UNIVERSITY OF LEEDS

This is a repository copy of *How do we optimize message matching interventions? Identifying matching thresholds, and simultaneously matching to multiple characteristics*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/204139/

Version: Accepted Version

Article:
Joyal-Desmarais, K. orcid.org/0000-0003-0657-8367, Rothman, A.J. and Snyder, M. (2020) How do we optimize message matching interventions? Identifying matching thresholds, and simultaneously matching to multiple characteristics. European Journal of Social Psychology, 50 (3). pp. 701-720. ISSN 0046-2772

https://doi.org/10.1002/ejsp.2645

© 2019 John Wiley & Sons, Ltd. This is the peer reviewed version of the following article: Joyal-Desmarais, K, Rothman, AJ, Snyder, M. How do we optimize message matching interventions? Identifying matching thresholds, and simultaneously matching to multiple characteristics. Eur J Soc Psychol. 2020; 50: 701–720, which has been published in final form at https://doi.org/10.1002/ejsp.2645. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

[Preprint Article]

How do we Optimize Message Matching Interventions? Identifying Matching Thresholds, and Simultaneously Matching to Multiple Characteristics

Keven Joyal-Desmarais[a], Alexander J Rothman[a], Mark Snyder[a]

[a] Department of Psychology, University of Minnesota, 75 E River Rd, Minneapolis, MN 55455

Author Notes

Contact Information. Correspondence concerning this article should be addressed to Keven Joyal-Desmarais. Contact: joyal008@umn.edu.

Open Science Foundation Project Page. A project page for this project can be accessed at osf.io/c7t3w. This page contains our supplemental materials, analysis script files, and the data we used.

Acknowledgements. This research was supported by a Social Sciences and Humanities Research Council (SSHRC) of Canada doctoral fellowship to KJD. Data collection for Part II was supported by the Center for the Study of the Individual and Society at the University of Minnesota. The authors did not receive any funds or grant to do this research from funding agencies in the public, commercial, or not-for-profit sectors. We extend special thanks to Midori Nishioka for her substantial feedback on this manuscript.

Conflict of Interest Statement: The authors declare no conflict of interest.

How do we Optimize Message Matching Interventions? Identifying Matching Thresholds, and Simultaneously Matching to Multiple Characteristics

Abstract

Matching the content of persuasive messages to the characteristics (e.g., motives, personality) of people receiving them is a widely used technique to improve persuasion. However, little is known about how to optimize matching beyond simply using the technique. We propose that matching interventions can be strengthened by matching messages to multiple characteristics at a time, and introduce the concept of *matching thresholds* to improve the way interventionists assign messages. Matching thresholds are defined as the points along characteristics where people change from being most responsive to one message type to another. We provide statistical and methodological tools to estimate matching thresholds, and evaluate these tools in two simulation studies. We then report an online experiment (N=568) where we find an advantage for simultaneously matching messages to two characteristics (promotion focus and interdependent self-construal) and provide estimates of the matching thresholds to guide the assignment of gain/loss frames, and independence/interdependence appeals.

Keywords: message matching; message tailoring; persuasion; matching thresholds; independent and interdependent self-construal; promotion and prevention focus

How do we Optimize Message Matching Interventions? Identifying Matching Thresholds, and Simultaneously Matching to Multiple Characteristics

Research on persuasion has a long-standing focus on identifying and optimizing strategies to improve persuasive communication, and one of the most well-documented strategies for increasing persuasion is *message matching*: altering the content of a message to *match* the characteristics, needs, concerns, and/or preferences of the individuals to whom they are delivered (Maio & Olson, 2000; Noar, Benac, & Harris, 2007). For example, messages emphasizing the career outcomes of volunteerism are more persuasive for individuals highly motivated by career concerns, whereas messages emphasizing interpersonal outcomes are more persuasive for those highly motivated by social concerns (Clary et al., 1998). Message matching ensures people receive personally relevant information, and has been used to influence many types of behaviors, including physical health (Noar et al., 2007), mental health (Lueck, 2018), consumer (Snyder & DeBono 1985), prosocial (Clary et al., 1998), pro-environmental (Tangari & Smith, 2012), and political behaviors (Voelkel & Feinberg, 2018). However, beyond the idea that matched messages are more persuasive, principles for how to design matched messages to maximize their effectiveness have yet to be specified (Hawkins, Kreuter, Resnicow, Fishbein, & Dijkstra, 2008; Updegraff & Rothman, 2013).

We address this limitation in several ways. First, we demonstrate that matching a message to multiple characteristics can provide improvements in persuasion over matching to one characteristic. Second, we argue that message matching can be optimized by explicitly identifying the point along a characteristic at which people change their message preferences, and present methodological and statistical procedures to estimate such a point. We validate and

demonstrate our procedures using simulation work and an empirical study.

Overview of Message Matching

Message matching techniques are widespread in social scientific research and go by many names. For example, *functional matching* research in psychology and marketing has long emphasized that messages are most persuasive to the degree that they appeal to *functions* (i.e., goals, needs, and motivations) that are important to individuals (Maio & Olson, 2000; Shavitt & Nelson, 2002). For example, Clary et al. (1998) demonstrated that pro-volunteerism messages are most effective when they match individuals' strongest motivations. Similarly, a large literature has emerged in behavioral medicine under the banner of *message tailoring*. Message tailoring takes a conceptually broader approach than functional matching, emphasizing how messages can be matched to *any* individual difference variable deemed meaningful (e.g., people's names, occupations, or cultural backgrounds; Hawkins et al., 2008; Huang & Shen, 2016; Noar, Harrington, Van Stee, & Aldrich, 2011). Although research on functional matching and message tailoring operate mostly independently, the ideas underlying the techniques are largely equivalent.

Taken together, research on functional matching and message tailoring provide ample evidence that matching messages increases persuasion (Carpenter, 2012; Noar et al., 2007); however, there are few established principles that specify *how* to optimize the design of message matching interventions. We therefore examine two principles for optimizing message matching. First, we examine the potential of matching to multiple characteristics at a time to achieve greater persuasion. Second, we discuss how estimating the point along a characteristic at which people change their message preferences can benefit message matching, and provide

methodological and statistical procedures for doing so.

Principle 1: Optimizing Messages by Matching to Multiple Characteristics

One promising principle in improving message matching effectiveness is to increase the number of characteristics to which a message is matched (e.g., Noar et al., 2007; Strecher et al., 2008). To date, a few studies have evaluated this idea by comparing messages matched to small numbers of characteristics with messages matched to substantially larger sets of characteristics. For example, Strecher and colleagues (2008) compared smoking cessation messages that matched only to a person's name and gender, to messages matched "not only to name and gender, but also to age, ethnicity, marital status, spouse's smoking status, number of cigarettes, biggest barrier to quitting, reason for wanting to quit, degree and type of social support, and whether the participant had children in the home, was physically active, or worked outside the home" (p.377). The authors found that the latter outperformed the former.

Although useful, studies of this type confound the *number* of and the *specific* characteristics matched to. For example, it is common to compare messages matched only to demographic characteristics (e.g., gender) to messages matched to demographic and psychological characteristics (e.g., gender and motivational goals). If the latter messages are more effective than the former, this is typically taken as evidence that matching to a larger number of characteristics is effective. This interpretation assumes that matching to gender or motivational goals each elicit matching effects, and that matching to both simply leads to an incremental effect over matching to each characteristic alone. However, the same pattern of results can arise if matching to motivational goals is effective but matching to gender is not. Consequently, the advantage of matching to two characteristics is confounded with matching to

motivational goals by itself. Confounding increases further if conditions differ by more than one matching effect at a time, as this introduces dependence between individual matching effects. Therefore, to reduce confounding, it is essential for study designs to ensure that each specific matching effect is represented at each level of the number of matching effects being implemented, and independent from all other matching effects under examination.

A second limitation is that comparing matching to smaller versus substantially larger numbers of characteristics does not allow an examination of the effects of *incrementally* increasing the number of matched characteristics. To date, constructing messages that match to larger numbers of characteristics seems beneficial (Noar et al., 2007), but we do not know whether matching to 2 characteristics performs better than to 1, to 3 than to 2, and so forth. This knowledge is essential as assessing extra characteristics and creating matched messages to each requires resources that could otherwise be spent on optimizing other aspects of interventions.

In this report, we propose a procedure to address the limitations above and provide an initial examination of whether matching messages to two characteristics outperforms matching to one characteristic.

Principle 2: Using Matching Thresholds to Classify and Assign Messages

Theoretically, message matching can be conceptualized on a continuum, with message features varying in the *degree* to which they match people's characteristics (e.g., the stronger a person's career goals, the more a message emphasizing career benefits is a match)[1]. Yet, practically, interventions must assign messages in a categorical manner. Assigning messages involves determining which message would lead to the best effect for a given person (i.e., identifying which message is a match) and assigning that message over alternatives. Optimal

message assignment requires interventionists to identify *matching thresholds*—points along characteristics' continua that indicate when messages shift from being matches to being mismatches (see Figure 1, left side). For example, when using a 7-point scale to measure career goal strength, it may be that career-oriented messages are more effective for people scoring above 5, but less effective for people scoring below 5. In this case, the matching threshold would be the value 5.

Little guidance exists for determining matching thresholds. Instead, interventionists rely on intuition and/or conventions to assume values, and if the assumed value is mistaken, systematic misclassification will occur. For example, it is common to use the median score on measured characteristics to determine whether a given message represents a match/mismatch (e.g., Snyder & Debono, 1985; Werth & Foerster, 2007). However, if the matching threshold lies half a standard deviation (SD) below the median, this convention will misclassify messages for around one fifth of people.

Given the importance of accurately identifying matching thresholds, can we design studies to directly estimate their values? We argue that this can be accomplished by building on another common research design in the literature. Specifically, in contrast to intervention studies that assign messages using categorical assessments, many studies assign messages in a way that is non-contingent on a person's characteristic score, and then operationalize message matching continuously. For example, a study might randomly allocate people to view a liberally-oriented or a conservatively-oriented message, and then define matching in proportion to people's political ideology (e.g., a liberal message is deemed a stronger match the stronger a person's liberal leaning; Kidwell, Farmer, & Hardesty, 2013). Importantly, this type of design makes no

assumption about matching thresholds and can be used to directly estimate their values. In this report, we delineate a means for doing so.

Our Goals

The current report is divided into two parts. In Part 1, we introduce statistical procedures to estimate matching thresholds and evaluate these procedures using Monte Carlo simulations. In Part 2, we describe a procedure to evaluate whether matching messages to multiple characteristics at a time outperforms matching messages to only one characteristic. This procedure systematically teases apart the effects of matching to incrementally larger numbers of characteristics from the specific characteristics targeted, thus overcoming limitations of past research. We then use our procedure to evaluate the benefits of simultaneously matching messages to individual differences in *regulatory focus* (the degree to which individuals are promotion- or prevention-focused in their motivational orientations; Higgins, 1997) and *self-construal* (the degree to which individuals adopt an independent or interdependent construal of the self; Markus & Kitayama, 1991). We also use the statistical procedure developed in Part 1 to estimate the matching thresholds for both regulatory focus and self-construal variables.

