J Behav Med (2021) 44:74–83
https://doi.org/10.1007/s10865-020-00163-0

Incremental criterion validity of message perceptions and effects perceptions in the context of anti-smoking messages

Sabeeh A. Baig[1] · **Seth M. Noar**[2,3] · **Nisha C. Gottfredson**[1] · **Allison J. Lazard**[2,3] · **Kurt M. Ribisl**[1,3] · **Noel T. Brewer**[1,3]

Received: 11 June 2019 / Accepted: 21 May 2020 / Published online: 9 June 2020

Abstract To select promising health messages, formative research has often relied on perceived message effectiveness (PME) scales assessing either of two related constructs, *message perceptions* (persuasive potential) and *effects perceptions* (potential for behavioral impact). We sought to examine their incremental criterion validity within a comparative framework. Participants were 703 U.S. adult smokers (ages ≥21) who received anti-smoking or comparable control (littering) messages on their cigarette packs for 3 weeks. Structural equation models examined both PME constructs as simultaneous correlates of outcomes from the UNC Tobacco Warnings Model. Message perceptions demonstrated incremental criterion validity with attention, an early behavioral antecedent ($\beta = 0.82$, $p < .001$). Effects perceptions demonstrated incremental criterion validity with later behavioral antecedents (range $\beta = 0.74$–0.87, all $p < .01$) and quitting behaviors ($\beta = 0.36$–0.66, all $p < .001$). Formative research on anti-smoking messages may benefit from focusing on effects perceptions to characterize potential for behavior change.

Keywords Effects perceptions · Formative research · Measurement · Perceived message effectiveness · Structural equation modeling

Introduction

Health communication efforts commonly employ brief messages designed to discourage health risk behaviors (Noar 2006). Researchers often use audience ratings of health messages or perceived message effectiveness (PME) to select the most promising ones for further clinical testing or dissemination (Noar et al., 2018b). For example, previous studies have used PME to identify potent messages about the chemicals in cigarette smoke (chemical messages) (Noar et al., 2018c) and evaluate HIV testing messages already out in the field (Davis et al., 2011). In the absence of behavioral data on message impact, PME ratings are a cost-effective means to efficiently evaluate many candidate messages.

While use of PME has steadily increased over the last two decades (Noar et al., 2018b), little is definitively known about the extent to which PME is indicative of message impact on behavior (Yzer et al., 2014). A major hurdle to evaluating the PME-behavior correspondence is the heterogeneity of PME measures (Noar et al., 2018a), which

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10865-020-00163-0) contains supplementary material, which is available to authorized users.

✉ Sabeeh A. Baig
sbaig1@alumni.jh.edu

Seth M. Noar
noar@unc.edu

Nisha C. Gottfredson
gottfredson@unc.edu

Allison J. Lazard
lazard@unc.edu

Kurt M. Ribisl
kurt_ribisl@unc.edu

Noel T. Brewer
ntb@unc.edu

[1] Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina, 325 Rosenau Hall CB 7440, Chapel Hill, NC 27599, USA

[2] Hussman School of Journalism and Media, University of North Carolina, Chapel Hill, NC, USA

[3] Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

have traditionally assessed either of two constructs, message perceptions and effects perceptions (Dillard et al., 2007a, b; Dillard & Ye, 2008; Noar et al., 2018b). *Message perceptions* are judgments about whether a message will promote further processing that leads to persuasion, and *effects perceptions* are judgments about a message's potential to change important antecedents of behavior or behavior itself. To our knowledge, only one study has formally examined the structure of both types of perceptions and found some evidence of bidimensionality (Dillard & Ye, 2008). However, due to their high correlation, the researchers ultimately concluded that message perceptions and effects perceptions could be used interchangeably in applied settings (Dillard & Ye, 2008).

Message perceptions and effects perceptions have important differences, however, which may preclude their interchangeable use in formative research. One difference is that message perceptions focus on characteristics that facilitate initial processing of the message itself. These characteristics may include message credibility (e.g., "How believable was the message in this?") (Donovan et al., 2006) and perceived argument strength (e.g., "This ad was convincing.") (Lee et al., 2011). In contrast, effects perceptions focus on the behavior that a message is designed to influence. Dimensions of effects perceptions may include perceived impact on behavioral motivation (e.g., "This ad makes me want to quit smoking.") (Niederdeppe et al., 2011) or an antecedent such as concern or worry (e.g., "This ad made me feel concerned about my smoking.") (Brennan et al., 2013). Based on this difference in focus (the message or behavior), we hypothesize that message perceptions and effects perceptions are distinct, although related, constructs (*bidimensionality hypothesis*).

Another difference is that effects perceptions items ask recipients of a message to think about how a message affects their own behavior using behavioral and personal referents. The use of behavioral referents shifts the meaning of effects perceptions away from general persuasive potential, as in the case of message perceptions, and toward motivation to change behavior. The use of personal referents further increases the specificity of effects perceptions to a message recipient's motivation to change their own behavior. The principle of compatibility, originally articulated in the context of the Theory of Reasoned Action, indicates that a construct must be measured at the same level of specificity as the behavior of interest for it to demonstrate correspondence with that behavior (Fishbein & Ajzen, 2010). Therefore, we hypothesize that effects perceptions are a better correlate of behavior than message perceptions (*behavioral compatibility hypothesis*).

By pursuing the bidimensionality and behavioral compatibility hypotheses, our study ultimately sought to shed light on the distinction between message perceptions and effects perceptions and which of these constructs is more relevant to developing messages for behavior change. We examined the incremental criterion validity of message perceptions and effects perceptions in the context of anti-smoking messages about the chemicals in cigarette smoke (chemical messages) as a case study.

