

Moving Beyond Assumptions:
The Reconceptualization and Measurement of Workplace Gossip

by
Daniel Brady

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Arts
in
Psychology

Waterloo, Ontario, Canada, 2015

© Daniel Brady 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Despite decades of research from other academic fields arguing that gossip is an important and potentially functional behavior, organizational research has largely assumed that gossip is malicious talk. This has resulted in the proliferation of gossip items in deviance scales, effectively subsuming workplace gossip research into deviance research. In this paper, it is argued that organizational research has traditionally considered only a very narrow subset of workplace gossip, focusing almost exclusively on extreme negative cases which are not reflective of typical workplace gossip behavior. Instead of being primarily malicious, typical workplace gossip can be either positive or negative in nature and may serve important functions. It is therefore recommended that workplace gossip be studied on its own, independent of deviance. To facilitate this, the workplace gossip construct is reconceptualized and then a series of general-purpose English- and Chinese-language workplace gossip scales are developed. Using 8 samples (including multisource, multiwave, and multicultural data), this research demonstrates the construct validity, reliability, cross-cultural measurement invariance, and acceptable psychometric properties of the workplace gossip scales. Relationships are demonstrated between workplace gossip and a variety of other categories of organizational variables and processes, including sensemaking, emotion validation, organizational justice, influence, self-enhancement, job performance, and turnover. Future directions in workplace gossip research are discussed.

Keywords: workplace gossip, deviance, measurement, social comparison

Acknowledgments

I would like to thank my supervisor Dr. Doug Brown for his support and guidance during this research. I would also like to thank my readers Dr. Winny Shen and Dr. James Beck for their valuable feedback, Dr. Lance Ferris for his friendly review, and Lindie Liang and Sam Hanig for their valuable assistance. Finally, I am grateful for the support of Laurie, without whom none of this would be possible.

Table of Contents

Author's Declaration.....	ii
Abstract.....	iii
Acknowledgments.....	iv
Table of Contents.....	v
List of Tables.....	vii
Introduction.....	1
Gossip.....	3
Workplace Gossip and Deviance.....	5
Workplace Gossip Reconceptualization.....	7
Study Preview.....	10
Phase 1: Qualitative Study of Workplace Gossip.....	12
Participants and Procedure.....	12
Results and Discussion.....	12
Phase 2: Item Generation and Reduction.....	14
Substantive Validity Assessment.....	14
Participants and Procedure.....	14
Results.....	15
Item Frequencies.....	15
Participants and Procedure.....	16
Results and Discussion.....	16
Final Item Reduction.....	16
Participants and Procedure.....	17
Results and Discussion.....	17
Phase 3: Psychometric Properties.....	19
Participants.....	19
Measures.....	20
Analytic Strategy.....	21
Results.....	22
Discussion.....	24
Phase 4: Construct Validity.....	25
Measures.....	30
Results.....	32
Discussion.....	35
Phase 5: Cross-Cultural Measurement Invariance.....	36
Participants and Procedure.....	36
Analysis.....	36
Results and Discussion.....	37
Phase 6: Toward Studying WG Independent of Deviance.....	38

Measures	38
Results and Discussion	38
General Discussion	40
Study Limitations.....	42
Future Research into Workplace Gossip.....	43
Conclusion	47
References.....	48
Appendices.....	65

List of Tables

Table		Page
1.	Overview of Measures Used for Validation Phases 3 and 4	60
2.	Descriptive Statistics, Correlations, and Alpha Reliabilities for Sample 5	61
3.	Descriptive Statistics, Correlations, and Alpha Reliabilities for Sample 6	62
4.	Descriptive Statistics, Correlations, and Alpha Reliabilities for Sample 7	63
5.	Cross-cultural Assessment of Measurement Equivalence/Invariance	64

Moving Beyond Assumptions:

The Reconceptualization and Measurement of Workplace Gossip

If researchers continue to dismiss gossip as essentially idle, malicious talk about unfortunate individuals, they may overlook some of its most important and valuable functions

– Baumeister, Zhang, and Vohs (2004, p. 120)

What is gossip? Scholars and layperson alike often assume that gossip is primarily idle or malicious talk (Baumeister et al., 2004; Foster, 2004). Interestingly, this decidedly negative view of gossip is not generally shared by gossip researchers. Instead, gossip researchers have taken a more balanced view of gossip, arguing that it is primarily a tool for gathering and validating information (Paine, 1967; Suls, 1977; Wert & Salovey, 2004). Recent research into gossip motivations supports this position. It found that the most prevalent self-reported reason to gossip was to gather and validate information, while the motivation to negatively influence another (i.e., with intent to harm or influence) was tied for the least prevalent reason to gossip (Beersma & Van Kleef, 2012). This finding reflects traditional theory and evidence from the broader gossip literature, but it directly contradicts common implicit assumptions of gossip as being malicious.

Baumeister and colleagues (2004) warned researchers that negative assumptions about gossip impair our ability to recognize the important and functional aspects of gossip. Nowhere does this warning seem to be more relevant than in the field of organizational behavior. Traditionally, organizational researchers have categorized workplace gossip (WG) as a form of minor interpersonal deviance (Robinson & Bennett, 1995). This categorization has resulted in gossip or gossip-like items being routinely included in deviance scales (e.g., “gossiped about my supervisor”, Mitchell & Ambrose, 2007), effectively subsuming WG research into deviance research. This has led to a glaring cross-disciplinary contradiction in how gossip is conceptualized. While other fields argue that gossip is an important and potentially functional

behavior, organizational behavior largely assumes that WG is deviant.

In this paper, we argue that WG does not cleanly fit a deviance conceptualization. Although some extreme and rare cases of WG may be malicious, *typical* WG can be either positive or negative in nature and may serve important organizational functions (Kurland & Pelled, 2000; Noon & Delbridge, 1993). Unfortunately, the traditional practice of treating WG as deviance has artificially impeded WG research by implying that WG is better studied as a part of other constructs (e.g., interpersonal deviance) rather than on its own. This has resulted in both an over-representation of gossip in deviance scales and a lack of general-purpose WG-only measurement instruments. With limited ability to study WG independent of deviance, the cross-disciplinary contradiction in how gossip is treated has been allowed to persist. If research from other disciplines is correct, and gossip is an important and fundamental human behavior (Dunbar, 2004), then research into why WG manifests and how it relates to other organizational variables could be essential for advancing our knowledge of organizational behavior.

The goal of this paper is to move beyond assumptions about WG and to facilitate future research into this important behavior. We begin by drawing on gossip research from multiple fields to reconceptualize the WG construct. We then address a long-standing WG scale deficiency by systematically developing and validating four general-purpose behavioral measures of WG (i.e., negative WG about a supervisor, positive WG about a supervisor, negative WG about co-workers, and positive WG about co-workers). In so doing, we construct an initial nomological network for WG, demonstrating relationships between WG and a variety of organizational variables and processes. We then demonstrate that WG is distinct from several measures of deviance and provide preliminary evidence that the traditional practice of including gossip items in deviance scales could lead to undesirable measurement issues, including an over-

reporting of deviance behavior. Finally, we chart a path forward for WG research and discuss new research directions that can result from studying WG as its own construct, independent of deviance. We begin by briefly reviewing the gossip literature.

Gossip

Gossip is formally defined as “evaluative talk about a person who is not present” (Eder & Enke, 1991, p. 494). Despite common assumptions that gossip is only negative, researchers have argued that gossip can be either positive (e.g., a compliment) or negative (e.g., a criticism) in nature (Foster, 2004). It has been suggested that the layperson’s view of gossip is often confused or even contradictory (e.g., Baumeister et al., 2004; Ben-Ze’ev, 1994; Foster, 2004; Spacks, 1982). On one hand, many individuals believe that gossip is generally bad and that it should be discouraged. On the other hand, almost everyone still gossips. It has been argued that this contradiction stems from an availability bias, where extreme and memorable cases of gossip (e.g., malicious talk) define the prototype of the gossip concept (Ben-Ze’ev, 1994). This is thought to interfere with our ability to recognize typical gossip manifestations (e.g., chitchat) as being gossip because typical cases of gossip differ substantially from the extreme prototype.

For nearly a century, anthropologists have observed and debated the role of gossip in social groups. Interest in this area has been so prevalent that anthropology has been referred to as “the social science of gossip” (Foster, 2004, p. 78). Anthropologists have documented gossiping behavior in a wide variety of cultures (e.g., Brenneis, 1984; Beisner, 1989; Cox, 1970; Gilmore, 1978; Haviland, 1977; Loudon, 1961; Radin, 1927; Wiessner, 2014) and have argued that gossip facilitates positive group functions, such as the promotion and maintenance of group unity, norms, and values (Gluckman, 1963). Gossip is thought to enable the construction of group identity, as gossip serves to delineate in-group from out-group members, and also conveys

information regarding what it means to be a member of the group. Anthropologists have argued that gossip is an important and required part of group membership.

At an individual level, gossip is thought to be a self-interested behavior that reflects an individual's need for information (Paine, 1967). Drawing on social comparison theory, researchers have argued that gossip is a social comparison process (Suls, 1977; Wert & Salovey, 2004) that can be used to validate an individual's opinions (Festinger, 1954), abilities, and emotions (Schachter, 1959). When gossipers provide evaluative (i.e. value-laden) information about a gossip subject, there is an implied upward or downward social comparison between the gossiper and the gossip subject. In effect, the gossip subject is judged relative to the gossiper. The gossip recipient's reaction to the gossip then either confirms or denies the social comparison, resulting in an exchange of information between the gossiper and gossip recipient. Building on this informational view of gossip, Baumeister et al. (2004) proposed that gossip is a form of social learning. Through gossip, individuals can learn of the previous successes or failings of others. This informs group members of how to successfully function in society without having to experience the pain of negative outcomes first-hand.

Evolutionary psychologists have taken a much more fundamental view of gossip, arguing that it is an evolutionary adaptation that enables the development of large social groups (Dunbar, 1993; 1996; 2004; Emler, 1994). Gossip is thought to facilitate social bonding, trust, and mutual obligation. It helps to regulate large groups, as it deters selfish behaviors through the exchange of reputational information. There is some evidence that supports a fundamental view of gossip: gossip has been observed in very young children (Fine, 1977), in a variety of cultures (Gluckman, 1963), and has been shown to physiologically affect how we see and cognitively process facial stimuli (Anderson, Siegel, Bliss-Moreau, & Barrett, 2011). Further, evidence has

shown that individuals will act on negative reputational gossip even if it directly contradicts first-hand observational evidence (Sommerfeld, Krambeck, Semmann, & Milinski, 2007).

Recent empirical gossip research has examined prosocial gossip, defined as negative gossip which protects group members from harm or exploitation (Feinberg, Willer, Stellar, & Keltner, 2012). Prosocial gossip is conceptualized as a functional form of gossip, as it facilitates the positive functioning of a group. It builds on previous gossip theory that gossip can be used to enforce group norms and manage reputations, thereby leading to positive outcomes. Research has shown that negative gossip increases in the presence of social loafing (Kniffin & Wilson, 2005) and that the threat of negative gossip constrains selfish behaviors, while increasing cooperation and group-beneficial behaviors (Beersma & Van Kleef, 2011; Feinberg, Willer, & Schultz, 2014; Feinberg et al., 2012; Piazza & Bering, 2008).

Workplace Gossip and Deviance

Despite many decades of research suggesting that gossip can be an important and functional behavior, the field of organizational behavior has historically taken a very negative view of gossip. With little empirical evidence, organizational practitioners have frequently drawn upon implicit theories of gossip to recommend that gossip be completely eliminated from organizations (Noon & Delbridge, 1993). Although there have been warnings that WG is misunderstood and should be studied further (Kurland & Pelled, 2000; Noon & Delbridge, 1993), there has been little resulting WG research. This is surprising, as gossip is likely to be particularly important in a workplace context due to the role that gossip is thought to play in fostering effective group functioning (Dunbar, 2004; Feinberg et al., 2014; Gluckman, 1963).

In their classic deviance typology, Robinson and Bennett (1995) categorized gossip as a form of deviance, defined as “voluntary behavior that violates significant organizational norms

and in so doing threatens the well-being of the organization, its members, or both” (p. 556).

While this gossip categorization was a very small part of their overall paper, it has had a profound impact on WG research, as it reinforced existing assumptions about gossip. Further, it suggested that WG is better studied as a part of other constructs rather than on its own. The result has been a proliferation of gossip or gossip-like items in deviance scales (e.g., “gossiped about my supervisor”, Aquino, Lewis, & Bradfield, 1999; “made fun of someone at work”, Bennett & Robinson, 2000; “talked bad about you behind your back”, Duffy, Ganster, & Pagon, 2002; “makes negative comments about me to others”, Tepper, 2000), thereby subsuming WG research into deviance research. Unfortunately, it is not clear if this deviance categorization is an accurate assessment of gossip behavior or rather a reflection of common negative implicit theories related to the word “gossip” (Baumeister et al., 2004).

We argue that while some extreme cases of WG could be deviant, *typical* WG does not fit a deviance conceptualization.¹ In contradiction to common assumptions about gossip, research suggests that gossip can manifest in either positive or negative forms (Foster, 2004). Positive gossip (e.g., saying something nice about someone) is clearly not deviant, providing initial support that not all gossip is deviant. For negative gossip, we draw upon three theoretical tests of deviance from the broader deviance literature to argue that negative gossip is also not deviant. First, in a statistical test of deviance, a behavior can be deemed deviant if the frequency of the behavior is a statistical outlier from normal behavior (Spreitzer & Sonenshein, 2004). Research has shown that gossip is a common behavior and that almost everyone gossips (Dunbar, Marriott, & Duncan, 1997; Emler, 1994). Gossip is therefore not statistically deviant. Second, in a reactive test of deviance, a behavior is deemed deviant if other group members observe the behavior and

¹ Although some deviance research has made the distinction between negative and positive deviance (Spreitzer & Sonenshein, 2004), WG has traditionally been viewed as negative or dysfunctional deviance in the organizational behavior literature. We argue that typical WG is different from deviance altogether (negative or positive).

condemn it (Spreitzer & Sonenshein, 2004). By definition, gossip always involves multiple individuals (i.e., the gossiper and gossip recipients). Qualitative research indicates that gossip recipients do not usually avoid or condemn gossip, but rather often respond with gossip of their own (Bergmann, 1993; Eder & Enke, 1991). Gossip, is therefore not reactively deviant.