Part 1: Estimating Matching Thresholds for a Target Characteristic

To determine the matching threshold of a characteristic, a study must evaluate how the effects of messages vary depending on where participants score on a given characteristic. This is most easily achieved by using a *continuous* measure of a characteristic, and *random assignment* to two message conditions. From this design, regression analyses can be used to calculate the linear relationship between the characteristic and persuasion outcomes (e.g., attitudes) for *each* message. This is represented by the simple slopes of the characteristic, conditional on each

message, and is depicted in Figure 1 (left side). The *matching threshold* is operationalized as the point at which the two simple slopes predict equal values. For any score on the characteristic below or above the matching threshold, the predicted value of the outcome favors one message over the other. Mathematically, the matching threshold is expressed as:

$$\mathrm{MT} = \frac{i_1 - i_2}{s_2 - s_1} \tag{1}$$

Where:

- MT is the value of the matching threshold (defined along the measured characteristic);
- i_1 is the intercept and s_1 the simple slope corresponding to the regression line conditional on one level of the message feature (e.g., a gain frame condition), and;
- i_2 and s_2 are the intercept and slope of the regression line conditional on the other level of the message feature (e.g., a loss frame condition).

These values can be easily obtained from fitting linear regressions (see our supplemental materials for detailed instructions).

Equation 1 is a straightforward manner to calculate the matching threshold. However, because this statistic has never been used before, it is important to evaluate how it performs. Consequently, we will use Monte Carlo Simulations to simulate the sampling distribution of the matching threshold. Doing so will allow us to examine whether Equation 1 provides an *unbiased* estimate that does not systematically diverge from the population value. For example, if the matching threshold in a population is equal to 0, the sampling distribution of the estimate given by Equation 1 should be centered around 0. Additionally, we can examine whether Equation 1 provides a *consistent* estimate; that is, whether estimates provided by Equation 1 become

increasingly concentrated around the population value as the number of data points used (e.g., sample size) increases. Finally, we can also examine the *shape* of the sampling distribution afforded by Equation 1 under different conditions (e.g., varying effect sizes, sample sizes, amount of error) to examine whether and when we can assume normality, and gain insight as to when Equation 1 may perform most adequately.

Should Equation 1 provide a reliable and consistent estimator, we may still need additional tools to gain insight into whether a given estimate is useful. Specifically, because point-based estimates tend to vary substantially from sample to sample, often deviating importantly from the population parameters they approximate, tools that give us confidence in how likely an estimate is to represent a population parameter can be invaluable. This report therefore compares two techniques for obtaining confidence intervals (CIs) around matching threshold estimates. The first technique constructs CIs directly around the matching threshold using a nonparametric bootstrapping procedure (e.g., the percentile method to calculate bootstrap CIs; Efron & Tibshirani, 1994). This CI corresponds to the range of values on a characteristic over which two message conditions *do not* perform significantly differently from one another, and for which values outside the CI correspond to a significant advantage of one message over the other. This dynamic is represented in Figure 1 (right side). For example, a CI of -.4 to .6 when matching to promotion focus may indicate that gain frames perform significantly better than loss frames above promotion focus scores of .6, that loss frames perform significantly better under scores of -.4, and that between scores of -.4 and .6, gain and loss frames do not perform significantly differently. The second method to construct CIs around the matching threshold employs the Johnson-Neyman (JN) technique to calculate the range of values along the

characteristic for which two message conditions have significantly different effects from each other (also known as evaluating a "region of significance", or the "floodlight" approach; Johnson & Neyman, 1936; see also Spiller, Fitzsimons, Lynch, & McClelland, 2013). Although the nonparametric bootstrap and the JN technique construct CIs using different rationales, their interpretations in this context are equivalent; they both evaluate the range of values on a characteristic for which two messages do not have a significantly different effect from one another. Importantly, when such CIs perform well, their widths provide insight into how confident we can be in our specific point-estimates of the matching threshold.

To date, neither of these techniques have been used to assess CIs around matching thresholds. Our bootstrap procedure is novel, and although the JN technique has been used to calculate when two messages are significantly different from one another (e.g., Reczek, Trudel, & White, 2018), it has not been used to evaluate matching threshold estimates. Thus, whether and the degree to which these methods perform well for this purpose is unknown.

We address these questions by conducting additional Monte Carlo simulations to compare the relative usefulness of the bootstrap and JN methods to create CIs around matching thresholds. In particular, we pay attention to the *accuracy* of the two techniques by comparing the coverage rate of their CIs; whether 95% and 90% CIs capture the true population value of the matching threshold at least 95% and 90% of the time, respectively. Further, we examine the relative *precision* of the bootstrap and JN techniques by comparing the narrowness of their CIs. Finally, we examine how sampling conditions (e.g., effect sizes, sample sizes) impact the relative performance of the two techniques.

Simulation Study 1: Monte Carlo Simulations of the Sampling Distribution of Matching

Thresholds

We conducted analyses in *R* (R Core team, 2017); we used the *boot* package (Canty & Ripley, 2017) to conduct Bootstrapping, and the *jtools* package (Long, 2018) to calculate JN intervals. Analysis scripts are available at osf.io/c7t3w.

Simulation procedure. To examine the sampling distribution matching thresholds given by Equation 1, we generated data using this formula:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + e \quad (2)$$

Where

- Y is an outcome variable (e.g., willingness to enact a behavior after seeing a persuasive message);
- β_0 is the intercept;
- β_1 is the main effect of a simulated message feature X_1;
- X_1 is the simulated message feature (e.g., *gain* vs. *loss-framed messages*). This is a dichotomous 2-level variable, with each level generated with equal probability;
- β_2 is the main effect of a characteristic X_2;
- X_2 is a set of simulated scores on a characteristic used for matching (e.g., *promotion focus* scores*),* with scores generated as a continuous normally distributed variable with a mean of 0 and SD of 1;
- β_3 represents the matching effect; that is, the interaction between the manipulated message feature (X_1; e.g., gain vs. loss frames) and the characteristic (X_2; e.g., promotion focus). Finally;

- e is an error term made to be normally distributed around a mean of zero and a standard deviation SD_e.

Simulation values for each parameter were systematically varied to examine how the sampling distribution of the matching threshold varies across different conditions researchers may encounter. First, we wished to represent the range of standardized effect sizes reported in the matching literature. Reviews and meta-analyses of message matching effects indicate that most matching effects operate around R^2 = .001 to .010 (Gallagher & Updegraff, 2012; Lustria et al., 2013; Noar et al., 2007), but it is not unusual for larger effect sizes to also be reported, with some effect sizes even reaching R^2 = .5 (Carpenter, 2012). We therefore chose values of β_3 (small = -.3; medium = -.6; or large = -1) and SD_e (low = .5; moderate = 1.5; high = 3), such that the R^2 attributable to the matching effect would vary between .001 and .5, covering a variety of values in between. Since it is common for studies to use sample sizes (N) between 100 and 1,000 (Noar et al., 2007), we simulated samples of N=100, 500, and 1,000. Combined with the values of β_3 and SD_e, these values allowed us to examine conditions in which the statistical power to obtain a statistically significant value of β_3 would vary from very high (approaching 1) to very low (less than .1).

The remaining parameters (β_0, β_1, β_2) usually take a secondary role in matching research and are not frequently interpreted. Consequently, we do not consider the impact of independently altering their values. For simplicity, β_0 was held at a constant value of 0, whereas β_2 took on a constant value defined in relation to β_3, such that interactions were always defined by one positive slope and one negative slope of similar magnitude (e.g., at β_3 = -1, β_2 = .5, the slopes were .5 and -.5). With these constraints in place, we manipulated the value of β_1 to examine how

matching threshold estimates perform when: (a) they are equal to the mean of X_2, or (b) they diverge significantly from the mean of X_2. The matching threshold was either set at 0 by setting $\beta_1 = 0$, at -1 by having $\beta_1 = \beta_3$, or at 1 by having $\beta_1 = -\beta_3$. The above variations led to 81 unique combinations of parameters. For each combination, we simulated 10,000 independent samples, generating a total of 810,000 datasets.

Findings. We provide a sample of our results in Table 1, and the full results are reported in Table S1 of the supplemental materials. We summarize the main patterns here.

First, we discuss the shapes of the sampling distributions. A large portion of the sampling distributions were characterized by high skewness, high kurtosis, and large SDs. Several distributions were also normal, but this generally appeared limited to cases when sampling conditions were conducive to high statistical power (i.e., larger N, larger β_3, lower SD_e). This pattern is expected as conditions associated with lower power can lead the denominator in Equation 1 to approach a value of 0, which would produce extreme positive and negative values. Because of these extreme values, the sampling distribution for Equation 1 may generally follow a Cauchy distribution rather than a normal distribution.

We next explore whether equation 1 provides an *unbiased* estimate of the matching threshold (i.e., ensuring it does not systematically deviate from the true population value from which it is drawn). To do this, we use the median of the distribution as an indicator of bias. When the true matching threshold was set to 0, medians were always very close to 0, across all sampling distributions. However, several cases of biased estimation occurred when the true matching threshold was distant from the mean value of the continuous variable X_2 (i.e., matching thresholds of -1 or 1). Specifically, when statistical power to detect the matching effect (β_3) was

very low (e.g., .5 and lower), the matching threshold appeared increasingly biased towards the mean value of X_2. However, when adequate power was present, there did not appear to be any indication of bias. Taken together, we can conclude that Equation 1 generally provides an *unbiased* estimate of the matching threshold so long as studies make sure to be adequately powered to detect their matching effects of interest (e.g., having power greater than .8). Further, it was also generally true that whenever the sample size and interaction effect size increased, or error decreased, the sampling distribution of the matching effect provided by Equation 1 became increasingly concentrated around the true population value of 0. This provides evidence that Equation 1 is also a *consistent* estimator of the matching threshold.

The evidence indicates that Equation 1 is a promising method of estimating matching thresholds—at least under conditions of repeated sampling. However, because researchers must typically draw inferences and make decisions using single or small numbers of samples, it is also necessary to establish a way to evaluate our certainty in the generalizability of our estimates. The construction of CIs can provide a valuable tool in achieving this goal.

Simulation Study 2: Comparing the Bootstrap and Johnson-Neyman (JN) Procedures to Construct Confidence Intervals (CIs) Around the Matching Threshold

Simulation procedure. In this section, we compare the relative usefulness of the nonparametric bootstrap and the JN procedure for building CIs surrounding the matching threshold. We again used a series of Monte Carlo simulations generating data according to Equation 2, altering sample sizes (N), the size of matching effects (β_3, and β_2), the degree of error present (SD_e), and the location of the matching thresholds (via β_1). Parameter values were similar to Simulation Study 1 and selected so that we could examine the performance of the two

CI building techniques under conditions ranging from very poor (small N, small β_3, large SD_e, matching threshold distant from mean of X_2) to ideal (large N, large β_3, small SD_e, matching threshold equal to mean of X_2). The specific set of parameters we used, as well as more details on our rationale for picking parameters, are available in the supplemental materials.

For each combination of parameters, we randomly generated 1,000 datasets. Then, for each dataset, we conducted nonparametric bootstrapping (using 2,000 bootstrap samples and the percentile method) to calculate 90 and 95% CIs. We also computed JN intervals with alphas of .10 and .05. We deemed 1,000 simulated samples per condition enough, given that regression models are low in complexity, and that the bootstrapping procedure is computationally demanding to simulate. Instead of drawing inferences on any given combination of parameters, we focus our inferences on patterns that emerged across the 31,000 simulated samples. For each combination of parameters, we evaluate the accuracy of the bootstrap and JN techniques by examining their coverage rate (i.e., the percentage of times each technique computed provides intervals that contain the true value of the matching threshold underlying the datasets). We also compared the precision of the two techniques by calculating the average width of the obtained intervals.

