Behavioral Variability Reduces the Harmful Longitudinal Effects of Partners’ Negative-Direct Behavior on Relationship Problems
Nickola C. Overall

CITATION
Behavioral Variability Reduces the Harmful Longitudinal Effects of Partners’ Negative-Direct Behavior on Relationship Problems

Nickola C. Overall
University of Auckland

Based on growing evidence that negative-direct behavior that addresses important contextual and situational demands is less harmful than negative-direct behavior that occurs irrespective of current demands, the current investigation tests whether the longitudinal impact of partners’ negative-direct behavior depends on whether that behavior is more variable versus stable across couples’ daily life and conflict interactions. In Studies 1 and 2, participants rated how much their partner behaved in critical and unpleasant ways every day for 21 days. In Study 3, couples were video-recorded discussing an important area of conflict, and independent coders rated how much partners expressed criticism and hostility within every 30-s segment of the discussion. In each study, the repeated assessments were used to calculate average levels (within-person mean across days or couples’ discussions) and variability (within-person SD across days or couples’ discussions) of partners’ negative-direct behavior. Participants also reported on the severity of their relationship problems and relationship satisfaction at the beginning of each study and then 9 months later (Studies 1 and 2) or repeatedly across the following year (Study 3). High mean levels of partners’ criticism and hostility predicted greater relationship problems (Studies 1–3) and lower relationship satisfaction (Study 3) when partners’ negative-direct behavior was stable across time (low within-person variability), but was less harmful when partners’ negative-direct behavior varied across time (high within-person variability). These novel results illustrate that behavioral variability offers a valuable way to understand and examine behavioral patterns that will be more helpful versus harmful in navigating the challenges of social life.

Keywords: behavioral variability, conflict, daily sampling, negative-direct behavior, relationship problems

Supplemental materials: http://dx.doi.org/10.1037/pspi0000231.supp

Even the best of partners can behave in hurtful and critical ways or express anger and hostility during conflict. These negative-direct behaviors risk damaging relationships across time (Gottman, 1998; Karney & Bradbury, 1995b), but there is growing evidence that criticism, anger, and hostility can sometimes improve relationship problems and preserve satisfaction, even for the person who is the target of partners’ negative-direct behavior (Overall & McNulty, 2017; Overall, 2018).1 Crucially, however, whether partners’ negative-direct behavior is more versus less harmful depends on a range of factors that determine the contextual meaning of partners’ negative-direct behavior (McNulty, 2016). For example, expressing criticism and hostility can help sustain relationship quality when negative-direct behavior is diagnostic of important problems that partners are invested to change, but partners’ negative-direct behavior exacerbates relationship problems when it reflects more general insecurities or unreasonable, disproportionate demands, the current investigation tests whether the longitudinal impact of partners’ negative-direct behavior varies across different taxonomies, including negative-direct (McNulty & Russell, 2010; Overall et al., 2009; Overall, 2018), hostility (see Woodin, 2011), active-destructive (Rusbult et al., 1991) and direct opposition (McNulty, 2016; Overall & McNulty, 2017). These behaviors are referred to as direct because they are explicit, overt, and confrontational, and these direct features account for why negative-direct behaviors can be more effective than negative-indirect behaviors (e.g., manipulation, supplication) in resolving problems across time (McNulty & Russell, 2010; Overall et al., 2009; Overall & McNulty, 2017). The term oppositional rather than negative has been sometimes used to ensure that the valence is not confused with the effects of these behaviors, which are not always detrimental (McNulty, 2016). Negative-direct is used here because it aligns more closely with the range of behavioral and affective science reviewed in support of this investigation. Negative refers to the unpleasant valence and affective elements of these behaviors, rather than the effects of these behaviors, which are clarified throughout the paper.

1 This class of negative behaviors has been labeled in various ways across different taxonomies, including negative-direct (McNulty & Russell, 2010; Overall et al., 2009; Overall, 2018), hostility (see Woodin, 2011), active-destructive (Rusbult et al., 1991) and direct opposition (McNulty, 2016; Overall & McNulty, 2017). These behaviors are referred to as direct because they are explicit, overt, and confrontational, and these direct features account for why negative-direct behaviors can be more effective than negative-indirect behaviors (e.g., manipulation, supplication) in resolving problems across time (McNulty & Russell, 2010; Overall et al., 2009; Overall & McNulty, 2017). The term oppositional rather than negative has been sometimes used to ensure that the valence is not confused with the effects of these behaviors, which are not always detrimental (McNulty, 2016). Negative-direct is used here because it aligns more closely with the range of behavioral and affective science reviewed in support of this investigation. Negative refers to the unpleasant valence and affective elements of these behaviors, rather than the effects of these behaviors, which are clarified throughout the paper.

Editor’s Note. Shelly Gable served as the action editor for this article—KK.

This research was supported by a Royal Society of New Zealand Marsden Fund Grant (UOA0811) and University of Auckland Science FRDF Grant (3626244) awarded to Nickola C. Overall, and a Victoria University Wellington grant (200759) awarded to Garth J. O. Fletcher. I thank Matthew Hammond, Garth Fletcher, Melissa Grouden, Helena Struthers, Rosabel Tan, Kelsey Deane, Yuthika Girme, Desmond Packwood, Briar Douglas, Phoebe Molloy, Shuai Han, Lucy Travaglia, David Pirie, and Jan Trayes for their contribution to data collection and behavioral observation coding. Data and syntax for the results presented in this paper are available on the Open Science Framework: https://osf.io/9s3yt/.

Correspondence concerning this article should be addressed to Nickola C. Overall, School of Psychology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. E-mail: n.overall@auckland.ac.nz

© 2020 American Psychological Association
ISSN: 0022-3514 http://dx.doi.org/10.1037/pspi0000231
ate responses to the varying problems couples encounter (see McNulty, 2016; Overall & McNulty, 2017 for reviews).

The strongest extant evidence for the varying effects of negative-direct behavior comes from dyadic research modeling how between-person differences in partners’ behavior predict changes in relationship problems and satisfaction across time, and testing whether these longitudinal effects are modified by between-person factors (see McNulty, 2016 and Overall & McNulty, 2017). The number and diversity of factors determining the relative costs of negative-direct behavior reveals that, rather than a single definitive contextual feature, partners’ negative-direct behavior will be less harmful when it varies according to the changing demands couples are facing (see McNulty, 2016). The aim of the current studies is to integrate and extend the factors that reduce or magnify the impact of partners’ negative-direct behavior by considering a new behavioral dimension—within-person variability—to assess temporal patterns that should emerge if behavior is sensitive to situational demands (more variable) or is rigid and less diagnostic of current demands (more stable).

The following sections provide the theoretical and empirical foundation for this novel application of within-person variability. The first section briefly describes the extensive body of work examining intrapersonal variability, which is theorized to represent contextually sensitive patterns of responding versus stable patterns of responding irrespective of situational demands. The second section describes why considering behavioral variability may pull together the array of factors shown to modify the relative harm of partners’ negative-direct behavior. The third section illustrates these theoretical connections by considering variable versus stable patterns of partners’ negative behavior during daily life and couples’ conflict interactions. The final section outlines three dyadic longitudinal studies that test whether partners’ negative-direct behavior predicts growing relationship problems across time when partners’ behavior is stable across daily life and couples’ conflict interactions. The final section outlines three dyadic longitudinal studies that test whether partners’ negative-direct behavior predicts growing relationship problems across time when partners’ behavior is stable across daily life and couples’ conflict interactions (low within-person variability), but will be less harmful when partners’ negative-direct behavior is more variable across daily life and couples’ interactions (high within-person variability).

The Application and Conceptualization of Intrapersonal Variability

Temporal dynamics in interpersonal behavior have not been previously examined, but variability versus stability of intrapersonal states has been widely investigated. One principal application involves examining the outcomes of emotional variability. The typical paradigm involves using experience sampling paradigms to repeatedly assess emotions across daily life and then creating an index of emotional variability across the sampling period, such as calculating the within-person standard deviation (SD) in emotions across time points (e.g., Eid & Diener, 1999; Gable & Nezlek, 1998; Gruber, Kogan, Quoidbach, & Mauss, 2013; Kuppens, Van Mechelen, Nezlek, Dossche, & Timmermans, 2007; Koval, Pe, Meers, & Kuppens, 2013; see Houben, Van Den Noortgate & Kuppens [2015] for a review). Greater emotional variability is theorized to reflect an inability to effectively deal with and regulate emotions in response to changes in the environment, which is supported by consistent associations between greater variability (higher within-person SD) and poorer psychological and physical health (Gruber et al., 2013; Hardy & Segerstrom, 2017; Houben, Van Den Noortgate, & Kuppens, 2015; Jenkins, Hunter, Richardson, Conner, & Pressman, 2019; Koval et al., 2013; Kuppens et al., 2007). Similarly, variability in relationship-related feelings and evaluations (satisfaction, security, commitment) is theorized to reflect more intense reactions and turmoil in response to relationship events, and is thus associated with poorer psychological and relationship wellbeing (Arriaga, 2001; Arriaga, Reed, Goodfriend, & Agnew, 2006; Campbell, Simpson, Boldry, & Rubin, 2010; Whitton & Whisman, 2010).

Investigations of intrapersonal variability have also shown that variability can combine with global or initial levels to determine psychological outcomes. For example, an extensive research program led by Kernis (2005) illustrated that variability in self-esteem emerges from greater reactivity to self-relevant events, whereas stable self-esteem represents equanimity in response to evaluative events across daily life. Accordingly, high levels of global self-esteem combined with greater stability in self-esteem (low within-person SD across repeated assessments) predicts more favorable personal and interpersonal outcomes, such as greater wellbeing and lower anger and hostility (e.g., Kernis, Gramman, & Barclay, 1989; Kernis, Grammenn, & Mathis, 1991; Paradise & Kernis, 2002). Girme et al. (2018) also recently demonstrated that the effects of intrapersonal variability in attachment security depend on initial levels. Greater variability in relationship-specific security (higher within-person SD across time) was associated with declines in relationship satisfaction for people who were initially higher in security, and thus expected the quality of their relationships to be stable and less affected by the inevitable ups and downs of relationships.

These distinct applications and outcomes of intrapersonal variability converge to show that variability in emotional states, as well as self and relationship evaluations, likely reflect greater reactivity to the changing situational demands people face across daily life. In contrast to variability in internal states, however, variability in behavior should be more beneficial. For example, variability in emotion regulation behavior has been theorized to reflect contextually sensitive behavioral patterns in which people adaptively shift behavior to situational demands (Aldao, Sheppes, & Gross, 2015). This conceptualization is founded on the growing consensus that specific emotion regulation strategies are not necessarily maladaptive or adaptive, but may have costs or benefits depending on whether strategies used meet changing contextual demands (Aldao et al., 2015; Bonanno & Burton, 2013; Cheng, 2018).

2 Other temporal dynamics of emotions have also been assessed (see Houben et al., 2015 for a comparative description and analysis). Emotional instability (indexed by the mean squared successive difference) quantifies change from one moment to the next rather than over an entire period as the SD does. Emotional inertia (indexed by the autocorrelation) quantifies how much emotions carry over or linger across time. These different indicators of emotional dynamics demonstrate similar patterns with wellbeing (see Houben et al., 2015), and the SD has independent and sometimes more consistent effects (e.g., Gruber et al., 2013; Koval et al., 2013; Jenkins et al., 2019). However, the SD is the best indicator of the pattern targeted in the current studies because it assesses the degree of variability vs. stability without making additional assumptions regarding the appropriate time course of that variability, which will differ across situations. This is one reason why, as discussed next, the use of the SD has been adopted in recent applications to examine variability in behavior (i.e., emotion regulation variability; Blanke et al., 2019; Eldesouky & English, 2018).
Lau, & Chan, 2014). The most recent tests of this proposition have assessed variability in emotion regulation behavior across daily life (Blanke et al., 2019; also see Benson et al., 2019; Cheng, 2001; Eldesouky & English, 2018). In four experience sampling studies, Blanke et al. (2019) found that greater variability in emotion regulation strategies (greater SD of strategy use) was associated with better emotional outcomes across the sampling period (lower negative affect). These results support that greater variability in emotion regulation behavior captures a context-dependent pattern—adjusting regulation strategies to the different demands of different situations—that is adaptive, whereas low variability reflects a stable, consistent use of emotion regulation strategies across situations that is maladaptive (Aldao et al., 2015; Blanke et al., 2019).

Although the investigation of behavioral variability is only emerging, these initial findings support the widely applied conceptualization that greater variability represents a more contextually sensitive pattern of responses versus a stable pattern of responding irrespective of situational demands. Variability in emotional and evaluative states is a marker of greater reactivity and poorer coping in response to situational challenges, which is supported by the consistent links between intrapersonal variability and poorer wellbeing. By contrast, greater variability in behavior has opposite implications by indicating that behavior is adaptively shifting along with situational demands, which is supported by the links between emotion regulation variability and better emotional outcomes. Based on this established application and conceptualization of variability, the current studies examine the degree to which variability versus stability can advance understanding of when interpersonal behavior—intimate partners’ critical, hostile and hurtful relationship behavior—has harmful interpersonal outcomes, including the relationship problems and satisfaction experienced by the individual receiving partners’ negative-direct behavior.

Variability in Partners’ Negative-Direct Behavior and Relationship Problems

It seems obvious that partners’ critical, hostile, or otherwise hurtful behavior will create problems in relationships. Perceiving greater negative behavior from partners often leads people to feel they are not cared about, valued, or appreciated, which creates problems with intimacy, trust, and closeness and thus erodes relationship quality (e.g., LaMotte, Khalilian, & Barry, 2017; Overall & Fletcher, 2010; Wieselquist, Rusbult, Foster, & Agnew, 1999). Partners’ criticism and hostility also tend to trigger defensiveness and reciprocal negativity, which impedes problem solving and reduces relationship quality (Gottman, 1998; Lemay, Overall, & Clark, 2012; Liu, Lemay, & Neal, 2018). Similarly, people often respond to partners’ negative-direct behavior with withdrawal and disengagement, creating distance and exacerbating problems in relationships (Christensen & Heavey, 1990; Heavey, Layne, & Christensen, 1993). Finally, critical and hostile behavior can escalate into verbal and physical aggression, creating more serious problems and damage to relationships (Lawrence & Bradbury, 2001; Murphy & O’Leary, 1989; Rogge & Bradbury, 1999).