Methods

Participants and procedures

A convenience sample of 719 U.S. adult smokers from the San Francisco Bay Area participated in a RCT that compared the impact of chemical messages versus anti-littering messages. Specifically, participants were ages 21 or older following the legal age of buying tobacco products in California, proficient in English, and current smokers (i.e., had smoked at least 100 cigarettes in their lifetime and now smoke every or some days) (Davis et al., 2009). Exclusion criteria were pregnancy, enrollment in an ongoing smoking cessation trial, living in the same household as another trial participant, and smoking fewer than seven cigarettes per week or only roll-your-own cigarettes.

Smokers attended visits at the trial office and brought an eight-day supply of cigarettes to all but the last visit. Participants were randomized to receive chemical messages (intervention) or messages about properly disposing cigarette litter (littering messages; control) on the sides of their cigarette packs each week for 3 weeks. An example of a chemical message is, "Cigarette smoke contains formaldehyde. This causes throat cancer"; and an example of a littering message is, "Cigarette butts don't biodegrade. Please do not litter." While participants in the intervention condition completed weekly surveys, study personnel applied a new chemical message on the side of their cigarette packs. Thus, the smokers in the intervention arm rotated through three chemical messages during the course of the RCT in a random order. The control arm followed identical procedures. Smokers received up to $300 for their participation in the trial. Data collection lasted from September 2016 to March 2017. Additional details on design and protocol are available in the main RCT paper (Brewer et al., 2018).

Measures

The baseline survey assessed participant demographics and smoking behavior. The final visit survey assessed message perceptions using an established scale with six items: "This message is worth remembering"; "This message grabbed my attention"; "This message is powerful"; "This message is informative"; "This message is meaningful"; and "This message is convincing," ($\alpha = .94$) (Davis et al., 2013). We

assessed effects perceptions also at Visit 5 using the UNC Perceived Message Effectiveness Scale with three items: "This message discourages me from wanting to smoke"; "This message makes smoking seem unpleasant to me"; and "This message makes me concerned about the health effects of smoking," ($\alpha = .90$) (Baig et al., 2018). The response options for both scales ranged from "strongly disagree" to "strongly agree" (coded as 1–5).

The UNC Tobacco Warnings Model (TWM) informed the selection of major construct validators (Brewer et al., 2018). These were attention to the labels, the early antecedent to behavior from the TWM; the two intermediate antecedents number of conversations about the labels in the past week and negative affect; and the three late antecedents thinking about the chemicals in cigarette smoke, thinking about the harms of smoking, and quit intentions. Given our behavioral compatibility hypothesis, we expected message perceptions to be positively associated with the early and intermediate antecedents and not be associated with the late antecedents. In contrast, we expected effects perceptions to not be associated with the early antecedents and be positively associated with the intermediate and late antecedents. The main behavioral outcomes were number of times butting out a cigarette in the past week, number of times forgoing a cigarette in the past week, and weekly recall of quit attempts. Given that quitting initiation exists on a spectrum (Partos et al., 2014), examining multiple quitting behaviors that differ in intensity is likely to add depth to the incremental criterion validity findings. We expected message perceptions to not be associated with and effects perceptions to be positively associated with the behavioral outcomes.

Because chemical messages can have a range of effects (Brewer et al., 2018), we also included recognition of the labels, avoidance of the labels, and seeking information about the chemicals in cigarette smoke as construct validators (Table 1) (Noar et al., 2016). Recognition conceptually overlaps with attention; avoidance is a largely affect-driven behavior that is productive for tobacco warning impact (Hall et al., 2018); and chemical information-seeking may reinforce some of the educational effects of tobacco warnings (Lambert & Loiselle, 2007). The final visit survey assessed all TWM constructs, behavioral outcomes, and other construct validators. Table 1 presents information on the measures for all constructs assessed in this study including ordinal reliability coefficients for multi-item measures.

Data analysis

Analyses used R (ver. 3.5.1) (R Core Team, 2000) with three add-on packages, psych (ver. 1.8.4) (Revelle, 2011) for calculating ordinal reliability coefficients, lavaan (ver. 0.6–2) (Rosseel, 2012) for estimating measurement and structural models, and ggplot2 (ver. 3.0.0) (Wickham, 2016) for creating figures.

Participants with complete missingness on both message perceptions and effects perceptions were dropped from the analyses resulting in an analytic sample of 703 smokers. Single-item outcomes did not have any missingness, and partial missingness on multi-item outcomes was negligible (< 3%). Therefore, we used pairwise deletion to handle missingness on the outcomes. Doing so enabled us to use ordinal estimation for all models without having to drop any additional participants. Analyses used data exclusively from the final visit, and statistical tests used a critical alpha of .01 (or a 99% confidence interval) to account for the possibility of artificially inflated validity associations due to assessing all constructs at a single time point.

Dimensionality

We examined the dimensionality of the message perceptions and effects perceptions scales by estimating unidimensional and bidimensional confirmatory factor analytic (CFA) models. The unidimensional model permitted the nine items from both scales to load onto a general factor for PME. The bidimensional model allowed the six message perceptions items and the three effects perceptions items to load onto two separate but correlated factors without any cross-loadings. We used the likelihood ratio (LR) test for nested models to compare both CFA models.