Finally, in a normative test of deviance, a behavior is deemed deviant if it violates group norms (Spreitzer & Sonenshein, 2004). When Robinson and Bennett (1995) categorized gossip as deviance, they argued that gossip violates organizational norms. In contrast, research from other disciplines has argued that gossip is used to construct, enforce, and communicate group norms (Gluckman, 1963). That gossip is thought to both be a norm violation and also necessary to construct norms is paradoxical. If we focus less on extreme cases of negative gossip, and more on *typical* gossip manifestations, then a very different picture of gossip emerges. Instead of being a norm violation, it is more likely that gossiping is a behavioral norm of its own. Gossip is thought to be a ubiquitous behavior (Foster, 2004) which “does not have isolated roles in community life, but is part of the very blood and tissue of that life” (Gluckman, 1963, p. 308). Qualitative research has demonstrated that individuals who do not gossip are marginalized from their group (Gluckman, 1963; Loudon, 1961). This is consistent with gossip being an obligation, not a norm violation. Typical gossip is therefore unlikely to be normatively deviant.

While some extreme cases of negative WG may be deviant, it is our position that *typical* WG fails all theoretical tests of deviance. This suggests that WG may have been misconceptualized in the organizational literature. We therefore propose that WG be studied on its own, independent of deviance. To facilitate a shift in thinking about WG, we borrow from other academic literatures to reconceptualize and clarify the WG construct.

Workplace Gossip Reconceptualization

Following earlier definitions of gossip (see Eder & Enke, 1991; Foster, 2004; Kurland & Pelled, 2000; Noon & Delbridge, 1993), we define WG as *an informal and evaluative (i.e., positive or negative) conversation between work colleagues about another member of the organization who is not present to hear what is said*. At its core, WG is conceptualized as a social exchange of information (Cropanzano & Mitchell, 2005) through which norms and reputations are understood, constructed, and communicated within an organization. As such, typical WG is not generally a violation of norms, but rather an information gathering and validation process which can be used to evaluate behaviors and reputations. While norms and reputations are managed through WG, this is not thought to be the primary reason that individuals gossip. Instead, individuals can gossip to satisfy a variety of contextual needs, including the need to gather and validate information (Paine, 1967).

To qualify as WG, a conversation must include value-laden information which leads to an explicit or implied evaluation of the gossip subject's behavior or reputation. For example, a complaint about something that a co-worker has done could be WG, as it is a negative evaluation of the co-worker's behavior against workplace norms. Similarly, discussions of how nice a co-worker is would also be WG, as this is a positive evaluation of a co-worker's reputation. Conversations which only discuss factual information about a co-worker are not WG because these conversations are not evaluative and therefore do not affect the gossiper's understanding of norms or reputations. For example, saying that a co-worker has the flu would not be WG as there is no evaluation of the co-worker. In contrast, venting or bitching (Jones, 1980) about a co-worker's excessive use of sick-days would be WG as it is evaluative of norms and reputation.

WG is a conversation and therefore includes an exchange of information. Gossipers will always provide some value-laden information about the gossip subject. In exchange, the gossip

recipient(s) can confirm, disconfirm, or help to assess the gossiper's evaluation, often while providing extra confirmatory or disconfirmatory evidence (Bergmann, 1993; Eder & Enke, 1991). Communication which does not allow for a free-flowing exchange of information is generally not WG. For example, insults and name calling would not, on their own, qualify as WG, as they do not include an exchange of information. Most non-spoken communication (e.g., email, memos, and message forums) would also not generally be true WG, as the communication medium restricts the dynamics of the information exchange.

WG is always informal. Formal conversations, such as reporting a co-worker's inappropriate behavior to a human resources manager do not qualify as WG. While such a complaint could be value-laden, it is not a process for evaluating norms and reputations, but rather a direct appeal for formal intervention. Further, in this case, the human resources manager is restricted by their formal role and will often not engage in a truly evaluative conversation. Similarly, conversations among large groups or in official company meetings do not qualify as WG, because formal role and behavioral restrictions inherent to the formal setting generally prevent there from being a free-flowing evaluative conversation.

Our conceptualization of WG includes two dimensions: gossip valence and gossip subject. The first dimension, gossip valence, refers to the positive or negative nature of the WG evaluation (see Foster, 2004). Positive WG is seen as talk about normative behaviors or positive reputations (i.e., positive evaluations) while negative WG is talk about norm-violations or negative reputations (i.e., negative evaluations). It is our position that both positive and negative WG are required elements for understanding and constructing norms and reputations. Each provides behavioral examples which can be used to form or modify their respective category prototypes, thereby facilitating subsequent categorization. Together, positive and negative WG

enable the categorical delineation between acceptable and non-acceptable behaviors.

Our second dimension, gossip subject, identifies the individual who is being discussed in the gossip conversation. While gossip can be about a variety of different gossip subjects, we have chosen to focus on WG about supervisors and WG about co-workers. These two gossip subjects were chosen so as to maximize the research flexibility of our scales and to reflect the distinction that the organizational literature has made between behaviors directed toward co-workers (e.g., interpersonal deviance; Bennett & Robinson, 2000) and toward supervisors (e.g., supervisor-directed deviance; Mitchell & Ambrose, 2007).

Study Preview

The objective of this paper is to move beyond assumptions about WG and to facilitate future WG research. Unfortunately, despite calls for better measurement instruments (Noon & Delbridge, 1993), research into WG is severely limited by the lack of valid and appropriate general-purpose behavioral measures of WG. While a limited number of gossip measures do exist (e.g., Erdogan, Bauer, & Walter, 2014), they have not been sufficiently validated and generally measure gossip disposition rather than gossip behavior. This runs counter to gossip theory, which argues that gossip is a behavioral reaction to a contextual need (Paine, 1967).

In this paper, we address the long-standing gossip measurement deficiency by systematically developing and validating four WG scales (i.e., two dimensions: gossip valence and gossip subject) following the procedures recommended by Hinkin (1995; 1998). Scale development is summarized in 6 scale development phases. In Phase 1, we gather qualitative WG data to test assumptions about gossip and to guide item writing. In Phase 2, we generate and then select WG items. In Phase 3, we test the psychometric properties of the WG scales (e.g., reliability, model fit, and susceptibility to method effects). In Phase 4, we test the construct

validity of the scales by hypothesizing and testing relationships with a variety of other organizational variables. In Phase 5, we demonstrate the cross-cultural measurement invariance of the WG scales using North American and Chinese samples. Finally, in Phase 6, we use the negative WG about a supervisor scale and a measure of supervisor-directed deviance to demonstrate that WG is distinctly different than deviance, and that including gossip items in deviance scales may result in undesirable measurement issues, including an over-reporting of deviance behavior.

Phase 1: Qualitative Study of Workplace Gossip

We began our research by conducting a qualitative pilot study. Our primary goal with this individual study was to qualitatively assess the common assumption that gossip is primarily malicious talk (i.e., with intent to harm). While some research has examined relative frequencies of motivations for social gossip (Beersma & Van Kleef, 2012), we are unaware of any research that has qualitatively examined motivations for WG, specifically. A secondary goal of this study was to gather example WG manifestations that could guide item development for the WG Scales.

Participants and Procedure. For Sample 1, we recruited 153 students (52% male; mean age = 20.53, $SD = 1.21$) from a large Canadian business school to participate in a paper-and-pencil qualitative study of informal workplace communication. Previous work experience was a stated participation requirement (mean work experience = 3.49 years, $SD = 2.12$). Participants were asked to remember and describe two informal workplace conversations that they started: one in which they talked about their non-present supervisor, and one in which they talked about a non-present co-worker. Participants were asked to describe both the conversation content and their reasons for starting the conversation, in detail using an open-ended question format. To avoid potential bias, the word “gossip” was not used in the survey.

Results and Discussion. In total, this procedure generated detailed descriptions of 306 conversations (153 about a supervisor and 153 about a co-worker). Each conversation was evaluated by two independent raters to determine if it met the definition of WG. Specifically, conversations which were either not evaluative or which did not occur in the workplace were deemed to not be WG and were excluded from analysis. After review, 263 conversations were deemed to fit the definition of WG (221 negative WG, and 42 positive; 100% rater agreement). Valid WG cases were then independently categorized by the two raters according to a researcher-

generated list of motivations that were derived from existing theory and a random sampling of the WG cases. Raters were instructed to select the single motivation that best summarized the written content. After a first-pass categorization, new categories were added to better describe some rare cases. Raters then performed a second-pass categorization. All rater disagreements were completely resolved through rater discussions.

The WG cases were categorized according to motivation as follows: gathering and validating information (32% of WG cases), emotion venting and coping (29%), building friendships (7%), entertainment or alleviating boredom (6%), seeking or giving social support (6%), seeking help from co-workers (6%), spreading information (5%), defending another's reputation (3%), warning others (2%), and impression management (2%). No cases were categorized as best reflecting an intentional undermining motivation (i.e., intent to harm), although raters identified 3 cases (out of 263) as having some element of undermining in them.

While previous organizational research has focused almost exclusively on negative WG, evidence from this study indicates that both positive and negative WG do occur in organizations. This strongly contradicts a strictly malicious or deviance interpretation of WG. Further, evidence regarding gossip motivations indicates that individuals engage in WG for a variety of different reasons, many of which do not appear to be malicious. Overall, this pattern of results is more consistent with a functional view of gossip than a purely deviant view.

Phase 2: Item Generation and Reduction

Potential WG scale items were then generated based on WG manifestations from the qualitative study and on existing items in deviance scales (e.g., “made fun of someone at work”; Bennett & Robinson, 2000). In total, 34 potential WG items were developed, with each item carefully worded for dual usage in both supervisor and co-worker WG scales (e.g., “vented to a work colleague about something that your supervisor [another co-worker] has done”).

Substantive Validity Assessment

As recommended by Hinkin (1998), we next assessed the substantive validity of our 34 items using a substantive validity item-sort task (Anderson & Gerbing, 1991). Substantive validity is defined as the degree to which an item assesses the intended construct rather than some other construct (Holden & Jackson, 1979). It is a necessary requirement of content validity. Our goal for this study was to identify and retain items which were substantively valid and eliminate items which were not. This would build confidence that our items assess WG rather than some other unintended construct.

Participants and Procedure. For Sample 2, student participants were recruited from a psychology research pool at a large Canadian university to participate in a pencil-and-paper laboratory task ($N = 24$; 75% female; mean age = 19.92 years, $SD = .83$). Previous work experience was a stated participation requirement (mean work experience = 2.41 years, $SD = 1.90$). Participants were given the list of 34 supervisor-worded WG items and 5 construct definitions written in non-technical language. Participants were instructed to choose one construct which best fit each item-described behavior: either WG, rumor, urban legend, voice, or factual information.² In an effort to reduce bias and increase rational decision-making, no

² Construct definitions: WG - “An informal conversation about another person who is not present. The conversation is evaluative in nature (either positive or negative)”, Rumor - “Statements that are unconfirmed (i.e., speculation,

construct names were provided. After completing the item-sort task, participants rated the written clarity of each item (7-point scale; 1 = *very unclear* to 7 = *very clear*).

Results. In this sorting task, items are deemed to be substantively valid if they are assigned significantly more often to the intended construct than to the next most frequently chosen construct (for a full description of this statistical test, see Anderson & Gerbing, 1991; p. 734). Using the recommended analysis, 23 supervisor-worded items were found to be substantively valid at a $p < .01$ level. Clarity ratings for these 23 items were very high ($M = 6.01$, $SD = .27$). One other item (i.e., “told a negative story about your supervisor while talking to a work colleague”) was deemed to be a borderline case (sorted at $p = .01$). Because this item had one of the lowest mean clarity ratings ($M = 5.33$), we decided to retain this item after rewording (i.e., “told an unflattering story about your supervisor while talking to a work colleague”). In total, 10 items were eliminated in this step, leaving us with a pool of 24 WG items.

Item Frequencies

As noted by Hinkin (1998), items with excessively low or high frequencies can negatively affect scale variance. This can result from either a poorly-chosen response scale or inappropriate scale items. Typically, when gossip items have been included in deviance scales, gossip behavior has been assessed over long time periods (e.g., one year). In this study, we set out to establish that gossip behavior could be measured over a shorter timeframe than this, as gossip is not thought to be a low base-rate behavior and assessments over shorter timeframes are more likely to assess episodic memories of behavior (see Robinson & Clore, 2002). We therefore

conjecture, not true, etc.) and in widespread circulation. These statements are often speculation about events.” (Foster, 2004; Rosnow, 1991; 2001), Urban Legend - “A modern-day myth (i.e., a fictional tale) about unknown people or improbable events.” (Guerin & Miyazaki, 2006), Voice - “Speaking up with a constructive idea regarding how to improve something. This is not just a criticism – it is a positive and realistic proposal for change.” (Van Dyne & LePine, 1998), Factual Information - “News about a person or event. It is purely informational and intended for widespread distribution. It is not evaluative in nature (i.e., it does not say anything good or bad about a person – it is just news).” (DiFonzo & Bordia, 2007).

assessed item frequencies using both one-year and one-month response scales. Our goal was to choose an appropriate response scale and eliminate items which negatively affect scale variance.