Given the slew of potential problems arising from partners’ negative-direct behavior, it is not surprising that meta-analytic reviews have concluded that critical, hostile behavior is associated with poorer relationship quality (Karney & Bradbury, 1995b; Woodin, 2011). Yet, longitudinal studies assessing how partners’ negative-direct behavior contributes to the development of relationships across time have produced inconsistent findings. Some studies suggest that negative-direct behavior has damaging effects on relationships (e.g., Huston, Caughlin, Houts, Smith, & George, 2001; see Karney & Bradbury, 1995b for review). Other studies provide evidence that partners’ negative-direct behavior has beneficial effects (e.g., Cohan & Bradbury, 1997; Gottman & Kroff, 1989; Heavey et al., 1993; Karney & Bradbury, 1997; Overall, Fletcher, Simpson, & Sibley, 2009). Others still indicate that partners’ negative behavior is not reliably associated with satisfaction across time (Lavner, Karney, & Bradbury, 2016).

These inconsistencies arise because, until recently, most research has simply tested whether between-person differences in levels of negative-direct behavior predict changes in relationship problems or quality. Yet, similar to the costs and benefits that can emerge for specific emotion regulation strategies (Aldao et al., 2015; Bonanno & Burton, 2013; Cheng et al., 2014), a growing body of evidence has shown that partners’ negative-direct behavior can have costs and benefits depending on whether that behavior is diagnostic of important contextual and situational demands that need to be addressed (McNulty, 2016). For example, criticism, anger, and hostility during conflict can express investment in improving problems, motivate change in targeted problems, and thus reduce problem severity and preserve relationship satisfaction (Cohan & Bradbury, 1997; Gottman & Kroff, 1989; Heavey et al., 1993; Heavey, Christensen, & Malamuth, 1995; Karney & Bradbury, 1997). Accordingly, partners’ criticism and hostility during conflict is less harmful when partners are trying to improve dissatisfying problems (Overall et al., 2009; Overall, 2018), problems are very serious and need to be changed (McNulty & Russell, 2010), and when targets are open to and feel able to change the problem (e.g., Baker & McNulty, 2015; Jayamaha, Antonellis, & Overall, 2016).

By contrast, partners’ criticism and hostility is associated with growing problems and declines in satisfaction when these negative-direct behaviors are enacted in resistance of improvement attempts (Overall et al., 2009; Overall, 2018), disproportionate to problem severity (McNulty & Russell, 2010), directed toward targets who are not confident or secure enough to be responsive (e.g., Baker & McNulty, 2015; Jayamaha et al., 2016), or likely represent more global negativity rather than dissatisfaction of specific problems (e.g., Jayamaha & Overall, 2015). This range of modifying effects indicate that partners’ negative-direct behavior can improve relationships when it is diagnostic of specific and serious problems that couples are invested in and able to change, but partners’ negative-direct behavior is likely to reinforce or intensify relationship problems and damage relationship quality when it is not diagnostic of important problems or contexts and instead signals disproportionate responses or persistent negativity irrespective of situational demands.

The factors that modify the impact of partners’ negative-direct behavior have primarily been identified in studies examining behavior during couples’ observed conflict interactions (see McNulty, 2016; Overall & McNulty, 2017). However, there is also evidence that the effects of partners’ daily negative-direct behavior should depend on the contextual relevance of that behavior. For example, partners’ daily negative emotions have a less immediate
impact on individuals’ own emotions when they arise in contexts of significant external stress (Thompson & Bolger, 1999). Acute stressors are likely to offer contextual meaning to partners’ negativity, including the importance of immediately supporting partners or resolving any relationship tension, which may help mitigate any long-term impact of partners’ daily negative-direct behavior as it can within specific conflicts (e.g., Cohan & Bradbury, 1997). However, this short-term tolerance of partners’ daily negative emotions is not evident when partners’ negativity is likely indicative of more enduring problems, such as prolonged aggression (Timmons, Arbel, & Margolin, 2017), or reflects frequent daily hassles that are unlikely to generate specific problem solving (Bodenmann, Ledermann, & Bradbury, 2007). Thus, when partners’ daily negativity is not diagnostic of current demands, stressors, or problems, but instead represents ongoing, persistent patterns of negative behavior, it will likely accumulate to amplify relationship problems across time.

In sum, just as greater variability in emotion regulation behavior is theorized to represent a context-sensitive pattern in which individuals adjust emotion regulation strategies to different situational demands (Aldao et al., 2015; Benson et al., 2019; Blanke et al., 2019), variability in partners’ negative direct behavior may also reflect a contextually sensitive pattern that reflects responding to the varying situational demands couples encounter as they negotiate conflict and daily life. Given that prior research has established that partners’ negative-direct behavior can help couples attend to and resolve serious problems when negative behavior is diagnostic of important problems that need to be addressed, such a variable pattern should be more likely to balance the costs of partners’ negative-direct behavior with the potential benefits of problem improvement. Moreover, greater variability in partners’ negative-direct behavior should reduce the presence of negative-direct behavior when it is not needed or not appropriate, such as in the absence of problems or stressors, when partners need support, or when couples have the opportunity to generate intimacy and maintain relationships (Gottman, 1998; McNulty, 2016). In short, the degree to which partners’ negative-direct behavior predicts growing relationship problems and dissatisfaction may be attenuated if partners’ behavior is more variable.

By contrast, just as low variability is understood to reflect a stable use of emotion regulation strategies across situations that is maladaptive (Aldao et al., 2015; Benson et al., 2019; Blanke et al., 2019), persistent negative-direct behavior irrespective of changing contexts or situational demands will be less diagnostic of current problems and be more damaging to targets. For example, high levels of invariant negative-direct behavior may arise from stable partner dispositions that create persistent interpersonal negativity, such as greater attachment insecurity, lower self-esteem, or higher neuroticism (Karney & Bradbury, 1995b; McNulty, 2008; Murray, Bellavia, Rose, & Griffin, 2003; Simpson & Rholes, 2012). Greater negative-direct behavior arising from these individual differences may reflect poorer regulation of negative emotions and insecurities that prevent recovery from the challenges that trigger negative-direct behavior (Campbell, Simpson, Boldry, & Kashy, 2005; Salvatore, Kuo, Steele, Simpson, & Collins, 2011; Shaver & Mikulincer, 2007). Regardless of the cause, chronic negativity renders the presence and intensity of negative-direct behavior less diagnostic of current needs (Forest, Kille, Wood, & Holmes, 2014), and thus greater negative-direct behavior by partners who tend to exhibit negativity across relationship contexts predicts more sustained relationship problems (Jayamaha & Overall, 2015). Thus, high and stable levels of partners’ negative-direct behavior should be most likely to produce problems in relationships.

Benefits and Examples of Assessing Variability in Partners’ Negative-Direct Behavior

Integrating the empirical and theoretical work investigating intrapersonal variability with the empirical and theoretical work showing that the impact of partners’ negative-direct behavior depends on a range of contextual and situational demands, the current studies test whether the longitudinal impact of partners’ negative-direct behavior depends on whether partners’ negative-direct behavior is more variable versus stable across couples’ daily life and conflict interactions. Consistent with the extensive application and conceptualization of intrapersonal variability reviewed above, the primary reason greater variability in partners’ negative-direct behavior was hypothesized to be less harmful is because partners’ negative-direct behavior that occurs in the context of important demands that need to be addressed has less costs. By contrast, partners’ negative-direct behavior that is more stable and persistent, and thus less diagnostic of current situational demands, should be more harmful.

One disadvantage of examining variability as a marker of situationally contingent behavioral patterns versus stable, persistent behavior irrespective of situational demands is that this approach does not test the underlying causes of variability versus stability or the relevant situational demands and contexts that variability might reflect. Yet, one key advantage of this approach is that it identifies general behavioral patterns across a range of conditions rather than limiting tests to specific contextual features from an array of demands that people need to respond to in their daily life and social interactions (Blanke et al., 2019; Houben et al., 2015). Indeed, partners’ negative-direct behavior has been shown to be less versus more harmful across a diverse set of factors, including person-level characteristics (e.g., partners’ or targets’ insecurity), couples’ interaction dynamics (e.g., targets’ resistance, motivation and ability to be responsive), features of the current problems (e.g., severity, controllability), and factors external to the relationship (e.g., stressful life circumstances; see McNulty [2016] for a taxonomy of these disparate factors). The increasing number of distinct moderators identified demonstrates that one specific factor does not determine the relative impact of negative-direct behavior, but rather partners’ negative-direct behavior that varies in response to an array of meaningful contexts and situational demands will not be as harmful as negative-direct behavior that occurs irrespective of the varying contexts and demands couples encounter. Moreover, although more stable negative-direct behavior may arise from a range of factors (e.g., partners’ poor emotion regulation ability, dispositional reactivity or insecurity, persistent stress), the degree to which these factors affect targets will, in large part, occur via targets’ experiences of high and stable levels of partners’ negative-direct behavior (Karney & Bradbury, 1995b; Kelley et al., 1983; McNulty, 2016; Overall & McNulty, 2017).

Thus, examining variability in partners’ negative-direct behavior is one valuable way of integrating the range of factors that moderate the effects of partners’ negative-direct behavior into a global index that captures patterns that should emerge if partners...
are adjusting behavior to situational demands (more variable) or are behaving in rigid, persistent ways that is less diagnostic of current demands (more stable). This global ingredient of behavior may also generalize across methods used to assess naturally occurring interpersonal behavior within couples’ interactions, including the level and variability of partners’ negative-direct behavior across couples’ daily life as well as across couples’ specific conflict interactions.

Variability in Partners’ Negative-Direct Behavior Across Days

To illustrate, the first two studies in the current research apply the same methods used to assess intrapersonal variability as a marker of contextually sensitive responding versus stability of emotions and behavior irrespective of daily demands. The top panel of Figure 1 presents actual data from Study 1, in which participants reported on their partners’ critical and unpleasant behavior across 21 days. Demonstrating a highly variable pattern, the participant represented by the dashed line reported that their partner behaved in very critical and unpleasant ways on a few days, mildly critical and unpleasant on some other days, and then did not behave negatively at all on many days. Such behavioral variability is likely a good indicator that the partner’s negative behavior on specific days were more closely tied to varying contexts or events relevant to relationship and personal dynamics, which as the research described above suggests should not be as harmful to relationships.

The participant represented by the solid line in the top panel of Figure 1 had the exact same average score of partners’ negative behavior across the 21 days, but a much lower standard deviation: this participant did not report that their partner behaved very negatively on any given day, but instead reported their partner was consistently mildly or moderately critical and unpleasant across days. Such patterns of continuous negative behavior that varies little across days likely indicates that partner’s negative behavior is not tied to specific problems, stressors, or daily demands, but instead reflects persistent expressions of negativity and dissatisfaction irrespective of current demands and needs. Thus, higher mean and stable levels of partners’ negative behavior should be more likely to create relationship problems.

Building on the established use of variability to assess contextually sensitive patterns in affective dynamics across daily life, these examples highlight how considering the mean and variability of relationship behavior offers a new way of capturing the degree to which partner’s negative-direct behavior varies in ways that would be expected if that behavior is responsive to the daily context versus invariant and nondiagnostic of current situational demands. Given the costs of receiving negative-direct behavior from partners, perceiving partners to behave in highly unpleasant and critical ways across daily life risks growing relationship problems and dissatisfaction. However, based on the documented contextual effects, high levels of partners’ negative behavior should be particularly harmful when that behavior reflects persistent and stable delivery of negative behavior (see the solid line in Figure 1), but this harm should be reduced if high levels of partners’ negative-direct behavior emerge from a highly variable pattern (see the dashed line in Figure 1).

Variability in Partners’ Negative-Direct Behavior Across Couples’ Conflict Interactions

The third study in the current research applies the same theorizing to examine behavioral variability within couples’ discussions of serious relationship conflicts, which is the principal method used in prior research examining when partners’ negative-direct behavior is more or less harmful. Although daily life may offer more contextual variation in the types of problems, stressors, and demands couples face, the relative meaning and impact of behavioral variability should have the same implications within a specific social interaction, including indicating that partners are responsive to the varying needs and demands that emerge across couples’ discussions, such as the changing content, structure, and tenor of the interaction.

Prior research has provided good evidence that partners’ negative-direct behavior when pushing for change during couples’
conflict discussions helps sustain relationships across time because partners’ negative-direct behavior is diagnostic of the importance and severity of the problem, the need for change, and partners’ investment in the relationship (McNulty, 2016; McNulty & Russell, 2010; Overall et al., 2009; Overall, 2018; Overall & McNulty, 2017). Yet, these potential benefits should be most likely to occur if partners are also responsive to situational demands, such as the targets’ needs, emotions, and behavioral responses across couples’ interactions. For example, targets’ anger and hostility, withdrawal and disengagement, or expressions of hurt and guilt, may indicate the need for partners to temporarily ease cumulating tension so that targets understand the importance and meaning of partners’ negative-direct behavior, while also remaining receptive and engaged in the interaction (e.g., Overall et al., 2013; Overall, Girme, Lemay, & Hammond, 2014). Partners’ may also need to respond to the changing topic of the conversation, pushing when the conversation is focused on the main points of contention, but listening and being receptive as the topic of the conversation shifts to less conflictual issues and points of agreement. Partners’ may also need to take opportunities to maintain closeness and offer or accept bids of repair (Gottman, 1998).