Findings from our simulations. We provide a representative sample of our results in Table 2, and our full results in Table S2 of the supplemental materials. Overall, when simulations were generated using sampling properties conducive to high power (i.e. larger samples, larger effects, lower error), the two methods gave very similar CIs. For instance, with N = 1000 [defined as large], β_3 = -1 [defined as large], and SD_e = .5 [defined as low], the average bootstrap and JN intervals were virtually identical and narrowly focused around the true value of

the matching threshold (e.g., average 95% CIs = .062 to .063 for both methods when the true value was 0). This supports the notion that the two methods are estimating equivalent intervals. Under the same conditions, both methods had high accuracy in their coverage rates (e.g., 95.4% of bootstrap CIs and 95.7% of JN intervals captured the true parameter value when alpha was set at .05). Moreover, when two of the manipulated components N, β_3, or SD_e (sample size, effect size, error) were optimal, the two methods gave very similar intervals and coverage rates that closely corresponded to desired alpha levels. However, the bootstrap CIs were slightly more precise/narrow than the JN CIs (see last two columns of Table 2). This pattern was consistent both when the matching threshold was equal to the mean of X_2 (0), or distant from it (1).

When conditions became less optimal (e.g., N = 100, with a large effect, but large error), the bootstrap CIs continued to show good coverage rates, being close to 90% with an alpha of .10, and close to 95% with an alpha of .05. However, several 90% CIs had coverage rates in the range of 94-96% (going as high as 98.9%) and many 95% CIs had an observed coverage rate closer to 98-99%. Overall, this indicates that bootstrap CIs are accurate, but may tend to be conservative. In contrast, JN intervals often fared poorly when evaluated under these conditions. JN coverage rates were unstable and occasionally substantially below acceptable levels. This was especially true when the matching threshold was set away from the mean of X_2 (i.e., at 1 SD from the mean). For instance, we observed a coverage rate as low as 63.6% using JN intervals with an alpha of .05, signaling very low accuracy. Additionally, a substantial proportion of JN CIs had very poor precision, sometimes even extending from negative to positive infinity. Although infinitely wide CIs inflate coverage rates (as they must capture the true value), coverage rates still remained low. For example, in one set of simulations (matching threshold at

1, N =1000, small matching effect size, large error), 43.3% of the JN intervals with an alpha = .10 had an infinitely large interval, but the coverage rate was only 69.6%. In other words, of the CIs that were not infinitely wide, only 46.4% contained the true population value. This indicates JN intervals can have unacceptably low accuracy and precision and be too liberal (prone to false positives).

Discussion of Simulation Results

The success of message matching depends on correctly deciding which message a person should receive. To do this, it is important to know the matching thresholds associated to targeted characteristics. Equation 1 allows researchers to estimate the matching threshold of characteristics directly. Our simulation findings show that Equation 1 provides a generally unbiased and consistent estimator of the matching threshold so long as studies are adequately powered (e.g., .8 or above) to detect a matching effect of interest.

However, because Equation 1 only provides a point estimate of the matching threshold, it conveys little information about the generalizability of findings. Therefore, we also offer two methods to construct CIs around the matching threshold estimate. With CIs, we are looking for methods that produce precise (i.e., narrow) intervals that are highly accurate—that is, that capture a true population parameter at a probability that is commensurate with a desired level (i.e., a 95% CI should contain the true population value at least 95% of the time). When CIs are accurate, the narrower the CI, the more confidence we have that a point estimate represents its corresponding population parameter. Our simulation work provides evidence that both a nonparametric bootstrap and the JN technique each show promise in constructing CIs under well-powered study conditions, with the nonparametric bootstrap affording slightly higher

accuracy and precision. However, under conditions of lower power (e.g., small sample sizes, small matching effects), the precision and accuracy of the bootstrap CIs considerably surpass JN CIs. In fact, the bootstrap CI remains highly accurate even under very underpowered circumstances. Overall, these findings lead us to generally recommend the use of our bootstrapping procedure over the JN technique to calculate CIs for the matching threshold, and to evaluate the point estimate in direct relation to the width of the bootstrap CI. The JN technique should only be used under conditions of high power, when it is expected to converge with the bootstrap CI. With this knowledge in mind, our hope is that researchers can use these tools to estimate and evaluate matching thresholds for a variety of characteristics targeted in message matching interventions.

In Part 2, we provide a demonstration of these tools with two characteristics commonly used in message matching research. We simultaneously explore a second idea for increasing the impact of message matching interventions.

Part 2: Using a Multi-Matched Messaging Procedure to Evaluate the Advantage of Simultaneously Matching Messages to Regulatory Focus and Self-Construal

Estimating matching threshold should improve interventions by helping interventionists decide *which* messages to administer to specific individuals. In Part 2, we examine a second tool for optimizing message matching: matching messages to *multiple characteristics* at once. Specifically, we test whether matching to two characteristics leads to stronger effects than matching to only one characteristic. In doing so, we also demonstrate how matching thresholds can be estimated when matching to multiple characteristics simultaneously.

Because the promise of matching to multiple characteristics assumes there are matching

effects to combine, we focus our empirical study on characteristics that already have a history of eliciting matching effects. Specifically, we evaluate the benefits of matching messages to individual differences in *regulatory focus* (the degree to which individuals are promotion- or prevention-focused in their motivational orientations; Higgins, 1997) and *self-construal* (the degree to which individuals adopt an independent or interdependent construal of the self; Markus & Kitayama, 1991). Regulatory focus and self-construal have each been used successfully in prior message matching work (for a discussion, see Sherman, Uskul, & Updegraff, 2011), but their joint influence has never been documented. Generally, previous research has found that highly promotion-focused individuals respond more favorably to *gain-framed* messages (messages emphasizing positive and/or beneficial outcomes), whereas highly prevention-focused individuals respond more favorably to *loss-framed* messages (messages emphasizing negative and/or costly outcomes; Cesario, Corker, & Jelinek, 2013). For self-construal, messages containing *interdependence*-related appeals have been found to be more persuasive for individuals with a more interdependent self-construal, whereas messages containing *independence*-related appeals are more persuasive for individuals with a more independent self-construal (Lalwani & Shavitt, 2009). Given the past success of studies matching to these characteristics, they are good candidates to examine the benefits of matching messages to two characteristics simultaneously.

Since our interest in matching to multiple characteristics is to strengthen an already effective persuasion strategy (i.e., message matching to single characteristics), we focus on a behavior that many individuals are reluctant to perform. Specifically, we examine the effects of message matching on individuals' willingness to consume edible insects. Edible insects are

highly nutritious and more environmentally sustainable than many other food sources (Abbasi & Abbasi, 2016). For these reasons, in 2013, the United Nations urged Western countries to take up greater consumption of edible insects (Van Huis et al., 2013), but acceptance of *entomophagy*—the practice of eating insects—has continued to be low in the Western world (Abbasi & Abbasi, 2016). Thus, entomophagy is a good candidate for a behavior which people are generally reluctant to adopt.

Hypotheses

We focus on willingness to consume edible insects as our primary outcome and include a measure of attitudes towards the practice as a secondary outcome. Our primary hypothesis of interest is:

H_1: Matching messages to self-construal should provide an *additive effect* over matching messages to an individual's regulatory focus, and vice-versa, such that messages are most persuasive (e.g., lead to higher willingness to eat insects) the more they match both an individual's self-construal *and* regulatory focus.

Because this hypothesis is conditioned on the expectation that matching to regulatory focus and self-construal are each able to induce increased persuasion effects on their own, we first examined whether the following conditions would be met in our data:

C_1: Individuals with stronger promotion focus scores would be more persuaded (i.e., display higher willingness to eat insects) by gain frames than loss frames

C_2: Individuals with stronger prevention focus scores would be more persuaded by loss frames than gain frames

C_3: Individuals with stronger independent self-construal would be more persuaded by

independence appeals than interdependence appeals.

C_4: Individuals with stronger interdependent self-construal would be more persuaded by interdependence appeals than independence appeals.

Our primary hypothesis can be examined as long as at least one regulatory focus effect (C_1 or C_2), and one self-construal effect (C_3 or C_4) are obtained. We do not formulate predictions about what the matching thresholds for regulatory focus and self-construal will be. Rather, our goal is to offer estimates of matching thresholds for these characteristics for future research to build on.

Methods

Multi-Matched Messaging Procedure. To examine the incremental effect of matching messages to multiple characteristic, we utilize and demonstrate an approach we refer to as *multi-matched messaging*. This approach consists of study design requirements complemented with a set of statistical steps that can be used to study matching to any number of characteristics. Our study focuses on the case of matching to two characteristics.

Design Requirement. To overcome limitations of past research, we suggest studies use designs that: (1) do not condition the presence of specific matching effects on the number of matching effects being implemented at a time; (2) ensure each matching effect is independent of each other, and; (3) allow smaller, gradual, increases in the number of matching effects to be examined. Several designs can meet these requirements and allow us to statistically evaluate the benefits of multiple matching. In the current study, we implement a factorial experiment, in which *only* message features are manipulated. Specifically, messages were designed and randomly distributed to participants such that features meant to appeal to any given characteristic (gain and loss frames) were independent of features manipulated to appeal to other

characteristics (independence versus interdependence appeals), and of the characteristics they target (regulatory focus and self-construal scores). This process ensured that each specific matching effect could operate regardless of the number of matching effects achieved (criterion 1), that the two matching effects operated independently from each other (criterion 2), and that participants could experience any degree of each or both matching effects (criterion 3).

Statistical Tests. We suggest that the incremental benefit of matching messages to multiple characteristics can be evaluated in three steps. In Step 1, we test the assumption that there are matching effects to be combined. In our study, this represents evaluating conditions C1-C4. To do this, separate statistical models (e.g., linear regressions) are computed to examine matching effects for *each* specific characteristic. Each matching effect is represented by the interaction term between a measured characteristic, and its corresponding manipulated message feature. For example, a significant interaction between promotion focus and message frame would provide support for one matching effect, whereas a significant interaction between independent self-construal and message appeal would provide support for a second. If statistical support is found for more than one matching effect, we proceed to Step 2.

Step 2 is to evaluate the *presence* of an incremental effect from matching to multiple characteristics. To do so, we evaluate a regression that simultaneously models each matching effect that was significant in Step 1. If each matching effect remains significant, we obtain evidence that matching to an additional characteristic provides an independent benefit over the other. If three or more matching effects were significant in Step 1, Step 2 can be expanded by running models with each pair of matching effects, each triple, and so forth.

In Step 3, we can *quantify* the degree to which matching to a larger number of

characteristics allows us to achieve increased persuasion, while accounting for the unique influence of each specific characteristic. We shift from determining whether an incremental effect exists (Step 2), to whether the incremental effect is sizeable enough to be meaningful. To do this, we examine how model fit (e.g., R^2, adjusted R^2) improves by modelling to a larger number of matching effects. When matching to two characteristics, we compare (a) a model containing all main effects of message features and characteristics; (b) models containing a single matching effect at a time, and; (c) a model containing both matching effects. Comparing models *a* and *b* evaluates the degree to which matching to one characteristic accounts for greater variance than matching to zero characteristics, accounting for all main effects. Comparing model *b* to each model in *c* then evaluates the degree to which matching to two characteristics accounts for more variance than matching to either characteristic by itself. Crucially, the pattern of results matters. For instance, we may find that the two matching effects complement each other equally (i.e., that matching to both is simply twice *each* individual match), which would be strong evidence in favor of combining them. Alternatively, we may find that one of the two effects drives most of the benefits. In such a case, whether we promote matching to both characteristics may depend on whether the smaller effect provides enough increased persuasion over the larger effect to be worthwhile. As with Step 2, Step 3 can be expanded to consider more than 2 matching effects.

