DUAL COGNITIVE PATHWAYS TO VOICE QUALITY

-FREQUENT VOICERS IMPROVISE, INFREQUENT VOICERS ELABORATE-

Wolsink, I. [1], Den Hartog, D. N. [2.], Belschak, F. D. [2.], and Sligte, I.G. [3]

1. University of Amsterdam, Faculty of Social and Behavioural Sciences, Department of Social Psychology
2. University of Amsterdam, Faculty of Economics and Business, Department of Leadership and Management
3. University of Amsterdam, Faculty of Social and Behavioural Sciences, Department of Brain and Cognition

Author Contributions

Theory and studies were designed by I. Wolsink, with input from all authors and specific help from I. G., Sligte on the design of study 2. All authors contributed with valuable comments during theory development, study design, and analysis. I. Wolsink acted as lead author, with all authors contributing to the writing by providing valuable comments and input to the text of the manuscript.

ABSTRACT

We investigate the involvement of Working Memory Capacity (WMC, the cognitive resource necessary for controlled elaborate thinking) in voice behavior (speaking up with suggestions, problems, and opinions to change the organization). While scholars assume voice requires elaborate thinking, some empirical evidence suggests voice might be more automatic. To explain this discrepancy, we distinguish between voice quantity (frequency of voice) and voice quality (novelty and value of voiced information) and propose that WMC is important for voice quality, but not voice quantity. Furthermore, we propose that people who voice often need less WMC to reach high voice quality than people who voice rarely. To test our ideas, we conducted three studies: a between-participant lab-study, a within-participant experiment, and a multi-source field-study. These studies confirmed that voice quantity is unrelated to WMC. Voice quality is positively related to WMC, but only for those who rarely voice. This indicates that the decision to voice (quantity) might be more automatic and intuitive than often assumed, whereas its value to the organization (quality), relies more on the degree of cognitive elaboration of the voicer. It also suggests that frequent and infrequent voicers use distinct cognitive pathways to voice high quality information: frequent voicers improvise, while infrequent voicers elaborate.

INTRODUCTION

~"Wise men talk because they have something to say; fools, because they have to say something" ~

Plato

We all know people in our work context who speak up frequently, fast, and fluently and individuals who are vocally more frugal. Yet, as Plato suggests, people who speak up rarely may offer great insight when they do, while those who speak up frequently may not always add novelty or value. Quantity thus does not always equal quality. This paper is about that difference between the quantity and the quality of speaking up, specifically in terms of the amount of thought that goes into these different aspects. In the literature, speaking up through the constructive communication of problems, challenging opinions, and suggestions for improvement, is labeled voice behavior (Van Dyne & LePine, 1998; Morrison, 2011). Voice is proactive, discretionary, planned behavior, with the intent to change something in the organization. Voice is important for organizations as management is often too far removed from the workplace to be able to oversee all haphazard, errors, inefficiencies, improvement opportunities, and customer complaints. Managers therefore need their employees to come forward with such potentially useful information in order to be aware of what needs to be changed (see Morrison, 2011).

Before people voice, they need something potentially important to say. But even if that is the case, individuals only engage in voice when they judge voicing to be sufficiently instrumental to achieve change and safe to show (e.g., Morrison, Wheeler-Smith, & Kamdar, 2011). For the voicer, there are thus many aspects to consider before speaking up, for example: 'Do we need change?', 'Does my suggestion make sense?', 'Will my supervisor be offended?' 'What is the best way to argue my case?' This implies that people likely think carefully and elaborately before they speak up, and thus need extensive cognitive resources to do so. However, the question of how much cognitive elaboration one needs for voice has not been empirically investigated to this date. Investigating the depth of cognitive elaboration in voice behavior is not only important for understanding the fundamental underlying cognitive mechanisms of voice, but it also has practical implications. If voice requires cognitive elaboration, people need the time and cognitive resources to do so. If voice is more habitual and automatic, a managers rational and elaborate plea for the safety of voicing has little chance of enhancing voice.

The goal of this paper, is thus to investigate the depth of cognitive elaboration in voice behavior. Dual process theories (Barrett, Tugade, & Engle, 2004; Evans, 2008; Kahneman, 2011; Petty & Wegener, 1999) propose that there are two modes of cognitive processing. The first (system 1) is fast, intuitive, heuristic, and automatic. The second (system 2) is slow, deliberate, systematic, and elaborate. In the voice literature, there is an ongoing debate on whether voice is cognitively elaborated or can also be automatic. On the one hand, scholars propose that voice is proactive, future focused, and challenging, and therefore requires systematic, elaborative cognitive processing (Chiaburu, Marinova, & van Dyne, 2008; Liang, Farh, & Farh, 2012). On the other hand, researchers argue that whether people voice or stay silent, may depend on automatic responses and schemas that operate below the level of conscious, rational decision making (Morrison, 2011). In support of that, empirical evidence suggests that part of voice might not depend on elaborate cognitive processing: individual differences in general cognitive ability are unrelated to voice (LePine & Van Dyne, 2001), and actively withholding voice seems to depend on fear and implicit beliefs about the effectiveness of voice rather than rational, well elaborated decision making (Detert & Edmondson, 2011; Kish-Gephart, Detert, Treviño, & Edmondson, 2009).

Here, we propose a way to reconcile both positions, by arguing that there are two components of voice that should be distinguished, because one requires more elaborate thought and cognitive resources than the other. More specifically, we argue that it is essential to distinguish between the act of voicing (‘*voice quantity’*, how often does one speak up), and the quality of the voiced content (‘*voice quality’*, does voice entail novel and valuable information). People may intend to voice something important, but high quality intention does not equal high quality output. A voiced message may be something that is already known, thus not adding knowledge, or something that cannot be used or implemented, thus not adding value. As Chiaburu and colleagues (2008) note, to the extent that the voiced message is complex, important, and changing, elaborate thinking and cognitive resources may matter. The issue here is that the complex, important, and changing attributes of voice refer more to the quality of voiced content than the behavior of voicing something per se.

Although broader measures of cognitive resources such as General Mental Ability have been studied in relation to the quantity of voice (e.g., LePine & Van Dyne, 2001) the quality of voice has not yet been related to cognitive resources. We therefore propose that to reveal where in the voice process elaborate cognitive processing matters most, both quantity and quality of voice should be studied simultaneously. Presumably, elaborate processing is more important for quality than for quantity. Previous research has shown that elaborate processing is important for original idea generation in experiments (De Dreu, Nijstad, Baas, Wolsink, & Roskes, 2012). The generation of original ideas is one aspect that is important for voice quality. Further, we argue that cognitive elaboration is also necessary for evaluating how valuable a message is, and how it should be communicated to others. Voicing high quality is more complex than the mere generation of an original idea. Voice is a proactive behavior (self-started, extra-role, challenging behavior in the social environment), and creativity studied in experimental research is reactive (requested, in-role, free brainstorming in an anonymous setting). People who voice high quality do not only generate original ideas upon request, they search for problems themselves, proactively generate ideas, select the best suggestion in order to solve the problem, and communicate this to others. For these reasons, cognitive elaboration should enhance voice quality beyond mere idea generation.

If voice quality requires more elaborate processing than voice quantity, then voice quality (and not voice quantity) should be linked to the cognitive resource in the brain that is central to elaborate cognitive processing, which is Working Memory Capacity. Elaborate and systematic cognitive thought is limited by Working Memory Capacity (Evans, 2003). Working Memory Capacity (WMC hereafter) refers to the amount of information one can actively perceive in the attentional field and use for thoughts or actions (Fukuda, Vogel, Mayr, & Awh, 2010). Working memory represents those thoughts that we are aware of (Evans, 2008), and the Capacity of Working Memory supports conscious control over our thoughts, feelings, and actions (Barret, Tugade & Engle, 2004). Both individual differences in WMC as well as manipulations of WMC by means of distraction of attention, have been used in previous research to study the role of elaborate thought and controlled attention in idea generation (De Dreu et al., 2012). We therefore propose that a difference between quantity and quality of voice in terms of cognitive processing depth, should be reflected in differential relationships between these voice components and WMC.

We present 3 studies using different methods to investigate how voice quantity and quality are related to individual differences in WMC and are causally affected by WMC load-manipulations. We contribute to the literature by offering new theoretical insights about the differential involvement of cognitive processing in different components of voice, hereby reconciling current opposing views in the voice literature. Specifically, Study 1 (lab study, between-person individual differences) and Study 2 (experiment, within-person situational differences) were designed to test the involvement of WMC in objective, behavioral measures of quantity and quality of voice. In Study 3 (multi-source survey field study), we investigated voice quantity and quality towards both colleagues and supervisors to replicate the lab-results in an organizational sample.

Studies 1 and 2

Voice in the Social Context

To express voice is to challenge the status quo with the intent of improving the situation (LePine & Van Dyne, 1998). Employee voice is an informal, voluntary way of constructively communicating ideas, suggestions, concerns, information about problems or opinions on work-related issues. Voice is proactive and constructive in the sense that employees who voice try to anticipate future problems in the organization as opposed to merely criticizing the current state of affairs (Morrison, 2014). Scholars have argued that this type of dissenting behavior helps organizations to learn from their mistakes (Edmondson, 1996), to avoid crises like the explosions of NASA-space shuttles ‘the Challenger’ and ‘the Columbia’ by breaking with groupthink (Ferraris & Carveth, 2003), and to increase team-innovation and effectiveness (De Dreu, 2002).