As with the number of contextual factors that qualify the degree of harm negative-direct behavior has, any given interaction will vary in the type, range, intensity, and relevance of various situational demands, and thus a benefit of assessing variability is that it reflects a global ingredient that should have different implications for relationships. To illustrate, the bottom panel of Figure 1 presents actual data involving observational ratings of partners’ hostile behavior rated each 30-s across a 7-min conflict discussion. The two different profiles reflect exactly the same mean score of negative-direct behavior across the conflict discussion, but differ in the degree to which the intensity of negative-direct behavior varies across the discussion. The partner depicted by the dashed line exhibited greater behavioral variability, showing high levels of criticism and hostility in some segments of the discussion, average levels in some segments, and very low levels in others. Such variability likely indicates that the partner is responding to the changing content and nature of the targets’ responses across the discussion. Accordingly, even within problem-solving discussions in which negative-direct behavior can lead to improvement in problems, a variable behavioral pattern of partners’ negative-direct behavior should be less harmful by balancing the need to push for change with the need to be sensitive to the changing nature of the discussion.

By contrast, the partner represented by the solid line exhibits a stable, invariable pattern in which moderate levels of criticism and hostility are exhibited more consistently throughout the discussion. This behavioral stability is likely to reflect a rigid delivery of hostility and criticism irrespective of the topic, content, emotional tenor, and behavioral dynamics within each segment of the interaction, undermining the degree to which partners’ negative-direct behavior conveys investment and is diagnostic of the specific issues discussed (Overall & McNulty, 2017; Overall, 2018). Instead, high stable levels of negative-direct behavior probably convey a lack of responsiveness and regard, global negativity beyond the specific issue, and perhaps attempts to dominate or control the interaction. As noted above, behavioral stability may emerge from more chronic dissatisfaction, insecurity or stressors, and related factors, such as poor emotion regulation or an inability to flexibly respond to the demands of the situation. Nonetheless, it is the resulting nature of partners’ negative-direct behavior that will shape the targets relative understanding of the seriousness and solvability of the problem, the partners’ regard and commitment, and other relevant interaction outcomes that go on to influence targets’ evaluations of the problem and relationship across time. (Karney & Bradbury, 1995b; Kelley et al., 1983; McNulty, 2016; Overall & McNulty, 2017).

Overview of the Current Studies

Three longitudinal studies test whether the degree to which high levels of partners’ negative-direct behavior predicts relationship problems and satisfaction across time depends on whether that behavior varies (high within-person SD) or is stable (low within-person SD) across couples’ daily life or specific conflict interactions. The first two studies adopted methods used to assess intrapersonal variability in the experience and regulation of emotions and self-evaluations across daily life. In two independent but virtually identical daily sampling studies, participants reported on the degree to which their partner behaved in critical, unpleasant, and hurtful ways every day for 21 days. The repeated daily assessments were used to calculate, for each participant, average levels of partners’ negative-direct behavior (mean across the 21 days) as well as the variability of partners’ negative-direct behavior (SD of partners’ behavior across the 21 days) across couples’ daily life. Prior to and nine months following the 3-week daily sampling procedure participants reported on the severity of their relationship problems and their relationship satisfaction.

Outside the typically examined context of acute conflicts or stressful events, high levels of daily negative-direct behavior from partners are likely to predict growing relationship problems and dissatisfaction across time. However, partners’ negative-direct behavior that varies in response to an array of meaningful contexts and situational demands should not be as harmful as negative-direct behavior that occurs irrespective of the varying contexts and demands couples encounter. Accordingly, high levels of partners’ negative-direct behavior across couples’ daily life should be less harmful when it arises from a variable temporal pattern (high average levels + high within-person SD) and more harmful when it arises from a stable pattern indicating persistent negativity less diagnostic of specific contextual and situational demands (high average levels + low within-person SD).

Study 3 applied the same approach at a more microlevel by assessing the mean and variability of partners’ negative-direct behavior exhibited during specific conflict discussions. Couples were video recorded having a 7-min. discussion regarding an important relationship problem the partner wanted to change. A team of independent coders then rated the degree to which partners expressed criticism, anger, and hostility within every 30-s segment of the discussion. These coder ratings were used to calculate, for each participant, average levels of partners’ negative-direct behavior (mean across the 14 segments) and the variability of partners’ negative-direct behavior (SD across the 14 segments) across the discussion. Prior to the discussion and repeatedly across the following year, participants reported on the severity of the problem discussed and their relationship satisfaction. As in Studies 1 and 2, the degree to which high levels of partners’ negative-direct behavior predicted problem resolution and relationship satisfaction...
across time was expected to depend on whether that behavior varied as couples’ discussed their problems (high average levels + high within-person SD) or was highly stable across couples’ discussion (high average levels + low within-person SD).

Studies 1 and 2

Studies 1 and 2 examined the longitudinal associations between the mean and variability of partners’ negative-direct daily behavior and relationship problems. In two independent daily sampling studies, participants rated the degree to which their partner behaved in critical, unpleasant or hurtful ways every day for three weeks. Reports of partners’ negative-direct behavior were assessed and analyzed because the ways in which partners’ behavior affects individuals’ relationship evaluations should primarily occur through individuals’ interpretation and experience of partners’ behavior. As reviewed above, partners’ negative-direct behavior can create relationship problems because it undermines felt regard, reduces closeness, and triggers defensive negativity or withdrawal—all of which will only occur if partners’ behavior is perceived as negative. Similarly, partners’ negative-direct behavior can improve relationships because it motivates individuals to change problems, which will only occur if partners’ behavior is perceived by individuals. Moreover, any effects of behavioral variability should also be most evident when assessing perceptions of partners’ behavior. The existing theory and research indicates that partners’ negative-direct behavior may be less harmful if it occurs in response to, and thus offers diagnostic information about, specific problems or stressors, but this requires individuals perceiving variation in partners’ behavior. Similarly, the meaning and impact of stable partners’ negative behavior that is not tied to specific situational demands should also primarily arise from individuals’ perceptions of persistent negative behavior.3

To test the longitudinal effects of perceptions of partners’ negative-direct behavior during daily life, participants completed assessments of the problems experienced in their relationship and their relationship satisfaction prior to and nine months following the daily sampling procedure. The more partners are perceived to behave negatively across daily life, the more likely individuals are to report growing relationship problems across time. However, the primary aim of Studies 1 and 2 was to test whether the longitudinal association between partners’ daily negative-direct behavior and problem severity depended on the variability of partners’ behavior across the daily sampling period. To calculate behavioral variability, the 21 daily reports were used to calculate both (a) average levels of partners’ negative-direct behavior across the sampling period (indexed by the person mean), and (b) variability in partners’ negative-direct behavior across the sampling period (indexed by the within-person SD). The same high average levels of partners’ negative-direct daily behavior were expected to be less harmful when arising from a variable behavioral profile (high mean levels + high within-person SD; see dashed lines in Figure 1), and more harmful when arising from a stable behavioral profile (high mean levels + low within-person SD; see solid lines in Figure 1).

Method

Studies 1 and 2 consisted of two independent studies collected at different research sites in different cities. Study 1 was designed and collected first, and then Study 2 conducted to replicate the design and procedure of Study 1. Given that each study followed the same procedures, the methods and results for Study 1 and Study 2 are presented jointly and integrative data analysis used to pool the data across studies to maximize power and directly test whether the effects replicated across studies (see Curran & Hussong, 2009; Hussong, Curran, & Bauer, 2013).

Participants. Both samples were recruited via paper and electronic announcements posted across university-based organizations (e.g., health centers, newsletters, childcare services). Couples were reimbursed $100NZD for completing a 3-week daily sampling procedure and a follow-up questionnaire 9 months later. Both samples have been used previously to examine predictors of daily emotional and behavioral dynamics, but the longitudinal associations between partners’ daily behavior and relationship problems and satisfaction across time have not been examined previously (see the online supplemental materials for more details). Each study received approval from the human participant ethics committee at the institution the study was collected.

Study 1. Seventy-eight heterosexual couples (N = 156 individuals) completed the daily sampling procedure. However, 27 couples did not complete the longitudinal assessment because of dissolution (N = 15) or opting out of participation (N = 13), leaving a sample of 100 participants on which the longitudinal analyses are based. Participants were 23.34 years old on average (SD = 5.48) and were involved in serious romantic relationships (49% married or cohabiting) that averaged 2.67 years in length (SD = 2.29).

Study 2. Seventy-three heterosexual couples (N = 146 individuals) completed the daily sampling procedure. However, 16 couples did not complete the longitudinal assessment due to dissolution (N = 8) or opting out of participation (N = 8), leaving a sample of 114 participants on which the longitudinal analyses are based. Participants were 23.69 years old on average (SD = 6.69)

3 One criticism of this approach, however, is that individuals might be biased in their reports, such as perceiving their partner to behave more negatively than warranted based on their partner’s reports (Hammond & Overall, 2013a; Overall & Hammond, 2013). Yet, partners are also motivated by self-serving biases and thus more likely to under-report their own negative-direct behavior. Accordingly, partner reports do not represent a more objective assessment. Moreover, as described above, the overall impact of partners’ negative-direct behavior will be determined by the way individuals experience and perceive that behavior, and factors that moderate this impact (such as variability) will also occur via the implications those factors have for individuals’ experiences and perceptions. Additional analyses demonstrated that perceptions of partners’ negative behavior corresponded to partners’ reports of that behavior. Multilevel analyses of the daily data revealed that perceptions of partners’ behavior tracked partners’ reports of their behavior across days (B = .41, 95% CI [.33, .49], t = 10.16, p < .001 and B = .49, 95% CI [.38, .60], t = 9.03, p < .001 for Study 1 and 2 respectively). Once aggregated, the mean and variability of perceptions of partners’ negative behavior were also correlated with the mean and variability of partners’ reported negative behavior (mean r = .44, p < .001 and variability r = .49, p < .001 pooled across Study 1 and 2). Nonetheless, consistent with the proposition that any longitudinal effects will ultimately result from how partners’ behavior is experienced or perceived, partners’ reported negative behavior did not reveal a main effect (B = 5.17, 95% CI [1.60, 11.95], t = 1.51, p = .133) or interaction effect with variability (B = .88, 95% CI [−10.46, 8.71], t = −1.18, p = .857) on relationship problems across time (integrated data analyses on pooled sample).
and were involved in serious romantic relationships (41% married or cohabiting) that averaged 3.35 years in length (SD = 3.94).

**Materials and procedure.** During an initial session, couples completed scales assessing the severity of problems in their relationship and their relationship satisfaction. Participants were then given detailed instructions for completing a 3-week daily sampling procedure involving completing an end-of-day record rating the degree to which they and their partner behaved negatively and positively toward each other that day. Nine months after completing the daily sampling procedure, participants completed an online questionnaire that included the same scales of problem severity and relationship evaluations assessed initially.

**Questionnaire assessments.**

- **Problem severity.** Participants completed the Marital Problems Inventory (Geiss & O’Leary, 1981), which asks participants to rate a list of 25 common points of disagreement according to how much each area is a problem in their relationship (e.g., communication, solving problems, power struggles, showing affection, sex; physical abuse; unrealistic expectations; 1 = not a problem, 7 = major problem). Given that couples can experience serious difficulties in some areas and little or no problems in other areas, the scores of all 25 items in the inventory are summed to assess overall severity of relationship problems (Geiss & O’Leary, 1981). Possible scores range from 25 to 175 (see Table 1), and greater problem severity scores are an established predictor of poorer relationship satisfaction across time (Lavner, Karney, Williamson, & Bradbury, 2017; McNulty & Russell, 2010).

- **Relationship satisfaction.** In Study 1, relationship satisfaction was assessed using the 5-item satisfaction subscale of Rusbult, Martz, and Agnew’s (1998) investment scale (e.g., “I feel satisfied with our relationship”; “My relationship is close to ideal”; 1 = strongly disagree, 7 = strongly agree). Items were averaged at both time points (as > .85). In Study 2, participants completed the seven-item Perceived Relationship Quality Components inventory at both assessments (Fletcher, Simpson, & Thomas, 2000), which measures the quality of relationships across similar dimensions, such as satisfaction, intimacy, trust and love (e.g., “How satisfied are you with your relationship?”; “How close is your relationship?”; 1 = not at all, 7 = extremely). Items were averaged at both time points (as > .81; see Table 1).

Relationship satisfaction assessments represent the only case in which measures were not identical across samples, but this situation does not rule out integrative data analysis if the measures assess similar constructs (see Hussong et al., 2013). The format and the length of the scales are similar, and both scales are widely used to assess relationship sentiments. Thus, the integrative data analysis approach was cautiously applied to models predicting relationship satisfaction for completeness, and because this approach tests for differences across studies that might arise due to differences in measurement (also see Footnote 7).

- **Attachment insecurity and self-esteem.** At the initial session, participants completed the Adult Attachment Questionnaire (AAQ; Simpson, Rholes, & Phillips, 1996) to assess attachment anxiety (M = 2.94, SD = 1.01 for Study 1; M = 2.95, SD = 1.12 for Study 2) and avoidance (M = 2.92, SD = 1.04 for Study 1; M = 2.83, SD = .91 for Study 2). Participants also completed the Rosenberg (1965) Self-Esteem Scale (M = 5.43, SD = .98 for Study 1; M = 5.50, SD = 1.06 for Study 2). These measures were used in additional analyses to ensure that the predicted effects were not attributable to relationship insecurities that produce more negative relationship perceptions and behaviors (see Murray, Holmes, & Collins, 2006; Simpson & Rholes, 2012).

- **Sampling of partners’ daily negative-direct behavior.** At the end of each day for 21 consecutive days, both dyad members completed a web-based record reporting on their (a) perceptions of their partner’s negative behavior, and (b) own negative behavior toward their partner each day. Participants completed an average of 19.6 (Study 1) and 19.5 (Study 2) of the requested 21 daily records.