Sample Recruitment and Demographics. Because effect sizes in message matching research vary considerably (Carpenter, 2012; Noar et al., 2007), we powered our study to detect small to moderate effect sizes. A power analysis for a 2x2 factorial design using *G*Power 3.1* (Faul, Erdfelder, Buchner, & Lang, 2009) suggested that a sample of around 550 individuals

would have a power of 0.92 to detect a medium effect (d = .3), and 0.78 for a small effect (d = .2), using an alpha level of 0.05. To ensure this sample size, we recruited 600 individuals using Amazon's Mechanical Turk to complete an online survey between January and February, 2016. We excluded thirty-two responses for not completing the full survey. All complete responses (N=568) were used in our analyses. Participants were predominantly female (56.0%), white/Caucasian (81.8%), with an average age of 39.8 years (SD = 13.6). Table 3 presents additional demographics. The study was approved by the institutional review board of the University of Minnesota, and all participants provided informed consent prior to participation.

Study Procedure. At the beginning of the survey, participants completed measures of regulatory focus and self-construal[2]. Scale order and the ordering of items within scales were randomized. Participants were then randomized to one of four persuasive message conditions using a computer algorithm. Following exposure to a persuasive message, participants completed a measure of willingness to eat various foods, including insect-based dishes, and a measure of attitudes towards eating insects. At the end of the study, participants completed demographics questions, and were then debriefed. Participants were blind to their experimental condition and to our hypotheses until debriefing.

Experimental Conditions. The experimental conditions were presented as six slides containing information promoting entomophagy. Each condition varied in the degree that it used gain-framed versus loss-framed messages, and interdependence versus independence appeals, for a 2 (Frame: Gain, Loss) by 2 (Appeal: Interdependence, independence) between-person factorial design. Table 4 summarizes strategies that were used to manipulate both message frame and appeal type. In total, 140 participants were randomized to the gain-framed interdependence

appeal message, 142 to the loss-framed interdependence appeal message, 145 to the gain-framed independence appeal message, and 141 to the loss-framed independence appeal message.

Measures.

Our supplemental files report the items making up the measures listed below, along with detailed results of confirmatory factor analyses (CFAs; Tables S3-S6) and Item Response Theory analyses evaluating each measure (see Tables S7-S16).

Regulatory focus. Regulatory focus was assessed using a measure developed by Fuglestad, Rothman, DeYoung, and Cunningham (2014; available at: osf.io/95gwn/) in response to criticisms of construct ambiguity in commonly used regulatory focus measures (Summerville & Roese, 2008). The measure was developed in consultation with Regulatory Focus experts who identified 30 items from the International Personality Item Pool (Goldberg, 1999) to assess both prevention and promotion focus. Fuglestad et al. (2014) analyzed the items using Item Response Theory, and 9 items were selected to reflect each of prevention focus and promotion focus. Each item presents participants with a descriptive statement (e.g., “am careful to avoid making mistakes”, “eagerly look forward to things to come”) to which participants indicate the degree to which the statement describes them (1=Strongly disagree; 4=Neither agree nor disagree; 5=Strongly agree). The authors report high correlations between their measure of promotion focus with the scales contained in the Regulatory Focus Questionnaire (RFQ; Higgins et al., 2001) and the Regulatory Focus Scale (Lockwood, Jordan, & Kunda, 2002), and the prevention scale was also moderately correlated with its counterpart in the RFQ, but not to the Lockwood et al. (2002) prevention focus measure (Fuglestad et al., 2014). In our study, we found Cronbach alphas of 0.86 for promotion and 0.84 for prevention focus. Further, a CFA confirmed each item

loaded significantly on its respective factor.

Self-construal. Independent and interdependent self-construal were assessed using Singelis' (1994) Self-Construal Scale, a 24-item instrument with two 12-item subscales; one for independent self-construal, and the other for interdependent self-construal. Items consists of statements (e.g., "my personal identity independent of others, is very important to me") to which participants indicate their agreement (1=Strongly disagree; 4=Neither agree nor disagree; 7=Strongly agree). In this study, we found Cronbach alphas of 0.85 for interdependent self-construal, and 0.82 for independent self-construal. A CFA indicated each item loaded significantly on its respective factor.

Willingness and attitudes toward eating insects. We developed an 8-item measure of willingness to eat insects and embedded it within a larger 30-item survey asking participants about their willingness to eat various foods on a 9-point Likert format (1 = "Extremely unwilling under any circumstance"; 5="Neither willing nor unwilling"; 9 = "Extremely willing under any circumstance"). Example insect-based food items include: dry roasted crickets, fried black ants, and fried rice with beetle larvae. The endpoints of the scale were labelled to reflect extreme positions as a small pilot test of the measure indicated that participants otherwise tended towards picking only the lowest response option for insect-based items. The reliability coefficient for this scale was 0.96. Scores were positively skewed, so a log transformation was applied.

Finally, we included two scales to assess participants' attitudes towards entomophagy. Ten items measured how participants felt about others eating insects using various terms (e.g., good, bad, risky, healthy), and ten items assessed how they felt about eating insects themselves (using the same adjectives). Although participants rated others eating insects slightly more

positively, the scales were highly correlated (r = .82) and were averaged into a single index (reliability alpha = 0.93).

Results

Correlations between each measured variables' raw scores are found in Table 5, along with means, medians, SDs, and skewness statistics. Before running inferential statistics, we standardized each variable into z-scores. The deidentified data and necessary script files to reproduce our analyses are available at osf.io/c7t3w. In presenting our results, we focus on providing 95% confidence intervals (CIs) over p-values to reinforce the notion that our findings are initial estimates to be refined by future research. To aid meta-analytic efforts, our supplemental files (Tables S17 to S20) provide additional correlation matrices and descriptive statistics, broken down by each message condition.

Step 1. Examining Matching Effects to Each Characteristics Individually (C_1 to C_4). We ran four regression models to examine whether message matching would increase willingness to eat insects using each of our measured characteristics individually. These regressions are summarized in Table 6 and depicted in Figure 2. Each interaction term represents a unique message matching effect. Overall, no interaction effect was observed between prevention focus and message frame, nor between independent self-construal and message appeal. However, we obtained a significant interaction between message framing and promotion focus (β= -.22, t(564) = -2.59, 95% CI [-.38, -.05]), such that for individuals in the loss frame condition, higher scores in promotion focus led to lower willingness (β= -.15, t(564) = -2.32, 95% CI [-.27, -.02]), whereas in the gain frame condition, higher promotion focus scores were positively, albeit non-significantly, related to willingness (β= 0.07, t(564) = 1.29, 95% CI [-.04,

.18]). We also obtained a significant interaction effect between message appeal and interdependent self-construal (β= .20, t(564) = 2.36, 95% CI [.03, .36]). For individuals who received the independence appeal, a higher interdependent self-construal score was significantly associated to lower willingness (β=-.12, t(564) = -2.05, [-.24, -.004]), whereas for those exposed to the interdependent appeals, a higher interdependent self-construal score was positively associated with willingness, although this effect was not significant (β= .07, t(564) = 1.27, 95% CI [-.04, .19]).

We also examined matching effects on attitudes towards eating insects. Each effect was in the same direction as with willingness, but no significant interaction effect was observed (see Table S21 of the supplemental files).

Step 2. Examining the Additive Effect of Matching Messages to Two Characteristics at a Time. Having found matching effects for both promotion focus and interdependent self-construal, we tested our primary hypothesis by simultaneously regressing willingness scores onto: (1) promotion focus; (2) message frame; (3) a promotion focus by message frame interaction; (4) interdependent self-construal; (5) message appeal, and; (6) an interdependent self-construal by message appeal interaction. In this model, summarized in Table 7, both interaction terms remained significant. Moreover, their magnitudes remained virtually unchanged from those obtained in step one (Table 6), offering evidence that these interactions show incremental effects over one another. The results are represented graphically in Figure 3.

Step 3. Quantifying the Effectiveness of Matching Messages to Two Characteristics at a Time. To examine the relative impact and usefulness of matching to two characteristics at a time, we examined the relative fit of models containing (1) all main effects [i.e., of message

frame, message appeal, promotion focus scores, and interdependent SC scores]; (2) all main effects and the matching effect to promotion focus; (3) all main effects and the matching effect to interdependent SC, and; (4) all main effects and the two matching effects. For each model, we extracted the R^2, the adjusted R^2, Akaike's information criterion (AIC), and the Bayesian information criterion (BIC). Fit indices are presented in Table 8. According to the R^2, adjusted R^2, and AIC, models with one matching effect were similar to one another and were preferable to a model with no matching effect; further, the model with the two matching effects performed best. Similar BIC values were obtained for all models. When examining R^2, on average, the models with 1 matching effect accounted for 2.76 times the variance of the model with no matching effect. The model with two matching effects, accounted for 1.62 times the variance of the average model containing only one matching effect. If we look at the adjusted R^2, the model with both matching effects accounted on average for 2.09 times the variance of the models with only one matching effect. These results indicate that the two matching effects contribute similarly to the overall additive effect, and that their combined effect represents a sizable increase in variance explained compared to each individual matching effect.

Calculating the Matching thresholds. Matching thresholds were calculated using the Step 1 model using a nonparametric bootstrap (with 100,000 samples), and the percentile method to derive 95% CIs. The matching threshold for the *promotion focus by message frame* interaction was -.38 (95% CI [-2.23, .51]) SDs from the mean. In terms of raw promotion focus scores, this places the matching threshold at 3.36 (95% CI [2.01, 4.01]) on the 5-point promotion focus measure. Using a region of significance interpretation of the CI, we can say that for raw promotion focus scores below 2.01, there is a significant advantage of loss-framed messages,

whereas for scores above 4.01, there is a significant advantage of gain-framed messages.

The matching threshold for the *interdependent self-construal by message appeal interaction* was -.58 (95% CI [-3.38, .39]) SDs from the mean. In raw scores, this represents a score of 4.20 (95% CI [1.48, 5.06]) on the 7-point interdependent self-construal measure. Using a region of significance interpretation, we can say that for raw interdependent self-construal scores below 1.48, there is a significant advantage of independence appeals, whereas for scores above 5.06, there is a significant advantage of independence appeals.

Matching thresholds when considering both matching effects simultaneously. The matching thresholds we calculated above correspond to and can be used when matching to either promotion focus or interdependent self-construal in isolation from one another. However, how might matching thresholds change if we match to both characteristics simultaneously?

To examine this, we re-estimated matching thresholds using the Step 2 model (matching thresholds were again calculated using Equation 1; however, see our supplemental files for how each intercept and slope were extracted from the Step 2 regression model). We used nonparametric bootstrapping (with 100,000 resamples) to construct 95% CIs around the estimated values.

The standardized results of our analyses are presented in Table 7. Expressed in raw scores, the matching threshold for the promotion focus by message frame interaction was 3.32 (95% CI [1.88, 3.97]), and the matching threshold for the interdependent self-construal by message appeal interaction was 4.13 (95% CI [1.32, 5.06]). These values are highly similar to those in the Step 1 model; however, they can be used to make observations and inferences regarding people's *simultaneous* preference for message frame and appeal type. For instance, we

can descriptively say that people with raw scores above 3.32 in promotion focus *and* above 4.13 in interdependent self-construal responded more favorably to a gain-framed interdependence appeal than to any other message frame by appeal combination. Using the region of significance interpretation of the CIs, we can further say that scores above 3.97 in promotion focus paired with scores above 5.06 in interdependent self-construal were associated to significantly higher willingness scores in the gain-framed interdependence appeal condition, than in any of the three other message conditions. Similar inferences can be made for other combinations of the matching thresholds and their CIs.

Robustness of Our Matching Effects. In addition to the main results we report above, we conducted analyses to examine the robustness of our findings. We summarize these analyses here. For a full report, see our supplemental materials.