Respecification of the preferred CFA model followed an iterative process in which we tested a select number of correlated errors based on substantial modification indices (MIs), beginning with the largest and sequentially moving to the smallest. Our previous study found greater amounts of measurement error in the six message perceptions items than the three effects perceptions items across three large samples and multiple measurement occasions (Baig et al., 2018). Therefore, we prioritized MIs involving message perceptions items and pursued corresponding modifications strictly in a confirmatory fashion. For each MI, we confirmed that estimating the associated path statistically significantly improved model fit via the LR test. After exhausting all substantial MIs, we examined the Bayesian Information Criterion (BIC; Bollen et al., 2012, 2014; Jarosz & Wiley, 2014), Root Mean Square Error of Approximation (RMSEA), and Comparative Fit Index (CFI) for adequacy (BIC < 0; RMSEA < .06; CFI > .95). We also substantively confirmed that suggested respecifications should be retained in the final model by examining relevant item wordings. Finally, we confirmed that the message perceptions and effects perceptions scales functioned similarly in the context of chemical and littering messages using standard procedures for testing

Table 1 Measures for all construct with ordinal reliability coefficients for multi-item scales

Construct, measure(s)	Response options (coding)	References
Effects perceptions ($\alpha = 0.90$)	Strongly disagree (1) to Strongly agree (5)	Baig et al., (2018)
This message discourages me from wanting to smoke		
This message makes smoking seem unpleasant to me		
This message makes me concerned about the health effects of smoking		
Message perceptions ($\alpha = 0.94$)	Strongly disagree (1) to Strongly agree (5)	Davis et al., (2013)
This message is worth remembering		
This message grabbed my attention		
This message is powerful		
This message is informative		
This message is meaningful		
This message is convincing		
Attention to the messages ($\alpha = 0.93$)	Not at all/Never (1) to Very much/All the time (5)	Nonnemaker et al., (2010), Fathelrahman et al., (2010)
How much did the labels grab your attention?		
How often did you notice the labels?		
How often did you read or look closely at the labels?		
Recognition of the messages	Correct recognition (1) Incorrect recognition (0)	–
Select the text from the label we put on the side of your cigarette packs at your last visit		
Number of conversations about the label in the past week	0–100 times	–
In the last week, how many times did you talk to other people about the label on your cigarette packs?		
Negative affect ($\alpha = 0.93$)	Not at all (1) to Extremely (5)	Keller & Block, (1996), Watson, Clark, & Telegen (1988), Nonnemaker et al., (2010)
How much did the labels on your cigarette packs make you feel anxious?		
How much did the labels on your cigarette packs make you feel sad?		
How much did the labels on your cigarette packs make you feel scared?		
How much did the labels on your cigarette packs make you feel guilty?		
How much did the labels on your cigarette packs make you feel disgusted?		
Thinking about the chemicals in cigarettes	Never (1) to All of the time (5)	–
In the last week, how much did you think about the chemicals in the smoke from your cigarettes?		
Thinking about the harms of smoking ($\alpha = 0.84$)	Not at all/Never (1) to Very much/All of the time (5)	Fathelrahman et al., (2010)
How much did the labels make you think about the health problems caused by smoking?		
In the last week, how much did you think about the harm your smoking might be doing to you?		
In the last week, how much did you think about the harm your smoking be doing to other people?		
Avoidance of cigarette pack labels ($\alpha = 0.89$)	Never (1) to All of the time (5)	Hyland et al., (2016)
How often did you try to avoid the labels on your cigarette packs?		
How often did you try to avoid the labels on your cigarette packs?		
How often did you put your cigarettes away because you didn't want others to see the labels on your packs?		
Seeking information about the chemicals in cigarette smoke	0 times (1) to 6 or more times (4)	Nelson et al., (2004)
In the last 3 weeks, how many times have you looked for information about the chemicals in cigarettes or cigarette smoke?		
Quit intentions ($\alpha = 0.96$)	Not at all [item stem] (1) to Very [item stem] (4)	Klein et al., (2009)
How interested are you in quitting smoking in the next month?		
How much do you plan to quit smoking in the next month?		
How likely are you to quit smoking in the next month?		
Number of times butting out a cigarette in the past week	Never (1) to 10 or more times (5)	Li et al., (2014)
In the last week, how often have you butted out a cigarette before you finished it because you wanted to?		

Table 1 (continued)

Construct, measure(s)	Response options (coding)	References
Number of times forgoing a cigarette in the past week In the last week, how often have you stopped yourself from having a cigarette because you wanted to?	Never (1) to 10 or more times (5)	Li et al., (2014)
Weekly recall of quit attempts In the last week, did you stop smoking for 1 day or longer because you were trying to quit smoking?	Yes (1) No (0)	CDC, (2008)

measurement invariance involving comparisons of increasingly constrained multiple-group versions of the respecified CFA model (Millsap & Yun-Tein, 2004).

Incremental criterion validity

We examined the incremental criterion validity of message perceptions and effects perceptions using structural equation modeling (SEM). For a given outcome, we estimated a structural model that simultaneously regressed the outcome on the latent variables for message perceptions and effects perceptions using the best fitting bidimensional CFA model as the measurement model. The resulting associations were estimated using the unique variances of message perceptions and effects perceptions and represented the extent to which one type of perception was associated with the outcome after controlling for the other type of perception. This was a departure from the small number of longitudinal and meta-analytic PME validation studies (Brennan et al., 2013; Davis et al., 2016; Dillard et al., 2007a, b; O'Keefe, 2018), which have focused on independent associations between PME constructs and outcomes, making our validation efforts inherently comparative. The analytic approach was identical for all outcomes except that we specified an outcome measured with a single item as a manifest variable and one measured with a multi-item scale as a latent variable. All multi-item scales demonstrated high reliability ($\alpha \geq 0.84$). We confirmed model fit using the BIC, RMSEA, and CFI based on previously mentioned criteria. Supplemental Table 1 provides bivariate correlations for all constructs modeled in this study; consistent with all other analyses that treated the corresponding measures as ordinal (or dichotomous), these were Spearman correlations. For all outcomes, the absolute differences in the bivariate correlations of message perceptions and effects perceptions largely corresponded to the results from the SEM models.