However, voice is not always appreciated by the individuals that the ‘voicer’ attempts to influence. For example, research shows that employees who voice do not always get favorable performance evaluations (Grant, Parker, & Collins, 2009). Evaluations by others play a crucial role because voice is inherently social. It always involves a target. Employees can voice towards different targets including colleagues (speaking out) or supervisors (speaking up), depending upon whom they feel might be able to set the desired change in motion (Liu, Zhu, & Yang, 2010). However, even though voice behaviors are improvement- and socially oriented (Van Dyne & LePine, 1998), the challenging nature of voice makes it interpersonally risky (Burris, 2012). Since voice entails social risk and is not always effective in initiating change (Detert, Burris, Harrison, & Martin, 2013), one might assume that people think carefully before they speak out or up.

Voice Quantity

Although it seems logical that people ‘think’ before (or while) they voice, it is not that clear yet how the voice decision process develops in the mind. Until recently, the predominant focus in the empirical voice literature has been on situational and personality factors that influence the act of voicing. Early voice measures tended to focus on whether or not people voice ideas, problems, or suggestions in organizations (e.g., ‘this employee communicates his / her opinion … even if others disagree’, Van Dyne & LePine, 1998, p.112). We refer to this as voice quantity since it assesses whether or not people speak up. More recent measures of voice do include more aspects pertaining to the content of voice, yet use items that possibly combine both quality and quantity thus making it difficult to disentangle the two (e.g., Liang et al., 2012; Mayne & Podsakoff, 2014). Studies show several contextual and trait variables to be related to voice, such as safety and self-esteem (LePine & Van Dyne, 1998; Morrison & Milliken, 2000), leadership (Detert & Burris, 2007; Liu, Zhu & Yang, 2010), personality (Crant, Kim, & Wang, 2011) and job autonomy (Tangirala & Ramanujam, 2008). Cognitive and experimental research on voice, however, is scarce and insight into the cognitive processes underlying voice decisions is limited to date.

Several theoretical ideas exist on the involvement of cognition in the voice process. For example, some authors assume that voice requires elaborate cognitive processing, because it is intentional, planned, and future focused (Chiaburu et al., 2008; Liang et al., 2012). Others argue that voice requires a risk-reward calculus: when judgements of efficacy and safety increase, people voice; when they decrease, voice involves risk, and thus people remain silent. Factors that influence evaluations of risk and efficacy include voice climate in groups, felt psychological safety and the perceived probability of success, organizational support, and trusting relationships with voice targets (Ashford, Rothbard, Piderit, & Dutton, 1998; Detert & Burris, 2007; Morrison, Wheeler-Smith, & Kamdar, 2011). However, even though there are many elements to consider when deciding to voice (‘how to voice, what to voice, will someone punish me, am I able to get the message across?’), that does not necessarily mean that everyone always elaborately and consciously considers all those elements.

Extensive research in other fields has shown that risky decisions are often based on heuristics and biases, reflecting system 1, automatic cognitive processing (Kahneman & Tversky, 1979). Other studies show that people judge whether they feel safe or at risk based on affective properties (Loewenstein, Weber, Hsee, & Welch, 2001). Reliance on affect in risky decisions cognitively forms a much more efficient way to navigate in a complex, uncertain world than elaborate processing, and people often remain unaware of their motivations for engaging in or refraining from risk taking (Mishra, 2014; Slovic et al., 2004; Slovic, Peters, Finucane, & MacGregor, 2005). Even complex decisions such as buying a house presumably depend largely on automatic processing (Dijksterhuis, Bos, Nordgren, & Baaren, 2006). So even though decisions seem highly complex, they are often taken based on automated social scripts (Dehaene, 2014).

‘The decision to voice might thus in part be emotional or based on deeply rooted beliefs and schemas that operate below the level of conscious, rational decision making’ (Morrison, 2011: 384). Similarly, other authors suggest that decisions to voice ‘depend on general as opposed to specific effortful cognitive appraisals of context’ (Dutton, Ashford, O’neill, Hayes, & Wierba, 1997, p. 419). Also, affect and implicit theories about the effectiveness of voice were shown to predict whether people withhold voice that might be useful for management (Detert & Edmondson, 2011; Kish-Gephart et al., 2009; Milliken, Morrison, & Hewlin, 2003). Evaluations of perceived risks and benefits depend on individuals' general affective evaluation of an activity. If people like an activity, they tend to judge risks as low and benefits as high. It might thus be that people who voice often, like the behavior and therefore calculate risk differently than people who voice little. This implies that affective experiences with the behavior shape implicit efficacy, expectancy, and risk beliefs.

When people voice frequently, it would not be efficient to use elaborative information processing every time one decides to voice. Individuals likely simply choose to engage in the behavior that felt good in the past. Speaking up or staying silent may thus be the result of learned scripts for decisions under risk, which often involve little elaborate cognitive processing. In line with this, the only study that we know of that investigated the relationship between more general cognitive resources (GMA) and voice (Le Pine & Van Dyne, 2001) found no association. This suggests that an ability to engage in complex thinking does not influence whether people engage in voice or not. Thus, we propose that voice decisions (voice quantity) are not the result of careful cognitive elaboration, but are rather made using fast and automatic cognitive processes.

Voice Quality

A more likely part of the voice process to involve elaborate cognitive processing, is the quality of voiced content. Although pioneering voice research has focused primarily on voice quantity (frequency), more recently, scholars have started to discriminate between types of voice on the basis of content. Voice content research has highlighted that different voice types have different antecedents and consequences. For example, people who feel obligated to change things at work, voice more *suggestions* (i.e. promotive voice) whereas people who feel safe, voice more potential *problems* (i.e. prohibitive voice) (Liang et al., 2012; Morrison, 2011). *Opinion*-focused voice, which is more focused on deviating from the status quo opinion, is hindered by self-monitoring (Morrison, 2012; Premeaux & Bedeian, 2003). Yet, we do not yet know what determines the quality of these different types of voice content.

We thus do not know what makes voiced information, whether it be a suggestion, a problem, or an opinion, of actual importance to the organization. Because even if people who voice intend to improve something, that does not mean that they always do. We expect voice quality and quantity to be distinct but inter-related components of voice behavior (e.g., frequent voicers may experience what is valued and thus learn to voice over time). While the early voice measures (Van Dyne & LePine, 1998) focus on frequency of voice, more recent content-specific measures such as promotive voice (Liang et al., 2012), constructive voice (Maynes & Podsakoff, 2014) or prosocial voice (Van Dyne, Ang, & Botero, 2003) presumably also capture quality to a certain extent. They fail however, to disentangle the two components of voice, which is, in our opinion, necessary for studying the role of cognitive elaboration.

We define voice quality as ‘the discretionary constructive communication of suggestions, problems, or opinions that add novel and useful information to the organization’. This relates to creativity, which is the generation of original yet appropriate ideas in any domain (Amabile, 1996; Runco & Charles, 1993). Although in experimental creativity, measures often focus on original ideas, usefulness is critical for organizational change because original ideas that do not convey implemental value, are hard to implement. Within the voice quality construct, there is thus equal emphasis on whether voice entails novel information, and on whether the information valuable to the organization. Here, we see voice as useful when a suggestion can be implemented, a problem is about an important issue, or an opinion changes the discussion in a meaningful way. Voice is adds novelty when a comment is challenging, refreshing, and thus successful at changing the status quo opinion.

Creativity and voice quality share attributes. For example, originality or novelty plays a role in both concepts. One other way to conceptualize voice quality would thus be to refer to creative voice, or proactive creativity (e.g., Unsworth, 2001). Specifically the way creativity is usually measured in organizations, with the use of supervisor ratings (e.g., Zhou & George, 2001; Zhou & Shalley, 2003) is similar to what we define as voice quality. However, we refer to voice quality for the purpose of contrasting the quantity and quality aspects of voice in this paper. This means that compared to these organizational creativity measures, the voice quality concept has a broader focus. For example, besides proactively sharing solutions and suggestions, it is also about pointing out important problems, and besides sharing novel ideas or challenging opinions, it also focuses on usefulness. Even though usefulness is in the definition of creativity, it is usually less well represented in creativity measures. Furthermore, in the currently available measures of creativity in organizations, there is no clear differentiation between proactive creativity (self-initiated), and reactive creativity (a response to a request from the supervisor). Since we are interested in the quality of voice as a proactive behavior, such a distinction is important.