Participants and Procedure. For Sample 3, Amazon Mechanical Turk (MTurk) was used to recruit full-time workers (U.S. or Canada) for an online survey. A total of 200 participants completed the study (53% female; mean age = 35.57, $SD = 11.37$; mean work experience = 16.06 years, $SD = 10.13$; mean hours per week = 40.73, $SD = 4.54$; mean organizational tenure = 4.77 years, $SD = 5.10$; mean time reporting to current supervisor = 2.59 years, $SD = 3.59$). Participants were randomly assigned to one of two study conditions: a one-year response-scale condition ($n = 98$; 1 = *never*, 2 = *once a year*, 3 = *twice a year*, 4 = *several times a year*, 5 = *monthly*, 6 = *weekly*, 7 = *daily*) or a one-month response-scale condition ($n = 102$; 1 = *never*, 2 = *once a month*, 3 = *2-3 times a month*, 4 = *once a week*, 5 = *2-3 times a week*, 6 = *once a day*, 7 = *more than once a day*). Participants assessed how frequently they had performed each WG behavior using the respective response scale.

Results and Discussion. There was no significant difference in the number of respondents who reported engaging in some versus no WG based on response scale timeframe ($p > .05$). We therefore concluded that the one-month response scale was appropriate for measuring the behavioral frequency of WG. We then assessed dichotomized scores (i.e., some versus no behavior) for individual items in the one-month condition. Four items were found to have low frequencies and were subsequently eliminated.³ No items were found to have excessively high frequencies. A total of 20 items were retained for further validation.

Final Item Reduction

Next, we collected a sample that could be used to finalize our selection of WG items.

³ Eliminated items: “told a work colleague that your supervisor is undeserving of his/her position”, “warned a work colleague about your supervisor”, “told a work colleague that you feel inspired by your supervisor's actions”, and “asked a work colleague if they have a positive impression of your supervisor”

Hinkin (1995; 1998) recommends that most constructs can be measured by 4 to 6 well-chosen items. Following this recommendation, our goal was to reduce our pool of 20 items to a final set of 4 to 6 items for each of our 4 scales (i.e., 2 dimensions: gossip valence and gossip subject).

Participants and Procedure. For Sample 4, MTurk was used to recruit an independent sample of full-time workers (U.S. or Canada). A total of 213 participants completed this study (51% female; mean age = 33.56 years, $SD = 10.55$; mean work experience = 15.11 years, $SD = 10.17$; mean hours per week = 40.31, $SD = 6.08$). Participant education was varied (high school: 22%; technical college: 16%; university: 49%; graduate degree: 13%). Individual income levels ranged from under \$20,000 to more than \$80,000 per year (median = \$30,000 to \$39,999). Participants were presented with 20 supervisor-worded items and 20 co-worker-worded items. WG behavior was assessed using the previously chosen 7-point one-month response scale.

Results and Discussion. An EFA (maximum likelihood with promax rotation) was performed on the WG items. Four factors with eigenvalues greater than 1.0 were identified (minimum eigenvalue = 2.33; total variance explained = 78%) and a scree-test verified this four-factor structure (see Fabrigar, Wegener, MacCallum, & Strahan, 1999). As expected, all items clearly loaded onto the intended factor (factor loadings = .70 to .99). Given this, we chose not to eliminate items by factor structure. Instead, items which maximized coverage of WG manifestations as seen in our qualitative study were retained. Items which were deemed to not be as conceptually important or which were statistically redundant to other items were eliminated. In total, 8 items were eliminated during this step.⁴

⁴ Eliminated items: “said something unflattering about your supervisor while talking to a work colleague”, “told a work colleague that your supervisor upset you”, “expressed disappointment about your supervisor while talking to a work colleague”, “complained to a work colleague about how your supervisor treats people”, “expressed frustration regarding your supervisor while talking to a work colleague”, “told a work colleague that your supervisor made you angry”, “spoke highly of your supervisor while talking to a work colleague”, and “told a work colleague that you admire something that your supervisor has done”.

A confirmatory factor analysis (CFA) was then performed on the remaining items (i.e., 4 factors with 6 items each) to assess model fit, $\chi^2(246, N = 213) = 655.45, p < .001$. Inspection of the residuals showed that a better fit could be obtained after the elimination of two more items.⁵ The remaining items (i.e., 4 factors with 5 items each) were found to have very good model fit, $\chi^2(164, N = 213) = 384.33, p < .001, CFI = .98, SRMR = .05$. All items loaded onto the intended factor at a significant level ($p < .001$), with factor loadings ranging from .79 to .92.

With this evidence, the items were deemed to be finalized, pending further validation. The result was four 5-item WG scales: Negative WG about a Supervisor (NWGS), Positive WG about a Supervisor (PWGS), Negative WG about Co-Workers (NWGC), and Positive WG about Co-Workers (PWGC). The final WG scales are shown in Appendix A.

⁵ Eliminated items: “told a work colleague that your supervisor deserves good things”, and “made fun of your supervisor while talking to a work colleague”

Phase 3: Psychometric Properties

We next evaluated the psychometric properties of the WG scales using three independent samples (i.e., Samples 5, 6, and 7). Factor structure, model fit and internal consistency were assessed for all three samples. Susceptibility to method effects stemming from Negative Affect (NA), Positive Affect (PA), and Impression Management (IM) were estimated using Sample 6.

Participants

Sample 5 was an independent, multi-wave, multi-source data sample that was recruited using physical advertisements in local-area businesses and online advertisements. Data were collected in a prescreen survey and three subsequent survey waves, each separated by one week. A total of 183 participants (52% female; mean age = 36.79, $SD = 8.63$; mean hours per week = 41.37, $SD = 5.67$; organization tenure = 6.08 years, $SD = 5.88$; mean years reporting to supervisor = 3.11, $SD = 2.64$) completed the first wave, while 177 completed the second and third waves (97% retention rate). Participants were employed in a variety of industries, including architecture and engineering (6%), business and financial operations (22%), computer industries (8%), education (10%), government (10%), healthcare (11%), office administration (4%), and production (9%). During the prescreen, participants were asked to nominate a work peer who could also fill out an online survey. A total of 125 peer-reports were completed (68% response rate; 49% male; mean age = 38.57, $SD = 10.19$).

Sample 6 was an independent, multi-wave data sample recruited using MTurk (employed full-time, U.S. or Canada). Potential participants were first given a prescreen questionnaire intended to verify English-language comprehension and full-time working status. A total of 479 respondents correctly responded to simple verbal comprehension questions, multiple attention checks, and indicated that they had regular interaction with co-workers. These

participants were then invited to participate in a follow-up two-wave survey. A total of 266 participants completed the first wave (56% response rate; 50% male; mean age = 37.84, $SD = 11.19$ years; organization tenure = 5.77 years, $SD = 6.11$; mean years reporting to supervisor = 3.38, $SD = 4.35$). One week later, 222 participants completed the second wave (83% retention rate). Participant education level was varied (high school = 18%; university or college = 63%; Master's degree = 15%; Doctoral degree = 3%). Participants were employed in a variety of industries, including business and financial operations (9%), computer industries (9%), education (10%), food services (5%), healthcare (7%), office administration (8%), production (5%), sales (14%), and transportation (4%). Modal household income was \$40,000 to \$49,999 per year.

Sample 7 was an independent, single-wave online sample recruited from a survey panel managed by Qualtrics, LLC. Participants were prescreened for four criteria: they had to be employed full-time, reside in either the United States or Canada, work at a company location, and interact with co-workers at least once per day. A total of 177 participants completed the survey (61% female; mean age = 43.09, $SD = 11.74$; organization tenure = 7.18 years, $SD = 7.45$). Participant education level was varied (high school = 23%; technical college = 21%; undergraduate degree = 42%; graduate degree = 13%).

Refer to Table 1 for a summary of the measures included in each sample.

Measures

Workplace Gossip (WG). In an effort to validate multiple anticipated usage patterns for the WG scales, WG was measured using multiple item-presentation formats. In Sample 5, the WG items were presented as four separate item-blocks spread across two survey waves (i.e., NWGS and PWGS in wave 2; NWGC and PWGC in wave 3). In Sample 6, the items were presented as two item-blocks, both in wave 1 (i.e., NWGS and PWGS items were randomly

interleaved; NWGC and PWGC items were randomly interleaved). In Sample 7, the items were presented as four separate item-blocks, all in a single wave. In all samples, items were randomly presented. See Appendix A for instructions and response scale.

Positive and Negative Affect were measured using the 20-item PANAS (Watson, Clark, & Tellegen, 1988). Participants assessed how much they felt certain emotions at the present moment (e.g., “distressed”; 5-point scale; 1 = *very slightly or not at all* to 5 = *extremely*).

Impression Management was measured using the impression management scale (Paulhus, 1991; 20-items; e.g., “I sometimes tell lies if I have to”; 7-point scale; 1 = *not true*, 4 = *somewhat true*, 7 = *very true*). Following the scale’s published scoring guidelines, items with low scores (i.e., 1 to 5) were rescored as a 0 while items with high scores (i.e., 6 or 7) were rescored as a 1. The overall scale score is a sum of the rescored items (out of 20).

Analytic Strategy

Our general analytic strategy was to first establish the four-factor structure of our scale items by conducting a series of CFAs across all three samples. With evidence of acceptable factor structure, we would then analyse the psychometric properties of the 4 individual WG scales. Following the recommendations in Hu and Bentler (1999), we adopted a two-index reporting strategy for model fit, with a recommended cut-off criteria of .95 or higher for the comparative fit index (CFI) and .08 and lower for the standardized root-mean-square residual (SRMR). Models that meet these criteria are considered to have very good model fit. All factor analyses were performed using LISREL 9.1.

We tested for the presence of method effects in Sample 6 using a measured latent-variable technique (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Williams & Anderson, 1994). The procedure specifies a structural model where paths are defined between each of the

method effects latent variables and from these variables to the WG indicators. Two nested models are tested: a baseline model with the path loadings from the method effects variables to WG indicators constrained to zero and a confounded-measurement model where these path loadings are freely estimated. The presence of method effects is indicated when the model fit of the confounded-measurement model is significantly better than the model fit of the baseline model. An estimate of the percentage of variance explained by the method effects can then be calculated by squaring the standardized path loadings from the latent method effects variables to the indicators of the WG scale. Following recommendations from previous research, we tested method effects resulting from NA, PA, and IM (Ferris, Brown, Berry, & Lian, 2008; Williams & Anderson, 1994). Bentler and Chou (1987) recommends that structural models have a participant to estimated parameter ratio of 5:1. To meet this recommendation, we randomly combined method effects items into four parcels per latent method effect variable (see Williams & Anderson, 1994; Williams, Gavin, & Williams, 1996). The WG indicators were not parcelled.

Results

Four-Factor Structure. In all 3 samples, CFAs supported our hypothesized four-factor structure (mean $\chi^2(164) = 409.72$; mean CFI = .97; mean SRMR = .06). Factor loadings ranged from .66 to .98 ($M = .87$) and every indicator loaded onto the intended factor ($p < .001$).

NWGS. The NWGS scale had very good model fit across all samples (mean $\chi^2(5) = 17.41$, mean CFI = .99, mean SRMR = .02). All factor loadings were significant ($p < .001$). Reliabilities were good ($\alpha = .92, .93, .94$) and the average corrected item-total correlations were high ($M = .79, .81, .76$). NWGS had a significant correlation with NA ($r = .34, p < .001$) and IM ($r = -.29, p < .001$), but not with PA. Method effects were present ($\Delta\chi^2(15, N = 266) = 63.50, p < .001$), but these effects were weak (NA: 6% variance explained; IM: 5%; PA: 0%).

PWGS. The PWGS scale had very good model fit in all three samples (mean $\chi^2(5) = 9.08$, mean CFI = 1.00, mean SRMR = .01). All factor loadings were significant ($p < .001$). Reliabilities were good ($\alpha = .97, .90, .95$) and average corrected item-total correlations were high ($M = .91, .75, .78$). PWGS was significantly related to PA ($r = .26, p < .001$) and NA ($r = .12, p < .05$), but not to IM. Significant method effects were present ($\Delta\chi^2(15, N = 266) = 51.37, p < .001$), but these effects were weak (PA: 6% variance explained; NA: 1%; IM: 0%).

NWGC. The NWGC scale had very good model fit in all three samples (mean $\chi^2(5) = 29.63$, mean CFI = .98, mean SRMR = .02). All factor loadings were significant ($p < .001$). Reliabilities were good ($\alpha = .95, .92, .95$) and average corrected item-total correlations were high ($M = .86, .78, .79$). NWGC was significantly related to NA ($r = .32, p < .001$) and IM ($r = -.31, p < .001$), but not PA. Method effects were present ($\Delta\chi^2(15, N = 266) = 82.62, p < .001$), but they were again weak (NA: 5% variance explained; IM: 5%; PA: 0%).

PWGC. The PWGC scale had very good model fit in all three samples (mean $\chi^2(5) = 10.95$, mean CFI = 1.00, mean SRMR = .01). All factor loadings were significant ($p < .001$). Reliabilities were good ($\alpha = .97, .92, .95$) and average corrected item-total correlations were high ($M = .92, .80, .80$). PWGC was significantly related to PA ($r = .25, p < .001$) and NA ($r = .16, p < .01$) but not IM. Method effects were present ($\Delta\chi^2(15, N = 266) = 45.56, p < .001$), but they were also weak (PA: 6% variance explained; NA: 2%; IM: 0%).

WG Correlations. Mean correlations between the WG scales (Samples 5 to 7, weighted by sample size; Hunter & Schmidt, 2004) were all significant (NWGS and NWGC: $\bar{r} = .63, p < .001$; PWGS and PWGC: $\bar{r} = .55, p < .001$; NWGS and PWGC: $\bar{r} = .27, p < .001$; PWGS and NWGC: $\bar{r} = .20, p < .01$; NWGC and PWGC: $\bar{r} = .35, p < .001$), except the relationship between NWGS and PWGS ($\bar{r} = .10, n.s.$). The NWGS and PWGS relationship is unique in that it is the

only correlation in which both measures assess WG about a single gossip subject.