- **Daily negative-direct behavior.** Each day, participants rated two items assessing how much their partner exhibited negative-direct behavior: “My partner was critical or unpleasant toward me” and “My partner acted in a way that was hurtful to me.” These items were intended to be general to capture the various ways that relevant negative-direct behaviors can be expressed in daily contexts and in a way that participants could easily reflect on with

### Table 1

**Studies 1–3: Descriptive Statistics of Dependent Measures**

<table>
<thead>
<tr>
<th>Dependent measures</th>
<th>Initial assessments</th>
<th>3-month</th>
<th>6-month</th>
<th>9-month</th>
<th>12-month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Study 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem severity</td>
<td>52.27</td>
<td>22.38</td>
<td>49.94</td>
<td>13.33</td>
<td></td>
</tr>
<tr>
<td>Relationship satisfaction</td>
<td>6.11</td>
<td>.69</td>
<td>5.92</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>Study 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem severity</td>
<td>44.97</td>
<td>13.42</td>
<td>48.41</td>
<td>18.24</td>
<td></td>
</tr>
<tr>
<td>Relationship satisfaction</td>
<td>6.33</td>
<td>.49</td>
<td>6.22</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Integrated data analysis: Study 1 and 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem severity</td>
<td>47.29</td>
<td>13.57</td>
<td>50.22</td>
<td>20.32</td>
<td></td>
</tr>
<tr>
<td>Relationship satisfaction</td>
<td>6.23</td>
<td>.60</td>
<td>6.08</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td>Study 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem severity</td>
<td>4.93</td>
<td>1.51</td>
<td>3.72</td>
<td>1.62</td>
<td>3.39</td>
</tr>
<tr>
<td>Relationship satisfaction</td>
<td>5.85</td>
<td>.92</td>
<td>5.83</td>
<td>1.04</td>
<td>5.79</td>
</tr>
</tbody>
</table>

**Note.** In Studies 1 and 2, problem severity was assessed by rating the severity of an inventory of common problems across relationships, with possible scores ranging from 25 to 175. In Study 3, problem severity was assessed by rating the severity of a specific problem couples discussed during an observed conflict discussion, with possible scores ranging from 1 to 7. Relationship satisfaction assessments across studies represent averages across items on 1–7 Likert-type scales, with possible scores ranging from 1 to 7.
regard to their partner’s behavior across the day (Overall & Sibley, 2009a, 2009b; Overall & Sibley, 2010; Overall et al., 2010; for examples of similar daily assessments see Lemay, Lin, & Muir, 2015; Liu et al., 2018; Timmons et al., 2017). The specific terms used were derived from established continuous scales and typologies capturing this category of behaviors through conflictual exchanges across relationship life, including unpleasant (Rusbult, Verette, Whitney, Slovik, & Lipkus, 1991), hurtful (Kerig, 1996), and critical (Canary, Cunningham, & Cody, 1988; Christensen & Heavey, 1990). These items have been used extensively in prior studies to assess negative-direct behavior, which has validated that this daily assessment captures the behaviors targeted by established scales (Overall & Sibley, 2010; also Overall & Sibley, 2009b; Overall et al., 2010) and replicates the effects of negative-direct behavior assessed during couples’ observed conflict discussions (e.g., Cross, Overall, Hammond, & Fletcher, 2017; Cross, Overall, Low, & McNulty, 2019; Overall, Fletcher, Simpson, & Filo, 2015; Overall, Hammond, McNulty, & Finkel, 2016).

The two items assessing partners’ behavior (r = .70 for Study 1 and .77 for Study 2) were highly correlated in both studies and averaged to index daily assessments of partners’ negative-direct behavior. Participants rated the same items reporting on their own behavior toward their partner each day (e.g., “I was critical or unpleasant toward my partner”), which were highly correlated in both studies (r = .69 for Study 1 and .76 for Study 2) and averaged to index daily assessments of individuals’ own negative-direct behavior.

**Daily positive behavior.** Participants also reported on their partners’ positive relationship behavior, which was used in analyses to rule out the possibility that the predicted effects of variability in partners’ negative-direct behavior was due to the presence of more positive behavior. Participants rated four items that capture daily positive behaviors proposed to be constructive ways of managing negativity in relationships (Rusbult et al., 1991; also see Overall & Sibley, 2009b, 2010), and have been shown in prior research to be associated with improved daily (Laurenceau, Barrett, & Pietromonaco, 1998) and longitudinal relationship outcomes (Karney & Bradbury, 1995b; Overall, 2018; Overall & McNulty, 2017; Sullivan, Pasch, Johnson, & Bradbury, 2010): “My partner was affectionate and loving toward me,” “My partner tried to maintain or improve the quality of our relationship,” “My partner shared and discussed his/her feelings and opinions with me,” and “My partner supported me” (1 = not at all, 7 = extremely). These items have been used to index positive daily behavior in prior research using data from Study 1 (Overall & Hammond, 2013), and are ways in which partners may potentially repair or balance negative relationship behavior (Rusbult et al., 1991; Gottman, 1998). Participants rated the same items reporting on their own behavior toward their partner each day (e.g., “I was affectionate and loving toward my partner,” “I tried to maintain or improve the quality of our relationship”). The items were averaged to index daily assessments of (a) partners’ positive behavior (α = .86 for Study 1 and .68 for Study 2) and (b) individuals’ own positive behavior (α = .83 for Study 1 and .63 for Study 2).

**Results**

All analyses were conducted separately for Study 1 and 2, and then integrated data analyses conducted, which pools the data across studies to maximize statistical power and directly tests whether the effects replicate across studies (Curran & Hussong, 2009; Hussong et al., 2013). Integrative data analyses provides more reliable and stable estimation of effects, particularly when there may be relatively low base rates of behavior, such as partners’ negative behavior in daily life, and because it reduces the potential for extreme observations to influence the data, which may occur when sample sizes are inevitably constrained, such as when using dyadic experiencing sampling procedures and gathering longitudinal assessments (Curran & Hussong, 2009; Hussong et al., 2013). Relevant to these benefits of integrative data analysis, the first section of the results outlines how the mean and variability of partners’ behavior was calculated and considers the distribution of these primary variables. The second section presents the results of the longitudinal effects of mean, variability, and Mean × Variability of partners’ daily negative behavior for Studies 1 and 2 and then the integrated data analysis. A final section examines the effects of potential alternative explanations.

**Calculating mean and variability of partners’ daily negative behavior.** Participants repeated ratings of partners’ daily behavior was used to create two measures. First, the person mean was calculated by averaging negative-direct behavior scores reported across the 21-day sampling period. Second, the within-person standard deviation was calculated to represent the relative stability (lower SD) versus variability (higher SD) in partners’ negative-direct behavior across couples’ daily life (e.g., Benson et al., 2019; Blanke et al., 2019; Eldesouky & English, 2018; Girme et al., 2018; Gruber et al., 2013; Houben et al., 2015; Koval et al., 2013).

Table 2 presents the descriptive statistics of these behavioral measures, including information about the characteristics of the distribution. As in prior research assessing variability in emotions, relationship evaluations, and emotion regulation, in both studies there was considerable range of variability in partners’ negative behavior that was normally distributed. Mean levels of partners’ negative behavior also demonstrated substantial range, given that high scores require the presence of high levels of negative behavior each day. Mean scores were normally distributed in Study 1, but less so in Study 2, in which more people experienced low levels of partners’ negative behavior across the 21-day sampling period. Demonstrating one of the advantages of integrated data

---

4 Given the affective nature of negative-direct behavior, and the emotional elements within the assessments of these behaviors, readers may wonder whether the behavioral effects presented in the paper are independent of the variability of partners’ or individuals’ negative emotions. In Studies 1 and 2, participants also completed four items assessing perceptions of their partners’ negative emotions (e.g., “My partner was angry at me”) and their own negative emotions (e.g., “I was hurt by my partner”; 1 = not at all, 7 = very much). Additional analyses rerunning the primary analyses simultaneously modeling the mean and variability of partners’ and (in separate analyses) individuals’ own negative emotions demonstrated that the effects of partners’ negative-direct behavior were distinct from partners’ and individuals’ negative emotions, which did not have any independent longitudinal effects (see the online supplemental materials for results). Similarly, in Study 3, participants reviewed their recorded discussions and, for every 30-second of the discussion, rated their experience of negative emotions during that portion of the discussion (e.g., angry, hurt; 1 = not at all, 7 = extremely). As in Studies 1 and 2, additional analyses revealed that the longitudinal effects of variability in partners’ negative-direct behavior were distinct from the mean and variability of partners’ and individuals’ negative emotions, which revealed no independent longitudinal effects (see the online supplemental materials for results).
analyses, skewness and kurtosis for mean levels of partners’ negative behavior were acceptable in the data pooled across Studies 1 and 2 (Gravetter & Wallnau, 2014).

An important consideration is the inevitable confound between mean levels and variability. Mean behavior scores near average levels have more room to vary relative to low or high mean behavior scores, such that very low mean levels of behavior will have very low or no variability and high mean behavior scores will also vary less. The result is an artifactually inflated positive correlation between mean level and variability which is evident in the associations between mean levels and variability in behavior shown in Figure 2.

Of importance, inspection of the scatterplots in Figure 2 reveals a reasonable distribution of low versus high levels of variability when partners were perceived to have mean or higher levels of negative-direct behavior, which is the area of contrast that is the focus of this investigation (rather than no or low levels of negative behavior where the confound most strongly occurs). Moreover, the scatterplots offer another illustration of the distributional benefits that occur when increasing observations by integrating the samples for analyses, thereby reducing the influence of any extreme observations (see bottom left of Figure 2).

Finally, the artificially inflated correlations across mean levels and variability, which is evident in all studies assessing variability (as reviewed across this paper), is dealt with in the literature by modeling a heterogeneous compound symmetry error structure (see the online supplemental materials for annotated syntax). The primary analyses involved testing whether the longitudinal associations across days between perceptions of partners’ negative-direct behavior and relationship problems across time was moderated by the variability of partners’ negative-direct behavior across the daily sampling procedure. To do this, the (a) mean of partners’ negative-direct behavior (mean-centered), (b) SD of partners’ negative-direct behavior (mean-centered), and the (c) interaction between mean and SD of partners’ negative-direct behavior were entered as predictors of individuals’ reports of problem severity assessed 9 months following the daily sampling procedure controlling for initial problem severity scores assessed prior to the daily sampling. Problem severity at the follow-up 9-month assessment was centered on initial levels of problem severity so that the predicted values represent whether problem severity increased, reduced or remained the same across the year. This centering strategy alters the intercept only, and the main and interaction effects of the mean and SD of partners’ negative-direct behavior were identical when predicting uncentered problem severity scores.

The results analyzing each study separately are shown in the first column of Table 3. In both studies, greater levels of partners’ negative behavior predicted greater relationship problems across time (see main effect of mean levels in boldface). However, the

5 One approach to this problem is to exclude participants who reported partners did not exhibit any negative-direct behavior across the entire daily sampling period, and therefore could not have variability in perceptions of partner behavior across daily life. This included only four participants in Study 1, but 14 participants in Study 2, which contributed to the skewed distribution for that sample. The longitudinal effects were stronger when excluding these participants, and in particular the pattern was clearer in Study 2 because the predicted values from the interaction were less skewed at low mean levels (see Figure 3). However, as discussed next, the integrated data analysis approach reduces the effect of extreme observations without excluding data that represent possible values in people’s actual experiences (i.e., no negative behavior by partners).

6 The main effects of gender were also modelled across analyses because in these heterosexual samples gender was the distinguishing factor across dyad members (see the online supplemental materials). Additional analyses including the interaction effects of gender revealed there were no gender differences in the effects of mean, SD or Mean × SD of partners’ negative-direct behavior presented in Table 2 (all ts > 1.68, ps > .098).
significant interaction between mean and variability of partners’ negative behavior also emerged in both studies (see interaction effects in boldface). As shown in Figure 3, perceiving higher mean levels of partners’ daily negative-direct behavior predicted growing problem severity across time (scores above zero), but this damaging effect was significantly more pronounced when partners’ negative-direct behavior was relatively stable across daily life (low SD; slope = 25.55, 95% CI [13.13, 37.83], t = 4.16, p < .001 for Study 1 and slope = 25.73, 95% CI [9.15, 42.32], t = 3.10, p = .002 for Study 2) compared with when partners’ negative-direct behavior varied across daily life (high SD; slope = 10.85, 95% CI [2.79, 18.92], t = 2.69, p = .008 for Study 1 and slope = 12.12, 95% CI [4.10, 20.14], t = 3.02, p = .003 for Study 2).

The results from the integrative data analysis on the pooled study data are shown in the first column of Table 4. These analyses apply the same models to the pooled data across studies, and include the main and interaction effects of study membership (−1 = Study 1, 1 = Study 2) to test whether the effects were the same or significantly differed across each study (effect coding ensured that the estimates and predicted values in Figure 3 represent the effects across studies). There were no significant differences across studies in any of the parameters (see bottom of Table 4, all ps > .53). This analyses revealed a clear pattern that supported the predicted effects of behavioral variability versus stability (see bottom panel of Figure 3).

Perceiving greater mean levels of partners’ negative-direct behavior predicted greater problem severity across time, but the significant interaction between the mean and SD illustrated that this longitudinal association depended on the variability of partners’ behavior across daily life (see bold interaction in Table 4). As shown in the bottom panel of Figure 3, perceiving higher mean levels of partners’ daily negative-direct behavior predicted growing problems across time, but this damaging effect was greater when partners’ negative-direct behavior was relatively stable across daily life (low SD; slope = 25.40, 95% CI [14.74, 36.06], t = 5.33, p < .001) compared with when partners’ negative-direct
behavior was more variable (high SD; slope = 11.21, 95% CI [5.47, 16.93], t = 3.91, p < .001).