The same proactive – reactive distinction applies for a clear differentiation of voice quality from creativity measures that are used in experimental research. Experimental research on the influence of cognition and attention on creativity uses tasks (Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; De Dreu et al., 2012; Fink, Benedek, Grabner, Staudt, & Neubauer, 2007a) that are reactive in nature: people are requested to generate creative ideas. They are not assumed to initiate their own creative responses and goals, and they do not socially interact with others. Investigating more proactive forms of creativity, in this case, the quality of voice, is thus important because highly creative people may generate great ideas in experiments, but never share them in a social environment. We want to study the involvement of cognitive processes not only in the generation phase, as has been done in previous work, but also within the process of deciding what to share with other people. Voice quality is thus a proactive, extra-role behavior: it refers to a voluntarily offered opinion, idea, or solution to a self-discovered problem. Voice quality is thus more social, more proactive, and contextually broader than the creativity that is typically studied in experimental research.

While creativity and voice quality are thus different, we assume that they are related because both require the generation of novel information, and are thus facilitated by elaborative cognitive processing. Research suggests that Working Memory Capacity facilitates the idea generation process (Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; De Dreu et al., 2012). Here, we argue that WMC is likely to influence the quality of voluntarily voiced suggestions, problems, and opinions not only because it allows for original idea generation (creativity), but also because it allows for developing, selecting, and communicating those suggestions, problems, and opinion that matter to the organization (voice quality).

WMC: Cognitive Elaboration Resource

‘We constantly overestimate the power of consciousness in making decisions, but in truth, our capacity for conscious control is limited’ (Dehaene, 2014: 47). The capacity for conscious control is reflected in the capacity of working memory (Barret et al., 2004), which is limited by the quantity of information it can contain. Imagine being in a secluded park with 3 children. Two of them are playing in the grass, but the third one attracts your attention because she gets in trouble climbing up a tree. Before you attend to the tree-climber, you glance at the other two, remembering their position. You keep this visual-spatial information in mind when you climb up the tree to get the third one, which allows you to quickly locate and check on the others while you are in the tree. Now imagine the same scenario, but in a petting zoo with 8 children during high-season. Most people would fail to remember the location of all remaining 7 children, especially considering the distraction of animals and other kids. The amount of children to be remembered (storage capacity) and the ability to rule out irrelevant children (distractor interference), varies greatly between people: there is a capacity limit to the memory-system. Working memory holds limited information ‘online’ to serve cognition (Cowan et al., 2005). How much we can keep attentively online at a time (Working Memory Capacity or WMC) ultimately defines the possible depth of our thoughts as well as our capacity for conscious control (Barret et al., 2004).

Since WMC is central to effortful and elaborate cognitive processing (Evans, 2008) and attention control (Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2002, 2003), it is important for performing any task that requires focused and elaborate thinking, such as planning, reasoning, and persistent future goal pursuit. More specifically, the number of items (storage capacity) one can retain in the visual working memory is assumed to be equal to the number of interrelationships between elements that can be kept active while reasoning (Halford, Cowan, & Andrews, 2007). The storage component of (visual) working memory is thus a key limiting factor in our ability to understand abstract relationships between novel items (Fukuda et al., 2010). This means that WMC is related to fluid intelligence (Cowan et al., 2005; Johnson et al., 2013; Vogel & Machizawa, 2004), but is nevertheless a different construct (Ackerman, Beier, & Boyle, 2005; Conway et al., 2002). WMC is a more basic, specific measure of elaborative processing capacity, which is not distorted by domain-specific, cultural, or knowledge-dependent skills. Visual working memory tasks are only about the number of items one can retain within the visual attention scope, not about reading ability, mathematical ability, or knowledge from long-term memory (crystallized intelligence).

Another way to measure the involvement of WMC in any behavior that specifically focusses on the part of WMC that controls attention, is to put load on working memory such that it cannot be used to execute other tasks (Baddeley, 2003). Because WMC is limited, once it is ‘full’, it is occupied and cannot be used to cognitively elaborate on other issues. In our studies, we therefore use both visual WMC measures as well as attentional load manipulations to study the involvement of elaborative processing in the quantity and quality of voice. We thus investigate WMC's involvement in voice both in terms of individual differences in evolutionary basic, visual WMC, and in terms of causality. As argued above, we expect that with regard to the quantity of voice, there is no elaborate cognitive processing involved, and we therefore propose the following hypothesis:

H1 - WMC is unrelated to voice quantity.

In contrast to voice quantity, we propose that the role of WMC is more important for the quality of the voiced message. Voicing with the intention to improve does not necessarily mean that the communicated message is original and useful, that is, of high quality. In support of a link between elaborate cognitive processing and idea development, experimental research has shown that elaborate processing is important for the generation of original ideas and the execution of original musical improvisation. De Dreu and colleagues (2012) showed that WMC facilitates idea generation because it allows for a persistent and perseverant focus of attention on the creative task. People high on WMC are able to stay focused on the generation process for longer periods of time and generate original ideas because they consider more possible alternatives within contextual categories. Similarly, Benedek and colleagues (2014) found effects of both WM storage capacity and WM distractor interference on creative performance, indicating that idea generation is fostered by the ability to filter out distracting information and the amount of items one can retain in working memory. If thus, elaborate processing facilitates the generation of original ideas, it may also be important for the evaluation, selection, and communication of high quality ideas, problems, and opinions. Being able to assess whether a message contains novel information presumably requires understanding of the status quo and a thorough thought process to check whether indeed, the idea, problem, or opinion is not a mere repetition of issues that were previously raised by others.

In addition to fostering the generation and communication of original ideas, we propose that cognitive processing is likely to help people select and communicate those ideas that matter in their specific context (usefulness). Sawyer (2008) describes the importance of understanding and mastering ‘the context’ in order for group improvisation to lead to innovations, instead of mere chaos. According to Sawyer, the innovation process is about selecting those ideas with potential, and making incremental changes in multiple rounds of improvisations. This incremental, implementation focused, usefulness element is important for voice because voice is social and needs to help the organization, as opposed to disrupting organizational functioning. Selection and communication of ideas, problems, and opinions that are not only original, novel, and challenging, but also useful, valuable, and implementable, requires understanding and shaping of the content. This makes the ability to keep a large quantity and variety of information active in the mind (WM storage capacity) important. Furthermore, as voice is social, the ability to filter out distracting environmental noise (WM distractor interference) while thinking about what to say and how to say it might be needed for selecting the best idea for communication.

The arguments above suggest that WMC is relevant to voice quality because the generation, selection, and communication of messages that are both novel and useful requires elaborate cognitive processing. We thus hypothesize that:

H2 - WMC is positively related to voice quality.

Study 1 tests our hypotheses by investigating individual differences in WMC whereas in Study 2, we manipulated WMC through distracting people during the selection of voice-messages.

STUDY 1 METHOD AND RESULTS

Design and Sample

This between-subjects lab study tested whether individual differences in cognitive resources (visual WMC) predict whether people voice (voice quantity, *H1*), and whether the content of people's voice is useful and takes new information into account (voice quality, *H2*). Participants were 72 businesses students (57,7% male, 42,3% female) in all years of enrollment with an average age of 21.07 (*SD* = .21, *range* = 18-26) and a part-time job of 11.84 hours a week on average (*SD* = 7.39). They were tested for visual WMC and worked in a dyad on a fictional business case together with a team-leader. Participants had the chance, but were not required or requested to voice opinions and suggestions (voice quantity) to solve the case. The quality of the solution reached served as a measure of voice quality.

Procedure, Tasks, and Measures

Independent variable WMC. The independent variable WMC was assessed by means of a delayed serial recognition task (see e.g., De Dreu et al., 2012; Rossion & Pourtois, 2004; Sligte, Vandenbroucke, Scholte & Lamme, 2010). Participants performed a series of 96 trials. On each trial, participants were presented with eight (8) pictures, appearing sequentially in the center of a laptop screen. Each stimulus remained on screen for 250 milliseconds (ms), extending the complete trial over a period of 2000 ms. After a series of eight stimuli, the screen went blank for 1000 ms during which participants had to keep all information in memory. Capacity of working memory was computed with the K formula: K = (% correct – chance)* *n* stimuli (1/0.875), resulting in an estimation of the maximum number of items individuals can retain in the working memory (Cowan, 2001). The amount of items participants could retain in working memory ranged from 4.8 to 7.3 (*M* = 6.46, *SD* = .50).

Dependent variables voice quality and quantity. Following this test, participants were introduced to their team-leader with whom they performed a filler task to get acquainted. Leaders were confederates and specifically instructed and trained for the study (namely a male and a female faculty employee who were blind to the study's hypotheses and uninvolved in the study as authors). Gender of the leaders was counterbalanced between groups. To measure voice, we provided the team (participant and leader) with an information sharing task adapted from Stasser and Titus (1985). This task is specifically useful to measure voice and initiative behaviors in the social environment, because team-members need to voice in order to get to the best solution. We specifically created a leader-follower situation because we wanted any voice from participants to be upward, discretionary, risky, challenging and extra-role: the responsibility for the task-execution was completely in the hands of the leader, and leaders had the power to distribute payment at the end of the experiment (or so we told the participants). Participants had the chance to voice, but were never specifically requested for their opinion or help during the task.