First, we examined whether adding covariates to our Step 2 model altered our inferences, and it did not. In a fairly stringent test, we controlled for prevention focus, independent self-construal, all of the demographic variables in Table 3 (gender, known allergies, dietary restrictions, ethnicity, education, employment status, and family income), as well as the interaction between each demographic variables and both message frame and message appeal. However, the matching effects of interdependent self-construal and promotion focus remained largely unchanged.

Second, although our hypotheses only specified that matching to regulatory focus and self-construal should have additive effects, it remains possible that the matching effects interact with one another, or that their combined effect is non-linear. We examined this possibility in an exploratory fashion using a smoothing spline ANOVA model. The results of these analyses,

along with a discussion of the results, are reported in full in our supplemental materials. The results suggest that the two matching effects interact with one another, but the general pattern of effects for who responded most positively to each message condition was similar to our Step 2 model. Therefore, our general pattern of findings held.

General Discussion

Matching the content of messages to a person's personal characteristics is a frequently used strategy to increase the persuasiveness of interventions. However, despite the wide use of message matching strategies, principles to optimize matching have yet to be fully delineated.

The current work describes two principles that can be used to improve message matching. The first is to match the content of messages to multiple personal characteristics at a time. The second is to improve the process by which interventions allocate message conditions through the estimation of a *matching threshold*—the point along a characteristic where people shift from holding a preference for one message type to another.

Optimizing Interventions by Matching to Multiple Characteristics at a Time

We found evidence that matching messages to both promotion focus and interdependent self-construal led to an additive effect over matching to either characteristic alone. Specifically, after accounting for the main effects of either characteristics, the two matching effects had effects sizes of $R^2 = .012$ (for promotion focus) and $R^2 = .010$ (for interdependent self-construal). When taken together, the benefits of matching messages to both characteristics simultaneously was around $R^2 = .022$; that is, twice the magnitude of matching to *either* characteristic alone.

To put these effects into perspective, we note the average matching effect reported in meta-analyses is usually between $R^2 = .001$ and .010 (Gallagher & Updegraff, 2012; Lustria et

al., 2013; Noar et al., 2007; O'Keefe & Jensen, 2006). Comparatively, the singular matching effects we report are on the larger side of this range, and the additive effect of matching to our two characteristics was greater than twice the magnitude of the average reported matching effect.

Matching messages to multiple characteristics appears to be a promising strategy, but the relative benefits need to be further documented, as our study was the first systematic test of the incremental benefit of matching to a single additional characteristic at a time (i.e., 2 over 1). Future research will need to confirm our findings regarding promotion focus and interdependent self-construal, and will benefit from exploring matching to other characteristics, as well as to larger combinations of characteristics. We intuit that there is an optimal number of characteristics interventionists ought to target, such that matching to more than two characteristics will continue to accrue benefits, but that these benefits will eventually level off. It is even possible that matching to too many characteristics could backfire if it leads participants to feel their privacy has been violated (as matching procedures can increase feelings of intrusiveness: e.g., van Doorn & Hoekstra, 2013). Consequently, finding the optimal number of characteristics to target will be a key endeavor for future research.

Optimizing Interventions by Estimating Matching Thresholds

In addition to optimizing interventions by matching messages to multiple characteristics, we suggest that research should consider more carefully the decision rules employed to determine which messages people should receive. Currently, the most common method is to choose a point along the target characteristic (usually the median), and assign messages depending on whether a person scores above or below that point. Implicit in this procedure, is that the chosen point ought to approximate the *matching threshold* of the target characteristic—

the point along the characteristic at which people shift from preferring one type of message to another. To date, the literature has not provided guidance on how to approximate the matching threshold, and interventionists are left to rely on either intuition or arbitrary norms. This creates a problem, as the extent to which a chosen point deviates from the actual value of the matching threshold, interventions will systematically misidentify which message people should receive. Our work addresses this gap by providing, validating, and demonstrating a procedure to directly estimate matching thresholds.

In Part 1, we showed that Equation 1 provides a meaningful, generally unbiased, and consistent estimator of the matching threshold. We then evaluated two ways of building CIs around the estimates and found that a nonparametric bootstrap procedure had both high accuracy and precision relative to the JN method.

In Part 2, we used Equation 1 with a bootstrap procedure to evaluate the matching thresholds for promotion focus and interdependent self-construal and found that the two characteristics had matching thresholds at -.38 and -.58 SDs below their respective means. Interestingly, these values are close to the midpoints of the measure we used and correspond closely to the response “neither agree nor disagree”. Theoretically, such midpoints are intuitive locations for a matching threshold, as they represent points of semantic shift in meaning; it seems logical that that as people shift from agreeing to disagreeing with values, their preferences for certain messages would shift accordingly. In contrast, the median values we observed reflect small-to-moderate levels of agreement, a less intuitive point for a threshold. That said, we note that our matching thresholds estimates had relatively large CIs, which encompassed the median/mean scores of the scales. Consequently, the above pattern may only be a coincidence.

Future studies will be needed to further validate and narrow our estimates, as well as to examine the degree to which our results generalize beyond our specific sample.

Our simulation studies in Part I suggest that although CIs around matching thresholds are very likely to capture true population values, the point estimates of matching thresholds are expected to vary substantially across samples. Further, much like matching effects themselves, matching thresholds are likely to depend on factors such as the specific characteristics targeted, the specific measure used to assess the characteristic, the way in which message components are manipulated, among other factors. Consequently, when choosing to use certain estimates for matching thresholds, interventionists will need to carefully evaluate the precision of past estimates as well as the correspondence between the context in which past estimates were derived and the context of the interventions they wish to implement.

In terms of the CIs themselves, we note that these carry additional meaningful information, especially when interpreted as *regions of significance;* regions lying outside the CIs represent scores along the characteristic where we are more certain that differences in preferences exist. Interpreting our bootstrap CIs in this manner is equivalent to the typical use of JN intervals (Spiller et al., 2013), but benefits from an increase in both accuracy and precision.

Using Our Principles to Identify Target Characteristics for Matching Messages

One of the most important decision message matching interventionists must make is choosing *which* characteristic(s) to target. For example, should one target personality traits and motivational orientations (e.g., Goldberg, 1990; Stevens & Fiske, 1995), or a person's behavioral stage (e.g., Prochaska, 2008)? To date, the only criteria researchers and interventionists have explicitly discussed is the magnitude of matching messages to specific types of characteristics a

time. For instance, it appears that stronger effects may be obtainable when matching to constructs explicitly reflecting motivational preferences (e.g., self-monitoring, attitude bases as utilitarian vs. social-adjustive: Carpenter, 2012) than to non-explicitly motivational constructs (e.g., ethnicity, stage of change: Noar et al., 2007). Yet, this evaluation relies on comparisons across rather than within studies and may be confounded with research designs typically used to target different types of characteristics.

In addition to the criteria above, we suggest that the ideas from the current work can be used to derive additional criteria as well. First, based on our finding that matching to multiple characteristics can viably increase message matching effectiveness, we argue that interventionists can shift their focus from considering the magnitude of matching to specific characteristics, to considering whether certain *combinations* of matching effects are more effective than others. For example, characteristics that lead to additive matching effects—e.g., promotion focus and interdependent self-construal as they operate in our study—are arguably better targets for interventions than characteristics that fail to provide incremental benefits over one another. Characteristics that independently lead to strong matching effects (e.g., explicitly motivational characteristics), are likely good candidates for achieving strong additive effects.

Alternatively, we can consider whether certain matching effects interact with one another. This could take the form of a synergistic (i.e., matching to X increases the benefits of matching to Y, and vice versa), attenuating (i.e., matching to X decreases the benefits of matching to Y, and vice-versa), or non-linear interaction (see our supplemental analyses). For example, imagine an interventionist promotes volunteerism by matching gain- and loss-framed messages to promotion focus scores. The volunteerism literature suggests individual differences

in what outcome domains people weigh more heavily. For example, one person may weigh career-related outcomes, but not interpersonal outcomes such as making friends (Clary et al., 1998). Consequently, it may be that matching to promotion focus elicits benefits only to the extent a message also matches outcome domains a person cares about, contributing to a synergistic pattern. To date, theories do not make specific predictions regarding these possible interactions (Rothman, Joyal-Desmarais, Lenne, in press), and it may be relatively early to assume a given form underlies the interaction between any two traits. Given this, along with practical limitations inherent to studying higher-order interactions, researchers may be limited in the extent to which they can currently make reliable and generalizable inferences about interactions. Therefore, we recommend that researchers engage in a first phase of exploratory research using non-parametric and/or descriptive analyses to document these interactions and inform future theory.

Second, we suggest that choosing characteristics can be informed by the *location* of matching thresholds. Overall, the usefulness of message matching depends on the idea that a population can be divided into large subgroups that differ in the messages to which they respond best. Ideally, one may want a matching threshold close to a population's median score on a given characteristic as this indicates that half of people prefer one message type, and the other half to prefer another. In contrast, if the matching threshold lies far away from the median score (e.g., at the 90th percentile), this indicates that the majority of people share the *same* preference. In such a situation, developing alternate messages based on that characteristic is likely to have low returns relative to using a single generic message designed to appeal to most people. As studies on matching thresholds accumulate, it will become possible to use this criterion to identify when to

use message matching, and what kinds of characteristics to target for specific populations.

Evaluating Our Singular Matching Effects

Although our discussion has centered on the two principles we outlined to improve message matching, we close with a note on the specific matching effects we observed. Although we observed matching effects for promotion focus and interdependent SC, we did not find corresponding effects for prevention focus or independent SC. This was surprising and may in part be owing to our use of a relatively new regulatory focus measure or due to sampling variation. However, it is important to note that we treated promotion and prevention focus, as well as independent and interdependent SC, as four distinct dimensions in our analyses. This conceptualization is consistent with theory and the psychometric properties of measures to assess the constructs (e.g., Fuglestad et al., 2014; Higgins et al., 2001, Singelis, 1994), but past research often simplifies the dimensions by using difference scores (e.g., subtracting prevention from promotion focus to get an overall regulatory focus score; e.g., Chang, 2017; Han, Park & Khang, 2018). Using difference scores unfortunately confounds differences between the constructs, making it difficult to evaluate how much our findings deviate from past studies.

Conclusion

Message matching is a popular method to increase the efficacy of persuasive messaging, but knowledge of principles by which to optimize this technique continues to be sparse. We delineate a procedure to more effectively study the principles underlying message matching techniques, focusing on the estimation of matching thresholds—the point along a characteristic when optimal messages shift from one type to another—and on examining the impact of matching messages to multiple characteristics at a time. In an initial test of our framework with

regulatory focus and self-construal measures, we find matching threshold nearer to the midpoint of our scales (i.e., “neither agree nor disagree”) than to the mean/median (which reflect some agreement). We also find that matching messages to both promotion focus and interdependent self-construal provides incremental benefits to matching messages to either personality characteristic alone.

References

Abbasi, T., & Abbasi, S. A. (2016). Reducing the global environmental impact of livestock production: the minilivestock option. *Journal of Cleaner Production, 112*, 1754-1766.

Canty, A., & Ripley, B. (2017). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-20

Carpenter, C. J. (2012). A meta-analysis of the functional matching effect based on functional attitude theory. *Southern Communication Journal, 77*(5), 438-451. doi: 10.1080/1041794X.2012.699989

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319. doi:10.1037/0022-3514.67.2.319

Cesario, J., Corker, K. S., & Jelinek, S. (2013). A self-regulatory framework for message framing. *Journal of Experimental Social Psychology, 49*(2), 238-249. doi:10.1016/j.jesp.2012.10.014

Chang, C. (2009). Enhancing the effectiveness of antismoking messages via self-congruent appeals. *Health Communication, 24*(1), 33-40.

Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., Stukas, A. A., Haugen, J., Miene, P. (1998). Understanding and assessing the motivations of volunteers: a functional approach. *Journal of Personality and Social Psychology, 74*(6), 1516-1530. doi:10.1037/0022-3514.74.6.1516

Covey, J. (2014). The role of dispositional factors in moderating message framing effects. *Health Psychology, 33*(1), 52.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*: New York, NY: CRC

press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. doi: 10.3758/BRM.41.4.1149

Fuglestad, P. T., Rothman, A. J., DeYoung, C. G., & Cunningham, W. A. (2014). *Using the international personality item pool to create a measure of regulatory focus*. Poster presented at the Annual Meeting of the Society of Southeastern Social Psychologists (SSSP), Athens, GA. Available at: osf.io/95gwn/

Gallagher, K. M., & Updegraff, J. A. (2012). Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. *Annals of Behavioral Medicine, 43*(1), 101-116. doi:10.1007/s12160-011-9308-7

Goldberg, L. R. (1990). An alternate "description of personality": the Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216-1229. doi:10.1037/0022-3514.59.6.1216

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe, 7*(1), 7-28.

Han, E. K., Park, C., & Khang, H. (2018). Exploring linkage of message frames with personality traits for political advertising effectiveness. *Asian Journal of Communication, 28*(3), 247-263.

Hawkins, R. P., Kreuter, M., Resnicow, K., Fishbein, M., & Dijkstra, A. (2008). Understanding tailoring in communicating about health. *Health Education Research, 23*(3), 454-466.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280. doi:10.1093/her/cyn004

Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: promotion pride versus prevention pride. *European Journal of Social Psychology, 31*(1), 3-23. doi:10.1002/ejsp.27

Huang, Y., & Shen, F. (2016). Effects of cultural tailoring on persuasion in cancer communication: a meta-analysis. *Journal of Communication, 66*(4), 694-715. doi:10.1111/jcom.12243

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs, 1*, 57–93.

Kidwell, B., Farmer, A., & Hardesty, D. M. (2013). Getting liberals and conservatives to go green: Political ideology and congruent appeals. *Journal of Consumer Research, 40*(2), 350-367. doi:10.1086/670610

Lalwani, A. K., & Shavitt, S. (2009). The “me” I claim to be: cultural self-construal elicits self-presentational goal pursuit. *Journal of Personality and Social Psychology, 97*(1), 88. doi:10.1037/a0014100

Lockwood, P., Jordan, C. H., & Kunda, Z. (2002). Motivation by positive or negative role models: regulatory focus determines who will best inspire us. *Journal of Personality and Social Psychology, 83*(4), 854. doi:10.1037/0022-3514.83.4.854

Long J. A. (2018). jtools: Analysis and Presentation of Social Scientific Data. R package version 1.1.0.

Lueck, J. A. (2018). Respecting the 'stages' of depression: Considering depression severity and readiness to seek help. *Patient Education and Counseling, 101*(7), 1276-1282. doi:10.1016/j.pec.2018.02.007

Lustria, M. L. A., Noar, S. M., Cortese, J., Van Stee, S. K., Glueckauf, R. L., & Lee, J. (2013). A meta-analysis of web-delivered tailored health behavior change interventions. *Journal of Health Communication, 18*(9), 1039-1069. doi:10.1080/10810730.2013.768727

Maio, G. R., & Olson, J. M. (Eds.). (2000). *Why we evaluate: Functions of attitudes*. Mahwah, NJ: Lawrence Erlbaum Associates.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224-253. doi:10.1037/0033-295X.98.2.224

Noar, S. M., Benac, C. N., & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin, 133*(4), 673-693. doi:10.1037/0033-2909.133.4.673

Noar, S. M., Harrington, N. G., Van Stee, S. K., & Aldrich, R. S. (2011). Tailored health communication to change lifestyle behaviors. *American Journal of Lifestyle Medicine, 5*(2), 112-122. doi:10.1177/1559827610387255

O'Keefe, D. J. (2013). The relative persuasiveness of different message types does not vary as a function of the persuasive outcome assessed: Evidence from 29 meta-analyses of 2,062 effect sizes for 13 message variations. *Annals of the International Communication Association, 37*(1), 221-249. doi:10.1080/23808985.2013.11679151

O'Keefe, D. J., & Jensen, J. D. (2006). The advantages of compliance or the disadvantages of noncompliance? A meta-analytic review of the relative persuasive effectiveness of gain-

framed and loss-framed messages. *Annals of the International Communication Association, 30*(1), 1-43. doi:10.1080/23808985.2006.11679054

Payne, C., Megido, R. C., Dobermann, D., Frédéric, F., Shockley, M., & Sogari, G. (2019). Insects as food in the global north–the evolution of the entomophagy movement. In *Edible insects in the food sector* (pp. 11-26). Springer, Cham.

Prochaska, J. O. (2008). Decision making in the transtheoretical model of behavior change. *Medical Decision Making, 28*(6), 845-849. doi:10.1177/0272989x08327068

R Core Team. R: A language and environment for statistical computing [computer program]. Vienna, Austria: R Foundation for Statistical Computing, 2017.

Reczek, R. W., Trudel, R., & White, K. (2018). Focusing on the forest or the trees: How abstract versus concrete construal level predicts responses to eco-friendly products. *Journal of Environmental Psychology, 57*, 87-98. doi:10.1016/j.jenvp.2018.06.003

Rothman, A., Joyal-Desmarais, K., Lenne, R. (in press). Moving from research on message framing to principles for message matching. In Andrew Elliott (Ed.), *Advances in Motivation Science*. Academic Press: Elsevier.

Shavitt, S., & Nelson, M. R. (2002). The role of attitude functions in persuasion and social judgment. In J. P. Dillard, & L. Shen (Eds.), *The SAGE handbook of persuasion: developments in theory and practice* (2 ed., pp. 137-153). Thousand Oaks, CA: SAGE Publications.

Sherman, D. K., Uskul, A. K., & Updegraff, J. A. (2011). The role of the self in responses to health communications: a cultural perspective. *Self and Identity, 10*(3), 284-294. doi:10.1080/15298868.2010.517029

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin, 20*(5), 580-591. doi:10.1177/0146167294205014

Snyder, M., & DeBono, K. G. (1985). Appeals to image and claims about quality: understanding the psychology of advertising. *Journal of Personality and Social Psychology, 49*(3), 586. doi:10.1037/0022-3514.49.3.586

Spiller, S. A., Fitzsimons, G. J., Lynch Jr, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research, 50*(2), 277-288. doi:10.1509/jmr.12.0420

Stevens, L. E., & Fiske, S. T. (1995). Motivation and cognition in social life: a social survival perspective. *Social Cognition, 13*(3), 189-214. doi:10.1521/soco.1995.13.3.189

Strecher, V. J., McClure, J. B., Alexander, G. L., Chakraborty, B., Nair, V. N., Konkel, J. M., . . . Wiese, C. J. (2008). Web-based smoking-cessation programs: results of a randomized trial. *American Journal of Preventive Medicine, 34*(5), 373-381. doi:10.1016/j.amepre.2007.12.024

Summerville, A., & Roese, N. J. (2008). Self-report measures of individual differences in regulatory focus: a cautionary note. *Journal of Research in Personality, 42*(1), 247-254. doi:10.1016/j.jrp.2007.05.005

Tangari, A. H., & Smith, R. J. (2012). How the temporal framing of energy savings influences consumer product evaluations and choice. *Psychology and Marketing, 29*(4), 198-208. doi:10.1002/mar.20515

Updegraff, J. A., & Rothman, A. J. (2013). Health message framing: moderators, mediators, and

mysteries. *Social and Personality Psychology Compass, 7*(9), 668-679. doi:10.1111/spc3.12056

van Doorn, J., & Hoekstra, J. C. (2013). Customization of online advertising: The role of intrusiveness. *Marketing Letters, 24*(4), 339-351. doi:10.1007/s11002-012-9222-1

Van Huis, A., Van Itterbeeck, J., Klunder, H., Mertens, E., Halloran, A., Muir, G., & Vantomme, P. (2013). *Edible insects: future prospects for food and feed security*. Food and Agriculture Organization of the United Nations.

Voelkel, J. G., & Feinberg, M. (2018). Morally reframed arguments can affect support for political candidates. *Social Psychological and Personality Science, 9*(8), 917-924. doi:10.1177/1948550617729408

Werth, L., & Foerster, J. (2007). How regulatory focus influences consumer behavior. *European Journal of Social Psychology, 37*(1), 33-51. doi:10.1002/ejsp.343

Footnotes

[1]There may be exceptions, such as matching to a purely categorical variable (e.g., a name). In this report, we focus only on the case of matching to theoretically continuous variables.

[2]Behavioral Inhibition and Activation System scales (BIS/BAS; Carver & White, 1994) were also administered. However, due to a clerical mistake, the BAS scale was incomplete, and we did not analyze these variables. Additionally, participants completed a measure of food-related values, but we do not discuss it, as it was not used to operationalize any message matching effect. Apart from these omissions, we report all measures and manipulations in our study.

Table 1

Sample Results of Monte Carlo Simulations of the Sampling Distribution of the Matching Threshold Parameter.

Parameters[a]				Results from 10,000 Monte-Carlo Simulations					
Sample Size (N)	Effect Size[b]	Error (SD_e)[c]	True Matching Threshold	Median (Bias)[d]	SD	Skew	Kurtosis	% tests significant at 95%	Matching Effect (partial R^2)[f]
100	Small	Low	0	0.0	1.0	28.8	1775.8	81.7	0.089
1000	Small	Low	0	0.0	0.1	0.0	0.1	100.0	0.083
100	Large	Low	0	0.0	0.1	0.0	0.1	100.0	0.495
1000	Large	Low	0	0.0	0.0	0.0	0.0	100.0	0.499
100	Small	High	0	0.0	151.6	-51.1	5170.5	6.0	0.013
1000	Small	High	0	0.0	192.8	87.2	8134.3	16.2	0.004
100	Large	High	0	0.0	65.6	-82.0	7866.4	37.0	0.037
1000	Large	High	0	0.0	0.2	0.1	0.7	99.9	0.028
100	Small	Low	1	1.0	3.7	-41.4	2571.9	81.6	0.089
1000	Small	Low	1	1.0	0.2	0.5	0.7	100.0	0.083
100	Large	Low	1	1.0	0.2	0.6	0.7	100.0	0.494
1000	Large	Low	1	1.0	0.1	0.2	0.1	100.0	0.500
100	Small	High	1	0.2	173.2	88.1	8483.1	7.8	0.013
1000	Small	High	1	0.6	83.1	-65.7	4841.9	15.7	0.002
100	Large	High	1	0.9	79.1	2.7	3547.3	36.6	0.037
1000	Large	High	1	1.0	0.3	1.2	3.7	100.0	0.023

Note. See Table S1 for full results of simulations.

[a]Monte Carlo simulations were generated to follow the form outlined by $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + e$, where e is normally distributed with a mean of 0, and a standard deviation of SD_e.

[b]This is the effect size corresponding to β_3. The matching effect size was manipulated to be small ($\beta_2 = .125; \beta_3 = -.3$), or large ($\beta_2 = .5; \beta_3 = -1$)

[c]Error was manipulated to be either low ($SD_e = .5$) or high ($SD_e=3$)

[d]These are the mean and median of each sampling distribution comprised of 10,000 randomly generated datasets. Deviations from the true matching threshold (0 or 1) was indicative of bias.

[e]This column represents the range of matching thresholds obtained in each set of 10,000 replications containing 95% of generated matching thresholds.