The same analytic strategy was used to test whether the mean and variability in partners' negative-direct behavior predicted relationship satisfaction across time. As shown in the right side of Table 3, despite the replicated effects on problem severity across Studies 1 and 2, mean levels of partners' negative-direct behavior did not predict changes in relationship satisfaction across time in Study 1 or 2 regardless of the stability or variability of partners' behavior. As shown in the right side of Table 4, the integrative data analysis also revealed that the mean and interaction effects did not differ across studies, and did not predict changes in relationship satisfaction.7 The meaning of these null effects is considered in the discussion of Study 1 and 2, and more powerful analysis of change in relationship satisfaction across time applied in Study 3.

**Additional analyses and alternative explanations.** A series of additional analyses were run to rule out alternative explanations for the effects of mean and variability of partners' negative-direct behavior on relationship problems across time. Given the integrative data analysis revealed that the effects did not differ across Study 1 and 2, and that integrative data analysis maximizes power, improves the stability of model estimation, and thus provides greater capacity for more complex models (see Curran & Hussong, 2009; Hussong et al., 2013), these additional analyses were conducted on the data pooled across Studies 1 and 2. Following Hussong et al. (2013), the nonsignificant interactions with study membership were trimmed from the models, which is maximally conservative, maintains parsimony, and supports stability in model estimation when additional parameters are needed for control analyses. As with the primary analyses, additional models revealed there were no study differences and the conclusions below were supported examining each study separately.

Each additional analyses involved two stages: (a) examining whether alternative factors could potentially explain the effects (e.g., produced the same effects as partners' negative-direct behavior), and then (b) testing whether the significant longitudinal effects of mean and Mean × SD of partners’ negative behavior presented in Table 4 and Figure 3 remained when controlling for rival alternative explanations (see Table 5).

**Own negative-direct behavior.** The aims and predictions focused on mean levels and variability of receiving partners' negative-direct behavior. To illustrate that these results were a function of individuals’ experience of their partners’ behavior, rather than individuals’ own negative-direct behavior, the model presented in Table 4 was rerun modeling the mean and variability of individuals’ own negative-direct behavior across the daily sampling period predicting change in relationship problems across time. Greater levels of individuals’ own negative-direct behavior predicted greater problems (B = 8.63, t = 3.11, p = .002), and a marginal interaction provided weaker support that this effect was reduced when individuals’ behavior was more variable (B = −5.95, t = −1.82, p = .07). Moreover, as shown in Table 5, modeling the mean, SD, and mean × SD interaction of both individuals’ own and partners’ negative-direct behavior did not change the significant longitudinal effects of partners’ daily negative-direct behavior (interaction effects for comparison presented in boldface in Table 5).

**Partners’ positive behavior.** A potential alternative explanation of the effects is that variability in partners’ negative behavior simply or only occurs when partners also behave more positively, such as enacting repair or relationship maintenance attempts on or following days partners behave negatively. Additional analyses did not support this alternative explanation. First, mean, variability and mean × variability of partners’ negative behavior did not predict positive behavior across the daily sampling period (ts < −1.02, ps > .30). Second, replacing variability in partners’ negative behavior with the form B = 8.63, t = 3.11, p = .002, we have a different outcome. This suggests that the variability of partners’ negative behavior is not just a by-product of their positivity, but a potential mediator of the effects on relationship satisfaction.
behavior with partners’ positive behavior when predicting relationship problems across time revealed that greater positive behavior did not attenuate the damage of partners’ negative behavior (M Partner Negative Behavior × M Partner Positive Behavior → problems: $B = -0.93$, $t = -0.45$, $p = .65$), which continued to predict growing problems across time (M partner negative behavior → problems across time: $B = 10.27$, $t = 4.80$, $p < .001$), indicating that the longitudinal effects shown in Figure 3 did not reflect the presence of greater positive behavior by the partner. Finally, modeling the mean, $SD$, and Mean × $SD$ interaction of both partners’ negative-direct and positive behavior revealed that partners’ positive behavior had no independent longitudinal effects, whereas the longitudinal effects of partners’ daily negative-direct behavior were unchanged (see Table 5).

**Own positive behavior.** Perhaps, instead of greater positive behavior by the partner, variability in partners’ negative behavior arises from individuals’ own repair or relationship maintenance efforts. Again, additional analyses did not support this alternative explanation. First, mean, variability, and Mean × Variability of partners’ negative behavior was not associated with individuals’ own positive behavior across the daily sampling period ($ts < -1.25$, $ps > .35$). Second, replacing variability in partners’ negative behavior with individuals’ own positive behavior when predicting relationship problems across time revealed that individuals’ own positive behavior did not attenuate the damage of partners’ negative behavior (M Partner Negative Behavior × M Partner Positive Behavior → problems: $B = -1.30$, $t = -0.61$, $p = .54$), which continued to predict growing problems across time (M partner negative behavior → problems across time: $B = 10.31$, $t = 4.95$, $p < .001$). Finally, modeling the mean, $SD$, and Mean × $SD$ interaction of both partners’ negative-direct behavior and individuals’ own positive behavior revealed that individuals’ positive behavior had no effects, whereas the main and interaction effects of partners’ daily negative-direct behavior on relationship problems across time were unchanged (see Table 5).

**Attachment insecurity and self-esteem.** Finally, variable versus stable behavioral patterns, or perceptions of these patterns, could emerge from individual differences that may create destructive and rigid behavior, such as attachment insecurity or low self-esteem. Although individuals own or partners’ self-esteem did not predict mean and variability of partners’ negative-direct behavior ($ts < 1.04$, $ps > .30$), individuals higher in attachment anxiety and avoidance perceived their partners to behave more negatively (anxiety: $B = .10$, $t = 2.46$, $p = .015$; avoidance: $B = .15$, $t = 3.77$, $p < .001$), and individuals higher in attachment avoidance perceived their partners’ negative-direct behavior to be more stable across the daily sampling period (controlling for average levels: $B = -.05$, $t = -2.29$, $p = .023$). Partners’ attachment anxiety was also positively associated with greater perceptions of partners’ negative behavior ($B = .08$, $t = 2.02$, $p = .045$). Nonetheless, the longitudinal effects of partners’ daily negative-direct behavior presented in Table 4 and Figure 3 remained strong and significant when controlling for individuals’ own or their partners’ attachment insecurity or self-esteem (see the online supplemental materials for longitudinal effects controlling for attachment insecurity and self-esteem).

---

**Figure 3.** The effects of mean levels and variability of perceptions of partners’ daily negative-direct behavior on problem severity nine months following the daily sampling (controlling initial levels of problem severity). $SD$ = standard deviation. Problem severity at the follow-up 9-month assessment was centered on initial levels of problem severity so that the predicted values represent whether problem severity increased, reduced or remained the same across the 9-month follow-up period.
Table 4
Integrated Data Analyses of Studies 1 and 2: The Effects of Mean and Standard Deviation of Perceptions of Partners’ Negative-Direct Behavior Across Couples’ Daily Lives on Problem Severity and Relationship Satisfaction 9 Months Later

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Predictors</th>
<th>Problem severity</th>
<th>Relationship satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>95% CI</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>6.29</td>
<td>[2.52, 10.06]</td>
</tr>
<tr>
<td>Initial problems/satisfaction</td>
<td></td>
<td>−0.89</td>
<td>[−1.08, −0.70]</td>
</tr>
<tr>
<td>M of partners’ negative behavior</td>
<td></td>
<td>18.30</td>
<td>[11.11, 25.49]</td>
</tr>
<tr>
<td>SD of partners’ negative behavior</td>
<td></td>
<td>−3.59</td>
<td>[−10.66, −3.47]</td>
</tr>
<tr>
<td>M × SD of partners’ negative behavior</td>
<td></td>
<td>−13.05</td>
<td>[−21.17, −4.93]</td>
</tr>
<tr>
<td>Study</td>
<td></td>
<td>1.00</td>
<td>[−2.77, 4.76]</td>
</tr>
<tr>
<td>Study × Initial problems/satisfaction</td>
<td></td>
<td>−0.06</td>
<td>[−0.25, 0.13]</td>
</tr>
<tr>
<td>Study × M of partners’ negative behavior</td>
<td></td>
<td>−0.74</td>
<td>[−7.93, 6.45]</td>
</tr>
<tr>
<td>Study × SD of partners’ negative behavior</td>
<td></td>
<td>1.74</td>
<td>[−5.33, 8.80]</td>
</tr>
<tr>
<td>Study × M × SD of partners’ negative behavior</td>
<td></td>
<td>0.04</td>
<td>[−1.12, 2.46]</td>
</tr>
</tbody>
</table>

Note. The significant longitudinal effects are presented in bold. The M × SD interaction is displayed in Figure 3 (bottom panel). CI = confidence interval. Effect sizes (r) were computed using Rosenthal and Rosnow’s (2007) formula: $r = \sqrt{t^2 / df}$. In these multilevel models, the Satterthwaite approximation is applied to provide specific degrees of freedom for each effect, which were used to calculate the effect sizes.

Discussion

Studies 1 and 2 involved two daily sampling studies using identical procedures. In both studies, receiving higher levels of partners’ negative-direct behavior predicted growing relationship problems, but this longitudinal association depended on the variability of partners’ negative-direct behavior across daily life. Integrative data analyses revealed that there were no differences in the effects across studies, and supported the hypothesized pattern: perceiving partners to behave in more unpleasant, critical, and hurtful ways predicted growing relationship problems 9-months later, but this harmful longitudinal effect was more pronounced when high levels of partners’ negative-direct behavior arose from a stable pattern of consistent negativity (high mean levels + low within-person SD across days), and less pronounced when partners’ negative-direct behavior varied across the daily sampling period (high mean levels + high within-person SD across days). Additional analyses demonstrated that these effects were due to the partners’ negative behavior, and not individuals’ own negative behavior, partners’ or individuals’ positive behavior, or relationship insecurities.

These results advance prior research identifying a range of between-person factors that modify the harmful effects of negative-direct behavior. Studies 1 and 2 represent the first application of techniques used with daily sampling data to capture within-person variability, and thus index variable behavioral patterns that would be expected if partners’ behavior was sensitive to varying contexts, events and situational demands across daily life versus stable behavioral patterns that are likely nondiagnostic of current daily demands. However, daily sampling methods rely on self-reports that could be subject to biases. Moreover, although daily sampling is particularly beneficial in capturing variation in a range of important contexts and situations that couples will encounter across daily life, this method does not precisely capture specific contexts in which partners’ negative-direct behavior has

Table 5
Integrated Data Analyses of Studies 1 and 2: The Effects of Mean and Standard Deviation of Perceptions of Partners’ Negative-Direct Behavior Across Couples’ Daily Lives on Problem Severity Controlling for Alternative Behaviors

<table>
<thead>
<tr>
<th>Controlling for alternative behaviors</th>
<th>Effects of alternative variable</th>
<th>Effects of partners’ negative behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>95% CI</td>
</tr>
<tr>
<td>Own negative behavior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M of behavior</td>
<td>−10.65</td>
<td>[−19.39, −1.92]</td>
</tr>
<tr>
<td>SD of behavior</td>
<td>8.96</td>
<td>[0.94, 16.97]</td>
</tr>
<tr>
<td>M × SD of behavior</td>
<td>4.64</td>
<td>[−3.18, 12.46]</td>
</tr>
<tr>
<td>Partners’ positive behavior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M of behavior</td>
<td>−1.72</td>
<td>[−4.28, 0.84]</td>
</tr>
<tr>
<td>SD of behavior</td>
<td>0.27</td>
<td>[−6.67, 7.20]</td>
</tr>
<tr>
<td>M × SD of behavior</td>
<td>2.35</td>
<td>[−2.56, 7.27]</td>
</tr>
<tr>
<td>Own positive behavior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M of behavior</td>
<td>−2.03</td>
<td>[−4.70, 0.63]</td>
</tr>
<tr>
<td>SD of behavior</td>
<td>−1.33</td>
<td>[−8.41, 5.74]</td>
</tr>
<tr>
<td>M × SD of behavior</td>
<td>1.18</td>
<td>[−4.15, 6.52]</td>
</tr>
</tbody>
</table>

Note. The longitudinal effects presented in bold correspond to the significant longitudinal effects displayed in Table 4 and Figure 3 (bottom panel) controlling for the longitudinal effects of each alternative variable. CI = confidence interval. Effect sizes (r) were computed using Rosenthal and Rosnow’s (2007) formula: $r = \sqrt{t^2 / df}$. In these multilevel models, the Satterthwaite approximation is applied to provide specific degrees of freedom for each effect, which were used to calculate the effect sizes.
been shown to be less harmful and potentially beneficial. Thus, Study 3 tested the longitudinal effects of variable versus stable patterns of partners’ behavior by gathering observational assessments of partners’ criticism and hostility in a specific situation in which partners’ negative-direct behavior can improve relationship problems—couples’ conflict discussions.

Moreover, despite the replicated longitudinal effects on relationship problems, mean levels and variability of partners’ negative behavior did not have significant longitudinal effects on relationship satisfaction in Studies 1 and 2. Problem severity represents a primary mechanism for how and when partners’ negative-direct behavior affects satisfaction across time (see Jayamaha & Overall, 2015; McNulty, 2010, 2016; McNulty & Russell, 2010; Overall et al., 2009; Overall & McNulty, 2017). Thus, the replicated effects on relationship problems across time are likely to have negative implications for relationship satisfaction, but the direct effects on satisfaction may have been diluted across time as is common in longitudinal designs when effects are transmitted through intervening variables, such as severity of relationship problems (Shrout & Bolger, 2002). The measure of relationship problems was also more extensive than relationship satisfaction in Study 1 and 2, which may have provided a more reliable assessment of change when relying on only one follow-up assessment. Only assessing change across two time points may have reduced the power to detect longitudinal change in satisfaction for relatively satisfied couples who remained intact across the 9-month assessment period. Study 3 involved a more powerful design involving repeated longitudinal assessments of relationship satisfaction in a larger sample of couples.