Discussion

Evidence confirmed our hypothesized 4-factor model, with each factor corresponding to a single WG scale. Model fit and reliabilities for each of the individual WG scales were very good. While method effects due to NA, PA, and IM were found, estimates of the size of these effects indicated that the effects were weak and consistent with findings in previous research (Ferris et al., 2008; Keeping & Levy, 2000; Munz, Huelsman, Konold, & McKinney, 1996; Williams et al., 1996; for a summary, see Podsakoff et al., 2003). With this evidence, all four WG scales were deemed to possess acceptable psychometric properties.

Phase 4: Construct Validity

We next evaluated the construct validity (Cronbach & Meehl, 1955; Hinkin, 1995) of the WG scales using the 3 samples described in the previous phase. This process included assessments of convergent, discriminant, and criterion-related validity. To assess convergent validity, we tested the relationship between the negative WG scales and the tendency to gossip scale (TTG; Erdogan et al., 2014). Whereas the WG scales are behavioral frequency scales, the TTG is a dispositional measure that distinguishes items by gossip content rather than by behavior (e.g., “At work, I talk about other people’s poor performance” and “At work, I talk about other people’s failures”). As previous research argues that gossip results primarily from contextual needs (Paine, 1967) and not from an individual’s disposition, we expected the WG and TTG correlations to be significant, but not so high that the scales measured the same construct.

To assess discriminant validity, we tested that the WG scales were distinct from all other measures to which they were significantly correlated with across all 3 samples. We also predicted that the 4 scales would not be significantly related to age or gender. This prediction reflects a universal and functional view of gossip and is consistent with previous research suggesting that gender differences in gossiping are primarily related to gossip content rather than overall gossip frequency (Levin & Arluke, 1985; Nevo, Nevo, & Derech-Zehavi, 1993).

Finally, we hypothesized and tested relationships between WG and other constructs to which WG should relate. Evidence from this procedure developed an initial nomological network for WG and established the criterion-related validity of the scales. We hypothesized relationships between WG and a variety of other classes of workplace variables and processes, including sensemaking, emotion validation, organizational justice, influence, self-enhancement, job performance, and turnover.

Sensemaking. Research has shown that the most common self-reported motivation for gossiping is to gather and validate information (Beersma & Van Kleef, 2012). In organizations, this motivation is closely related to sensemaking, defined as a process “through which people work to understand issues or events that are novel, ambiguous, confusing, or in some other way violate expectations” (Maitlis & Christianson, 2014, p. 57). While both positive and negative gossip can be used to gather information, gossip researchers have predicted that negative gossip will be more effective at gathering threat-relevant information (Wert & Salovey, 2004). With this in mind, we predicted that negative WG would be related to three different organizational variables that should be related to a need for sensemaking.

Hypothesis 1: NWGS and NWGC will be positively related to job ambiguity, job insecurity, and workplace anxiety

Emotion Validation. Gossip is thought to be a social comparison process (Suls, 1977; Wert & Salovey, 2004) that can be used to gather information to validate emotions (Schachter, 1959). In effect, gossip can be used to help an individual clarify emotions and test attributions (Kelley, 1973) of the causes of those emotions. In cases of emotion validation, we would expect gossip valence to match affect valence, as this is a direct and informative test of an emotional hypothesis. This relationship should be most evident for emotions with high levels of arousal, as arousal has been shown to be related to the social transmission of information (Berger, 2011).

Hypothesis 2: NWGS and NWGC will be positively related to negative job-related affective well-being (high arousal), while PWGS and PWGC will be positively related to positive job-related affective well-being (high arousal).

Organizational Justice. Organizational justice refers to an individual’s subjective perceptions of fair treatment by their organization or supervisor (Colquitt, 2001). It is comprised

of four factors: distributive justice (i.e., fairness of outcomes), procedural justice (i.e., fairness of procedures), interpersonal justice (i.e., politeness and dignity in treatment), and informational justice (i.e., information regarding why or how something happened). Researchers have predicted that gossip will be related to justice, as individuals can use gossip to validate their opinions of fair treatment (Wert & Salovey, 2004). For example, individuals may gossip about a co-worker's wages as a way to assess the fairness of their own wages (i.e., distributive justice).

Hypothesis 3: NWGS will be negatively related to, and PWGS will be positively related to, distributive, procedural, interpersonal, and informational justice.

Influence. Anthropologists have long argued that negative gossip can be used as an influence tactic to enforce group norms of acceptable behavior through sanction and reputation management (Gluckman, 1963). We would therefore expect negative WG to occur in response to others violating group norms. Specifically, we predict that WG will relate to co-worker incivility, defined as co-worker behaviors that are rude or discourteous and in violation of norms for respect (Andersson & Pearson, 1999).

Hypothesis 4a: NWGS and NWGC will be positively related to co-worker incivility.

Gossip has been described as a tool for advancing personal interests (Paine, 1967), building friendships, and influencing others (Foster, 2004). In an organization, the ability to influence others is related to an individual's political skill (Ahearn, Ferris, Hochwater, Douglas, & Ammeter, 2004) and one component of political skill is the ability to network with other individuals (Ferris et al., 2005). As gossip is a tool for building relationships and influencing others, it should be related to networking ability. Positive WG, specifically, should be related to networking ability, as positive gossip is less likely than negative gossip to elicit hostility in others.

Hypothesis 4b: PWGS and PWGC will be positively related to networking ability.

Self-enhancement. Gossip researchers have used social comparison theory to argue that gossip can be a functional tool for self-enhancement and bolstering self-esteem (Suls, 1977; Wert & Salovey, 2004). To do this, an individual with low self-esteem would use negative gossip to construct and then validate a downward social comparison between themselves and a less fortunate other, thereby enhancing subjective well-being and self-esteem (Wills, 1981).

Hypothesis 5: NWGS and NWGC will be negatively related to self-esteem.

Job Performance. Job performance comprises three different categories of behavior: in-role performance, citizenship behaviors, and deviant behaviors (Rotundo & Sackett, 2002). In-role performance (i.e., task performance) includes behaviors which are formally recognized as advancing the technical core of the job (Borman & Motowidlo, 1993). Social comparison theory states that when objective information about performance is not available, individuals will construct social comparisons to better assess their own performance abilities (Festinger, 1954). As gossip is thought to be a social comparison process, we would expect it to be related to in-role performance. Specifically, we would expect individuals who are under-performing or not sure of how they are performing to use WG to gather information about acceptable performance. Poor performers could also use negative WG to feel better about their performance, similar to how negative gossip can be used to enhance self-esteem.

Hypothesis 6a: NWGS, PWGS, NWGC, and PWGC will be negatively related to in-role performance.

Organizational citizenship behavior (OCB) is defined as “performance that supports the social and psychological environment in which task performance takes place” (Organ, 1997, p. 95) and is often seen as encompassing interpersonal helping behaviors (e.g., “go out of the way

to make newer employees feel welcome in the work group” (Lee & Allen, 2002). Gossip is frequently conceptualized as a tool used to deliver information (Baumeister et al., 2004) and to create and maintain relationships (Foster, 2004). In situations where trust between conversation partners is low or non-existent, like when conversing with a new hire, we would expect employees to use positive WG as a way to inform other employees of the social environment.

Hypothesis 6b: PWGS and PWGC will be positively related to OCB.

Deviance is defined as a violation of organizational norms (Robinson & Bennett, 1995). It is seen as a form of counterproductive workplace behavior (CWB; Sackett, 2002), as it acts against an organization’s interests. Traditionally, WG has been seen as a form of deviance, with gossip items being included in deviance scales. We have argued that although extreme cases of WG may be deviant, typical WG is not generally deviant. However, because WG is an information gathering process, WG can be used to make causal attributions (Kelley, 1973) which could affect decisions to engage in deviant behavior (Barclay, Skarlicki, & Pugh, 2005). We therefore predicted WG to be related to, but distinct from, deviance and CWB.

Hypothesis 6c: NWGS and PWGS will be positively related to supervisor-directed deviance. NWGS, PWGS, NWGC, and PWGC will be positively related to interpersonal deviance, instigated workplace incivility, and CWB.

Turnover. The decision to quit is often accompanied by difficult and complex emotions. Drawing upon previous gossip research which theorizes that gossip is an effective tool for validating opinions and emotions (Wert & Salovey, 2004), we expect that WG will be related to turnover intentions and job-search behaviors. In particular, information gathered from negative WG about a supervisor will be especially valuable when deciding to quit, as poor supervisor-subordinate relationships and negative supervisor behavior are related to turnover intentions

(Ferris, 1985; Graen, Liden, & Hoel, 1982; Schyns & Schilling, 2013; Tepper, 2000).

Hypothesis 7: NWGS will be positively related to turnover intentions and job-search behaviors.

Measures

Workplace Gossip (WG) was assessed as described in Phase 3.

Tendency to Gossip (TTG) was measured with Erdogan et al.'s (2014) dispositional measure of WG (4-items; e.g., "At work, I talk about other people's failures"; 7-point scale; 1 = *strongly disagree* to 7 = *strongly agree*).

Job Ambiguity was measured using the job ambiguity scale (Breugh & Colihan, 1994; 9-items; e.g., "I know what my supervisor considers satisfactory work performance"; 7-point scale; 1 = *disagree strongly* to 7 = *agree strongly*). For readability, this scale was reverse-coded (i.e., 1 = *low ambiguity* to 7 = *high ambiguity*).

Job Insecurity was measured using the job insecurity scale (Boswell, Olson-Buchanan, & Harris, 2014; modification of Ashford, Lee, & Bobko, 1989; 6-items). Respondents were asked to evaluate the likelihood of certain negative events happening to their employment (e.g., "lose your job and be laid off permanently"; 5-point scale; 1 = *very unlikely* to 5 = *very likely*).

Workplace Anxiety was assessed using the 2-item Rodell and Judge (2009) measure. Participants assessed the extent that they feel "nervous" or "anxious" at work, in general (5-point scale; 1 = *very slightly* to 5 = *extremely*).

Job-Related Affective Well-being was measured using two 5-item subscales (negative affect high arousal and positive affect high arousal) from the Van Katwyk, Fox, Spector, and Kelloway (2000) measure. Respondents assessed their emotions (e.g., "my job made me feel excited") over the last month (5-point scale; 1 = *never* to 5 = *extremely often or always*).

Organizational Justice was measured using the four Colquitt (2001) justice measures (distributive: 4-items; procedural: 7-items; interpersonal: 4-items; informational: 5-items). Participants indicated their agreement with justice statements (e.g., “Has he/she treated you with dignity?”; 5-point scale; 1 = *to a small extent*, 3 = *neutral*, 5 = *to a large extent*).

Networking Ability was assessed using the Ferris et al. (2005) networking ability subscale (6-items; e.g., “I am good at building relationships with influential people at work”; 7-point scale; 1 = *strongly disagree* to 7 = *strongly agree*).

Co-worker Incivility was measured using the Cortina, Magley, Williams, and Langhout (2001) measure. Participants assessed co-worker incivility (7-items; e.g., “Put you down or were condescending to you?”) over the last four months (5-point scale, 1 = *never* to 5 = *many times*).

Self-esteem was measured using Rosenberg’s (1965) measure (10-items; e.g., “On the whole, I am satisfied with myself”; 6-point scale; 1 = *strongly disagree* to 6 = *strongly agree*).

In-Role Performance was assessed using the in-role behavior scale (Williams & Anderson, 1991). Work peers assessed the focal employee’s performance (7-items; e.g., “meets formal requirements of the job”; 7-point scale; 1 = *strongly disagree* to 7 = *strongly agree*).

Supervisor-Directed Deviance (SDD) was assessed using the Mitchell and Ambrose (2007) measure. Participants assessed how often they had engaged in specific behaviors (10-items; e.g., “Made an obscene comment or gesture toward my supervisor”) over the last year (7-point scale; 1 = *never* to 7 = *daily*).

Interpersonal Deviance was assessed using the Bennett and Robinson (2000) interpersonal deviance scale. Participants assessed their behavior (7-items; e.g., “Acted rudely toward someone at work”) over the last year (7-point scale; 1 = *never* to 7 = *daily*).

Instigated Workplace Incivility was assessed using the Blau and Andersson (2005)

measure. Participants assessed their behavior (7-items; e.g., “Put down others or were condescending to them in some way”) over the last year (5-point scale; 1 = *never* to 5 = *most of the time*).

Organizational Citizenship Behaviors (OCB) was assessed using two techniques. First, three 6-item OCB subscales from Dalal, Lam, Weiss, Welch, and Hulin (2009) were combined (i.e., OCB toward supervisor, co-workers, and organization; mean subscale $\bar{r} = .72, p < .001$) into a single measure. Work-peers rated how often the focal employee engaged in OCBs (e.g., “Tried to help his/her coworker”) during the last year (7-point scale; 1 = *never* to 7 = *daily*). Second, we used the Lee and Allen (2002) interpersonal OCB sub-scale to assess self-reported behaviors (8-items; e.g., “Assist others with their duties”) over the last month (7-point scale; 1 = *never* to 7 = *always*).

Counterproductive Workplace Behaviors (CWB) was assessed by combining three 6-item CWB scales (Dalal et al., 2009; i.e., CWB toward supervisor, co-workers, and organization; mean subscale $\bar{r} = .85, p < .001$) into a single measure. Work-peers rated how often the focal employee engaged in CWBs (e.g., “Tried to harm his/her supervisor”) during the last year (7-point scale; 1 = *never* to 7 = *daily*).

