisions, and behavior. *Health Education Quarterly*, 10(1), 3–29.
- Levy, D.J., Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, 22(6), 1027–38.
- Lieberman, M.D. (2010). Social cognitive neuroscience. In: Fiske, S., Gilbert, D., Lindzey, G., editors. *Handbook of Social Psychology*, 5th edn, 143–93. New York: McGraw-Hill.
- Liu, X., Hairston, J., Schrier, M., Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 35(5), 1219–36.
- Mokdad, A.H., Marks, J.S., Stroup, D.F., Gerberding, J.L. (2004). Actual causes of death in the United States, 2000. *JAMA*, 291(10), 1238–45.
- Myung, S.K., McDonnell, D.D., Kazinets, G., Seo, H.G., Moskowitz, J.M. (2009). Effects of Web- and computer-based smoking

- cessation programs: meta-analysis of randomized controlled trials. *Archives of Internal Medicine*, **169**(10), 929–37.
- NHS. (2004). The effectiveness of public health campaigns. *Health Development Agency Briefing*, 7. Available: <http://www.nice.org.uk/niceMedia/documents/CHB7-campaigns-14-7.pdf> (accessed 8 October 2013).
- Noar, S.M. (2006). A 10-year retrospective of research in health mass media campaigns: where do we go from here? *Journal of Health Communication*, **11**, 21–42.
- Northoff, G., Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, **8**(3), 102–7.
- Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage*, **31**(1), 440–57.
- Peters, G.J., Ruiter, R.A., Kok, G. (2013). Threatening communication: a critical re-analysis and a revised meta-analytic test of fear appeal theory. *Health Psychology Review*, **7**(Suppl 1), S8–31.
- Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, **10**(2), 59–63.
- R Development Core Team. (2014). R: A Language and Environment for Statistical Computing. Vienna, Austria.
- Richardson, A., Graham, A.L., Cobb, N., et al. (2013). Engagement promotes abstinence in a web-based cessation intervention: cohort study. *Journal of Medical Internet Research*, **15**(1), e14.
- Rosenstock, I.M. (1974). The health belief model and preventive health behavior. *Health Education Monographs*, **2**(4), 354–86.
- Schmitz, T.W., Johnson, S.C. (2006). Self-appraisal decisions evoke dissociated dorsal-ventral aMPFC networks. *Neuroimage*, **30**(3), 1050–8.
- Sherman, D.K. (2013). Self-affirmation: understanding the effects. *Social and Personality Psychology Compass*, **7**(11), 834–45.
- Strecher, V.J., Rosenstock, I.M. (1997). The health belief model. In: Baum, A., Newman, S., Weinman, J., West, R., McManus, C., editors. *Cambridge Handbook of Psychology, Health and Medicine*, 113–7. Cambridge: Cambridge University Press.
- Thrasher, J.F., Arillo-Santillan, E., Villalobos, V., et al. (2012). Can pictorial warning labels on cigarette packages address smoking-related health disparities? Field experiments in Mexico to assess pictorial warning label content. *Cancer Causes Control*, **23**(Suppl 1), 69–80.
- Wakefield, M.A., Loken, B., Hornik, R.C. (2010). Use of mass media campaigns to change health behaviour. *Lancet*, **376**(9748), 1261–71.
- Wang, A.L., Romer, D., Elman, I., Turetsky, B.I., Gur, R.C., Langleben, D.D. (2015). Emotional graphic cigarette warning labels reduce the electrophysiological brain response to smoking cues. *Addiction Biology*, **20**(2), 368–76.
- Wang, A.L., Ruparel, K., Loughead, J.W., et al. (2013). Content matters: neuroimaging investigation of brain and behavioral impact of televised anti-tobacco public service announcements. *Journal of Neuroscience*, **33**(17), 7420–7.
- Ward, B.D. (2000). Simultaneous Inference for FMRI Data (AlphaSim) (Version AlphaSim: AFNI version = AFNI_2011_12_21_1014 (Jan 28 2012)): Biophysics Research Institute, Medical College of Wisconsin. Retrieved from <http://afni.nimh.nih.gov/pub/dist/doc/manual/AlphaSim.pdf>
- Witte, K., Allen, M. (2000). A meta-analysis of fear appeals: implications for effective public health campaigns. *Health Education & Behavior*, **27**(5), 591–615.
- Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, **8**(8), 665–70.
- Zhao, X., Strasser, A., Cappella, J.N., Lerman, C., Fishbein, M. (2011). A measure of perceived argument strength: reliability and validity. *Communication Methods and Measures*, **5**(1), 48–75.