[f]This column represents the average *partial* R^2 value corresponding to the β_3 regression-based parameter estimate across simulations

Table 2

Simulations Comparing the Performance of Bootstrap Confidence Intervals (CIs) and Johnson-Neyman (JN) Intervals.

Parameters (1000 simulated datasets per combination)				Bootstrap CIs (2,000 bootstrap samples per dataset)		Johnson-Neyman (JN) Intervals				Relative Precision (width of Bootstrap CI / width of JN Interval)[a]	
				90% CI	95% CI	90% Intervals		95% Intervals			
Sample Size	Effect Size	Amount of Error	True Matching Threshold	Coverage (%)	Coverage (%)	Coverage (%)	Infinite Intervals (%)	Coverage (%)	Infinite Intervals (%)	90%	95%
100	Smaller	Low	0	94.3	97.7	89.8	24.3	90.2	15.4	NA	NA
1000	Smaller	Low	0	89.5	94.5	89.3	0.0	94.7	0.0	1.000	.998
100	Larger	Low	0	90.0	94.8	89.8	0.0	95.0	0.0	.989	.988
1000	Larger	Low	0	90.9	95.4	91.3	0.0	95.7	0.0	1.000	1.000
100	Smaller	High	0	98.8	99.8	90.4	83.1	94.6	70.1	NA	NA
1000	Smaller	High	0	96.7	98.8	87.5	63.2	90.4	49.1	NA	NA
100	Larger	High	0	96.7	98.7	88.3	52.1	88.6	53.5	NA	NA
1000	Larger	High	0	89.5	94.5	89.3	0.0	94.7	0.0	.997	.989
100	Smaller	Low	1	93.1	96.9	75.8	5.5	70.6	2.8	NA	NA
1000	Smaller	Low	1	89.6	94.5	89.8	0.0	94.6	0.0	1.000	1.001
100	Larger	Low	1	89.5	94.6	89.9	0.0	95.1	0.0	.992	.993
1000	Larger	Low	1	90.3	95.9	90.5	0.0	95.6	0.0	1.000	1.000
100	Smaller	High	1	94.3	98.0	82.7	80.6	89.8	66.4	NA	NA
1000	Smaller	High	1	95.3	98.1	69.6	43.3	71.6	31.1	NA	NA
100	Larger	High	1	95.4	97.9	68.3	29.2	66.2	18.2	NA	NA
1000	Larger	High	1	90.0	94.8	90.2	0.0	94.7	0.0	.994	.985

CI = Confidence Interval. Monte Carlo simulations were generated to follow the form outlined by $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + e$, where e is normally distributed with a mean of 0, and a standard deviation of SD_e. The matching effect size was manipulated to be small ($\beta_2 = .125; \beta_3 = -.25$), or large ($\beta_2 = .5; \beta_3 =$-1). Error was manipulated to be either low ($SD_e = .5$) or high (SD_e=3).

[a]The relative precision was calculated as a ratio between the width (i.e., range) of the intervals given by the bootstrap and the JN procedures. A value lower than 1 indicates that the bootstrap CI was more precise, and a value greater than 1.000 indicates that the JN interval was more precise. If the cells indicates "NA", a value could not be computed as the width of some JN intervals was infinite.

Table 3

Sample Demographics (N = 568)

Variable	N	%	Variable	N	%
Gender			Education		
Male	247	43.5	High School or Less	128	22.5
Female	318	56.0	Vocational/College	132	23.1
Other	3	0.0	Bachelors	221	39.3
Known Allergies (yes)	72	12.7	Master's/Doctoral/Professional	84	14.8
Dietary Restrictions[a]			Employment Status		
Any	231	40.7	Full-time	336	59.2
Vegetarian/Vegan/semi-vegetarian	84	14.8	Part-Time	95	16.7
Low calorie/low carb/low fat	131	23.1	Not-Working	74	13.0
High Protein	56	9.9	Retired	42	7.4
Other	65	11.4	Other/No answer	21	3.7
Ethnicity[a]			Family Income		
White/Caucasian	464	81.8	Less than 30,000	176	31.1
Black/African American/African	41	7.2	30,000 - 59,999	182	32.1
Asian/Asian American	40	7.1	60,000 - 89,999	129	22.8
Latino/Hispanic	38	6.7	90,000 and above	70	12.3
Other	8	1.4			

[a]Participants could select more than 1 answer.

Table 4

Manipulations used for message frame and message appeal type.

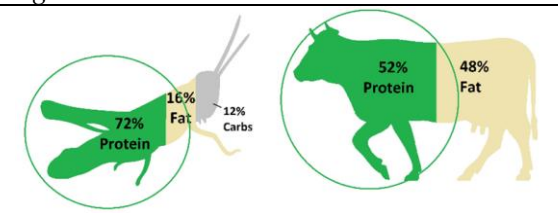
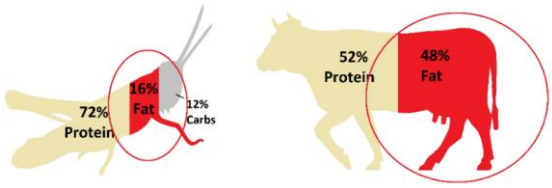
Strategy	**Factor/Level**	**Example**
Reframing content of sentences.	Gain Frame	"Adding insects to your diet can thus help you *obtain better health as well as a longer life expectancy.*"
	Loss Frame	"Adding insects to your diet can thus help you *avoid bad health and a shorter life expectancy.*"
Keeping the content frames identical, but changing the relative emphasis of the sentence.	Gain Frame	"Producing 1 pound of beef protein requires over 1000 gallons of water. *However, producing 1 pound of grasshopper protein requires only 1 gallon.*"
	Loss Frame	"Producing 1 pound of grasshopper protein requires 1 gallon of water. *But producing 1 pound of beef requires over 1000 gallons.*"
Altering an image to draw attention to gains versus losses associated to edible insects in relation to traditional livestock.	Gain Frame	
	Loss Frame	
Altering the pronouns used in sentences.	Independence Appeal	"Adding insects to *your diet* can thus help *you* obtain better health as well as a longer life *expectancy.*"
	Interdependence Appeal	"Adding insects to *our diets* can thus help *us* and our families avoid bad health and shorter life *expectancies.*"
Altering themes present in the message to reflect themes outlines by self-construal theory.	Independence Appeal	Themes of individual choice, uniqueness, and power. "Entomophagy is personal choice that can send a strong message about your personal beliefs and values on sustainability and health."
	Interdependence Appeal	Themes of family and societal concerns, conformity, responsibility and harmony. "It is our chance to do something for the good of society and ensure a safe, healthy, and sustainable life for future generations."
Using images focusing on either an individual, or a group.	Independence Appeal	Using an image of a single individual (a man running).
	Interdependence Appeal	Using an image of individuals interacting (a family).

Table 5

Overall first-order correlation between measures, and descriptive statistics on score distribution (using raw variable scores).

	1.	2.	3.	4.	5.	Mean / Median / Scale[a]	S.D.	Skew
1. Promotion Focus						3.64 / 3.78 / 5	0.73	-0.61
2. Prevention Focus	.49					3.98 / 4.00 / 5	0.65	-0.46
3. Interdependent Self-Construal	.22	.27				4.69 / 4.67 / 7	0.95	-0.15
4. Independent Self-Construal	.54	.40	.18			5.06 / 5.08 / 7	0.92	-0.23
5. Willingness	-.01	-.19	-.03	.00		2.93 / 2.00 / 9	2.16	0.95
6. Attitudes	.08	-.08	-.01	.06	.61	2.90 / 2.90 / 5	0.78	0.13

[a]Likert scale upper anchor value for each measure. Lowest value was always 1.

Table 6

Summary of multiple regression models examining simple matching effects on willingness to eat insects (N = 568).[a]

Model & Effect Source	Effects on Willingness to Eat Insects[b]		
	β	95% CI	SE
1. Matching to Prevention Focus			
Intercept	.053	[-.061, .167]	.058
Prevention Focus	-.203	[-.319, -.087]	.059
Loss Frame[c]	-.107	[-.269, .054]	.082
*Prevention Focus*Frame*	-.021	[-.183, .141]	*.082*
2. Matching to Promotion Focus			
Intercept	.041	[-.075, .156]	.059
Promotion Focus	.072	[-.038, .181]	.056
Loss Frame[c]	-.084	[-.248, .081]	.084
*Promotion Focus*Frame*	-.219	[-.384, -.053]	*.084*
Matching Threshold[d]	-.382	[-2.228, .508]	NA
3. Matching to Interdependent SC			
Intercept	-.054	[-.170, .061]	.059
Interdependent Self-Construal	-.124	[-.242, -.005]	.060
Interdependence Appeal[e]	.115	[-.049, .279]	.084
*Interdependent SC*Appeal*	.197	[.033, .362]	*.084*
Matching Threshold[d]	-.582	[-3.382, .394]	NA
4. Matching to Independent SC			
Intercept	-.055	[-.172, .061]	.059
Independent Self-Construal	-.049	[-.163, .066]	.058
Interdependence Appeal[e]	.113	[-.053, .278]	.084
*Independent SC*Appeal*	.025	[-.141, .190]	*.084*

The interaction representing the matching effect in each model is indicated using *italics*. β = Standardized regression coefficient; CI = confidence interval; SE = Standard error; SC = Self-Construal; NA = not applicable

[a]All measured variables were standardized using z-scores prior to analyses.

[b]This variable was corrected for skewness using a log-transformation prior to standardization. Pattern of results remains the same without using this correction.

[c]Loss Frame (coded as 1) compared to a gain-framed message (coded as 0).

[d]Matching threshold confidence intervals were computed using nonparametric bootstrapping (100,000 bootstrap samples, and the percentile method) and are only presented for significant matching effects

[e]Interdependence appeal (coded as 1) compared to an independence appeal (coded as 0).

Table 7

Summary of our multi-matched message regression model on willingness to eat insects (N=568).[a]

Effect Source	Effects on Willingness to Eat Insects[b]		
	β	95% CI	SE
Intercept	-.009	[-.150, .132]	.072
Promotion Focus	.075	[-.036, .186]	.056
Loss Frame[c]	-.095	[-.259, .069]	.083
*Promotion Focus*Frame*	*-.216*	*[-.381, -.051]*	*.084*
Interdependent Self-Construal	-.126	[-.246, -.006]	.061
Interdependence Appeal[d]	.116	[-.048, .279]	.083
*Interdependent SC*Appeal*	*.196*	*[.032, .360]*	*.083*
Matching Thresholds[e]	Estimate	95% bootstrap CI	
For matching to promotion focus	-.439	[-2.408, .454]	NA
For matching to interdependent SC	-.591	[-3.554, .388]	NA

The interactions representing matching effects are indicated using *italics*. β = Standardized regression coefficient; CI = confidence interval; SE = Standard error; SC = Self-Construal; NA = not applicable.

[a]All measured variables were standardized using z-scores prior to analyses.

[b]Scores on the outcome measure, willingness to eat insects, were corrected for skewness using a log-transformation prior to standardization. The pattern of results remains the same without using this correction.

[c]Loss Frame (coded as 1) compared to a gain-framed message (coded as 0).

[d]Interdependence appeal (coded as 1) compared to an independence appeal (coded as 0).

[e]Matching threshold confidence intervals were computed using nonparametric bootstrapping (100,000 bootstrap samples, and the percentile method)

Table 8.