The team-leaders task was to find the best new dean for the faculty. There were three candidates (A, B, and C) who all had different properties, which were selected via a pre-test of 283 students on favorable and unfavorable dean-attributes. The candidate profiles were composed in such a way that based on all available information, C would be the best candidate for the job, and A and B would both be unfavorable. However, the information that was provided to the individual team-members was (as it often is in practical situations of decision making), not identical. Based on only the *leaders'* information, B would be the best option, and based on the *participants'* information, A would be best. The team thus needed to share and integrate all information to get to the optimal decision, C (high quality).

After carefully reading all information on the sheet, the leader said: 'based on my information here, it is quite obvious that we should go for option B. Here is my information, confirming my decision'. If, at that point, the participant had proactively tried to find the best candidate for the job, s/he would have found out that from his/her perspective, A was the best choice. If the participant noticed this and wanted to mention it, s/he needed to challenge the leader and voice another option. The participant could now choose *(1)* not to voice and comply with the leader, or *(2)* oppose the leader and suggest another option (voice). Whether participants voiced or not, was our measure of *voice quantity*. *Voice quality* was assessed by the relative quality of the option participants voiced. As explained previously, the participant could either choose the less optimal option A, thus staying with his or her previous preference, or derive a new perspective from the leaders' information and voice option C, the best option from the combined viewpoints. The outcome difference between these suggestions, is assumed to be the result of the evaluation and selection process for the voice message that integrates novel information, and is most valuable to the organization (or in this case the team). Voice quality was thus measured as a dichotomous variable with low (option A) and high (option C) quality.

Results

Our data thus consisted of two groups: participants who voiced, and participants who did not voice. Because we hypothesized that WMC is unrelated to voice quantity *(H1)*, we expected no WMC differences between participants who voiced and who did not voice. Indeed, we found no WMC difference between these groups, $F(71) = .83$, $p = .367$, $\eta^2 = .012$. Within the group who voiced, there were participants who voiced low and high quality. Because we hypothesized a positive relationship between WMC and voice quality *(H2)*, we expected that WMC would be highest in the group who voiced high quality. As predicted, WMC was higher in the group who voiced high quality ($N = 22$, $M = 6.66$, $se = .10$, $CI_{95} = [6.46, 6.85]$) than in the group who voiced low quality ($N = 27$, $M = 6.37$, $se = .09$, $CI_{95} = [6.18, 6.56]$), $F(49) = 4.92$, $p = .031$, $\eta^2 = .095$. To visualize these effects, we split the sample into a Low- and High WMC group (using the median). The Low WMC group did not voice more than the High WMC group (Figure 1a, $\chi^2 = .58$, $p = .448$), whereas the High WMC group voiced more high than low quality solutions, and Low WMC group voiced more low than high quality solutions (Figure 1b, $\chi^2 = 6.20$, $p = .01$). This supports the idea that WMC is related to voice quality *(H2)*, but not to voice quantity *(H1)*.

Figure 1. Study 1 – Between Participants

Insert Figure 1 about here

----

a. Voice Quantity (N participants who voiced / did not voice) by WMC.

b. Voice Quality (N participants who voiced high / low voice quality) by WMC.

STUDY 2 METHOD AND RESULTS

Design and Sample

In order to show that our findings are not limited to between-person differences, but that they are similar within-persons when elaborate cognitive processing is either possible or limited, we designed another experiment using a within-participant design. One of the ways to study causal effects of WMC, is to put load on working memory. Taxing working memory by another activity (such as, for example, counting backwards), reduces its capacity to be used by another activity (for example, voice). We thus designed two within- person conditions in this experiment, one with, and one without distraction of working memory. All participants got the opportunity to voice in both conditions. Further, to assess voice quality, after the experiment, all voice was rated by three raters in terms of whether or not voice provided novel solutions or useful implementations superior to the status quo.

Participants were 77 (we aimed for 80, but lost 2 participants due to incomplete data, and 1 because average idea generation was more than 2.5 *SD* below the mean) primarily business (55%) and psychology (36%) students (67% female, Age $M =$ 21.20, $SD =$ 3.08) who participated in exchange for money. They were welcomed by the experiment leader and seated behind a computer. The experimenter explained that payment depended on task performance and that performance would be monitored and evaluated at distinct moments in the experiment.

Procedure, Manipulations, Tasks, and Measures

Participants could engage in two types of payment dependent behaviors. Their primary job was the working memory load task (counting backwards), always resulting in a €2,50 salary if performed correctly. Secondary to this task, there were (extra-role) opportunities to voice, which could result in €1,- bonus (for high quality) or €1,- loss (for low quality) for every voiced suggestion, problem, or opinion depending on the leaders evaluation of voice quality. The experiment started with an exercise of the working memory-load task (counting backwards from 100, in steps of 2) followed by performance feedback. We explained that this task would be repeated several times during the experiment, and that this would always be their primary task with a standard reward of €2,50.

Idea Generation Phase. Our main aim was to find out whether WMC is only important in the idea generation phase (creativity, as previously found by De Dreu et. al, 2012), or also important in the 'selection for communication' phase of voice. To make sure participants would have a pool of messages to choose from, they first anonymously generated 10 problems, opinions, or suggestions for improvement of their faculty (adopted from De Dreu, Baas, & Nijstad, 2008) in a period of 10 minutes (for the remainder of the text, we call these: ideas). We told them explicitly that these ideas would not be evaluated or used for analysis, but that this was just an anonymous practice round. We needed 10 ideas to be able to fill both conditions (No WMC Load – High WMC Load) with 5 of the participants' own ideas to potentially voice. During generation, the experimenter digitally classified the ideas (low/high quality). The computer used these classifications for random-weighted message distribution across the WMC Load conditions so both conditions contained an equal amount of high and low quality messages.

Practice phase with simultaneous load manipulation. The generation phase was followed by two combined load and selection practice tasks. We designed two tasks (one simple, one more complex) in which participants needed to select items from a pool of 10 words (the 3 largest animals) or 10 sentences (5 true statements about animals) while they performed a working memory load task (counting backwards). The purpose of these tasks was to familiarize participants with the combination of counting while selecting messages for voice, limiting possibilities for technical confusion during the real opportunity for voice.

Communication (voice) phase and simultaneous load manipulation. Following the practice phase, participants got the instructions that they would get to see a random selection of their own ideas, as well as a random selection of other peoples' ideas. We added 5 ideas from a database of relevant ideas to each Load condition because voice takes place in a social context. In the social environment, the evaluation of the quality of one's own ideas also depends on the evaluation of the ideas of others. These other-ideas were kept constant between participants: each Load condition contained the full quality (low – high) range of ideas as coded by two independent judges. During the 'window for voice', participants thus saw 10 messages from which they could select those ideas they found 'worthy' to voice: 5 of their own, and 5 generated by others.

We told them that they would get the opportunity to voice any idea to the experiment-leader, who would evaluate it and pay them accordingly. Voice (as opposed to the generation phase) was thus no longer anonymous and moreover, financially risky: voicing a good idea would result in a €1,- gain, voicing a bad idea in a €1,- loss. Participants were not pushed to voice any ideas, because their primary and most profitable task was to count backwards, either from 80 in steps of 2 (High Load) or from 100 in steps of 0, thus repeating the same number vocally (No Load). Participants performed two trials with 10 messages each and a random order of High vs. No Load. Thus, participants had the opportunity to voice their previously generated ideas, *while* performing a WMC-load task. Voicing ideas was done digitally, by selecting and sending the message to the experimenter, who would evaluate the idea and calculate payment accordingly.

Dependent variables voice quality and quantity. *Voice quantity* was measured by counting the number of voiced ideas in each Load condition. Specifically, we counted how many suggestions, problems, or opinions participants sent to the experimenter (quantity) to be evaluated. Across conditions, participants voiced 2.87 ideas on average (*range* = 0 - 10, *SD* = 2.29). Further, whereas Study 1 measured voice quality as an objective operationalization of value and insight, we now used expert ratings of voice quality. After the experiment, all ideas from the generation phase were evaluated by three independent judges, who rated whether the ideas provided novel solutions or useful implementations superior to the status quo, on a scale from 0 (not at all) to 6 (highly) (adopted from De Dreu, Nijstad, & Baas, 2011). Inter-rater reliability was reasonable for the generated quality (*ICC =.659*). Voice quality was the average quality of those ideas that were voiced by the participants, which was higher (M_{voiced} = 3.06, se = .036, $_{within}CI_{95}$ = [2.99,3.13]) than generated quality ($M_{\text{generated}}$ = 2.86, se = .036, $_{within}CI_{95}$ = [2.78,2.93]), F (1,64) = 7.76, p = .007, η^2 = .108, Greenhouse-Geisser corrected. Within-participant confidence intervals were calculated using the method developed by Loftus and Masson (1994), as revised by Cousineau (2005). The same method will be used to report all within subjects 95% confidence intervals throughout the rest of the paper.