**Study 3**

In Study 3, couples were video-recorded discussing two relationship problems arising from one partner (the agent of desired change) desiring some type of change in the other partner (the target of desired change). Prior research has shown that partners’ criticism and hostility can have beneficial effects when partners are agents of change because a negative-direct approach can motivate targets to change serious problems (McNulty & Russell, 2010; Overall et al., 2009), signal relationship investment to targets (Overall, 2018), and thus facilitate the satisfaction of both targets to change serious problems (McNulty & Russell, 2010; Overall et al., 2009; Overall & McNulty, 2017). Thus, the replicated effects on relationship problems across time are likely to have negative implications for relationship satisfaction, but the direct effects on satisfaction may have been diluted across time as is common in longitudinal designs when effects are transmitted through intervening variables, such as severity of relationship problems (Shrout & Bolger, 2002). The measure of relationship problems was also more extensive than relationship satisfaction in Study 1 and 2, which may have provided a more reliable assessment of change when relying on only one follow-up assessment. Only assessing change across two time points may have reduced the power to detect longitudinal change in satisfaction for relatively satisfied couples who remained intact across the 9-month assessment period. Study 3 involved a more powerful design involving repeated longitudinal assessments of relationship satisfaction in a larger sample of couples.

In Study 3, couples were video-recorded discussing two relationship problems arising from one partner (the agent of desired change) desiring some type of change in the other partner (the target of desired change). Prior research has shown that partners’ criticism and hostility can have beneficial effects when partners are agents of change because a negative-direct approach can motivate targets to change serious problems (McNulty & Russell, 2010; Overall et al., 2009), signal relationship investment to targets (Overall, 2018), and thus facilitate the satisfaction of both targets to change serious problems (McNulty & Russell, 2010; Overall et al., 2009; Overall & McNulty, 2017). Thus, the replicated effects on relationship problems across time are likely to have negative implications for relationship satisfaction, but the direct effects on satisfaction may have been diluted across time as is common in longitudinal designs when effects are transmitted through intervening variables, such as severity of relationship problems (Shrout & Bolger, 2002). The measure of relationship problems was also more extensive than relationship satisfaction in Study 1 and 2, which may have provided a more reliable assessment of change when relying on only one follow-up assessment. Only assessing change across two time points may have reduced the power to detect longitudinal change in satisfaction for relatively satisfied couples who remained intact across the 9-month assessment period. Study 3 involved a more powerful design involving repeated longitudinal assessments of relationship satisfaction in a larger sample of couples.

By contrast, partners who exhibit rigid and inflexible negative-direct behavior across the discussion may fail to provide diagnostic communication that conveys investment and motivates change in targeted problems. Instead, highly stable negative-direct behavior may simply express global hostility that is nondiagnostic of what and how problems need to change, generally demotivate targets, and thus maintain problems and reduce satisfaction. Thus, as in daily life, a stable behavioral pattern of partners’ negative-direct behavior (high mean levels + low within-person SD; see solid lines in Figure 2) should be more harmful.

To determine whether the longitudinal effects of partners’ negative-direct behavior depended on the relative variability versus stability of the partners’ behavior, trained coders rated the degree to which each partner exhibited hostile communication every 30-s of the conflict discussion. Following the procedures in Studies 1 and 2, the repeated assessments of hostile communication were used to calculate (a) average levels of each partner’s negative-direct behavior (the person mean across the discussion), and (b) variability in each partner’s negative-direct behavior (the within-person SD) across the discussion. As in Studies 1 and 2, the primary analyses tested whether the degree to which partners’ negative-direct behavior predicted longitudinal changes in problem severity and relationship satisfaction depended on whether that behavior arose from a variable temporal pattern (high within-person SD) versus an invariable temporal pattern (low within-person SD).

Prior to couples’ conflict discussion, and then via post every three months for the following year, participants rated the severity of the problem discussed and completed assessments of relationship satisfaction. These repeated assessments allowed the application of growth curve analyses to model the trajectory of change in problem severity and satisfaction across time. Assessing partner behavior in a specific context, measuring the severity of a specific relationship problem, and modeling trajectories of change across
time rather than predicting single longitudinal assessments, mean that the longitudinal pattern representing more versus less harmful outcomes differs than that in Studies 1 and 2.

First, the severity of specific problems discussed in the laboratory typically decline across time as couples improve or accommodate specific issues (Overall et al., 2009). Thus, compared with Studies 1 and 2 in which more harmful outcomes were indexed by greater increases in problem severity, a poor outcome in growth curve models predicting changes in severity of a specific problem discussed in the laboratory are shown by problem severity remaining stable versus reducing across time (also see McNulty & Russell, 2010). Second, trajectories of repeated assessments of satisfaction across time tend to show a negative or declining trajectory on average. Thus, more harmful outcomes are indexed by stronger reductions in satisfaction whereas less harmful or beneficial outcomes reflect flat trajectories showing maintenance of satisfaction across the year (Lavner & Bradbury, 2010).

In the context of these general trajectory patterns, the prediction that high mean levels of partners’ negative-direct behavior would be more harmful when arising from a stable behavioral pattern would be supported if high mean levels and low SD of partners’ negative-direct behavior predicted maintenance of problem severity and reductions in relationship satisfaction. By contrast, the prediction that partners’ negative-direct behavior would be less harmful when arising from a variable behavioral pattern would be supported if high mean levels and high SD of partners’ negative-direct behavior predicted reductions in problem severity and maintenance of relationship satisfaction.

Method

Participants. Study 3 was designed to overcome the limitation of typically small samples of dyads in behavioral observation and longitudinal studies by recruiting a large sample of couples via paper and electronic announcements posted across a large city-based university and associated organizations (e.g., employment agencies and health centers). Of the original sample of 180 heterosexual couples, 161 couples (N = 322 individuals) provided longitudinal data, with each couple providing data from two conflict discussions, resulting in a sample that is considerably larger than most prior longitudinal behavioral observation studies (see the online supplemental materials for more sample information). Couples were mostly married (43%) or cohabiting (32%), with the remainder in serious dating relationships. Mean relationship length was 3.01 years (SD = 2.23), and mean age was 23.21 years (SD = 4.27). Couples were paid NZ$70 for the initial laboratory-based session, and then NZ$30 for each follow-up assessment. This procedure was designed to differentiate conflict roles in serious problems involving one partner (the agent) wanting change in the other (the target; also see Heyman & Hunt, 2007; Klinetob & Smith, 1996). Prior research has shown that partners’ negative-direct behavior as agents can have less harmful or more beneficial longitudinal effects. The current study focuses on whether the variability of partners’ behavior moderates these outcomes.

Following a warm-up discussion, couples had two 7-min discussions involving (a) the most serious problem the male partner identified, and (b) the most serious problem the female partner identified (order counterbalanced across couples). Couples were instructed to talk about the issue as they normally would. The problems discussed reflected ongoing issues that couples had already discussed, and participants rated the discussions as typical of the way they normally discussed these and similar issues (see the online supplemental materials for further details).

During the next 12 months, each couple member was contacted four times at 3-month intervals via e-mail and/or telephone to assess the status of their relationship (intact vs. dissolved). Participants from intact couples were posted a questionnaire consisting of the same scales used at the initial session to assess problem severity and relationship satisfaction. Sixteen couples dissolved during the year, but the majority of participants completed all four longitudinal assessments (67.5%), with smaller groups completing three (10%), two (13.5%), or one follow-up (9%). Power was maximized by including all longitudinal data. The multilevel analyses used to test the longitudinal effects reliably assess within-person change when number of data points vary and accounts for sample attrition by weighting individual parameter estimates by their reliability so that participants with more data points contribute more to the final sample-level estimates (see Karney & Bradbury, 1995a).

Self-report assessments.

Relationship satisfaction. As in Study 1, relationship satisfaction was assessed using the satisfaction subscale of Rusbult et al.’s (1998) investment scales. Items were averaged to construct an overall measure of satisfaction at each time point (α > .80; Table 1).

Problem severity. Prior to the discussions, the dyad member who was being targeted for change rated the degree to which they perceived the issue was a serious problem, including the degree to which the topic/issue to be discussed was a serious problem in the relationship (1 = not at all serious, 7 = extremely serious). Participants rated the same item to report on the severity of the same problem at each follow-up time point (see Table 1).

Attachment insecurity and self-esteem. At the initial session, participants completed the scales used in Studies 1 and 2 to assess attachment anxiety (M = 3.03, SD = 1.07, α = .81), avoidance (M = 2.82, SD = 0.95, α = .76), and self-esteem (M = 5.27, SD = 1.05, α = .89).

Observational assessment of partners’ negative-direct behavior. Six trained coders independently rated the extent to which each dyad member exhibited different types of conflict behavior using a coding scheme that specified different categories of communication behaviors that varied in valence and directness and were consistent across the most commonly used coding schemes (see Overall et al., 2009; Overall, 2018). Coders were given detailed descriptions of each category of behaviors, and then globally rated the degree to which each participant exhibited each type of behavior for each 30-s segment of the discussion (1–2 =
low, 3–5 = moderate, 6–7 = high). The behaviors exhibited by men and women were coded in separate viewings. For half of the discussions, women were coded first; for the other half, men were coded first. Two to four coders from the team of six rated each participant. Coders’ ratings were averaged for each 30-s segment (ICCs > .85). See the online supplemental materials for more information on the development of the observational coding scheme and the training procedure and instructions for coders.

**Negative-direct conflict behavior.** The negative-direct behaviors coded were selected for their consistency across major coding systems and established ability to predict problem resolution and relationship quality (see Gottman, 1998; Heyman, 2001; Overall et al., 2009), including: (a) criticizing, derogating or blaming, (b) threatening or commanding, (c) expressing anger and irritation, and (d) being rejecting, invalidating or domineering. Although traditionally conceptualized as harmful to relationships, these behaviors have been shown to be less harmful when partners are agents of change trying to improve serious problems (e.g., Overall et al., 2009; Overall, 2018; McNulty & Russell, 2010).

**Positive-indirect behavior.** As in Studies 1 and 2, positive behaviors were also assessed and examined to rule out the possibility that any beneficial effects of variability in partners’ negative behavior reflects partners behaving more positively in an effort to repair or balance negativity. Positive-indirect behaviors include (a) attempts to soften conflict, (b) express affection and positive affect (e.g., humor), (c) recognize and validate their partner’s efforts and point of view, and (d) focus on positive aspects of the partner/relationships. These partner behaviors have been shown to reduce targets’ hostility and defensiveness within conflict interactions (e.g., Overall et al., 2009; Overall et al., 2013) and thus enable tests of whether any longitudinal effects of variability in partners’ negative-direct behavior could be accounted for by attempts to soften conflict and repair relationships during the interaction.10

**Results**

As in Studies 1 and 2, the results are presented in three sections. The first section considers the calculation and resulting distribution of the behavioral variables. The second section presents growth curve analyses testing the longitudinal effects of partners’ negative conflict behavior. The final section examines the effects of potential alternative explanations.

**Calculating mean and variability of partners’ conflict behavior.** As with the repeated daily assessments examined in Studies 1 and 2, the coder ratings of each person’s negative-direct behavior for each 30-s segment of the discussion was used to create two measures. The person mean was calculated by averaging negative-direct behavior scores exhibited across the 14 segments of the discussion. The within-person SD was calculated to represent the relative stability (lower SD) versus variability (higher SD) in negative-direct behavior across the discussion. Table 6 presents the descriptive statistics of these behavioral measures. As in prior research assessing negative conflict behavior, and variability in negative emotions and emotion regulation, there was a good range in the mean and variability of partners’ negative conflict behavior.

As in Studies 1 and 2, and prior research on variability, the association between mean levels and variability in behavior was inflated because of very low or negative conflict behavior con-founded with very low or no variability (see bottom right panel in Figure 2).11 However, inspection of the scatterplot in Figure 2 reveals a reasonable distribution of low versus high levels of variability when partners exhibited mean or higher levels of negative-direct behavior, which is the area of contrast that is the focus of this investigation. Moreover, as in Studies 1 and 2, the data analytic approach testing the main and interactive effects of variability controls for and tests the incremental prediction of variability above and beyond mean levels (also see Girme et al., 2018 for similar distributions and analytic approach).

**Testing the longitudinal effects of mean and variability of partners’ negative behavior.** Dyadic growth curve analyses were conducted to assess the degree to which partners’ negative-direct behavior during couples’ conflict discussion was associated with longitudinal change in the repeated assessments of problem severity and relationship satisfaction across the year (see Karney & Bradbury, 1995a). Using the MIXED procedure in SPSS 25, these growth curve analyses treated individuals’ scores as repeated measures within the dyad crossed with repeated measures of time, which accounts for dependence of observations within couples and across time (Kenny et al., 2006; see the online supplemental materials for annotated syntax).

The first stage in growth curve analyses is to specify the trajectory of problem severity/relationship satisfaction across time by modeling the multiple assessments across the year as a function of an intercept and a slope representing time. Time was coded 0 at the initial session through to 4 for the final 12-month follow-up; thus, the intercept represents initial levels of problem severity/relationship satisfaction and the slope of time represents whether problem severity/satisfaction increased, decreased, or remained stable across the year. The intercept and slope of time were modeled as random and allowed to covary.

---

10 Coders also rated the presence of positive-direct behaviors, including explaining and clarifying the problem, outlining possible causes and consequences; reasoning with the partner, and exploring solutions. These behaviors have beneficial longitudinal effects that are independent of both negative-direct and positive-indirect behaviors (Overall, 2018; Overall et al., 2009). The additional analyses focused on positive-indirect behaviors that are most relevant to the alternative explanation that variability in partners’ negative behavior may simply be attributable to the presence of more positive, relationship repairing, behaviors. Positive-direct behaviors involve pushing for change, and thus can create the same short-term reactance as negative-direct behaviors, rather than softening conflict (e.g., Overall et al., 2009). Nonetheless, for completeness, the additional analyses were also run examining the effects of positive-direct behaviors, which revealed that positive-direct behaviors did not account for the longitudinal effects of partners’ variability in negative-direct behavior (see the online supplemental materials). Consistent with prior research (Overall, 2018; Overall et al., 2009), greater mean levels of partners’ positive-direct behavior predicted sustained relationship satisfaction across time. Nonetheless, the longitudinal effect of Mean × Variability of partners’ negative behavior on relationship satisfaction displayed in Table 7 and Figure 5 was unaltered.