Results

More than a third of the participants were white (37.4%) or African-American (35.6%; Table 2). Nearly a quarter were gay, lesbian, or bisexual (25.3%). A sizable minority of the participants smoked only some days (22.9%).

Dimensionality

Message perceptions and effects perceptions, although conceptually distinct, represented related constructs (Fig. 1). In a bidimensional measurement model, the message perceptions and effects perceptions items loaded strongly (0.78–0.92) onto separate, but highly correlated factors ($r = 0.82$). The message perceptions and effects perceptions items also loaded strongly (0.80–0.90) onto a general factor for perceived message effectiveness in a unidimensional measurement model. However, the bidimensional model had better fit to the data than the unidimensional model ($\chi^2 = 136$, $p < .001$). Model fit indices improved in the bidimensional model (BIC = 106; RMSEA = .12; CFI = .99) as compared to the unidimensional model (BIC = 517; RMSEA = .19; CFI = .98) thereby confirming the result of the LR test.

The bidimensional model needed to account for correlated errors (i.e., local dependence) mostly in the message perceptions items to achieve adequate global fit. Thus, we correlated the errors on the *attention-grabbing* and *powerful*, the *informative* and *meaningful*, and the *meaningful* and *convincing* message perceptions items. LR tests confirmed that these three modifications improved model fit (range $\chi^2 = 19.9$–56.4, all $p < .001$). We also correlated the errors on the *discouragement* and *unpleasant* effects perceptions items, which improved model fit ($\chi^2 = 21.6$, $p < .001$). The final bidimensional model had adequate fit (BIC = − 64.2; RMSEA = .061; CFI = 1.00).

The message perceptions and effects perceptions scales functioned similarly among participants who received chemical or littering messages. More specifically, LR tests for measurement invariance found that the final bidimensional model held exactly among participants who received

Table 2 Participant characteristics

	$n = 703$ %
Age (years)	
21–29	22.6
30–39	22.0
40–49	18.9
50–59	25.6
60 +	10.9
Gender	
Male	50.8
Female	44.8
Transgender (includes other gender identity)	4.4
Gay, lesbian or bisexual	25.3
Race	
White	37.4
Black or African-American	35.6
Asian	8.4
American Indian or Alaska Native	4.8
Native Hawaiian or other Pacific Islander	3.3
Other	10.5
Latinx	13.8
Education	
High school graduate or less	21.8
Some college	37.8
Bachelor's degree	33.1
Graduate degree	7.3
Household income, annual	
$0–$24,999	45.2
$25,000–$49,999	24.0
$50,000–$74,999	12.4
$75,000 +	18.4
Low income, < 200% of federal poverty level	57.6
Smoking frequency	
Daily	77.1
Non-daily	22.9

either type of message except that the latent means for message perceptions and effects perceptions differed for chemical and littering messages ($\chi^2(2) = 16.6$, $p < .001$). Common fit indices for variously constrained multiple-group models confirmed this finding by indicating that the corresponding model had the best fit (BIC = − 420; RMSEA = 0.055; CFI = 1.00).

Incremental criterion validity

The SEM model for attention to the labels had marginally acceptable fit (BIC = 52.5; RMSEA = .10; CFI = .99). In comparison, the SEM models for the other 11 outcomes had more than adequate fit (BIC median [range] = − 90.7 [− 198, − 27.8]; RMSEA = .06 [.05, .08]; CFI = 1.00 [.99, 1.00]). To reiterate, these models estimated validity associations using the unique variances of message perceptions and effects perceptions.

Message perceptions demonstrated incremental criterion validity with the TWM construct that is farthest from behavior, while effects perceptions demonstrated incremental criterion validity with constructs that are conceptually closer to behavior and with behavior itself. Specifically, more positive message perceptions were associated with greater attention ($\beta = 0.82$, $p < .001$; Fig. 2). Neither message perceptions nor effects perceptions were associated with conversations about the labels (both $p \geq .27$). More positive effects perceptions were associated with more negative affect, more thinking about the chemicals in cigarette smoke, more thinking about the harms of smoking, and stronger quit intentions (range $\beta = 0.74$–0.87, all $p < .01$). Finally, more positive effects perceptions were associated with more frequent butting out a cigarette ($\beta = 0.36$, $p < .001$), more frequent forgoing a cigarette ($\beta = 0.53$, $p < .001$), and being more likely to engage in a quit attempt ($\beta = 0.66$, $p < .001$). Unexpectedly, more positive message perceptions were also associated with weaker quit intentions ($\beta = -0.29$, $p = .006$) and being less likely to engage in a quit attempt ($\beta = -0.35$, $p = .003$).

A similar pattern of distinct correlates for message perceptions and effects perceptions held for the other outcomes we examined. More positive message perceptions were associated with greater recognition of the labels ($\beta = 0.35$, $p = .003$). More positive effects perceptions were associated with greater message avoidance ($\beta = 0.78$, $p < .001$) and more information seeking about chemicals ($\beta = 0.44$, $p < .001$). Unexpectedly, more positive message perceptions were also associated with weaker avoidance ($\beta = -0.27$, $p = .004$). Analyses stratified by trial arm yielded similar findings for all outcomes, although some associations lost statistical significance due to smaller sample size.