Relative Fit Indices of Different Models to Quantify the Usefulness of Multi-Matching

Models	R^2	Adj. R^2	AIC	BIC
All main effects	0.006	-0.001	1619.402	1645.455
Main effects and matching to promotion focus	0.018	0.009	1614.610	1645.005
Main effects and matching to interdependent SC	0.016	0.007	1615.709	1646.104
Mains effects and both of the matching effects	0.028	0.017	1611.071	1645.808

The main effects included that of message frame (gain vs. loss frame), message appeal (independence vs. interdependence appeal), promotion focus scores, and interdependent self-construal.

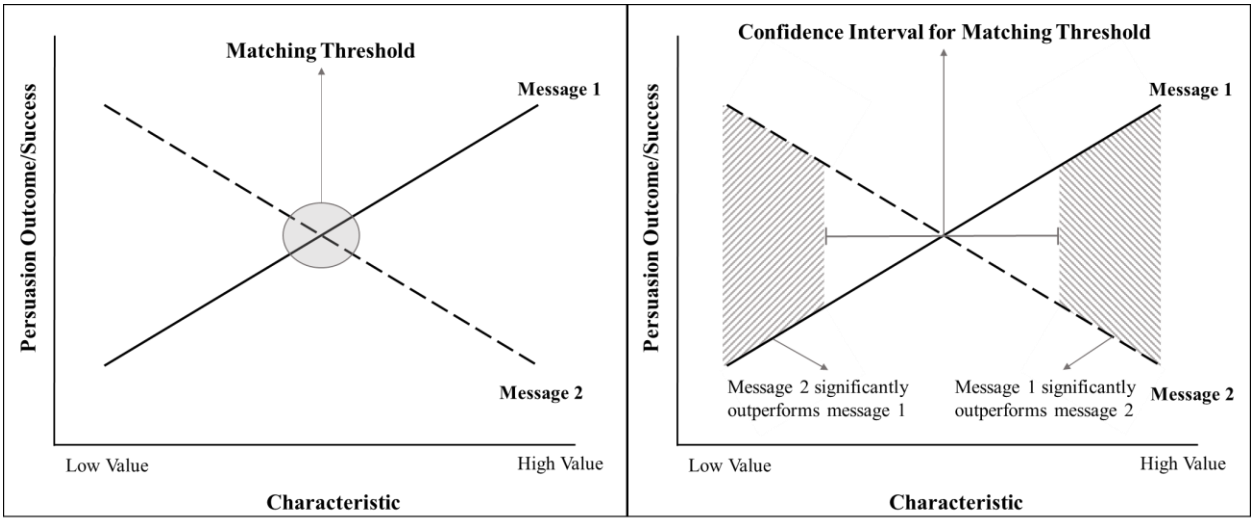


Figure 1. A visual example of a matching threshold (left side) and a confidence interval around it (right side). The matching threshold is operationally defined as the point at which two simple slopes intersect; in practical terms, this signifies the point along the characteristic for which reception of message 1 or 2 lead to equal levels of an outcome. The confidence interval around the matching threshold signals uncertainty around the point estimate of the matching threshold. Scores on the characteristic that overlap with the confidence interval are not associated with significant differences between the performance of the two messages (i.e., the difference between the 2 messages is not significantly different from 0); however, scores that lie beyond the confidence interval are associated with significant differences between the two messages. In this example, message 1 outperforms messages 2 at high values of the characteristic, but message 2 outperforms message 1 at low values of the characteristic; further, the matching threshold lies somewhere near the midpoint of the hypothetical characteristic.

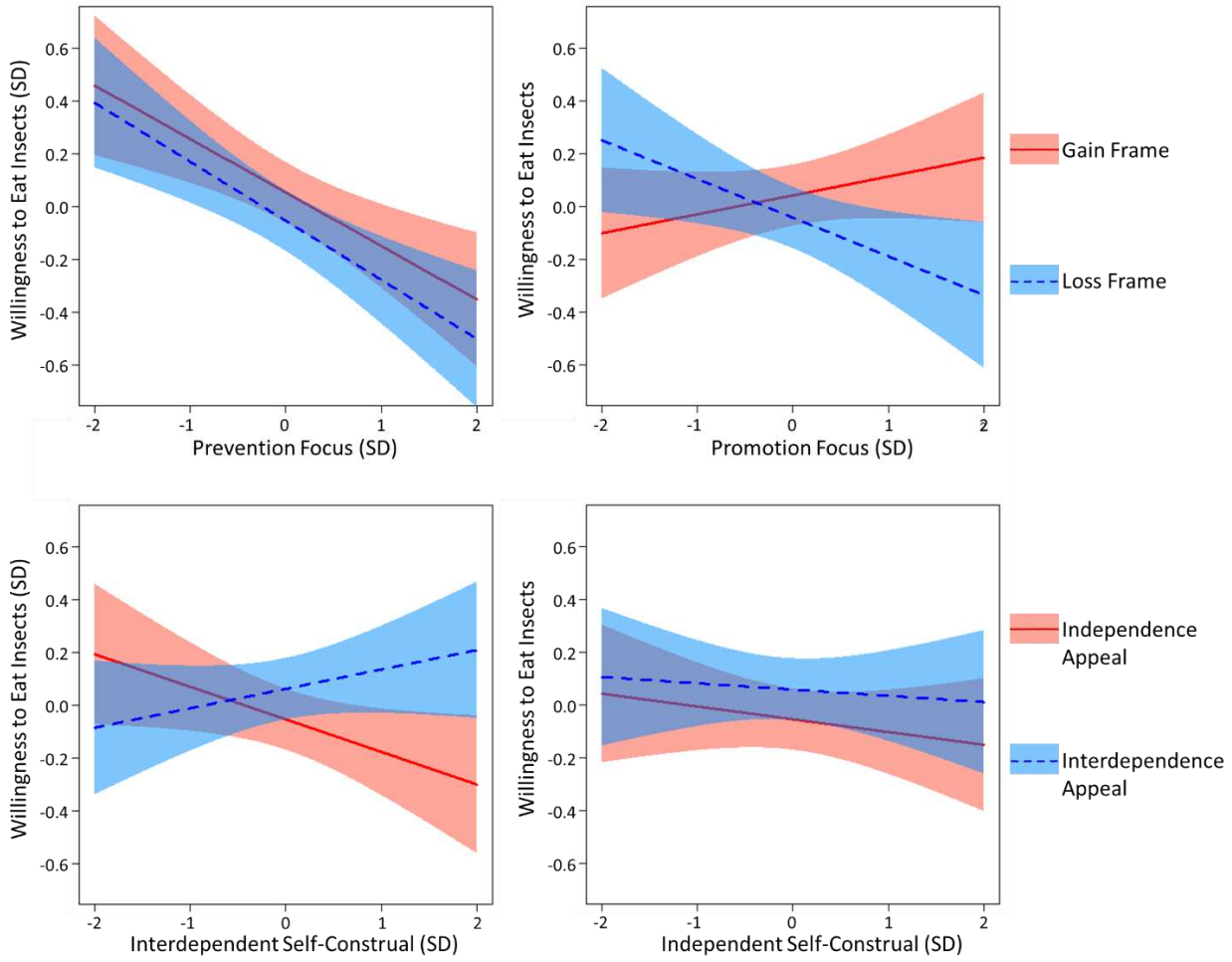


Figure 2. Visual depiction of matching messages to: Prevention focus (top left); promotion focus (top right); interdependent self-construal (bottom left), and; independent self-construal (bottom right). Units are in standard deviations (SD), and a 95% confidence band around the regression estimate is depicted.

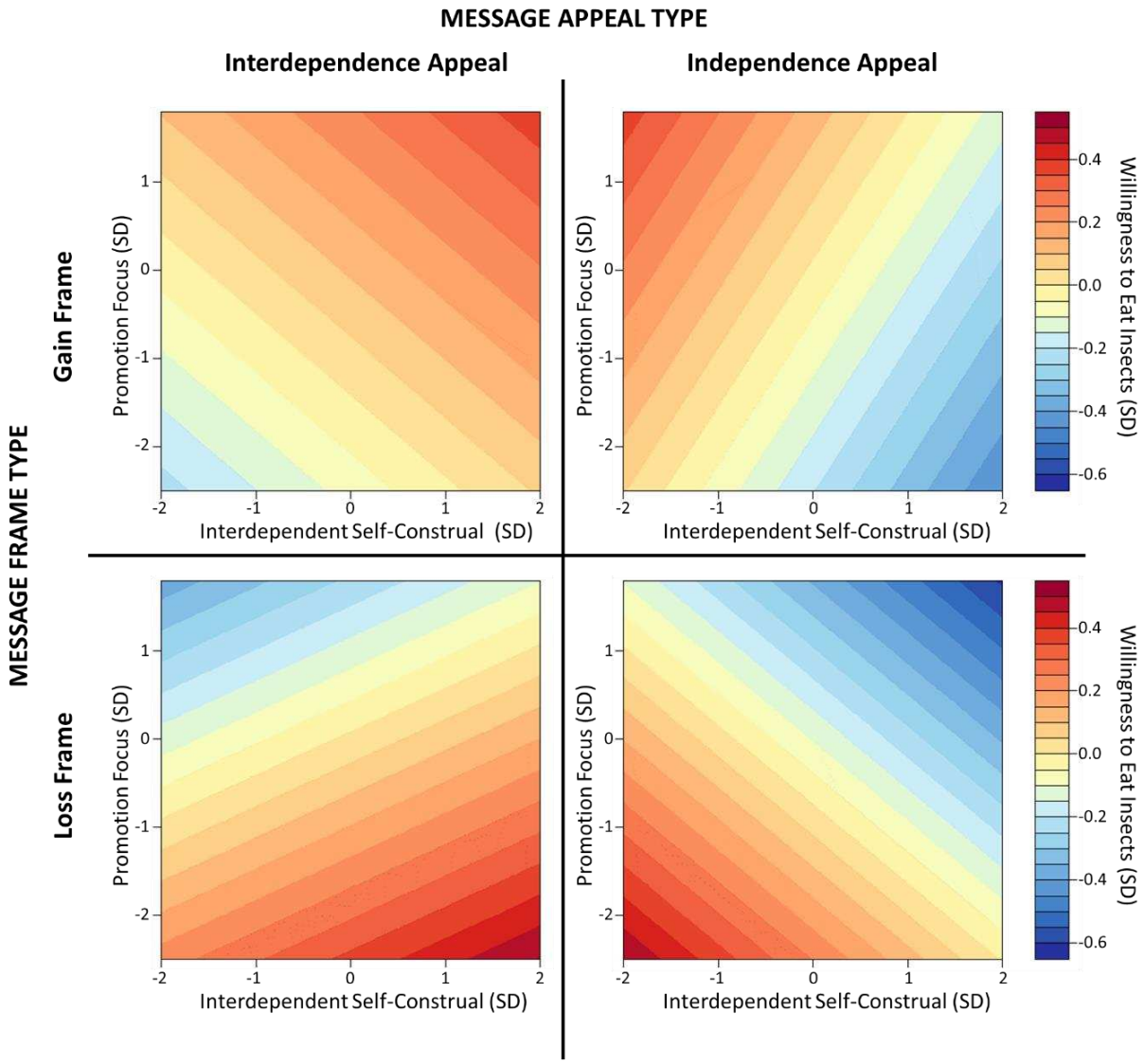


Figure 3. Visual depiction of the effects of matching messages to both promotion focus, and interdependent self-construal simultaneously (step 3 of our analyses). The y-axis within each quadrant represents promotion focus scores, and the x-axis, interdependent self-construal scores. The color gradient represents degree of willingness expressed towards eating insects, and each unit is expressed in standard deviations. The quadrants depict results for gain-framed interdependence appeals (top left), gain-framed independence appeals (top right), loss-framed interdependence appeals (bottom left), and loss-framed interdependence appeals (bottom right).