Results

We ran all main analyses using 2-level (No-Load vs. High-Load) repeated measures ANCOVA's, controlling for the order of the load tasks (No-Load or High-Load first). Furthermore, as both parts of voice are intertwined (voice quality cannot be evaluated without voice quantity), we controlled for voice quality which we entered as a continuous standardized covariate when testing for load effects on voice quantity (*H1*). Please note that since some participants never voiced, we could not compute their voice quality, which is modeled as a covariate. This influences the degrees of freedom. Similarly, we controlled for voice quantity, entered as a continuous covariate, when testing for load effects on voice quality (*H2*). We expected that load would not influence voice quantity, and thus that average voice quantity would be the same across the Load conditions. Indeed, we found little difference in voice quantity between the Load conditions. In the No Load condition, participants voiced slightly more (M_{No} = 1.54, *se* = .082, ${}_{within}CI_{95}$ = [1.38, 1.71]) than in the High Load condition (M_{High} = 1.34, *se* = .082, ${}_{within}CI_{95}$ = [1.18, 1.51]), but this effect was not significant, F (1,61) = .270, p = .61, η^2 = .004, Greenhouse-Geisser corrected. To increase power, we repeated the analysis using average generated quality (instead of average voice quality, to increase the number of cases included in the analysis) as a covariate. This did not change the outcome of the results, (M_{No} = 1.53, *se* = .069, ${}_{within}CI_{95}$ = [1.39, 1.67]), (M_{High} = 1.36, *se* = .069, ${}_{within}CI_{95}$ = [1.22, 1.50]), F (1,72) = .61, p = .438, η^2 = .008, Greenhouse-Geisser corrected. Similar to Study 1, this suggests that the involvement WMC in deciding whether to voice and how often to voice, is limited.

In contrast, we did expect that load would influence voice quality, and thus anticipated a difference in voice quality between Load conditions. We expected higher quality of voiced ideas in the No Load condition compared to the High Load condition, because load on working memory should impair elaborate processing and should thus impair effective selection of high quality ideas. As expected, we found the predicted load effect on voice quality, but only when controlling for voice quantity. Participants voiced lower quality when under High Load (M_{High} = 3.03, *se* = .054, ${}_{within}CI_{95}$ = [2.92, 3.14]), compared to when under No Load (M_{No} = 3.19, *se* = .054, ${}_{within}CI_{95}$ = [3.08, 3.30]), F (1,41) = 8.87, p = .005, η^2

= .178, Greenhouse-Geisser corrected. Similar to Study 1, this suggests that WMC is important for deciding *what* to voice. However, the fact that the main effect was only significant when controlling for quantity, suggested that there might be an interaction effect. Interestingly, we found a significant interaction between load and overall voice quantity on voice quality, F (1,41) = 6.88, p = .012, η^2 = .144, Greenhouse-Geisser corrected. To shed light on the direction of this effect, we split (based on the median: 3) our sample into participants who voiced little (3 times or less, N = 21) and participants who voiced a lot (4 times or more, N = 24). We observed that there was no effect of load ($M\Delta_{NO\text{-}High}$ = .09) in the group who voiced a lot, F (1,23) = .52, p = .480, η^2 = .022. Frequent voicers showed equal voice quality under High Load (M_{High} = 3.16, se = .075, $_{within}CI_{95}$ = [3.00, 3.31]) and No Load (M_{No}= 3.06, se = .075, $_{within}CI_{95}$ = [2.91, 3.21]). In contrast, participants who voiced little showed impaired voice quality due to load ($M\Delta_{NO\text{-}High}$ = .44), F (1,20) = 6.05, p = .023, η^2 = .232. For infrequent voicers, voice quality was significantly lower in the High Load condition (M_{High} = 2.89, se = .080, $_{within}CI_{95}$ = [2.73, 3.05]) than in the No Load condition (M_{No} = 3.33, se = .080, $_{within}CI_{95}$ = [3.17, 3.49]). It thus appears that WMC is particularly important for voice quality of those who voice little (Fig. 2).

Figure 2. Study 2 - Within Participants

Interaction WMC * Voice Quantity on Voice Quality.

Insert Figure 2 about here

The results of this within-participant experiment support the idea that WMC is uninvolved in voice quantity (*H1*), but is involved in voice quality (*H2*). The findings suggest that WMC is not only related to voice quality because it fosters idea generation (as found in previous research), but also because WMC helps individuals to select the best ideas, solutions, problems, and opinions to voice. The interaction effect between voice quantity and working memory load suggests that there might be differences in the involvement of WMC between people who voice often and people who voice little. Presumably, people who voice often, do so more automatically and frequently in general. They might be more familiarized with cognitively evaluating and selecting ideas due to practice and therefore need less WMC to do so. In contrast, people who do not voice often, are not experienced at the evaluation and selection process, and therefore need more space in working memory to do so. In the third study, we set out to replicate our findings in a field setting and to test the idea that the role of WMC differs for those who voice often and those who voice rarely.

STUDY 3

Our next step in investigating (WMC) differences between voice quantity and quality, was to test our hypotheses in a field setting. The first challenge was to create a survey measure that would clearly distinguish between voice quantity and quality. While voice quantity and quality are likely related in organizations, we argued previously that someone who voices often does not necessarily always voice the best ideas or opinions. Although scholars do assume that voice has qualitative and innovative properties (e.g., van Dyne and Le Pine, 1998; Van Dyne, Ang and Botero, 2003), voice measures typically do not incorporate these properties. Voice measures such as the often used one developed by Van Dyne and Le Pine (1998), thus seem to tap more into the frequency of voice, than its quality. We tested this assumption in a pilot study among 53 employees and their supervisors, by comparing the voice measure by Van Dyne and LePine (1998) to several newly developed voice quantity and quality items. We found that the traditional voice scale was only significantly related to voice quantity but not to voice quality (please see the supplementary materials for a full report on this pilot study).

Since our pilot study showed that the traditional voice measure was indeed unrelated to items tapping voice quality (but related to items focusing on voice quantity), our next step was to develop these new measures further in order to be able to clearly differentiate between voice quality and quantity in the field. We propose that this distinction is relevant because these two components are (in part) driven by different antecedents, and in particular show differential relationships with WMC. We also explored the relationships of voice quality and quantity with other proactivity constructs, such as proactive personality (Bateman & Crant, 1993), personal initiative (Frese et al., 1996) and taking charge (Morrison & Phelps, 1999), and more generally performance. The results of this multi-source pilot field study are reported in the supplementary materials (Table A4) which presents the nomological network of voice quality and quantity. In summary, the results indicate that voice quality is more strongly related to performance (rated by managers), personal initiative (rated by colleagues), and taking charge (rated by managers) than voice quantity. Further, our Studies 1 and 2 already confirmed the relevance of this distinction for more objective measures of voice quantity and quality specifically concerning WMC. The first goal of our field study (Study 3) was therefore to replicate those findings in organizations with our further developed voice quantity and quality scales using evaluations of supervisors and coworkers.

Our second goal was to investigate whether the reliance on WMC and thus the involvement of cognitive processing in voice quality, might differ between people. Our previous study (Study 2) showed that infrequent voicers were more dependent on WMC to reach high quality voice than those who voiced a lot. There could thus be a difference in the way people cognitively approach the act of speaking up. Some people may think elaborately before they voice, organizing thoughts beforehand, only to speak up after elaboration with the best comments. Others may be more prone to speak while they think, improvising on the way. Presumably, people get better at selecting and communicating good ideas as they practice. If practice indeed decreases the involvement of WMC in voice quality, we should be able to find similar effects in real organizations comparing people who typically voice frequently to those who do not.

To support the idea that practice may decrease WMC involvement in high quality voice, we draw on examples from literature on creativity in musical improvisation. De Dreu and colleagues (2012, Study 3) showed that WMC is beneficial to creativity of improvisational musical performance of undergraduates in classical music performance who were not trained at improvisation. In contrast, Beaty, Smeekens, Silvia, Hodges, and Kane (2013) found that in a sample of jazz undergraduates who *were* skilled at improvisation (jazz musicians practice improvisation, classical musicians do not), WMC was not related to creativity of the improvisational performance. This indicates that WMC might serve to overcome a lack of experience in improvising, by increasing creative quality of the improvisation through cognitive elaboration. In the same study, Beaty et al. (2013) found a strong effect between WMC and an idea generation task, something that the jazz musicians where presumably less trained at; this confirms the idea that practice reduces the role of WMC in creative performance.

There is more musical evidence to support this line of thinking. Bengtsson, Csíkszentmihályi and Ullén (2014) showed that the right dorsolateral prefrontal cortex, which has been linked to working memory processes in earlier research (Crone, Wendelken, Donohue, van Leijenhorst, & Bunge, 2006; Curtis & D'Esposito, 2003; Jolles, Kleibeuker, Rombouts, & Crone, 2011), was activated during improvisations of professional concert pianists. Again, these classically educated pianists are highly skilled in musical performance, but not at improvisation. Limb and Braun (2008) find the opposite pattern for jazz-experts (who are experienced in improvising). They showed decreased activity in brain-areas that are associated with consciously monitoring goal-directed behaviors. Extensive training in a creative domain thus seems to be associated with different patterns of brain activation (Sawyer, 2011). In sum, creating novel musical patterns may either be achieved through persistent cognitive elaboration (supported by WMC), or by more relaxed flexible and intuitive cognitive processing (supported by improvisational practice). We propose that these findings generalize also to other behaviors, and voicing works in a similar vein: those who voice often learn to improvise, whereas those who voice rarely need to think before they speak up. Both can reach high quality, but through different cognitive pathways.