Discussion

In a diverse sample of adult smokers, message perceptions demonstrated incremental criterion validity with attention to and recognition of the messages. In contrast, effects perceptions demonstrated incremental criterion validity with negative affective reactions to the messages, thinking about the chemicals in cigarette smoke, thinking about the harms of smoking, avoidance of the messages, seeking chemical information, quit intentions, butting out a cigarette, forgoing a cigarette, and quit attempts. Thus, message perceptions were associated with constructs most distal to behavior, and effects perceptions were associated with constructs more proximal

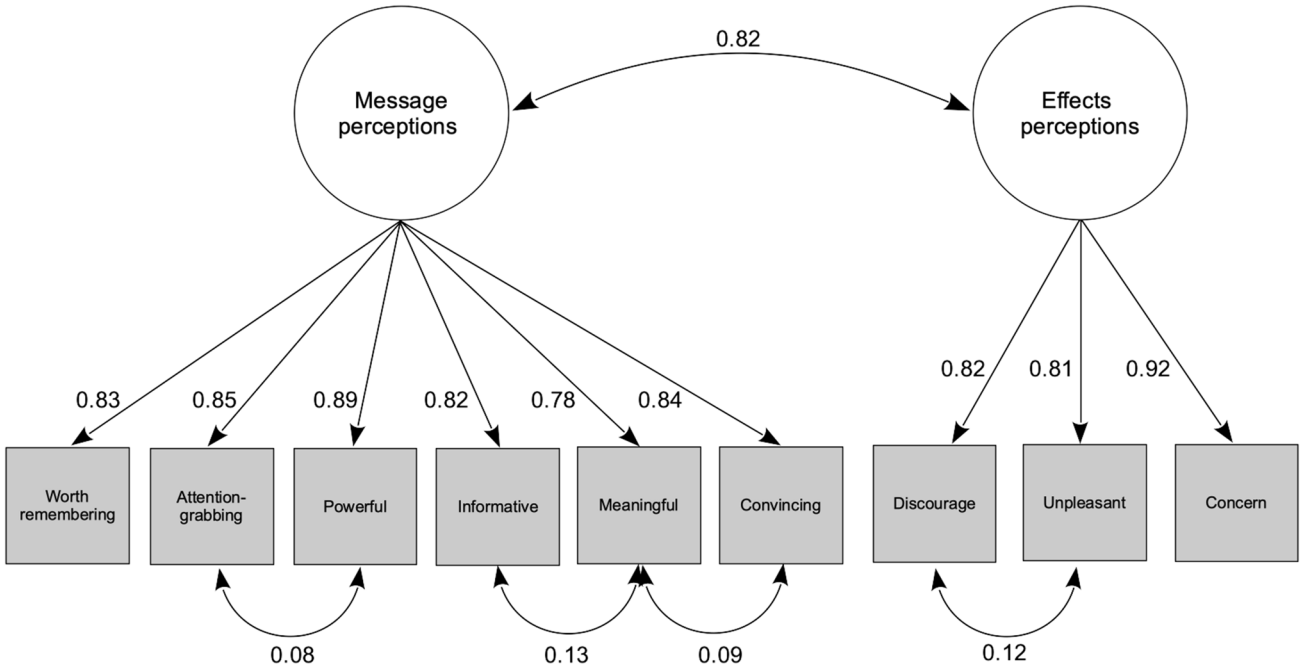

Fig. 1 Confirmatory factor analytic model of message perceptions and effects perceptions. Bayesian Information Criterion = −64.2; Root Mean Squared Error of Approximation = .061; and Comparative Fit Index = 1.00

to behavior as well as behavior itself (quitting behaviors and other behaviors that are productive for successful quitting). Formative research may use message perceptions in the earliest stages of message development and shift to effects perceptions when the focus is explicitly on behavior change.

Our findings are partially consistent with the bidimensionality hypothesis. Bidimensional confirmatory factor analytic models that represented message perceptions and effects perceptions as separate constructs were more consistent with our data than unidimensional models. This finding provides some support for the bidimensionality hypothesis. However, message perceptions and effects perceptions were highly correlated, suggesting that any bidimensionality is at the margins of substantially overlapping constructs. In addition, message perceptions were relevant in the earliest stage of the TWM while effects perceptions were relevant in the later stages, which is consistent with the behavioral compatibility hypothesis. It is important to note that the corresponding associations were estimated using the unique variances of message perceptions and effects perceptions. The unique variances of both types of perceptions were considerably smaller than their shared variance given that they were highly correlated. For a given outcome, an association for effects perceptions, for example, and a lack of association for message perceptions usually meant that effects perceptions were associated with the outcome above and beyond the extent to which message perceptions were. Therefore, the distinct patterns of incremental criterion validity generally indicated that effects perceptions were a better correlate of behavior and behavioral antecedents than message perceptions and that message perceptions were a better correlate of attention.

The effects perceptions scale used behavioral and personal referents while the message perceptions scale did not. This differential use of referents may have contributed to the bidimensionality between message perceptions and effects perceptions. The differential use of behavioral and personal referents may have also shifted the meaning of effects perceptions closer to the respondent's behavior as compared to message perceptions, thereby enhancing the correspondence of effects perceptions with behavioral antecedents and behavior. These observations raise questions about whether focusing on the message recipient and referencing the target behavior are defining features of effects perceptions. Furthermore, some existing message perceptions items may be readily adapted to use personal referents (e.g., "This message is meaningful to me.") (Davis et al., 2013) or behavioral referents (e.g., "This ad was informative about the harms of smoking.") (Davis et al., 2013). It is unknown how incorporating referents into a measure of message perceptions would change the meaning of the underlying construct. Future studies could use explanatory item response modeling (Chalmers, 2015) or generalizability theory (Vispoel et al., 2017) to better understand the sources of bidimensionality in and differences in the criterion validity of message perceptions and effects perceptions. These psychometric methods allow researchers to identify the psychological processes