Turnover Intentions were measured using the turnover intentions scale (Wayne, Shore, and Liden, 1997; 5-items; e.g., “As soon as I can find a better job, I’ll leave my company”; 7-point scale; 1 = *strongly disagree* to 7 = *strongly agree*).

Job Search Behavior was measured using the job search behavior scale (Blau, 1994). Participants assessed their behavior (12-items; e.g., “filled out a job application”) over the last 6 months (5-point scale; 1 = *never (0 times)* to 5 = *very frequently (at least 10 times)*).

Results

Convergent Validity. As predicted, there was a significant positive mean correlation (Samples 5 and 6) between the tendency to gossip scale and the negative WG scales (NWGS: $\bar{r} = .38, p < .001$; NWGC: $\bar{r} = .52, p < .001$), providing convergent validity evidence.⁶

Discriminant Validity. Discriminant validity tests were performed for every significant correlation between WG and another construct. CFAs showed that in every case, a 2-factor model was a better model fit than a 1-factor model (smallest $\Delta\chi^2(1) = 133.93$, all $ps < .001$; see Anderson & Gerbing, 1988; Bagozzi, & Phillips, 1982). No correlation confidence interval (plus and minus 2 standard errors) included 1.0. The Fornell and Larcker (1981) test showed that, in every case, the average variance explained (AVE) by both factors was greater than the squared correlation between the constructs (AVE Samples 5 to 7; NWGS: .70, .69, .76; PWGS: .86, .64, .78; NWGC: .79, .68, .78; PWGC: .87, .71, .80; smallest AVE for other scales = .48; largest $r^2 = .39$). Evidence from all three tests demonstrated that the WG constructs were distinct from every other construct. As predicted, WG was not, on average, significantly correlated with respondent age (NWGS: $\bar{r} = -.09, n.s.$; PWGS: $\bar{r} = -.05, n.s.$; NWGC: $\bar{r} = -.12, n.s.$; PWGC: $\bar{r} = .04, n.s.$) or gender (NWGS: $\bar{r} = -.03, n.s.$; PWGS: $\bar{r} = .01, n.s.$; NWGC: $\bar{r} = -.04, n.s.$; PWGC: $\bar{r} = .03, n.s.$).

Sensemaking. As predicted in Hypothesis 1, both negative WG scales were significantly positively related to job ambiguity (NWGS: $r = .33, p < .001$; NWGC: $r = .23, p < .001$), job insecurity (NWGS: $r = .42, p < .001$; NWGC: $r = .37, p < .001$), and workplace anxiety (NWGS: $r = .35, p < .001$; NWGC: $r = .27, p < .001$).

Emotion Validation. As predicted in Hypothesis 2, the negative WG scales were

⁶ In a follow-up test, multiple regressions were performed to verify that the WG scales had incremental predictive ability beyond the tendency to gossip scale (TTG). Tests showed that for 29 of 31 hypothesized relationships (Samples 5 and 6), the WG scales predicted significantly beyond TTG. Only the NWGC to self-esteem and NWGC to job-related affective well-being (negative) relationships became non-significant when TTG was included. Conversely, the TTG scale was found to have no significant predictive ability beyond the negative WG scales for relationships with supervisor-directed deviance, interpersonal deviance, in-role performance, CWB, turnover intentions, and job search behaviors.

significantly positively related to negative job-related affective well-being (NWGS: $r = .45, p < .001$; NWGC: $r = .23, p < .001$), and the positive WG scales were significantly positively related to positive job-related affective well-being (PWGS: $r = .36, p < .001$; PWGC: $r = .19, p < .01$).

Organizational Justice. As predicted in Hypothesis 3, justice was significantly negatively related to NWGS and significantly positively related to PWGS: distributive justice (NWGS: $r = -.25, p < .01$; PWGS: $r = .33, p < .001$), procedural justice (NWGS: $r = -.21, p < .01$; PWGS: $r = .28, p < .001$), interpersonal justice (NWGS: $r = -.33, p < .001$; PWGS: $r = .24, p < .01$), and informational justice (NWGS: $r = -.27, p < .001$; PWGS: $r = .23, p < .01$).

Influence. As predicted in Hypotheses 4a and 4b, both negative WG scales were significantly positively related to co-worker incivility (NWGS: $r = .53, p < .001$; NWGC: $r = .45, p < .01$), while both positive WG scales were significantly positively related to networking ability (PWGS: $r = .39, p < .001$; PWGC: $r = .22, p < .01$).

Self-Enhancement. As predicted in Hypothesis 5, negative WG was significantly negatively related to self-esteem (NWGS: $r = -.29, p < .001$; NWGC: $r = -.22, p < .01$).

In-Role Performance. Hypothesis 6a predicted that all four WG scales would be negatively related to in-role performance. Three of four WG scales were found to be significantly negatively related to peer-reports of in-role performance (NWGS: $r = -.35, p < .001$; PWGS: $r = -.20, p = .03$; NWGC: $r = -.37, p < .001$; PWGC: $r = .01, n.s.$).

Citizenship Behaviors. Hypothesis 6b predicted that both positive workplace scales would be significantly positively related to OCB. This was supported in self-reports of interpersonal OCB (PWGS: $r = .30, p < .001$; PWGC: $r = .30, p < .001$), but only PWGC was significantly related to peer-reports of OCB (PWGS: $r = .00, n.s.$; PWGC: $r = .28, p < .01$).

Deviance. As predicted in Hypothesis 6c, supervisor WG was significantly positively

correlated with supervisor-directed deviance (NWGS: $r = .62, p < .001$; PWGS: $r = .30, p < .001$). All four WG scales were significantly positively related to interpersonal deviance (NWGS: $r = .50, p < .001$; PWGS: $r = .25, p = .001$; NWGC: $r = .66, p < .001$; PWGC: $r = .18, p = .02$) and peer-reported CWB (NWGS: $r = .41, p < .001$; PWGS: $r = .26, p < .01$; NWGC: $r = .33, p < .001$; PWGC: $r = .19, p = .04$). Three of four WG scales were significantly positively related to instigated workplace incivility, while the fourth was marginally significant (NWGS: $r = .46, p < .001$; PWGS: $r = .12, p = .08$; NWGC: $r = .52, p < .001$; PWGC: $r = .24, p < .001$).

Turnover. As predicted in Hypothesis 7, NWGS was significantly positively related to turnover intentions (NWGS: $r = .33, p < .001$) and job search behaviors (NWGS: $r = .41, p < .001$).

Discussion

Results from three independent samples provided evidence of the construct validity of the WG scales. In three tests of discriminant validity, the WG scales were shown to be distinct from every other measured construct, including measures of supervisor-directed deviance, interpersonal deviance, instigated workplace incivility, and counterproductive workplace behaviors. The WG scales were also unrelated to both age and gender. Evidence confirmed our hypothesized nomological network, showing that WG was related to a variety of important organizational variables. In summary, there was considerable evidence which supported the construct validity of the WG scales.

Phase 5: Cross-Cultural Measurement Invariance

In an effort to increase the accessibility of WG research, we next sought to empirically establish that WG can be studied in non-North American populations. To do this, a cross-cultural measurement invariance test (Vandenberg & Lance, 2000) of WG was performed using North American and Chinese samples. If supported, measurement invariance would indicate that samples from both cultures had a similar conceptualization of WG in mind when they responded to the WG scales. Although gossip is often argued to be a universal behavior (Dunbar, 2004), we are unaware of any previous cross-cultural tests of gossiping behavior.

Participants and Procedure. Sample 8 was an independent sample of Chinese participants recruited from an online survey panel managed by Qualtrics, LLC. Participants were prescreened for four criteria: they had to reside in China, be employed full-time, work at a company location, and interact with co-workers at least once per day. A total of 211 participants completed the survey (65% male; mean age = 37.56, $SD = 9.43$; mean organization tenure = 7.54 years, $SD = 7.08$; mean time reporting to current supervisor = 4.08 years, $SD = 4.40$). All survey materials were presented in Chinese. To facilitate this, the WG scales were translated from English into Chinese and then back-translated into English by two independent bilinguals (see Brislin, 1980). The accuracy of the translation was then independently verified by a third bilingual. The scales were presented with random scale and item ordering. See Appendix B for the Chinese version of the WG scales.

Analysis. The cross-cultural measurement invariance test was performed on Sample 7 (North America) and Sample 8 (China) using the procedure recommended by Vandenberg and Lance (2000).⁷ In this procedure, there are three required tests of measurement invariance:

⁷ To evaluate measurement invariance, a series of nested models are assessed using a multigroup CFA. Each model introduces new model constraints that force specific parameters to be equal between samples. The change in model

configural invariance (i.e., equivalent factor structure), metric invariance (i.e., equivalent factor loadings), and scalar invariance (i.e., equivalent item intercepts). We assessed model fit for each of the three required invariance tests. All analyses were performed using LISREL 9.1.

Results and Discussion. The procedural steps and estimates of model fit are shown in Table 5. In each step of the procedure, subsequent models had a negligible change in model fit, meeting recommended guidelines. These results show that all four WG Scales had full configural, metric, and scalar invariance across the two samples.⁸ This suggests that North American and Chinese respondents had a similar conceptual frame of reference when they responded to the WG scales. While preliminary, this evidence indicates that it may be appropriate to study WG in non-North American populations (China specifically). Further, these results suggest that there may not be meaningful differences between WG in individualistic (e.g., North America) and collectivistic (e.g., China) cultures. This is consistent with gossip theories which argue that gossip is universal and cross-cultural (Dunbar, 2004; Gluckman, 1963).

fit between subsequent models is evaluated to determine if the introduced constraints have a negligible impact on fit. Steps which result in a change in model CFI of less than or equal to -.01 are deemed to be negligible model changes, implying invariance between the two samples for that particular test (Cheung & Rensvold, 1999; 2002). (For a description of measurement invariance, the reader is referred to Vandenberg & Lance, 2000; also Milfont & Fischer, 2010, and Steenkamp & Baumgartner, 1998)

⁸ Seeking further evidence of the cross-cultural invariance of the WG scales, we also performed three optional tests of invariance, assessing error-variance invariance (i.e., invariant uniqueness), factor-variance invariance, and factor-mean invariance (Vandenberg & Lance, 2000). NWGS and PWGS were found to possess partial error-variance invariance (NWGS: items 4 and 5 were unconstrained; PWGS: items 3 and 4 were unconstrained), while NWGC and PWGC were fully error-variance invariant. All four scales were both full factor-variance invariant and full factor-mean invariant.

Phase 6: Toward Studying WG Independent of Deviance

In this paper, we have argued that typical WG does not fit a deviance conceptualization. Evidence from Phase 4 supports this, demonstrating that all 4 WG scales are distinct from a variety of deviance scales. In this study phase, we assess the common practice of including gossip items in deviance scales. Specifically, we examine the supervisor-directed deviance (SDD; Mitchell & Ambrose, 2007) and NWGS scales to better understand the effect that including gossip items in a deviance scale can have on measurement. The SDD scale was chosen for this test because it appears to include two gossip items. The first item (i.e., “made fun of my supervisor at work”) appears to assess WG manifesting as humor (Morreall, 1994) while the second item (i.e., “gossiped about my supervisor”) directly assesses WG. In our analysis, we first verify that these items assess WG rather than SDD and then examine whether assessing higher-frequency gossip behaviors together with statistically deviant (i.e., low base-rate) behaviors could result in false positives and the over-reporting of deviance behavior.

Measures

All measures were assessed using Samples 3, 4, and 5 from earlier study phases.

Supervisor-Directed Deviance (SDD) was measured as described in Phase 4.

Negative Workplace Gossip about a Supervisor (NWGS). In Sample 3, NWGS was assessed in both the one-month and one-year response-scale conditions. In Samples 4 and 5, NWGS was measured using a one-month response scale (see Appendix A).

Results and Discussion

Using Sample 5, we tested our prediction that the SDD scale includes 2 items which assess WG. To do this, the SDD scale was split into two subscales: an 8-item deviance-only subscale (SDD_{DEV}) and a 2-item scale with the SDD gossip items (SDD_{WG}). A non-nested CFA

model comparison was then performed. Raftery (1995) suggests that when two models are compared, a difference in BIC values of at least 10 provides strong evidence of a better fitting model. Using this guideline, there was strong evidence that a model with combined NWGS and SDD_{WG} (i.e., NWGS and SDD_{WG} in factor 1, and SDD_{DEV} in factor 2; $\chi^2(89, N = 177) = 446.07$, $p < .001$, BIC = 480.13) had a better model fit than a model with combined SDD_{WG} and SDD_{DEV} (i.e., NWGS in factor 1, and SDD_{WG} and SDD_{DEV} in factor 2; $\chi^2(89, N = 177) = 552.79$, $p < .001$, BIC = 586.85), $\Delta\text{BIC} = 106.72$. This confirms that the 2 SDD_{WG} items better assess WG than deviance-only SDD. When these 2 items are removed from SDD, the mean correlation (Samples 3 to 5) between SDD and NWGS is reduced from $\bar{r} = .61$, $p < .001$, to $\bar{r} = .44$, $p < .001$, indicating that deviance-only SDD and NWGS are very different constructs.