There are three factors that seem important when considering how we cognitively prepare to voice in the workplace: feedback, task-novelty, and time-on task. Time-on task is important because of the distractor interference function of WMC. Unexperienced high WMC musicians only became more creative over time, presumably because they were better able to focus their attention on the creative task (De Dreu et al., 2012). Similarly, employees who do not voice often might thus think about voice and how others might react to it longer and more elaborately before they share their thoughts. This could lead to an advantage for high WMC employees because they are better able to control attention and thus elaborate on the content of voice without being distracted by other environmental demands. Frequent voicers might be better able to anticipate reactions of others due to previous experience, and may therefore act more intuitively during conversations, making the advantages of WMC less important.

Task-novelty is important because of the storage capacity function of WMC. People who voice often may not need the same degree of WMC to form complex connections between concepts, because these already exist in long term memory and automatized patterns. Because they have experience with voice (and thus thinking about what to say), they can chunk more information, making it easier to keep track of previously mentioned information, maintaining quality over time. People who voice little however, cannot rely on long term memory because to them, message evaluation and selection is novel, and thus their storage capacity determines whether or not they are able to voice high quality. The last way to learn what is novel and valued, is by receiving formal or informal feedback. Voicing often might help people efficiently organize and store information that is novel, relevant, and appreciated in a given context because voice evokes reactions from others. By voicing more often, employees are exposed to these reactions, which teaches them what quality is in their organizational context. Feedback may also help to apply previously voiced ideas to other subject areas, which, in turn, decreases dependency on WMC to organize novel information.

Thus, people who typically voice often may be less reliant on their cognitive resources to reach high quality voice as they may have learned how to voice well and usefully in a given context. Those who do not voice often or who have little experience in a given context need to organize all information in their mind before they speak up to be able to come up with a good comment, and they need WMC to do that. This implies that in practice, there may be two routes to voice quality: A cognitive elaboration route facilitated by WMC, and an intuitive improvisation route that is paved by voice experience (and independent of WMC). We thus hypothesize:

H3 - Voice quantity moderates the positive relationship between WMC and voice quality such that the relationship between WMC and voice quality is strongest if employees voice rarely.

Design and Sample

Participants were 158 employees ranging in age between 16 and 61 (*M* = 33.57, *SD* = 12.33) and their 79 supervisors who were 41.92 old on average (*range* = 20 – 65, *SD* = 11.81). The participants worked in a variety of organizations such as SMES (39.7%), consultancy (17.8%), retail (15.8%), governmental institutions (4.8%), non-profit (6.8%), education (6.8), healthcare (8.2%). Gender was equally distributed in the employee sample

(50/50), the majority of supervisors were male (62/38). Average employee tenure was 6.7 years (*SD* = 8.62), for supervisors, it was a little longer (*M* = 11.2, *SD* =10.5).

Procedure, Tasks, Measures

We collected visual WMC and voice data from 158 employees in 79 teams of unique matched triads. Each employee was tested for visual WMC, after which they completed a questionnaire to rate one of their colleagues' voice quantity and quality (and the colleague did the same about the focal employee). The team-supervisor (*N* = 79) rated both employees on voice quantity and quality. All tests and questionnaires were administered with a researcher present. Researchers made an appointment with each team, personally explained and distributed all measures, and left after all data were collected.

Independent variable WMC. Participants first took a visual WMC test (in Presentation version 17.3) which was a combination of change localization and storage tasks (adopted from Luck & Vogel, 1997; Conway et al., 2002). There were three different basic tasks: change localization, serial spatial memory, and simultaneous spatial memory. All basic tasks were executed with pure storage attributes as well as distractor interference resulting in 8 different types of trials. In all trial-types, participants had to store objects (circles and rectangles) in their working memory (remember them). During this target phase, objects were displayed for 250 ms in a 5x5 grid. The location of all objects should be remembered during a delay period of 500 ms after which they disappeared. Objects were green (targets) and yellow (distractors) rectangles (120x40 pixels, 0.89 visual degree) and circles (diameter 120 pixels, 2.68 visual degree). After a 500 ms delay (objects disappeared), participants either had to recall the location of all circles or localize one rectangle that changed. The test started easy (1 object), adding objects gradually over trials, and only ended when the individual capacity limit within all trial-types was reached. The test was reliable (α =.73). WMC was calculated using the same K –formula presented in Study 1 (Cowan, 2001).

Dependent variables voice quality and quantity. Colleagues and supervisors rated focal employee voice quantity (15 items) and quality (15 items) with a questionnaire (see supplementary materials Tables A3a and A3b for details of the development of the scales). To control whether raters understood the behavior we intended to measure was *a)* discretionary, proactive, and constructive and *b)* distinct in the sense that quantity does not imply quality and vice versa, we first explained this followed by control questions. Voice quality was meant to reflect novelty and usefulness of voiced information (suggestions, problems, and opinions). Voice quantity was reflected in the frequency of voiced information (suggestions, problems, and opinions). One voice quantity item was deleted due to poor factor loadings. Cronbach's' alphas were high for colleague ($\alpha_{quality}$ = .93, $\alpha_{quantity}$ = .89) and supervisor ratings ($\alpha_{quality}$ = .96, $\alpha_{quantity}$ = .91). Descriptives and correlations can be found in Table A4 (supplementary materials).

Measurement model. Our first aim was to replicate our pilot-study findings (supplementary materials A1) that voice quality and quantity are separate constructs. If voice quality and quantity are indeed distinct, a 4-factor (2 sources x 2 types) measurement model should fit the data better than a 2-factor (2 sources) model which represents voice as one construct. We tested this with a CFA in Mplus. As expected, the 4-factor model with a separation of quantity and quality provided a good fit (χ^2 = 167.07, df = 101, N=144, p < .001, CFI = .97, TLI .96, RMSEA =.07, 90% CI [.05 , .08], AIC = 4055, BIC = 4206). We compared this model to a 2 factor-model, with quantity and quality as one factor, and the source (colleague vs. supervisor) as separate factors, which showed a significantly poorer fit ($\Delta\chi^2$ = 397.53, df =2 , p < .001), (χ^2 = 564.60, df = 103, N=144, p < .001, CFI = .78, TLI = .74, RMSEA = .18, 90% CI [.16 , .19] , AIC = 4448, BIC = 4594). This again suggests that voice quality and quantity are distinct constructs, not only when rated by supervisors, but also by colleagues. More details and in depth analyses can be found in the supplementary materials, factor loadings (range .74 – .95) and scale details can be found in the supplementary materials, tables A3a and A3b.

Results

In the previous studies, we showed that WMC predicts the quality of voice *(H2)*, but not the quantity of voice *(H1)*. Although our hypotheses are at the individual level, our study has a two-level data structure, with employees nested in 3-person teams with one supervisor. Employees rated each other, whereas supervisors rated both employees. To take between-group and within-group variance into account when using combined ratings of colleagues and supervisors, we applied multilevel structural equation modelling (MPlus) to control for team-level variance. Since our measurement model indicated that colleague and supervisor ratings are distinct factors and the literature implies the same, we first tested our hypotheses in separate multiple regressions. However, we expected to find the same pattern of results in voice towards colleagues and supervisors.

The regression analysis of our first model (colleague ratings) showed no significant effect of WMC on the quantity of voice ($\beta_{controlled\ for\ quality}$ = .02, p = .832, se = .07, $\beta_{simple\ effect}$ = .10, p = .225, se = .08) whereas the effect of WMC on voice quality was positive ($\beta_{controlled\ for\ quantity}$ = .14, p = .04, se = .07, $\beta_{simple\ effect}$ = .18, p = .016, se = .08). Employees with high WMC thus voiced better suggestions, opinions, and problems to their colleagues than employees with low WMC *(H2)*. However, they did not voice more often *(H1)*. As expected, we found the same pattern of results for supervisor ratings. Employee WMC did not significantly predict voice quantity ($\beta_{controlled\ for\ quality}$ = -.03, p = .622, se = .05, $\beta_{simple\ effect}$ = .12, p = .095, se = .07), but did positively predict voice quality towards supervisors ($\beta_{controlled\ for\ quantity}$ = .13, p = .033, se = .06, $\beta_{simple\ effect}$ = .21, p = .011, se = .08). Results for the combined sources are depicted in Figure 4 and show the same pattern: WMC is unrelated to voice quantity ($\beta_{controlled\ for\ quality}$ = -.04, p = .418, se = .05) and positively related to voice quality ($\beta_{controlled\ for\ quantity}$ = .15, *p* = .005, *se* = .05). The capacity of working memory thus seems to be important for generation, selection, and communication of high quality voice, regardless of the voice target.