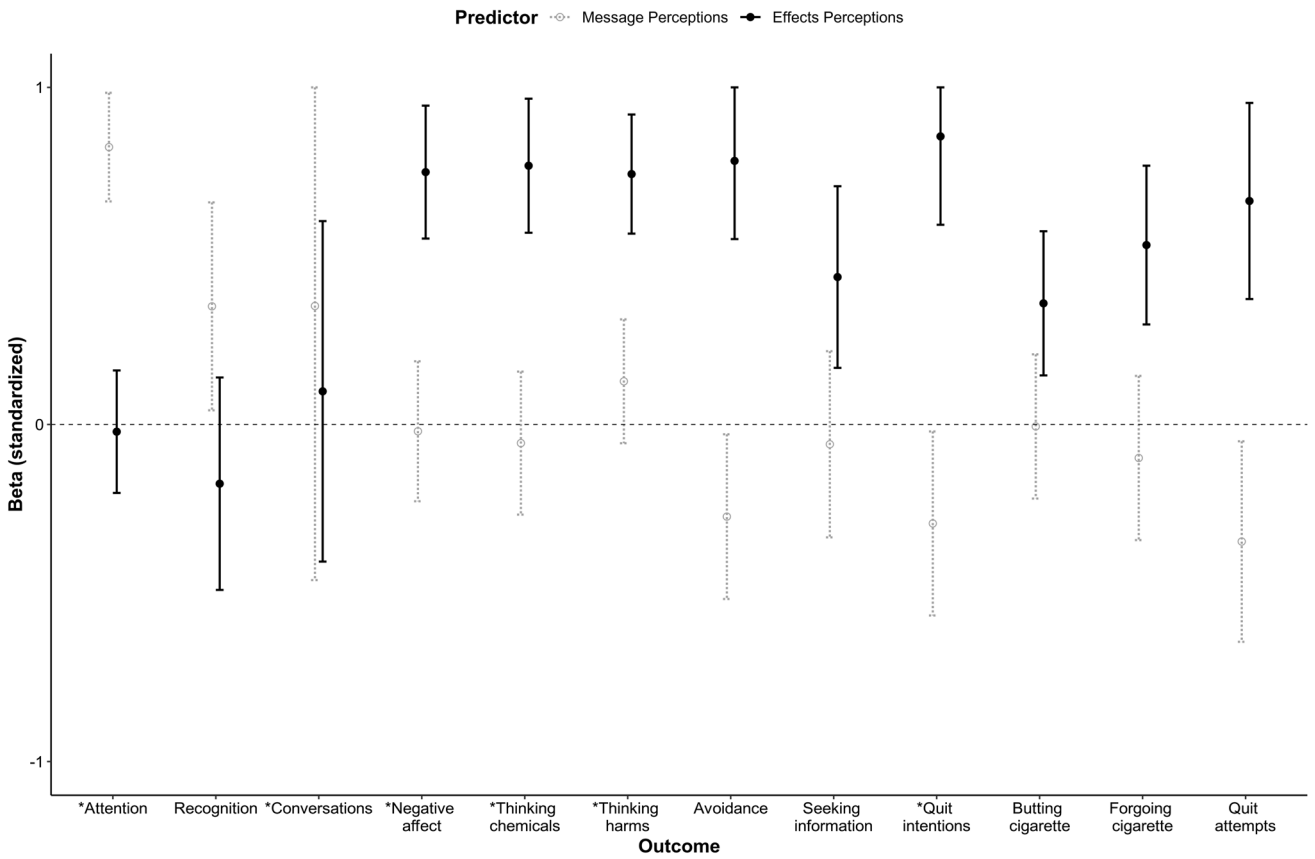

Fig. 2 Incremental criterion validity of message perceptions and effects perceptions with behavioral antecedents from the UNC Tobacco Warnings Model (*), quitting behavior, and related constructs. Error bars denote 99% confidence intervals

underlying responses to self-report measures that may be crucial to further differentiating between message perceptions and effects perceptions.

Our study also found that more negative message perceptions were unexpectedly associated with stronger message avoidance, stronger quit intentions, and being more likely to engage in a quit attempt (after controlling for effects perceptions). These associations may represent statistical anomalies from controlling for effects perceptions, always a possibility when controlling for highly correlated constructs. In addition, the confidence intervals for these unexpected associations included practically null values unlike for other statistically significant associations. Thus, these marginal findings should be interpreted with caution. Given these caveats, we offer a speculation on the unexpected findings. Perhaps audience members who think a message is problematic (e.g., do not find the message to be convincing) spend more time on it, leading to greater central processing and thereby greater behavioral motivation. This would be a dual process model in which promising messages motivate behavior, but their impact is somewhat undercut by additional attention to problematic messages.

While our study strongly supports formatively evaluating candidate messages for behavioral impact using effects perceptions measures, the actual validity associations with behavioral antecedents were strong and with quitting and other behaviors were weak to moderate. Thus, effects perceptions did not explain substantial amounts of variance in some constructs of interest. This suggests that a researcher could not conclude that messages with adequate effects perceptions will change behavior (i.e., adequate effects perceptions are not a sufficient condition for message impact as theorized by a number of researchers) (Dillard et al., 2007a, b; Noar et al., 2018a; Yzer et al., 2014). However, a researcher could conclude that messages without adequate effects perceptions will not change behavior (i.e., adequate effects perceptions are a necessary condition for message impact). This points to the practical value of evaluating candidate messages on effects perceptions: doing so can help researchers identify potent messages to further test for behavioral impact in field trials or longitudinal experiments. Effects perceptions being closer to behavior provides a conceptual basis for their use in formative research in this capacity. Future studies should explore what constitutes adequate effects perceptions for

message exclusion (Dillard & Ha, 2016). Future studies should also identify contexts in which it is useful to assess message perceptions, given that our findings do not allow us to rule out their use in formative research altogether.