11 As in Studies 1 and 2, one approach to this particular confound at low levels of negative behavior is to exclude data from partners who did not exhibit any negative-direct behavior across the conflict discussion and therefore did not have variability in negative-direct conflict behavior. This included 19 participants in Study 3. Analyses excluding these participants revealed the same or, in the case of problem severity, slightly stronger longitudinal effects. However, the analyses are presented with the full data, which captures people’s actual experiences of no negative behavior by partners.
The second stage of growth curve analyses involves assessing whether independent variables, in this case partners’ negative-direct conflict behavior, predict the trajectory of problem severity/relationship satisfaction across time. The effects of (a) mean of partners’ negative-direct behavior (mean-centered), (b) SD of partners’ negative-direct behavior (mean-centered), and the (c) interaction between mean and SD of partners’ negative-direct behavior were entered as predictors of the intercept (initial levels) and the slope of time (change across the year) of problem severity/relationship satisfaction.

The results are displayed in Table 7. The first row shows the average intercept (initial levels), and the second line shows the average slope of time (change across the year), for problem severity (left side) and relationship satisfaction (right side). On average, both problem severity and relationship satisfaction significantly declined over time, which is typical in studies modeling trajectories of relationship evaluations. In the context of this average decline, more harmful longitudinal outcomes involve maintenance of problem severity versus declines in problem severity. By contrast, more harmful longitudinal outcomes for relationship satisfaction involve declines of satisfaction versus a flat trajectory showing maintenance in satisfaction across the year (see Lavner & Bradley, 2010).

The next three rows of Table 7 present the effects of mean, SD and Mean × SD of partners’ negative-direct on initial levels of problem severity and relationship satisfaction, and the final three rows of Table 3 present the longitudinal effects of mean, SD and Mean × SD of partners’ negative-direct on changes in problem severity and relationship satisfaction across time. The longitudinal effects of behavioral variability are presented in bold. First focusing on problem severity, there was a significant longitudinal effect of variation in partners’ negative-direct behavior (SD of Partners’ Negative Behavior × Time) indicating that, controlling for mean levels, greater variability of partners’ negative-direct behavior across couples’ discussions predicted steeper declines in problem severity, whereas more stable negative-direct behavior predicted maintenance of problem severity across time.

The Mean × SD × Time interaction also provided support that the longitudinal effects of mean levels of partners’ negative-direct behavior depended on the variability of partners’ behavior. As shown in the left panel of Figure 4, low mean levels of partners’ negative-direct behavior predicted declines in problem severity across time regardless of whether partners’ negative-direct behavior was stable (low SD; slope = −0.39, 95% CI [−.49, −.30], t = −8.17, p < .001) or more variable (high SD; slope = −.46, 95% CI [−.67, −.25], t = −4.39, p < .001) across the discussion. However, as shown in the right panel of Figure 4, when partners’ exhibited high mean levels of negative-direct behavior, the variability of that behavior determined whether problems reduced or remained serious problems across the year. High mean levels of partners’ negative-direct behavior predicted maintenance of problem severity when partners’ behavior was relatively stable across the discussion (low SD; slope = −.06, 95% CI [−.32, .20], t = −0.44, p = .654). However, as when partners’ negative-direct behavior was low, high mean levels of partners’ negative-direct behavior predicted reductions in problem severity when partners’ negative-direct behavior varied more across the discussion (high SD; slope = −.45, 95% CI [−.54, −.35], t = −9.53, p < .001).

Behavioral variability had corresponding longitudinal effects on relationship satisfaction (see Table 7). The longitudinal effect of variation in partners’ negative-direct behavior (SD of Partners’ Negative Behavior × Time) provided support that greater variability of partners’ negative-direct behavior across couples’ discussions predicted lower declines in satisfaction, whereas more stable negative-direct behavior predicted steeper declines. The significant interaction between mean levels and variability predicting changes in relationship satisfaction (Mean × SD × Time) is presented in Figure 5. As shown in the left panel of Figure 5, low mean levels of partners’ negative-direct conflict behavior were associated with sustained relationship satisfaction, regardless of whether partners’ negative-direct behavior was stable (low SD; slope = −0.04, 95% CI [−.10, .01], t = −1.66, p = .098) or variable (high SD; slope = −.06, 95% CI [−.16, .04], t = −1.17, p = .243) across the discussion. By contrast, as shown in the right panel of Figure 5, high mean levels of partners’ negative-direct behavior predicted significant reductions in relationship satisfaction when partners’ negative behavior was stable across the discussion (low SD; slope = −.19, 95% CI [−.32, −.06], t = −2.92, p = .004). However, high mean levels of partners’ negative-direct behavior was associated with maintenance of relationship satisfaction when partners’ behavior varied more across the discussion (high SD; slope = .01, 95% CI [−.05, .06], t = 0.19, p = .848).

In sum, high mean levels of partners’ negative-direct behavior had harmful longitudinal effects when partners’ negative-direct behavior was highly stable across couples’ conflict discussion,
including predicting sustained problem severity and reduced relationship satisfaction across time. By contrast, the same high levels of partners’ negative-direct behavior was not harmful when partners’ behavior was more variable across the discussion, revealing the same longitudinal trajectories as low levels of partners’ negative-direct behavior, including reductions in problem severity and maintenance of satisfaction.

Additional analyses and alternative explanations. A series of additional analyses were run to rule out alternative explanations for the longitudinal effects of mean and variability of partners’ negative-direct behavior, including adding alternative factors into the models presented in Table 7 to test whether the effects of partners’ negative-direct behavior remained, whereas individuals’ own negative-direct behavior had no independent longitudinal effects with one exception: individuals’ own behavioral variability predicted greater relationship satisfaction across time (see Table 9), but the significant Mean × SD of partners’ negative-direct behavior shown in Figure 5 was unaltered.

![A. Low Mean Partner Negative Behavior](image1.png)

![B. High Mean Partner Negative Behavior](image2.png)

**Figure 4.** The effects of mean levels and variability of partners’ negative-direct conflict behavior on severity in the problem discussed across time. SD = standard deviation. The slopes across the x axis represent the linear change in problem severity from the initial session (Time 0) to 12 months later (Time 4).

**Table 7**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Problem severity</th>
<th>Relationship satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>95% CI</td>
</tr>
<tr>
<td>Intercept</td>
<td>4.51</td>
<td>[4.29, 4.73]</td>
</tr>
<tr>
<td>Time</td>
<td>−.34</td>
<td>[−.43, −.26]</td>
</tr>
<tr>
<td>M of partners’ negative behavior</td>
<td>−.15</td>
<td>[−.39, .10]</td>
</tr>
<tr>
<td>SD of partners’ negative behavior</td>
<td>.41</td>
<td>[−1.4, .96]</td>
</tr>
<tr>
<td>M × SD of partners’ negative behavior × Time</td>
<td>−.21</td>
<td>[−.51, −.05]</td>
</tr>
</tbody>
</table>

Note. The longitudinal effects of behavioral variability are presented in bold. CI = confidence interval. Effect sizes (r) were computed using Rosenthal and Rosnow’s (2007) formula: $r = \sqrt{(r^2/t^2 + df)}$. In these multilevel models, the Satterthwaite approximation is applied to provide specific degrees of freedom for each effect, which were used to calculate the effect sizes.
(B = −.52, t = −4.38, p < .001). Second, replacing variability in partners’ negative behavior with partners’ positive behavior in the models presented in Table 7 revealed that greater positive behavior did not predict more beneficial outcomes as variability in partners’ negative behavior did (main and interaction effects of partners’ positive behavior ts < −1.27, ps > .20). Finally, modeling the mean, SD, and Mean × SD interaction of both partners’ negative-direct behavior and partners’ positive behavior on change in problem severity and relationship satisfaction across time revealed that only partners’ negative-direct behavior, and not partners’ positive behavior, had independent longitudinal effects (see Tables 8 and 9).

**Own positive behavior.** Additional analyses also ruled out the possibility that the longitudinal effects of variability in partners’ negative-direct behavior was attributable to individuals’ own positive conflict behaviors. First, greater mean levels of partners’ negative-direct behavior was associated with individuals enacting lower positive behavior (B = −.24, t = −3.72, p < .001). Second, replacing variability in partners’ negative-direct behavior with individuals’ own positive behavior in the models presented in Table 7 revealed that greater own positive behavior did not predict more beneficial outcomes as variability in partners’ negative-direct behavior did (main and interaction effects of own positive behavior ts < .99, ps > .32). Finally, the mean, SD, and Mean × SD
interaction of individuals’ own positive behavior did not predict change in problem severity or relationship satisfaction across time, and controlling for these effects only had a minimal effect on the longitudinal effects of partners’ negative-direct behavior (see Tables 8 and 9).

**Attachment insecurity and self-esteem.** Additional analyses also ruled out the possibility that the results were attributable to variable versus stable behavioral patterns emerging from individual differences that may create more reflexive or rigid responding. Unlike Studies 1 and 2, neither individuals’ own nor their partners’ attachment insecurity were significantly associated with average or variability of partners’ negative-direct behavior (t < 1.40). Self-esteem was also not associated with average or variability of partners’ negative-direct behavior (t < 1.33). Thus, as in Studies 1 and 2, the longitudinal effects of variability in partners’ negative-direct behavior on problem severity and relationship satisfaction remained or became stronger when controlling for individual’s own or their partners’ attachment anxiety, attachment avoidance, or self-esteem (see results presented in the online supplemental materials).

**Discussion**

Study 3 provided a conceptual replication and extension of the results in Studies 1 and 2 by examining variability in partners’ negative-direct behavior during couples’ conflict discussions. Consistent with the harmful outcomes shown in Studies 1 and 2, partners’ criticism and hostility predicted persistence of serious relationship problems and declines in relationship satisfaction across the following year, but only when partners’ exhibited stable levels of negative-direct behavior across couples’ conflict discussion (high mean levels + low within-person SD). By contrast, consistent with the benefits of partners’ negative-direct behavior identified by prior investigations of couples’ conflict behavior, these harmful effects were eliminated when partners’ negative-direct behavior was more variable across couples’ conflict discussions. Thus, by uniquely applying techniques used with daily sampling data to capture behavioral variability within social interactions, the results of Study 3 (a) illustrate that variability across a specific social interaction identifies important behavioral patterns that have distinct longitudinal implications, (b) demonstrate that the longitudinal effects of these patterns replicate the reduced costs that arise when partners’ negative behavior varies across daily life (shown in Studies 1 and 2), and (c) provide novel support for theoretical models specifying that the effects of negative-direct behavior depend on whether behavior varies in ways that likely reflect sensitivity or responsiveness to situational or contextual demands.

**General Discussion**

Determining how social behavior affects social partners, and identifying how contextual factors shape these interpersonal dynamics, are cornerstones of social psychology. Relationship science has offered important advances in the study of social behavior because the dyadic and ongoing nature of most intimate relationships offers a unique social milieu to assess how one person’s behavior impacts another person’s outcomes across time. Yet, the comprehensive methods used to assess behavioral dynamics within relationships overlook what the current results suggest is an important determinant of the longitudinal impact of social behavior—behavioral variability. By repeatedly assessing partners’ behavior across days (Studies 1 and 2) and across a specific social interaction (Study 3), and calculating the average (within-person mean) and variability (within-person SD) of negative-direct behavior, the current studies revealed that average levels and variability index distinct theoretically relevant behavioral profiles that have distinct longitudinal interpersonal effects.

The profile of partners’ negative-direct behavior that was less harmful for relationships involved greater behavioral variability. Behavioral variability should be beneficial because behaviorally relevant demands vary across and within situations, and thus being responsive to changing situational demands should produce variability in behaviors across time (also Aldao et al., 2015; Blanke et
Variability in Social Behavior: Advances and Implications

Demonstrating the divergent outcomes associated with behavioral variability versus stability makes multiple advances across different areas. With regard to understanding behavior in close relationships, the results highlight that one reason that hundreds of studies examining negative-direct conflict behavior have produced inconsistent findings is because the same average levels of negative-direct behavior, even within a single interaction, have different longitudinal implications depending on the variability of that behavior. Standard behavioral investigations within the close relationships field have considerable strengths, including detailed daily and observational assessments of behavior and tracking changes in relationship outcomes across time (as in the current studies). The resulting longitudinal effects that arise from these methods indicate that such behavioral assessments provide a window into how couples typically respond to important situations that culminate across time to shape relationships. However, the current results illustrate that an additional temporal feature—variability versus stability in behavior—will change the meaning and interpretation of the varying demands partners encounter across daily life and specific social interactions. In the case of partners’ negative-direct behavior, low behavioral variability should be harmful because persistent, stable criticism and hostility will fail to diagnose important contexts, events, and situational needs, and instead likely communicates global and continual hostility that amplifies relationship problems. Accordingly, high average levels of partners’ negative-direct behavior were more harmful when partners’ behavior was highly stable across days and couples’ conflict discussions, including predicting greater increases in a range of problems across couples’ relationships (Studies 1 and 2), a lack of improvement in a serious problem causing conflict (Study 3), and declines in relationship satisfaction (Study 3), across time.