We hypothesized that due to practice, people who typically voice often rely less on WMC than people who tend to voice rarely *(H3)*. Consequently, we expected negative interactions between voice quantity and WMC when predicting voice quality: WMC should have the strongest positive effect on voice quality when voice quantity is low. We first tested this assumption concerning all ratings of voice (supervisor and colleague ratings combined) in MPlus using a robust maximum likelihood estimator. We started with the total model (both sources), and continued with robustness tests with any possible combination of sources in the model. If practice indeed increases automaticity of voice and in turn decreases the influence of cognitive resources (WMC), this pattern should exist in every model, regardless of the source rating quantity and quality. As expected, we found a negative interaction effect (β = -.17, *p* = .003, *se* =.06). Simple slope analysis (based on a split through the mean) showed that the relationship between WMC and voice quality was positive in the group who voices little ($\text{Low}_{Quantity}$: β = .40, *t* = 3.33, *p* = .001) and non-significant in the group who voices often ($\text{High}_{Quantity}$: β = .15, *t* = 1.30, *p* = .199). This interaction (see Fig. 3) is in line with our findings in Study 2 (Fig. 2), that WMC load affected voice quality only for participants who voiced little. Next, we proceeded with the same analyses for every possible combination of sources. We consistently found the same pattern of results in all models, which can be seen in Table 1.

TABLE 1
Model Summaries for all Interaction Models with Dependent Variable Voice Quality

DV Voice Quality	Independents	*β*	*se*	*p*
	Voice Quantity Total	.65**	.06	.000
Total	WMC	.17**	.05	.001
	Quantity*WMC	-.17**	.06	.003
	Voice Quantity to Supervisors	.66**	.05	.000
Supervisor Ratings	WMC	.14**	.05	.008
	Quantity*WMC	-.14**	.05	.007
	Voice Quantity to Colleagues	.44**	.09	.000
Colleague Ratings	WMC	.15*	.07	.029
	Quantity*WMC	-.23**	.07	.001
	Voice Quantity to Supervisors	.33**	.09	.000
Colleagues Ratings	WMC	.17*	.07	.017
	Quantity*WMC	-.22*	.10	.040
	Voice Quantity Colleagues	.30**	.07	.000
Supervisor Ratings	WMC	.18*	.08	.018
	Quantity*WMC	-.12†	.06	.055

Note.
† p < .10, * p < .05, ** p < .01, *** p < .001.

Figure 3. Study 3 - Between Participants

Interaction WMC * Voice Quantity on Voice Quality.

Insert Figure 3 about here

To summarize all results of Study 3, in Figure 4, we provide a visual overview of all hypothesized direct and interaction effects. This model shows that WMC predicts voice quality *(H2)*, but not voice quantity *(H1)*, and that quantity of voice (which is positively related to quality), moderates the relationship between WMC and voice quality *(H3)*.

Figure 4. Study 3

Model summary of all hypothesized effects.

Insert Figure 4 about here

GENERAL DISCUSSION

This paper aimed to incorporate knowledge from experimental psychology, cognitive science and decision making under risk to further our understanding of the cognitive processes underlying voice behavior (speaking up with suggestions, problems, and opinions to help the organization). Some previous conceptual work emphasized that voice is a function of elaborative cognitive processes (Chiaburu, Marinova, & Van Dyne, 2008), while several empirical studies suggested that decisions not to voice are often implicit and affect-driven and might therefore be more automatic (Detert & Edmondson, 2011; Kish-Gephart et al., 2009). Here, we aimed to provide further insight in which parts of voicing are more and which parts are less driven by elaborate cognitive processing. We proposed and found that voice quality and quantity differ in that respect. As hypothesized, the results show that voice behavior in and of itself (i.e., voice quantity or frequency) does not necessarily require either the ability (WMC) or the opportunity (no load on working memory) to engage in elaborative thinking. Generating, selecting, and communicating high quality messages to voice (voice quality), however, does require the cognitive resources to think thoroughly. Studies 2 and 3 highlight that this is especially likely to be the case when the person who voices has little voice experience in a given context.

Based on the relationship between WMC and creativity (Benedek, et al 2014; De Dreu et al., 2012), we hypothesized that WMC plays a role in reaching high quality voice content through facilitating persistent, elaborate thought. The more information one can actively retain in awareness (storage capacity), and the better one’s capacity to control attentional resources (distractor interference), the more elaborate and complex thinking is possible. This should increase chances to come up with new and useful ideas, problems, and opinions, but should also help to evaluate, select and communicate the best information (voice quality). Indeed, we found positive relationships between the cognitive resource WMC and objective (Study 1) and multi-source perceptual measures (Study 3) of voice quality. Further, hindering elaborate systematic thinking (through WMC-load, Study 2), impaired the selection of high-quality ideas for voice. People presumably need space in their working memory for voice, but mostly when they generate information and when they try to select high-quality messages to convey.

Further, our results showed that voice quantity (in contrast to quality) requires less cognitive resources. Voice is socially risky behavior, and previous research has shown that assessment of risk is often not rational, but rather based on affect and heuristics (Finucane, Alhakami, Slovic, & Johnson, 2000; Grossberg & Gutowski, 1987; Slovic et al., 2005). When people decide to voice, they might just do what felt good in the past, relying on social schema’s and affect and thereby leaving cognitive elaboration unnecessary to decide whether to voice or not.

Furthermore, even though voice is a proactive behavior, and proactive behavior is assumed to be discretionary and agentic, the existence of constructs like proactive personality (Bateman & Crant, 1993) and personal initiative (Frese et al., 1996) indicate that there is a trait-like, automatized component to proactivity that is relatively stable across situations. If people are consistently proactive (or un-proactive), they might not always carefully deliberate whether they want to be proactive or not. Our findings indicate that when it comes to proactive *voice* behavior, this might indeed be the case as in all three studies we found no evidence for the engagement of WMC in decisions to voice. Whether and how often people voice opinions, problems, or suggestions to colleagues or to supervisors was independent of individual differences in cognitive resources (Study 1 and 3) but also independent of within-person changes in cognitive resources (Study 2).

Our findings add to previous indications of automaticity in voice behavior which showed that silence is often not the result of carefully elaborated decisions but of fast, implicit, and affective reactions (Detert & Edmondson, 2011; Kish-Gephart et al., 2009). However, silence is not the opposite of voice (Van Dyne, Ang, & Botero, 2003): where silence means that the employee knowingly withholds information, not voicing might imply that there simply is no information to share. People high in WMC may pick up more information and therefore theoretically may have more to share, ultimately resulting in more voice. On the other hand, our first lab-study in which participants had to detect and voice information inconsistencies, showed that WMC was equal between the group who voiced and the group who did not voice. So even if high WMC increases the capacity to detect and solve problems and inconsistencies, this does not mean that this capacity is always used, and it does not increase the likelihood that people will voice. Our results thus add more nuance, indicating that not only the decision to remain silent, but also the decision to voice might be automatic rather than elaborate and cognitively complex.

The results also point at individual differences in using elaborate cognitive processing for voice behavior. It is too simplistic to state that elaborative processing is not involved when engaging in voice. For one, it plays a clear role in voice quality in all three studies. Additionally, the involvement of cognitive resources in voice does seem to differ between people who voice often and people who voice little. We argued that voicing often in a given context would result in experience with evaluation and selection of messages and with typical reactions, and that therefore people who are consistently more proactive than others (i.e., people who voice often but presumably also people high on proactive personality and personal initiative) decrease their reliance on cognitive resources for high quality voice. Indeed, the data indicates that the involvement of WMC is strong for people who voice little, but nonexistent for people who voice often. Cognitive elaboration thus seems to play less of a role for people who voice often, and a significant role for people who voice less, especially with regard to the quality of what they voice.

If not through elaborative thinking, how do frequent voicers reach high quality? Our findings point in two directions with similar implications, *(1)* a proactive trait direction, and *(2)* a cognitive processing direction. The latter implies that there are two cognitive routes to voice quality, an elaborate one (aided by WMC) and an improvisational one – (aided by voice-experience). Perhaps a certain amount of experience and automaticity (voicing in that specific organizational context) is important to reach novel and useful voice. Practicing how to voice one's messages more effectively presumably increases the likelihood that the receiver will value future voice. Previous experiences may shape the way messages are communicated and evaluated. Those who never voice may less easily learn what type of voice is valued and what type is not. If such experience is lacking, cognitive resources become important. This means that high quality voice can be trained (independent of cognitive resources) and shaped through practice.

The other, more trait-like interpretation of our results, indicates that there are two types of people: Those who 'have to say something' and those who 'have something to say'. We started our introduction with Plato's words on this: -"***Wise men*** *talk because they have something to say;* ***fools,*** *because they have to say something*" - Perhaps, people do not (only) differ in their use of working memory resources because of experience, but also because of personality. For example, introverts voice less than extraverts (Tornau & Frese, 2013), but they are not necessarily less creative (Feist, 1998). It could be that people who voice rarely only speak up when they feel sure they have something important to say, while people with high WMC have higher chances of having something novel and useful to say, resulting in the strong relationship between WMC and voice quality in this particular group. The other group just 'has to say something', meaning they will voice more automatically and regardless of their outcome expectancies because of their personality. Contradicting Plato, however, our results suggest that these people are not necessarily fools: they also get high quality ratings for voice. They reach quality either by practice over time, as described in the ‘cognitive processing’-explanation), or by a ‘quantity breeds quality’- principle: if one voices a lot of ideas, the chances that one message will be of high quality increase. Our results leave both of these explanations open for debate. However, it does seem to be the case that there are dual pathways to voice quality, whether it be dual cognitive pathways, or dual personality pathways.