Strengths of our study are the large number of theoretically informed validators, including early to late behavioral antecedents and behavioral outcomes, as well as the use of message perceptions and effects perceptions scales with adequate psychometric properties and latent variable models suitable to evaluating criterion validity in a comparative framework. The cross-sectional assessment of both PME constructs and all validators is a limitation of our study. Future studies should replicate our findings using appropriate longitudinal data, which would permit further differentiation between message perceptions and effects perceptions in terms of their predictive validity. Another limitation is that we could not examine whether PME represented a higher-order construct that influences message perceptions and effects perceptions, a possibility originally raised by Dillard & Ye (2008) that could also account for the bidimensionality observed in this study. This is largely because second-order CFA models are empirically indistinguishable from corresponding bidimensional CFA models and, as a result, cannot adequately be used to test for the existence of higher-order factor structures. Future studies should examine message perceptions and effects perceptions alongside other constructs that are relevant to message testing (e.g., message reactance, source credibility) to test for the existence of higher-order PME and related constructs. Two other limitations are that our study only used one message perceptions scale and one effects perceptions scale and included only chemical and littering messages. Therefore, our findings may not be generalizable to other PME measures and/or message testing scenarios. Future validation studies should focus on multiple message perceptions and effects perceptions scales and include diverse messages about many health behaviors and relevant antecedents.

Conclusions

While message perceptions and effects perceptions are highly correlated, our study shows that they can have distinct patterns of association with behavior and behavioral antecedents. In the context of tobacco use, message perceptions are more important in the earliest stages of message processing while effects perceptions appear to be more important in the intermediate and late stages leading to behavior change. Therefore, we recommend focusing on effects perceptions when evaluating messages for their potential to change smoking behavior and potentially other behaviors.

Acknowledgements Research reported in this publication was supported by Grant Number P50CA180907 from the National Cancer Institute and FDA Center for Tobacco Products. Research reported in this publication was also partially supported by the National Institute on Drug Abuse through grant funding awarded to Dr. Gottfredson (K01DA0351523). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Food and Drug Administration.

Compliance with ethical standards

Conflict of interest None of the authors have received funding from tobacco product manufacturers. Dr. Brewer has served as a paid consultant in litigation against tobacco companies. Dr. Ribisl has served as an expert consultant in litigation against cigarette manufacturers and Internet tobacco vendors. Dr. Ribisl is also a member of the Tobacco Products Scientific Advisory Committee for the FDA Center for Tobacco Products. The views expressed in this paper are his and not those of the FDA. The authors declare no conflict of interest.

Human and animal rights and Informed consent All procedures followed were in accordance with ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all individual participants included in the study.

References

Baig, S. A., et al. (2018). UNC Perceived Message Effectiveness: Validation of a brief scale. *Annals of Behavioral Medicine*. https://doi.org/10.1093/abm/kay080

Bollen, K. A., Harden, J. J., Ray, S., & Zavisca, J. (2014). BIC and alternative Bayesian information criteria in the selection of structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(1), 1–19. https://doi.org/10.1080/10705511.2014.856691

Bollen, K. A., Ray, S., Zavisca, J., & Harden, J. J. (2012). A comparison of Bayes factor approximation methods including two new methods. *Sociological Methods & Research, 41*(2), 294–324. https://doi.org/10.1177/0049124112452393

Brennan, E., et al. (2013). Assessing the effectiveness of antismoking television advertisements: Do audience ratings of perceived effectiveness predict changes in quitting intentions and smoking behaviours? *Tobacco Control, 23*(5), 412–418. https://doi.org/10.1136/tobaccocontrol-2012-050949

Brewer, N. T., et al. (2018). Cigarette pack messages about toxic chemicals: A randomised clinical trial. *Tobacco Control*. https://doi.org/10.1136/tobaccocontrol-2017-054112

CDC. (2008). Cigarette smoking among adults—United States, 2007. *MMWR. Morbidity and Mortality Weekly Report, 57*(45), 1221–1226.

Chalmers, R. P. (2015). Extended mixed-effects item response models with the MH-RM algorithm. *Journal of Educational Measurement, 52*(2), 200–222. https://doi.org/10.1111/jedm.12072

Core Team, R. (2000). *R language definition*. Vienna, Austria: R Foundation for Statistical Computing.

Davis, S., et al. (2009). State-specific prevalence and trends in adult cigarette smoking-United States, 1998–2007. *Morbidity and Mortality Weekly Report, 58*(9), 221–226.

Davis, K. C., et al. (2011). Exploring African American women's perceptions of a social marketing campaign to promote HIV

testing. *Social Marketing Quarterly, 17*(3), 39–60. https://doi.org/10.1080/15245004.2011.595536

Davis, K. C., et al. (2013). Perceived effectiveness of cessation advertisements: The importance of audience reactions and practical implications for media campaign planning. *Health Communication, 28*(5), 461–472. https://doi.org/10.1080/10410236.2012.696535

Davis, K. C., et al. (2016). Perceived effectiveness of antismoking ads and association with quit attempts among smokers: Evidence from the TIPS From Former Smokers campaign. *Health Communication, 32*(8), 931–938. https://doi.org/10.1080/10410236.2016.1196413

Dillard, J. P., & Ha, Y. (2016). Interpreting perceived effectiveness: Understanding and addressing the problem of mean validity. *Journal of Health Communication, 21*(9), 1016–1022. https://doi.org/10.1080/10810730.2016.1204379

Dillard, J. P., Shen, L., & Vail, R. G. (2007a). Does perceived message effectiveness cause persuasion or vice versa? 17 consistent answers. *Human Communication Research, 33*(4), 467–488. https://doi.org/10.1111/j.1468-2958.2007.00308.x