Next, we tested whether combining gossip items with statistically deviant items could lead to excessive false-positives. To test this, SDD scale and subscale scores were dichotomized according to some versus no behavior. In Samples 3 to 5, 318 (out of 590) participants reported some SDD behavior. Of those 318, 140 or 44% reported some SDD_{WG}, but no SDD_{DEV}. This indicates that 44% of reported SDD cases could be false positives due to the inclusion of gossip items in the SDD scale. This is highly problematic, as our evidence suggests that deviance-only SDD and NWGS have only 19% shared variance on average. With this in mind, it is our recommendation that WG and deviance not be measured together. Instead, WG should be measured as its own construct, independent of deviance. This would both facilitate future research into WG and potentially reduce false-positives in deviance measurement.⁹

⁹ In a follow-up test, Sample 5 was used to assess the interpersonal deviance (IDEV; Bennett & Robinson, 2000) and NWGC scales. We expected that one IDEV item (“made fun of someone at work”) would better assess WG than deviance. A non-nested model comparison provided strong evidence that supported this prediction (model 1 - NWGC and IDEV in separate factors: $\chi^2(53, N = 177) = 284.85$, BIC = 934.33; model 2 - NWGC and one IDEV item in factor 1, deviance-only IDEV in factor 2: $\chi^2(53, N = 177) = 250.20$, BIC = 899.69), $\Delta\text{BIC} = 34.64$. We then tested for possible false positives. In total, 31% of all cases of reported IDEV were solely due to the single gossip item.

General Discussion

Despite theory and evidence from other academic disciplines that gossip is an important and potentially functional behavior (Baumeister et al., 2004; Dunbar, 2004; Gluckman, 1963; Wert & Salovey, 2004), the field of organizational behavior has largely ignored workplace gossip (WG) as a specific field of study, choosing instead to categorize WG as a form of interpersonal deviance (Robinson & Bennett, 1995). This represents a serious cross-disciplinary contradiction in the way that gossip is conceptualized and studied. We have argued that this contradiction may have resulted from common implicit theories of gossip as malicious talk (Baumeister et al., 2004). While some extreme and rare cases of WG may be malicious, qualitative evidence presented in this research suggests that *typical* WG is not generally malicious and can result from a variety of different motivations. This finding contradicts a primarily gossip-as-deviance conceptualization of WG but is consistent with research from the broader gossip literature (Beersma & Van Kleef, 2012; Paine, 1967; Wert & Salovey, 2004).

To facilitate future research into WG, we have developed and validated 4 WG scales (2 dimensions: gossip valence and gossip subject). Using 8 samples, including multiwave, multisource, and multicultural data, we demonstrated that the WG scales are valid and reliable behavioral measures of WG. Further, evidence showed that the WG scales had very good model fit and that method effects due to negative affect, positive affect, and impression management were minimal. A cross-cultural measurement invariance test showed that the WG scales were full-configural, metric, and scalar invariant between North American and Chinese samples. This finding is consistent with research from other academic disciplines suggesting that gossip is present in a variety of different cultures (e.g., Gluckman, 1963) and provides preliminary evidence that WG can be studied in non-North American populations (i.e., in China specifically).

In developing the WG scales, an initial nomological network for WG was developed. Previous gossip research has theorized that gossip is a fundamental social comparison process (Suls, 1977; Wert & Salovey, 2004) that can be used to validate opinions, abilities, and emotions. Consistent with this social comparison theory explanation, negative WG was found to be positively related to variables associated with a need for sensemaking (i.e., job ambiguity, job insecurity, and workplace anxiety), negative affective well-being, and turnover (i.e., turnover intentions and job-search behaviors). Negative WG was also negatively related to organizational justice and self-esteem, also consistent with social comparison theory predictions. Gossip research has argued that gossip can be used to enforce norms of acceptable behavior through sanction and reputation management (Dunbar, 2004; Gluckman, 1963). Consistent with this prediction, our research found that negative WG was positively related to co-worker incivility.

While common assumptions are that gossip content is primarily negative, gossip research has suggested that both negative and positive gossip can occur (Foster, 2004). Our evidence confirmed this, showing that positive WG was positively related to organizational justice, positive affective well-being, networking ability, and OCBs. Further, both negative and positive WG were found to be negatively related to in-role performance (i.e., task performance) and positively related to deviance (i.e., supervisor-directed deviance, interpersonal deviance, instigated workplace incivility, and CWBs). These findings challenge a purely gossip-as-deviance conceptualization, as positive WG (i.e., saying good things about others) is clearly not deviant. And yet, the pattern of relationships is consistent between both positive and negative WG and in-role performance and deviance. While some readers may intuit that positive and negative WG simply co-occur, this is not supported by our evidence, as positive and negative WG about a single subject (i.e., NWGS and PWGS) was not significantly related.

Evidence from discriminant validity tests supported the argument that WG and deviance are distinctly different constructs. Tests showed that the WG scales were statistically distinct from all other measured constructs, including supervisor-directed deviance (SDD), interpersonal deviance, instigated workplace incivility, and CWB. This further challenges a gossip-as-deviance conceptualization and also calls into question the traditional practice of including gossip items in deviance scales. To assess this issue further, follow-up tests of a measure of SDD and NWGS were performed. Evidence from CFAs confirmed that the SDD scale measured 2 items that were better classified as WG than deviance. When these items were removed, the relationship between NWGS and SDD was only moderately strong, consistent with the argument that WG and deviance are very different constructs. Further, evidence from 3 samples indicated that the inclusion of these gossip items in the SDD scale may lead to excessive false positives, as 44% of all reported SDD cases only measured WG behavior and not deviance.

Study Limitations

As with any research, there are some limitations with our research that should be addressed. First, it is important to remember that the evidence presented herein is not able to effectively speak to the causality of relationships between measured variables. While we have been careful not to hypothesize causal relationships, some argumentation leading to those predictions was inherently causal in its language. It is important to remember that our goal was to develop and validate a series of WG scales, not to establish relationship causality. For scale development, only significant correlations are required to establish a scale's convergent validity, criterion-related validity, and nomological network (Hinkin, 1995; 1998). In this sense, a lack of causality should not be seen to affect the acceptability of the WG scales. If researchers are interested in the directionality of specific relationships, then we would advise the use of other

research methodologies, such as longitudinal or panel designs.

Another potential limitation of this research is the use of a single-wave sample in the construct validation (i.e., Sample 7). While this could lead to common method bias in this sample, it is important to remember that this was only one of three samples that were tested during this scale development phase. Wherever possible, we have attempted to use a variety of different data collection strategies in our research. Our objective was to build confidence that the WG scales are appropriate for as many different real-world usage patterns that we could test. Toward this end, we have collected single-wave, multiwave, multisource, and multicultural data samples. We have also measured the WG scales in the same wave, in different waves, using multiple item-presentation strategies (i.e., two versus four item-blocks), and using two different response scales. Whenever possible, we employed strategies to minimize method bias, including randomized scale and item order (Podsakoff et al., 2003). In all cases, evidence demonstrated that the WG scales had acceptable psychometric properties.

Future Research into Workplace Gossip

With theory and evidence both supporting the argument that WG is not best conceptualized as deviance, it is our recommendation that WG be studied on its own, independent of deviance. We see this as an exciting opportunity to both eliminate the cross-disciplinary contradiction in how gossip is treated and to also develop a field of research into what could be a very rich organizational construct. As such, it is our opinion that there is considerable opportunity for meaningful research contributions in the study of WG. Not only is it a common organizational behavior, but our research also suggests that it is related to a variety of important organizational variables and processes.

However, to truly advance our understanding of WG, we must reorient our thinking of

the construct. While WG has traditionally been treated as a criterion (i.e., deviance), our reconceptualization considers WG to be an information gathering and validation *process*. As such, we do not consider WG to be an end in and of itself, but rather a means to satisfy individual needs. Self-determination theory suggests that individuals have basic needs for competence, autonomy, and relatedness (Ryan & Deci, 2000). In our qualitative study of WG, we enumerated several motivations for WG which map onto these basic needs, including needs to gather and validate information, vent and cope, and build friendships. As basic need satisfaction is thought to relate to a variety of positive outcomes (e.g., intrinsic motivation, social development, and well-being), future research should examine the role of WG as a mediator between basic needs and positive workplace outcomes.

While it may be counter-intuitive for some that gossip could have positive outcomes, it is a natural outcome of conceptualizing WG from within social comparison (Festinger, 1954) and social exchange (Cropanzano & Mitchell, 2005) theoretical frameworks. Previous social comparison research has linked social comparisons to positive affect (Greenberg, Ashton-James, & Ashkanasy, 2007), performance (Festinger, 1954), self-improvement (Wood, 1989), self-enhancement (Wills, 1981), and coping with stress and negative events (Buunk & Gibbons, 2007; Schachter, 1959; Wood, Taylor, & Lichtman, 1985). As WG is thought to be a social comparison process (Suls, 1977; Wert & Salovey, 2004), research should examine how WG relates to these important variables. In particular, we believe that very meaningful work could be done regarding how WG both stems from and possibly manages negative workplace emotions and stress. Instead of WG being toxic, as is often assumed, WG could be an important tool for managing negative emotions and stress. Future research should investigate this possibility, as well as how WG affects relationships between individuals. Social exchange theory states that

repeated exchanges lead to high quality relationships (Cropanzano & Mitchell, 2005). As WG is a social exchange process, we would expect that WG could be used to build trust and affiliate with co-workers. Our preliminary evidence is consistent with this prediction, as positive WG was found to be related to networking ability.

While we have presented evidence that WG is distinct from deviance, future research should further investigate how WG behavior relates to deviance, including the role that WG could play as an antecedent to deviant behaviors. For example, previous research has shown that attributions of blame mediate the relationship between organizational justice and retaliation (Barclay et al., 2005). If WG is a social exchange of information, as we have argued, then we would expect that WG could be a fundamental process for making causal attributions (Kelley, 1973). WG could therefore affect both the frequency and form of deviance. But instead of simply recommending that WG be eliminated - an impossible recommendation (Noon & Delbridge, 1993), researchers should study WG to better understand when it manifests, when it is serious, and how to respond to it. Unlike some forms of deviance which are solitary acts, WG always involves multiple parties. As such, researchers could investigate WG interventions with the goal of affecting attributions and reducing the frequency of deviance. Further, researchers should also investigate moderators of WG to determine if contextual variables could reduce deviance.

WG is thought to be a process through which group norms and reputations are understood, constructed, and communicated. As such, WG is a tool for promoting norms of acceptable behavior, cooperation, and positive group functioning (Beersma & Van Kleef, 2011; Dunbar, 2004; Feinberg, et al., 2014; Feinberg et al., 2012; Piazza & Bering, 2008). Consistent with this, our research has provided initial evidence that co-worker incivility is related to negative WG. Future research should investigate this further, focusing on the role that WG plays

in response to observed deviant behaviors by others. Traditionally, practitioners have assumed that gossip is deviant and recommended that it be eliminated from organizations (Noon & Delbridge, 1993). If, instead of being deviant, WG actually deters deviance, then recommendations to eliminate gossip could potentially lead to an increase in overall deviance. Future research needs to examine this possibility, thereby enabling us to make evidence-based recommendations concerning WG.

Conclusion

...it is short-sighted to assume that gossip is destructive. It reflects a search for meaning as people negotiate their relationships and positions within the organization.

– Noon and Delbridge (1993, p. 34)

It is our opinion that WG has been misconceptualized in the organizational behavior literature. Negative assumptions about gossip have been allowed to persist for decades, despite conflicting research in other academic disciplines. The evidence that we have presented indicates that WG does not fit a deviance categorization and that WG should be studied on its own, independent of deviance. It is our hope that the WG scales that have been developed in this paper can both encourage and facilitate systematic research into this important construct. We must remind readers, however, to heed Baumeister and colleagues' (2004) warning that researchers can have difficulty seeing past implicit theories of gossip as malicious talk to recognize other, potentially functional, aspects of the construct. With this in mind, we urge researchers to move beyond negative assumptions about WG and to take a balanced view of the construct, based on typical manifestations rather than extreme cases. It is not our intention to argue that WG is always good, but rather that it is not typically as bad as people assume.

References

- Ahearn, K. K., Ferris, G. R., Hochwarter, W. A., Douglas, C., & Ammeter, A. P. (2004). Leader political skill and team performance. *Journal of Management*, 30(3), 309-327.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3), 411-423.
- Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*, 76(5), 732-740.
- Andersson, L. M., & Pearson, C. M. (1999). Tit for tat? The spiraling effect of incivility in the workplace. *Academy of Management Review*, 24(3), 452-471.
- Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The visual impact of gossip. *Science*, 332, 1446-1448.
- Aquino, K., Lewis, M. U., & Bradfield, M. (1999). Justice constructs, negative affectivity, and employee deviance: A proposed model and empirical test. *Journal of Organizational Behavior*, 20, 1073-1091.
- Ashford, S. J., Lee, C., & Bobko, P. (1989). Content, cause, and consequences of job insecurity: A theory-based measure and substantive test. *Academy of Management Journal*, 32(4), 803-829.
- Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly*, 27, 459-489.
- Barclay, L. J., Skarlicki, D. P., & Pugh, S. D. (2005). Exploring the role of emotions in injustice perceptions and retaliation. *Journal of Applied Psychology*, 90(4), 629-643.
- Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of General Psychology*, 8(2), 111-121.

- Beersma, B., & Van Kleef, G. A. (2011). How the grapevine keeps you in line: Gossip increases contributions to the group. *Social Psychological and Personality Science*, 2(6), 642-649.
- Beersma, B., & Van Kleef, G. A. (2012). Why people gossip: An empirical analysis of social motives, antecedents, and consequences. *Journal of Applied Social Psychology*, 42(11), 2640-2670.
- Besnier, N. (1989). Information withholding as a manipulative and collusive strategy in Nukulaelae gossip. *Language in Society*, 18(3), 315-341.
- Ben-Ze'ev, A. (1994). The vindication of gossip. In Goodman, R. F., Ben-Ze'ev, A. (Eds.), *Good gossip* (pp. 11-24). Lawrence, KS: University Press of Kansas.
- Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85(3), 349-360.
- Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16(1), 78-117.
- Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, 22(7), 891-893.
- Bergmann, J. R. (1993). *Discreet indiscretions: The social organization of gossip*. Hawthorne, NY: Walter de Gruyter, Inc.
- Blau, G. (1994). Testing a two-dimensional measure of job search behavior. *Organizational Behavior and Human Decision Processes*, 59(2), 288-312.
- Blau, G., & Andersson, L. (2005). Testing a measure of instigated workplace incivility. *Journal of Occupational and Organizational Psychology*, 78(4), 595-614.