The profile of partners’ negative-direct behavior that was more harmful involved low variability or behavioral stability. Whereas high levels of behavioral variability likely indicate sensitivity to shifting contextual and situational demands, behavioral stability likely reflects ongoing, persistent behavior that occurs irrespective of the varying demands partners encounter across daily life and specific social interactions. In the case of partners’ negative-direct behavior, low behavioral variability should be harmful because persistent, stable criticism and hostility will fail to diagnose important contexts, events, and situational needs, and instead likely communicates global and continual hostility that amplifies relationship problems. Accordingly, high average levels of partners’ negative-direct behavior was more harmful when partners’ behavior was highly stable across days and couples’ conflict discussions, including predicting greater increases in a range of problems across couples’ relationships (Studies 1 and 2), a lack of improvement in a serious problem causing conflict (Study 3), and declines in relationship satisfaction (Study 3), across time.

The current studies also complement and extend recent research using variability across daily life to assess situationally dependent emotion regulation patterns (Blanke et al., 2019; also Cheng, 2001; Eldesouky & English, 2018). Consistent with the current results, those prior applications suggested behavioral variability is beneficial as indicated by better emotional outcomes (Blanke et al., 2019). Expanding prior applications, the current results demonstrate that variability can be applied to understand behavior within specific social interactions in addition to daily life, and that behavioral variability has interpersonal (across dyad rather than within individual) effects that shape outcomes across time (longitudinal rather than cross-sectional associations). The interpersonal and longitudinal designs applied in the current studies could advance investigations of emotion regulation behavior, which also has implications for social partners and relationships (e.g., Butler et al., 2003; Impett et al., 2012; Peters & Jameson, 2016). In particular, adopting these designs would enable examination of whether variability in emotion regulation strategies promote the wellbeing of individuals and their partners by enhancing emotion regulation across different situations.

Integrating the within-individual and wellbeing focus of prior research examining intrapersonal variability may also enhance understanding of the outcomes of behavioral dynamics in relationships. In the current studies, partners’ and individuals’ positive behavior did not independently predict longitudinal change in problems, regardless of variability. This is consistent with prior research showing that positive-indirect behaviors often produce null effects on relationship problems because the benefits of easing conflict in the short-term is balanced by the costs of withholding negativity that conveys the importance and severity of problems and associated need to make relationship changes (see Overall & McNulty, 2017). Moreover, it is often the case that social behaviors that can benefit partners and relationships can potentially incur costs to personal wellbeing. Although compromising, sacrificing, and forgiving may help maintain relationships, these behaviors can also lead to lower self-worth and greater depressed mood if partners are not responsive (Baker, McNulty, Overall, Lambert, & Fincham, 2013; Luchies, Finkel, McNulty, & Kumashiro, 2010) or
if these behaviors are motivated by fears of rejection or avoidance of conflict (Impett, Gable, & Peplau, 2005; Lemay & Dudley, 2011; Righetti & Visserman, 2018). Thus, variability in relationship maintenance behaviors, indexing sensitivity to an array of important contextual and situational demands (e.g., issue importance, relative needs across partners, equity, partners’ reciprocation), may reflect a healthier pattern of balancing both personal and relationship wellbeing. By contrast, a persistent prioritization of the relationship that does not vary across different contexts and situational demands may predict poorer personal wellbeing over time because it reflects a blanket submission of one’s own needs for the partner or the relationship irrespective of current or chronic needs.

These possibilities and implications of behavioral variability also highlight that person-level characteristics and motivations are likely to create different behavioral profiles that may account for why some individual differences often produce relationship problems. Additional analyses indicated that attachment insecurity and self-esteem did not consistently predict differences in behavioral variability, with one exception: highly avoidant individuals were more likely to perceive their partners’ daily negative behavior as more stable across days (Studies 1 and 2), which is consistent with avoidant individuals’ expectations that others will be consistently neglectful and exploitative (Mikulincer & Shaver, 2003). Given this pattern of partner behavior predicted the worse outcomes across time, perceiving partners to consistently behave negatively may be a key reason why avoidant people experience greater problems in their relationships. Although the way relationship insecurities promote different perceptions and behaviors is central to models outlining how both attachment and self-esteem impact relationship quality (Murray et al., 2006; Simpson & Rholes, 2012), the mass of work examining these processes has focused on average levels of behavior or perceptions. The current results highlight that behavioral variability is also theoretically relevant, and likely consequential, to attachment processes, which aligns with theoretical propositions that the development of insecurities, such as attachment avoidance, arise from consistently negative or harsh environments (Mikulincer & Shaver, 2003; Simpson & Rholes, 2012).

Other individual differences and motivations might be more relevant to the enactment, rather than the perception, of variability in negative-direct behavior. For example, partners high in social dominance or need for power may be more likely to show a rigid, domineering behavioral pattern of high and stable levels of negative-direct behavior (Felson & Outlaw, 2007; Whitaker, 2013). By contrast, partners low in power or high in relational dependence may show greater variation in a range of behaviors because they need to be more attuned to situationally relevant information, such as their partners’ perspective, to get their needs met (Gordon & Chen, 2013; Righetti et al., 2015). These patterns are also relevant to levels of dominance and power in nonintimate contexts (Galinsky et al., 2006; Keltner, Gruenfeld, & Anderson, 2003; Maner & Mead, 2010), and perhaps personality traits that may create more persistent negative behaviors across social contexts (e.g., neuroticism). Thus, the identification of variability versus stability as a consequential feature of social behavior offers fruitful directions for examining how social motivations and individual differences shape temporal patterns, rather than just levels, of social behavior, and the way these behavioral profiles may be more or less effective at navigating interpersonal interactions.

Strengths, Caveats, and Future Research Directions

The role of behavioral variability in determining the relative harm of partners’ negative-direct behavior was supported across three longitudinal studies using different methods to repeatedly assess naturally occurring relationship behavior. The most consistent effects emerged for relationship problems, including the degree to which partners’ negative-direct behavior predicted increases in a range of problems across couples’ relationships (Studies 1 and 2) and the maintenance of specific problems that couples discussed in the laboratory (Study 3). However, partners’ daily negative-direct behavior did not directly predict relationship satisfaction (Studies 1 and 2) as variability in partners’ negative-direct behavior during conflict interactions did (Study 3). The stronger effects on relationship problems is consistent with prior research and theory highlighting the particular relevance of partners’ negative-direct behavior for understanding and managing relationship problems. Prior research has established that partners’ negative-direct behavior can improve relationships when it is diagnostic of, and promotes improvement, in serious problems that need to be addressed, but maintains specific problems or produces broader problems when not matched to the relative demands couples are facing (see McNulty, 2010, 2016; McNulty & Russell, 2010; Overall et al., 2009; Overall & McNulty, 2017). Thus, problem severity represents a primary mechanism explaining how and why partners’ negative-direct behavior affects satisfaction across time (also see Footnote 8). Thus, the relative impact of variability in partners’ negative-direct behavior on relationship problems shown across the current studies has established implications for relationship satisfaction (also see Lavner et al., 2016).

There may be other methodological differences across the studies that account for the inconsistent effects on satisfaction, including the more extensive measure of problems compared with satisfaction in Studies 1 and 2, and the more extensive repeated assessments of satisfaction in Study 3. The different methodological contexts also may explain differences in the main effects of mean levels versus variability in negative-direct behavior. Greater mean levels of partners’ negative-direct behavior across couples’ daily life had a strong harmful effect on general relationship problems across time, which was most pronounced when partners’ behavior was more stable versus variable across days. By contrast, rather than mean levels, stability versus variability in partners’ negative-direct behavior during conflict interactions predicted more harmful outcomes across time, perhaps because the localized assessment of a specific conflict interaction in which negative-direct behavior is likely to emerge makes the stability rather than the mean-level presence and intensity of that behavior more meaningful. Nonetheless, the theoretically consistent Mean × Variability interaction pattern that emerged across studies highlights that behavioral variability versus stability is most meaningful when partners’ negative-direct behavior is present at a level of intensity that diagnoses important, potentially problematic, situational demands that need to be attended to.

Longitudinal designs are the strongest way to examine how behaviors naturally shape the course of relationships, but they are nonetheless correlational and thus cannot establish that behavioral
variability is playing a causal role. Perhaps greater relationship problems also create more stable levels of negativity, and these bidirectional associations are reinforcing and strengthen over time. In that case, the results continue to have important implications for identifying and targeting more or less problematic behavioral patterns that have interpersonal effects across partners. Furthermore, although additional analyses ruled out some alternative explanations, such as enduring relationship insecurities and partners' positive behavior, there may be other contributing variables, such as persistent stress or ongoing demands that create stable behavioral patterns. Nonetheless, the consistent and important longitudinal outcomes associated with stability versus variability in partners' negative-direct behavior highlight that uncovering what causes these patterns is a valuable direction for future research.

As with the growing body of work recognizing the importance of variability in psychological phenomenon, the central caveat involves equivocality in what behavioral variability represents. Emotional variability is theorized to reflect an inability to effectively deal with and regulate emotions in response to changes in the environment (see Houben et al., 2015). Similarly, variability in relationship evaluations is theorized to reflect more intense reactions and turmoil in response to relationship events (see Campbell et al., 2010). Yet, variability itself does not provide information about whether changes in emotions and evaluations are produced by environmental challenges, are appropriate to those challenges, or are reflective of regulation failure. Instead, the consistent findings that greater variability in emotions and evaluations is associated with lower personal and relationship wellbeing provide supporting evidence that these temporal patterns are indicative of poorer coping with personal and relationship challenges as theorized (see Houben et al., 2015 for discussion).

Behavioral variability has opposite implications than emotional variability, but the same criteria used in that body of work can be applied to evaluate the theoretical basis and value of behavioral variability. Variability in emotion regulation behavior is theorized to reflect an adaptive, contextually dependent behavioral pattern involving adjusting regulation strategies to different situations, and this conceptualization is supported by associations between emotion regulation variability and better emotional outcomes (Aldao et al., 2010). Applying the same logic, variability in partners' negative-direct behavior should capture a temporal pattern that would emerge if partners' behavior varied along with current situational demands. Given that the longitudinal outcomes in the current studies correspond to contemporary theoretical models and evidence specifying that negative-direct behavior that is diagnostic of important contextual and situational demands will not be as harmful, the current results provide support that greater behavioral variability likely captures patterns of responding that are more sensitive to the varying demands couples face across their daily life and social interactions. By contrast, invariable, highly stable, negative-direct behavior likely reflects persistent negativity irrespective of current situational demands.

Nonetheless, although behavioral variability should emerge if partners are responding in ways that are sensitive to situational demands, variability does not provide any information about those relevant demands or whether behavior is appropriate for current demands. Other methods used to examine contextually sensitive responding, such as those used to test psychological or regulatory flexibility (Bonanno & Burton, 2013; Kashdan & Rottenberg, 2010), requires defining in advance what appropriate and adaptive responses to particular contexts should be, such as specific changes in affect to daily stressors (e.g., Hardy & Segerstrom, 2017). Like prior research examining how contextual factors moderate the effects of negative-direct (and other relationship) behavior, these approaches are necessarily narrow by prioritizing specific contexts and situations over a range of important contexts and situations that help determine whether a given behavior is relevant, appropriate or diagnostic. Moreover, these methods require defining context-behavior profiles that are assumed to be generally adaptive, yet the meaning of any particular context, situation or behavior will also be shaped by within-person and within-dyad differences and relationship histories that are impossible to fully account for in any given investigation. Thus, behavioral variability is a valuable way to index broader patterns that should emerge if partners are adjusting behavior to common and idiosyncratic situational demands (more variable) or are behaving in rigid, persistent ways that is less diagnostic of current demands (more stable).

Despite these differences, both methodological approaches are valuable in understanding and identifying the behavioral patterns that best help sustain relationships. Studies identifying particular contextual/situational factors or specific situation-behavior signatures that help determine when negative-direct (and other relationship) behavior is more or less harmful offer specific guidance regarding how people could alter their behavior in response to specific challenges in ways that will enhance their relationships and wellbeing. By contrast, investigations focusing on behavioral variability without the constraints of defining specific context- or situation-behavior linkages may be particularly useful in revealing a general behavioral profile that can be fostered more broadly in social relationships, such as promoting the expression of diverse social behaviors to align with the diverse demands people face in their relationships. Focusing on variable versus stable behavioral profiles may also be useful in detecting problematic behavioral sets that require different interventions, such as targeting entrenched vulnerabilities and motivations that produce behavioral rigidity, which if resolved will likely improve relationship dynamics.

Finally, the key points regarding behavioral and contextual variability also apply to the relative generalizability of these profiles and longitudinal outcomes to other samples. Couples who are experiencing very serious relationship problems and facing extremely difficult environments may not show the same effects as emerged with the relatively satisfied couples in the current studies. Persistent stress may produce what appears to be rigid, stable negative-direct behavior, which may require disengagement rather than behavioral variability due to a greater risk of compounding problems (Ross, Karney, Nguyen, & Bradbury, 2019). Moreover, even situationally appropriate negative-direct behavior requires the skills and motivation to deal with the challenges at hand, and extreme acts of aggression will by uniformly harmful regardless of the context they are enacted. Indeed, other noncontextual sources of variability, such as unpredictable behavior arising from dysregulation or psychopathology, should amplify the harm of negative social behavior. Yet, the potential for different results to emerge in different contexts and samples highlights that variability is an important, overlooked, dimension that shapes the interpersonal impact of social behavior.
Conclusions

Social behavior has important interpersonal consequences, including creating versus improving interpersonal problems. Understanding these consequences, however, requires not only assessing levels of important social behaviors, such as criticism and hostility, but also assessing how people’s behavior varies across days or social interactions. In the domain of intimate relationships, partners’ critical and hostile behavior predicted greater problems (Studies 1–3) and lower relationship satisfaction (Study 3) when partners’ negative-direct behavior was stable across time (low within-person variability), but was less harmful when partners’ negative-direct behavior varied across time (high within-person variability). Behavioral variability versus stability should be equally consequential whenever the meaning and impact of social behavior varies across different diagnostic contexts and situational demands, and provides a new way of understanding and measuring behavioral patterns that will be more helpful versus detrimental in navigating the challenges of social life.

References


Received May 9, 2019
Revision received October 25, 2019
Accepted November 28, 2019