Dillard, J. P., Weber, K. M., & Vail, R. G. (2007b). The relationship between the perceived and actual effectiveness of persuasive messages: A meta-analysis with implications for formative campaign research. *Journal of Communication, 57*(4), 613–631. https://doi.org/10.1111/j.1460-2466.2007.00360.x

Dillard, J. P., & Ye, S. (2008). The perceived effectiveness of persuasive messages: Questions of structure, referent, and bias. *Journal of Health Communication, 13*(2), 149–168. https://doi.org/10.1080/10810730701854060

Donovan, R. J., Jalleh, G., & Carter, O. B. J. (2006). Tobacco industry smoking prevention advertisements' impact on youth motivation for smoking in the future. *Social Marketing Quarterly, 12*(2), 3–13. https://doi.org/10.1080/15245000600721644

Fathelrahman, A. I., Omar, M., Awang, R., et al. (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 and Public Health, 7*(11), 4089–4099. https://doi.org/10.3390/ijerph7114089

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

Hall, M. G., et al. (2018). Why smokers avoid cigarette pack risk messages: Two randomized clinical trials. *Social Science & Medicine, 213,* 165–172.

Hyland, A., et al. (2016). Design and methods of the Population Assessment of Tobacco and Health (PATH) study. *Tobacco Control, 26*(4), 371–378. https://doi.org/10.1136/tobaccocontrol-2016-052934

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *Journal of Problem Solving*. https://doi.org/10.7771/1932-6246.1167

Keller, P. A., & Block, L. G. (1996). Increasing the persuasiveness of fear appeals: The effect of arousal and elaboration. *Journal of Consumer Research, 22*(4), 448. https://doi.org/10.1086/209461

Klein, W. M. P., Zajac, L. E., & Monin, M. M. (2009). Worry as a moderator of the association between risk perceptions and quitting intentions in young adult and adult smokers. *Annals of Behavioral Medicine, 38*(3), 256–261. https://doi.org/10.1007/s12160-009-9143-2

Lambert, S. D., & Loiselle, C. G. (2007). Health information-seeking behavior. *Qualitative Health Research, 17*(8), 1006–1019. https://doi.org/10.1177/1049732307305199

Lee, S., et al. (2011). Smoking cues, argument strength, and perceived effectiveness of antismoking PSAs. *Nicotine & Tobacco Research, 13*(4), 282–290. https://doi.org/10.1093/ntr/ntq255

Li, L., et al. (2014). Smoking-related thoughts and microbehaviours, and their predictive power for quitting. *Tobacco Control, 24*(4), 354–361. https://doi.org/10.1136/tobaccocontrol-2013-051384

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479–515. https://doi.org/10.1207/s15327906mbr3903_4

Nelson, D., et al. (2004). The Health Information National Trends Survey (HINTS): Development, design, and dissemination. *Journal of Health Communication, 9*(5), 443–460. https://doi.org/10.1080/10810730490504233

Niederdeppe, J., et al. (2011). Socioeconomic variation in recall and perceived effectiveness of campaign advertisements to promote smoking cessation. *Social Science and Medicine, 72*(5), 773–780. https://doi.org/10.1016/j.socscimed.2010.12.025

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*(1), 21–42. https://doi.org/10.1080/10810730500461059

Noar, S. M., et al. (2016). Pictorial cigarette pack warnings: A meta-analysis of experimental studies. *Tobacco Control, 25*(3), 341–354. https://doi.org/10.1136/tobaccocontrol-2014-051978

Noar, S. M., Barker, J., & Yzer, M. (2018a). Measurement and design heterogeneity in perceived message effectiveness studies: A call for research. *Journal of Communication*, *68*(5), 990–993. https://doi.org/10.1093/joc/jqy047

Noar, S. M., et al. (2018b). Perceived message effectiveness measures in tobacco education campaigns: A systematic review. *Communication Methods and Measures*, *12*(4), 295–313. https://doi.org/10.1080/19312458.2018.1483017

Noar, S. M., et al. (2018c). Identifying principles for effective messages about chemicals in cigarette smoke. *Preventive Medicine*, *106*, 31–37. https://doi.org/10.1016/j.ypmed.2017.09.005

Nonnemaker, J., et al. (2010). Experimental study of graphic cigarette warning labels. *Final Results Report. RTI Project*, *7*

O'Keefe, D. J. (2018). Message pretesting using assessments of expected or perceived persuasiveness: Evidence about diagnosticity of relative actual persuasiveness. *Journal of Communication, 68*(1), 120–142. https://doi.org/10.1093/joc/jqx009

Partos, T. R., et al. (2014). The predictive utility of micro indicators of concern about smoking: Findings from the International Tobacco Control Four Country Study. *Addictive Behaviors, 39*(8), 1235–1242. https://doi.org/10.1016/j.addbeh.2014.04.001

Revelle, W. (2011). An overview of the psych package. Department of Psychology Northwestern University. Retrieved 3 March, 2012.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 1,* 2. https://doi.org/10.18637/jss.v048.i02

Vispoel, W. P., Morris, C. A., & Kilinc, M. (2017). Practical applications of generalizability theory for designing, evaluating, and improving psychological assessments. *Journal of Personality Assessment*. https://doi.org/10.1080/00223891.2017.1296455

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Berlin: Springer. https://doi.org/10.1007/978-3-319-24277-4

Yzer, M., LoRusso, S., & Nagler, R. H. (2014). On the conceptual ambiguity surrounding perceived message effectiveness. *Health Communication, 30*(2), 125–134. https://doi.org/10.1080/10410236.2014.974131