- Borman, W. C., & Motowidlo, S. M. (1993). Expanding the criterion domain to include elements of contextual performance. In Schmitt, N., Borman, W. D. & Associates (Eds.), *Personnel Selection in Organizations* (pp. 71-98). San Francisco, CA: Jossey-Bass.
- Boswell, W. R., Olson-Buchanan, J. B., & Harris, T. B. (2014). I cannot afford to have a life: Employee adaptation to feelings of job insecurity. *Personnel Psychology*, *67*(4), 887-915.
- Breaugh, J. A., & Colihan, J. P. (1994). Measuring facets of job ambiguity: Construct validity evidence. *Journal of Applied Psychology*, *79*(2), 191-202.
- Brenneis, D. (1984). Grog and gossip in Bhatgaon: style and substance in Fiji Indian conversation. *American Ethnologist*, *11*(3), 487-506.
- Brislin, R. W. (1980). Translation and content analysis of oral and written material. In Triandis, H. C., & Berry, J. W. (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 349–444). Boston, MA: Allyn & Bacon.
- Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, *102*(1), 3-21.
- Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, *25*(1), 1-27.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, *9*(2), 233-255.
- Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*(3), 386-400.
- Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D. (2001). Incivility in the workplace: incidence and impact. *Journal of Occupational Health Psychology*, *6*(1), 64-80.

- Cox, B. A. (1970). What is Hopi gossip about? Information management and Hopi factions. *Man*, 5(1), 88-98.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
- Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management*, 31(6), 874-900.
- Dalal, R. S., Lam, H., Weiss, H. M., Welch, E. R., & Hulin, C. L. (2009). A within-person approach to work behavior and performance: Concurrent and lagged citizenship-counterproductivity associations, and dynamic relationships with affect and overall job performance. *Academy of Management Journal*, 52(5), 1051-1066.
- DiFonzo, N., & Bordia, P. (2007). *Rumor psychology: Social and organizational approaches*. Washington, DC: American Psychological Association
- Duffy, M. K., Ganster, D. C., & Pagon, M. (2002). Social undermining in the workplace. *Academy of Management Journal*, 45(2), 331-351.
- Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16(04), 681-694.
- Dunbar, R. I. M. (1996). *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.
- Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, 8(2), 100-110.
- Dunbar, R. I. M., Marriott, A., & Duncan, N. D. C. (1997). Human conversational behavior. *Human Nature*, 8(3), 231-246.

- Eder, D., & Enke, J. L. (1991). The structure of gossip: Opportunities and constraints on collective expression among adolescents. *American Sociological Review*, *56*(4), 494-508.
- Emler, N. (1994). Gossip, reputation, and social adaptation. In Goodman, R. F., Ben-Ze'ev, A. (Eds.), *Good gossip* (pp. 117-138). Lawrence, KS: University Press of Kansas.
- Erdogan, B., Bauer, T. N., & Walter, J. (2014). Deeds that Help and Words that Hurt: Helping and Gossip as Moderators of the Relationship Between Leader–Member Exchange and Advice Network Centrality. *Personnel Psychology*, *68*(1), 185-214.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299.
- Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science*, *25*(3), 656-664.
- Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology*, *102*(5), 1015-1030.
- Ferris, G. R. (1985). Role of leadership in the employee withdrawal process: A constructive replication. *Journal of Applied Psychology*, *70*(4), 777-781.
- Ferris, D. L., Brown, D. J., Berry, J. W., & Lian, H. (2008). The development and validation of the Workplace Ostracism Scale. *Journal of Applied Psychology*, *93*(6), 1348-1366.
- Ferris, G. R., Treadway, D. C., Kolodinsky, R. W., Hochwarter, W. A., Kacmar, C. J., Douglas, C., & Frink, D. D. (2005). Development and validation of the political skill inventory. *Journal of Management*, *31*(1), 126-152.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117-140.

- Fine, G. A. (1977). Social components of children's gossip. *Journal of Communication*, 27(1), 181-185.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Foster, E. K. (2004). Research on gossip: Taxonomy, methods, and future directions. *Review of General Psychology*, 8(2), 78-99.
- Gilmore, D. (1978). Varieties of gossip in a Spanish rural community. *Ethnology*, 17(1), 89-99.
- Gluckman, M. (1963). Gossip and scandal. *Current Anthropology*, 4(3), 307-316.
- Graen, G. B., Liden, R. C., & Hoel, W. (1982). Role of leadership in the employee withdrawal process. *Journal of Applied Psychology*, 67(6), 868-872.
- Greenberg, J., Ashton-James, C. E., & Ashkanasy, N. M. (2007). Social comparison processes in organizations. *Organizational Behavior and Human Decision Processes*, 102(1), 22-41.
- Guerin, B., & Miyazaki, Y. (2006). Analyzing rumors, gossip, and urban legends through their conversational properties. *The Psychological Record*, 56, 23-34.
- Haviland, J. B. (1977). Gossip as competition in Zinacantan. *Journal of Communication*, 27(1), 186-191.
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of management*, 21(5), 967-988.
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods*, 1(1), 104-121.
- Holden, R. R., & Jackson, D. N. (1979). Item subtlety and face validity in personality assessment. *Journal of Consulting and Clinical Psychology*, 47(3), 459-468.

- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Hunter, J. E., & Schmidt, F. L. (Eds.). (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- Keeping, L. M., & Levy, P. E. (2000). Performance appraisal reactions: measurement, modeling, and method bias. *Journal of Applied Psychology*, 85(5), 708-723.
- Jones, D. (1980). Gossip: Notes on women's oral culture. *Women's Studies International Quarterly*, 3(2), 193-198.
- Kelley, H. H. (1973). The processes of causal attribution. *American psychologist*, 28(2), 107-128.
- Kniffin, K. M., & Wilson, D. S. (2005). Utilities of gossip across organizational levels: Multilevel selection, free-riders, and teams. *Human Nature*, 16(3), 278-292.
- Kurland, N. B., & Pelled, L. H. (2000). Passing the word: Toward a model of gossip and power in the workplace. *Academy of Management Review*, 25(2), 428-438.
- Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: the role of affect and cognitions. *Journal of Applied Psychology*, 87(1), 131-142.
- Levin, J., & Arluke, A. (1985). An exploratory analysis of sex differences in gossip. *Sex Roles*, 12(3-4), 281-286.
- Loudon, J. B. (1961). Kinship and crisis in South Wales. *British Journal of Sociology*, 12(4), 333-350.
- Maitlis, S., & Christianson, M. (2014). Sensemaking in organizations: Taking stock and moving forward. *The Academy of Management Annals*, 8(1), 57-125.

- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111-130.
- Mitchell, M. S., & Ambrose, M. L. (2007). Abusive supervision and workplace deviance and the moderating effects of negative reciprocity beliefs. *Journal of Applied Psychology*, 92(4), 1159-1168.
- Morreall, J. (1994). Gossip and humor. In Goodman, R. F., Ben-Ze'ev, A. (Eds.), *Good gossip* (pp. 56-64). Lawrence, KS: University Press of Kansas.
- Munz, D. C., Huelsman, T. J., Konold, T. R., & McKinney, J. J. (1996). Are there methodological and substantive roles for affectivity in job diagnostic survey relationships?. *Journal of Applied Psychology*, 81(6), 795-805.
- Nevo, O., Nevo, B., & Derech-Zehavi, A. (1993). The development of the tendency to gossip questionnaire: Construct and concurrent validation for a sample of Israeli college students. *Educational and Psychological Measurement*, 53(4), 973-981.
- Noon, M., & Delbridge, R. (1993). News from behind my hand: Gossip in organizations. *Organizational Studies*, 14(1), 23-36.
- Organ, D. W. (1997). Organizational citizenship behavior: It's construct clean-up time. *Human performance*, 10(2), 85-97.
- Paine, R. (1967). What is gossip about? An alternative hypothesis. *Man*, 2(2), 278-285.
- Paulhus, D. L. (1991). Measurement and control of response bias. In Robinson, J. P., Shaver, P. R., Wrightsman, L. S. (Eds.), *Measures of social psychological attitudes: Vol. 1. Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

- Piazza, J., & Bering, J. M. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. *Evolution and Human Behavior*, 29(3), 172-178.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903.
- Radin, P. (1927). *Primitive Man as a Philosopher*. Oxford, England: Appleton.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111-163.
- Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal*, 38(2), 555-572.
- Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6), 934-960.
- Rodell, J. B., & Judge, T. A. (2009). Can “good” stressors spark “bad” behaviors? The mediating role of emotions in links of challenge and hindrance stressors with citizenship and counterproductive behaviors. *Journal of Applied Psychology*, 94(6), 1438-1451.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosnow, R. L. (1991). Inside rumor: A personal journey. *American Psychologist*, 46(5), 484-496.
- Rosnow, R. L. (2001). Rumor and gossip in interpersonal interaction and beyond: A social exchange perspective. In R. M. Kowalski (Ed.), *Behaving badly: Aversive behaviors in interpersonal relationships* (pp. 203-232). Washington, DC: American Psychological Association.

- Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: a policy-capturing approach. *Journal of Applied Psychology*, 87(1), 66-80.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- Sackett, P. R. (2002). The structure of counterproductive work behaviors: Dimensionality and relationships with facets of job performance. *International Journal of Selection and Assessment*, 10(1-2), 5-11.
- Schachter, S. (1959). *The psychology of affiliation*. Palo Alto, CA: Stanford University Press.
- Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, 24(1), 138-158.
- Sommerfeld, R. D., Krambeck, H. J., Semmann, D., & Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences*, 104(44), 17435-17440.
- Spacks, P. M. (1982). In praise of gossip. *The Hudson Review*, 35(1), 19-38.
- Spreitzer, G. M., & Sonenshein, S. (2004). Toward the construct definition of positive deviance. *American Behavioral Scientist*, 47(6), 828-847.
- Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-90.
- Suls, J. M. (1977). Gossip as social comparison. *Journal of Communication*, 27(1), 164-168.
- Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of management journal*, 43(2), 178-190.

- Van Dyne, L., & LePine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal*, 41(1), 108-119.
- Van Katwyk, P. T., Fox, S., Spector, P. E., & Kelloway, E. K. (2000). Using the job-related affective well-being scale (JAWS) to investigate affective responses to work stressors. *Journal of Occupational Health Psychology*, 5(2), 219-230.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, 3(1), 4-70.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
- Wayne, S. J., Shore, L. M., & Liden, R. C. (1997). Perceived organizational support and leader-member exchange: A social exchange perspective. *Academy of Management journal*, 40(1), 82-111.
- Wert, S. R., & Salovey, P. (2004). A social comparison account of gossip. *Review of General Psychology*, 8(2), 122-137.
- Wiessner, P. W. (2014). Embers of society: Firelight talk among the Ju/'hoansi Bushmen. *Proceedings of the National Academy of Sciences*, 111(39), 14027-14035.
- Williams, L. J., & Anderson, S. E. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, 17(3), 601-617.

- Williams, L. J., & Anderson, S. E. (1994). An alternative approach to method effects by using latent-variable models: Applications in organizational behavior research. *Journal of Applied Psychology, 79*(3), 323-331.
- Williams, L. J., Gavin, M. B., & Williams, M. L. (1996). Measurement and nonmeasurement processes with negative affectivity and employee attitudes. *Journal of Applied Psychology, 81*(1), 88-101.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90*(2), 245-271.
- Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin, 106*(2), 231-248.
- Wood, J. V., Taylor, S. E., & Lichtman, R. R. (1985). Social comparison in adjustment to breast cancer. *Journal of Personality and Social Psychology, 49*(5), 1169-1183.
- .

Table 1
Overview of Measures Used for Validation Phases 3 and 4

Measure	Sample 5	Sample 6	Sample 7
NWGS	X ₂	X ₁	X ₁
PWGS	X ₂	X ₁	X ₁
NWGC	X ₃	X ₁	X ₁
PWGC	X ₃	X ₁	X ₁
PANAS (Watson et al., 1988)		X ₁	
Impression Management (Paulhus, 1991)		X ₁	
Tendency to gossip (Erdogan et al., 2014)	X ₁	X ₂	
Job ambiguity (Breugh & Coplin, 1994)			X ₁
Job insecurity (Boswell et al., 2014; Ashford et al., 1989)			X ₁
Anxiety (Rodell & Judge, 2009)			X ₁
Job-related affective well-being (Van Katwyk et al., 2000)		X ₂	
Justice (Colquitt, 2005)			X ₁
Networking ability (Ferris et al., 2005)		X ₂	
Co-worker incivility (Cortina et al., 2001)	X ₁		
Self-esteem (Rosenberg, 1965)	X ₀		
In-role performance (Williams & Anderson, 1991)	X _p		
Organizational citizenship behaviors (Dalal et al., 2009)	X _p		
Counterproductive workplace behaviors (Dalal et al., 2009)	X _p		
Supervisor-directed deviance (Mitchell & Ambrose, 2007)	X ₃		
Interpersonal deviance (Bennett & Robinson, 2000)	X ₃		
Instigated workplace incivility (Blau & Andersson, 2005)		X ₂	
Organizational citizenship behaviors interpersonal (Lee & Allen, 2002)		X ₂	
Turnover Intentions (Wayne et al., 1997)		X ₂	
Job search behaviors (Blau, 1994)		X ₂	

Note. NWGS = negative WG about supervisor; PWGS = Positive WG about supervisor; NWGC = negative WG about co-workers;

PWGC = positive WG about co-workers.

X₀ = scale was assessed at prescreen. X_{1, 2, 3} = scale was assessed in wave 1, 2, or 3. X_p = scale was assessed by work-peers.

Table 2
Descriptive Statistics, Correlations, and Alpha Reliabilities for Sample 5

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Age (years)	36.79	8.63																
2. Gender	1.52	.50	.19**															
3. NWGS	1.89	1.20	-.03	-.02	.92													
4. PWGS	2.69	1.64	-.13	-.02	.07	.97												
5. NWGC	2.06	1.36	-.06	-.05	.61***	.11	.95											
6. PWGC	3.18	1.63	.01	.05	.15	.42***	.29***	.97										
7. Tendency to gossip	3.13	1.67	-.06	-.03	.46***	-.13	.52***	-.01	.95									
8. Co-worker incivility	1.57	.85	-.06	-.06	.53***	.12	.45***	.01	.44***	.93								
9. Self-esteem	4.62	.85	.26***	.10	-.29***	-.01	-.22**	.05	-.25**	-.12	.88							
10. SDD	1.44	.88	-.14	-.12	.62***	.30***	.59***	.18*	.30***	.56***	-.14	.94						
11. SDD _{WG}	1.95	1.42	-.01	-.03	.73***	.09	.65***	.20**	.47***	.50***	-.17*	.77***	.81					
12. SDD _{DEV}	1.31	.87	-.17*	-.14	.50***	.35***	.49***	.15	.20**	.51***	-.12	.96***	.57***	.98				
13. Interpersonal deviance	1.45	.89	-.13	-.19*	.50***	.25**	.66***	.18*	.40***	.55***	-.19*	.78***	.61***	.75***	.92			
14. In-role performance ^a	5.91	1.07	.21*	.08	-.35***	-.20*	-.37***	.01	-.22*	-.30**	.20*	-.33***	-.16	-.36***	-.31**	.90		
15. OCB ^a	5.26	1.21	.13	-.03	-.14	.00	-.07	.28**	-.05	-.18*	.14	-.08	-.04	-.09	-.05	.52***	.96	
16. CWB ^a	2.14	1.40	-.21*	-.12	.41***	.26**	.33***	.19*	.17	.30**	-.17	.35***	.28*	.33***	.38***	-.53***	-.19*	.98

Note. *n* ranges from 120 (peer-report) to 183 (self-report); Cronbach's alpha reliabilities are on the diagonal in bold. For gender, 1 =

male, 2 = female. NWGS = negative WG about supervisor; PWGS = Positive WG about supervisor; NWGC = negative WG about co-workers; PWGC = positive WG about co-workers; SDD = supervisor-directed deviance; SDD_{WG} = 2-item WG-only SDD subscale (see Phase 6); SDD_{DEV} = 8-item deviance-only SDD subscale (see Phase 6); OCB = organizational citizenship behaviors; CWB = counterproductive workplace behaviors.

^a Scale was assessed by work peer.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3
Descriptive Statistics, Correlations, and Alpha Reliabilities for Sample 6

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Age (years)	37.84	11.19																	
2. Gender	1.49	.50	.03																
3. NWGS	2.04	1.26	-.08	.01	.93														
4. PWGS	2.72	1.18	-.05	.08	.02	.90													
5. NWGC	2.30	1.19	-.15*	-.02	.57***	.17**	.92												
6. PWGC	3.09	1.25	.04	.07	.35***	.53***	.45***	.92											
7. NA	1.23	.45	-.16**	.06	.34***	.12*	.32***	.16**	.91										
8. PA	2.85	.89	.20**	-.01	-.04	.26***	-.02	.25***	-.03	.92									
9. Impression management	6.83	4.67	.12	.08	-.29***	.00	-.31***	-.03	-.26***	.17**	.86								
10. Tendency to gossip	3.04	1.56	-.16*	-.05	.31***	-.10	.52***	.14*	.23***	-.24***	-.45***	.96							
11. JAWS (negative)	1.90	.80	-.06	-.01	.45***	-.07	.23***	.12	.37***	-.11	-.27***	.30***	.85						
12. JAWS (positive)	2.60	.89	.14*	.02	-.13*	.36***	-.03	.19**	-.01	.43***	-.03	-.14*	-.28***	.90					
13. Networking ability	4.32	1.18	.04	.05	-.12	.39***	.08	.22**	-.06	.34***	.06	-.16*	-.15*	.47***	.91				
14. Instigated incivility	1.68	.59	-.14*	-.04	.46***	.12	.52***	.24***	.40***	-.04	-.41***	.51***	.52***	-.07	-.02	.86			
15. OCB interpersonal	4.64	.99	.24***	.12	-.05	.30***	.01	.30***	-.11	.33***	.26***	-.23**	-.15*	.33***	.34***	-.17*	.88		
16. Turnover intentions	3.39	1.74	-.06	-.03	.33***	-.23**	.07	-.05	.07	-.20**	-.05	.14*	.49***	-.59***	-.27***	.20**	-.17*	.93	
17. Job search behaviors	1.82	.78	-.12	.09	.41***	.07	.13*	.08	.25***	.09	-.07	-.03	.34***	-.05	.00	.27***	-.06	.49***	.92

Note. *n* ranges from 221 to 266; Cronbach's alpha reliabilities are on the diagonal in bold. For gender, 1 = male, 2 = female. NWGS = negative WG about supervisor; PWGS = Positive WG about supervisor; NWGC = negative WG about co-workers; PWGC = positive WG about co-workers; NA = negative affectivity; PA = positive affectivity; JAWS = job-related affective well-being scale.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4

Descriptive Statistics, Correlations, and Alpha Reliabilities for Sample 7

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age (years)	43.09	11.74													
2. Gender	1.60	.49	-.02												
3. NWGS	2.03	1.33	-.18*	-.12	.94										
4. PWGS	2.72	1.45	.01	-.04	.24**	.95									
5. NWGC	2.20	1.44	-.13	-.07	.74***	.34***	.95								
6. PWGC	3.26	1.49	.06	-.06	.27***	.70***	.26***	.95							
7. Job ambiguity	1.93	.96	-.27***	-.12	.33***	-.10	.23**	-.15*	.95						
8. Job insecurity	1.78	.80	-.19*	.04	.42***	.11	.37***	.04	.36***	.91					
9. Anxiety	1.81	.99	-.25**	-.07	.35***	.02	.27***	.09	.46***	.48***	.85				
10. Distributive Justice	3.62	1.12	.07	-.10	-.25**	.33***	-.17*	.25**	-.37***	-.19*	-.25**	.94			
11. Procedural Justice	3.39	.97	.13	-.04	-.21**	.28***	-.15*	.28***	-.40***	-.17*	-.26***	.69***	.91		
12. Interpersonal Justice	3.89	1.08	.05	-.01	-.33***	.24**	-.19*	.14	-.36***	-.21**	-.22**	.63***	.66***	.93	
13. Informational Justice	3.54	1.06	.06	-.02	-.27***	.23**	-.17*	.14	-.32***	-.21**	-.26***	.64***	.71***	.71***	.91

Note. $n = 177$; Cronbach's alpha reliabilities are on the diagonal in bold. For gender, 1 = male, 2 = female. NWGS = negative WG

about supervisor; PWGS = Positive WG about supervisor; NWGC = negative WG about co-workers; PWGC = positive WG about co-workers. To increase readability, the job ambiguity scale is presented with reversed scoring (i.e., 1 = low ambiguity to 7 = high ambiguity).

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 5
Cross-cultural Assessment of Measurement Equivalence/Invariance

	χ^2	<i>df</i>	CFI	SRMR	Comparison	Δ CFI	Decision
NWGS							
Step 1: Test for full configural invariance	45.64	10	.99	.01	---	---	Accept
Step 2: Test for full metric invariance	69.17	14	.98	.04	1 vs. 2	-.01	Accept
Step 3: Test for full scalar invariance	103.77	18	.97	.04	2 vs. 3	-.01	Accept
PWGS							
Step 1: Test for full configural invariance	10.59	10	1.00	.01	---	---	Accept
Step 2: Test for full metric invariance	20.08	14	1.00	.03	1 vs. 2	.00	Accept
Step 3: Test for full scalar invariance	50.01	18	.99	.03	2 vs. 3	-.01	Accept
NWGC							
Step 1: Test for full configural invariance	20.84	10	1.00	.01	---	---	Accept
Step 2: Test for full metric invariance	34.50	14	.99	.04	1 vs. 2	-.01	Accept
Step 3: Test for full scalar invariance	73.23	18	.98	.04	2 vs. 3	-.01	Accept
PWGC							
Step 1: Test for full configural invariance	47.95	10	.99	.03	---	---	Accept
Step 2: Test for full metric invariance	49.91	14	.99	.04	1 vs. 2	.00	Accept
Step 3: Test for full scalar invariance	65.70	18	.98	.04	2 vs. 3	-.01	Accept

Note. This comparison was between Sample 3 (North America; $n = 177$) and Sample 4 (China; $n = 211$); NWGS = negative WG about supervisor; PWGS = Positive WG about supervisor; NWGC = negative WG about co-workers; PWGC = positive WG about co-workers.

Model fit and reliabilities (Sample 8): 4-factor $\chi^2(164) = 448.61$, CFI = .98, SRMR = .05; NWGS $\chi^2(5) = 20.18$, CFI = .99, SRMR = .01, $\alpha = .97$; PWGS $\chi^2(5) = 2.99$, *n.s.*, CFI = 1.00, SRMR = .01, $\alpha = .97$; NWGC $\chi^2(5) = 17.19$, CFI = .99, SRMR = .01, $\alpha = .96$; PWGC $\chi^2(5) = 42.62$, CFI = .97, SRMR = .03, $\alpha = .96$.

Appendix A

Final WG Scales

INSTRUCTIONS: The following questions are about workplace conversations in which you talked about your supervisor when he/she was not present to hear what was said.

In the last month, how often have you...

Negative WG About a Supervisor (NWGS)

1. asked a work colleague if they have a negative impression of something that your supervisor has done
2. questioned your supervisor's abilities while talking to a work colleague
3. criticized your supervisor while talking to a work colleague
4. vented to a work colleague about something that your supervisor has done
5. told an unflattering story about your supervisor while talking to a work colleague

Positive WG About a Supervisor (PWGS)

1. complimented your supervisor's actions while talking to a work colleague
2. told a work colleague good things about your supervisor
3. defended your supervisor's actions while talking to a work colleague
4. said something nice about your supervisor while talking to a work colleague
5. told a work colleague that you respect your supervisor

INSTRUCTIONS: The following questions are about workplace conversations in which you talked about a co-worker when he/she was not present to hear what was said. The co-worker could be any co-worker who is not your supervisor.

In the last month, how often have you...

Negative WG About Co-workers (NWGC)

1. asked a work colleague if they have a negative impression of something that another co-worker has done
2. questioned a co-worker's abilities while talking to another work colleague
3. criticized a co-worker while talking to another work colleague
4. vented to a work colleague about something that another co-worker has done
5. told an unflattering story about a co-worker while talking to another work colleague

Positive WG About Co-workers (PWGC)

1. complimented a co-worker's actions while talking to another work colleague
2. told a work colleague good things about another co-worker
3. defended a co-worker's actions while talking to another work colleague
4. said something nice about a co-worker while talking to another work colleague
5. told a work colleague that you respect another co-worker

SCORING INSTRUCTIONS: Items are evaluated on a 7-point response scale (1 = *never*, 2 = *once a month*, 3 = *2 - 3 times a month*, 4 = *once a week*, 5 = *2 - 3 times a week*, 6 = *once a day*, 7 = *more than once a day*). Item scores are combined (e.g., mean item score) into a composite score for each respective scale.

Appendix B

Chinese Translation of the WG Scales

INSTRUCTIONS: 以下问题有关你在工作场合中聊到上司的谈话内容。这些谈话发生在你上司本人不在场，听不到谈话内容的情况下。

在过去一个月中，你多经常。。。

Negative WG About a Supervisor (NWGS)

1. 问过同事是否对你上司做过的某些事情持有负面的看法。
2. 在与同事交谈时质疑过你上司的能力。
3. 在与同事交谈时批评过你的上司。
4. 针对你上司做过的事情向一个同事发牢骚。
5. 在与同事交谈时讲过你上司的一件不光彩的事情。

Positive WG About a Supervisor (PWGS)

1. 在与同事交谈时称赞过你上司的所作所为。
2. 跟同事说过你上司的好话。
3. 在与同事交谈时为你上司的所作所为辩护。
4. 在与同事交谈时说过某些关于你上司的好话。
5. 跟同事说过你尊重你的上司。

INSTRUCTIONS: 以下问题有关你在工作场合中聊到同事的谈话。这些谈话发生在这些同事本人不在场，听不到谈话内容的情况下。这些同事可以是你的任何一位同事，但是不能是你的上司。

在过去一个月中，你多经常。。。

Negative WG About Co-workers (NWGC)

1. 问过同事是否对另外一位同事做过的某些事情持有负面的看法。
2. 在与同事交谈时质疑过另外一位同事的能力。
3. 在与同事交谈时批评过另外一位同事。
4. 针对某同事做过的事情向另外一位同事发牢骚。
5. 在与同事交谈时讲过另外一位同事的一件不光彩的事情。

Positive WG About Co-workers (PWGC)

1. 在与同事交谈时称赞过另一位同事的所作所为。
2. 跟同事说过另一位同事的好话。
3. 在与同事交谈时为另一位同事的所作所为辩护。
4. 在与同事交谈时说过某些关于另一位同事的好话。
5. 跟同事说过你尊重某位同事。

SCORING INSTRUCTIONS: Items are evaluated on a 7-point response scale (1 = “从来没有”, 2 = “一个月一次”, 3 = “一个月2到3次”, 4 = “一周一次”, 5 = “一周2到3次”, 6 = “一天一次”, 7 = “一天超过一次”). Item scores are combined (e.g., mean item score) into a composite score for each respective scale.