Regarding the influence of voice quantity on others, Parnes and Noller (1972) already pointed out that if people are exposed to a variety of challenging ideas, their individual idea generation also increases. Presumably, this also increases voice quality through vicarious learning (Manz & Sims, 1981): when people observe other people voice often, they learn what is appreciated by others, and they learn what to select and communicate for voice. Voice quantity might thus fuel innovation in groups through facilitating voice quality in individuals, but also in team-mates. Other research on the effects of dissenting opinions and suggestions by minorities in groups (De Dreu & West, 2002; Van Dyne & Saavedra, 1996) shows that team-innovation is facilitated by exposure to deviating comments of others. Voice quantity might thus not only be important for the individual who voices, but also for others who can use their messages to enhance their own generation, selection and communication process.

Strengths and Limitations, and Future Research

We tested our core assumptions about the relationship between WMC and voice quality and quantity in three different designs: a between-subjects lab design, a within-subjects experimental design, and a multi-source field study. The proposed (un)involvement of WMC in quantity and quality of voice was found and replicated in all three studies. Study design only allowed us to test (and confirm) the negative interaction between WMC and voice quantity in Studies 2 and 3. It would be interesting to replicate and extend this using multiple indicators of experience and personality. Investigating differences in WMC involvement during voice for employees whose core task involves innovation and from other (non-innovation trained) departments, between teams who had proactivity training or not (Scott, Leritz, & Mumford, 2004), between individuals from teams with differing climates for voice (Morrison et al., 2011), or people high and low in personal initiative, would increase in-depth insight into this interaction.

One point of note, is that the correlation between supervisor-ratings of voice quantity and quality was rather high (higher than for colleague ratings). This could indicate that supervisors are less able or inclined to separate the two. However, it could also imply that if employees choose to voice they may be more selective in what they voice to supervisors compared to what they voice towards colleagues. First, if the employee feels that s/he has something valuable or useful to say, the likelihood of voice to a supervisor presumably increases. Second, if voice quality towards supervisors is low, negative feedback might influence quantity of voice more severely than in the case of voice towards colleagues. Furthermore, even though the relationship between supervisor rated quantity and quality of voice was relatively strong, WMC was still only related to voice quality, not to voice quantity. So even though voice towards supervisors might be more risky than voice towards colleagues, WMC does not influence voice quantity to supervisors more than to colleagues. This is in line with our idea that deciding to voice is not necessarily elaborated because risky decisions rely on automaticity rather than on cognitive elaboration.

Another limitation concerns the relatively high levels of WMC in our studies samples. Whereas Cowan (2010) argues that WMC in adults is limited to 3-5 items on average, mean WMC-scores in our student and field sample were 6.5 and 5.3 respectively. This means that we have no insight yet into how people with WMC below the population mean reach high quality voice. Lastly, one could argue that since the results from Studies 1 and 3 are correlational, we cannot conclude that elaborate systematic thinking is or is not involved in quantity and quality of voice. The results only point in such a direction, but are strengthened by the experimental design in Study 2 where impairment of cognitive resources harmed only voice quality, as well as the negative interaction effects we found between voice quantity and WMC. Replication of this experiment and longitudinal studies are necessary to further strengthen our assumptions and add nuance. Presumably, for people who voice very little, WMC is also important for voice quantity, but this relationship might be mediated by creativity. Those high in WMC may generate more ideas that are original than those low in WMC, which means they have a larger pool of good ideas to choose from, increasing their voice quantity.

Implications for Practice

The findings of our studies offer important implications for practice. Cognitive neuroscience has shown that visual WMC might be a fundamental basis for complex thinking (Fukuda et al., 2010; Vogel, McCollough, & Machizawa, 2005). Visual working memory tests are faster, cheaper, evolutionarily more basic and less culturally biased than the more general mental ability tests that are common in management literature and selection procedures. They do not require domain specific knowledge and thus does not obstruct people with for example, dyslexia or other language (due to being from a different culture) disadvantages. In terms of enhancing diversity in organizations, this might thus be a better way to assess whether people are able to engage in effortful complex thinking.

Regarding our dual pathways to voice quality, there are two important issues to consider for managers (if they wish to enhance quality voice). First, some people (infrequent voicers) need cognitive resources to voice high quality. Measuring WMC before hiring people may thus be a way to select those. However, these people also elaborate to reach high quality voice, and elaboration takes time and effort. Merely hiring high WMC people is thus not enough, there needs to be time and space to elaborate. Second, some people (frequent voicers) need to be able to practice voice in order to learn how to improvise. Since quantity of voice seems to be more automatic (not well elaborated), trying to influence this process with rational arguments might not be very effective. People either voice often because it is in their personality, or because in their team, there is an implicit feeling that voice is safe. If people in a team generally voice rarely, it may be more useful to try to enhance voice by focusing on the implicit cues the team leader is sending, than explicitly telling people they can always share their concerns.

Theoretical Contributions

The first contribution of this study is the distinction between voice quality and quantity. It opens up new avenues for studying the multidimensionality of voice behavior and emphasizes the difference between a decision to say something and the quality of what is said. This focus on content relates to both the work on creativity, such as the combination of originality and usefulness for truly innovative ideas (Woodman, Sawyer, & Griffin, 1993) , and to voice research, such as the (complexity) differences between suggestions, opinions, and problems (Morrison, 2011), and the differences between promotive versus prohibitive voice (Liang et al., 2012). This study strengthens the claims of those studies that it is important to consider the content of voice if we want to know more about how effective voice is generated, selected and communicated.

Our second contribution is the investigation of the involvement of cognitive resources in voice behavior. Morrison (2011) called for more research in the cognitive elaboration versus automaticity area, which we answered by challenging the idea that voice behavior in and of itself is the result of a complex cognitive calculation. We attempted to sharpen the hypotheses about the involvement of elaborate cognitive processing in voice by showing that the arguments of Chiaburu and colleagues (2008) about cognitive elaboration hold, but only for the quality of voice and especially when employees don't voice often. Our results are also in line with the idea that choosing (not to) voice depends on affect and implicit theories about voice (Detert & Edmondson, 2011; Kish-Gephart et al., 2009). Our findings thus reconcile both views by showing that indeed, voice quality (sometimes) requires elaborate processing, whereas voice quantity relies on automatic processing. The results of this study emphasize the importance of a cognitive-processes stream in the voice literature. Automaticity in a self-initiated, discretionary proactive behavior such as voice may seem like a paradox, but often, cognitive elaboration is less important than we might think.

Third, we contributed new designs for testing causal relations in proactive voice behavior. Although we did not use our first lab-setting for causal purposes, the design does offer that possibility. In the present paper, we used two different designs (within- and between subjects), which makes a wide range of causal research questions about individual differences as well as situational differences about voice and initiative possible. Since experimental research in voice and proactivity literature is still scarce, we hope that this helps other researchers to tackle causal research questions concerning proactive behavior.

ACKNOWLEDGEMENTS

The authors thank Brent Scott, Antje Schmitt, Renske van Geffen, Maarten Wolsink and Mats Nagel for reading earlier versions of the manuscript, Jasper Wijnen for his help with computer programming, and Bram Kuijken and Monika Kackovic for their role as confederates in study 1.

REFERENCES

Amabile, T. M. (1996). Creativity in context: Update to" the social psychology of creativity.". Westview press.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, 131(1): 30-60.

Ashford, S. J., Rothbard, N. P., Piderit, S. K., & Dutton, J. E. (1998). Out on a Limb: The Role of Context and Impression Management in Selling Gender-Equity Issues. *Administrative Science Quarterly*, 43(1): 23–57.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4(10): 829–839.

Bagozzi, R. P., & Edwards, J. R. (1998). A General Approach for Representing Constructs in Organizational Research. *Organizational Research Methods*, 1(1): 45–87.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind. *Psychological Bulletin*, 130(4): 553–573.

Bateman, T. S., & Crant, J. M. (1993). The Proactive Component of Organizational Behavior: A Measure and Correlates. *Journal of Organizational Behavior*, 14(2): 103–118.

Beaty, R. E., Smeekens, B. A., Silvia, P. J., Hodges, D. A., & Kane, M. J. (2013). A First Look at the Role of Domain-General Cognitive and Creative Abilities in Jazz Improvisation. *Psychomusicology: Music, Mind & Brain December 2013*, 23(4): 262–268.

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, 46: 73–83.

Bengtsson, S. L., Csíkszentmihályi, M., & Ullén, F. (2014). Cortical regions involved in the generation of musical structures during improvisation in pianists. *Journal of Cognitive Neuroscience*, 19(5): 830–842

Burris, E. R. (2012). The Risks and Rewards of Speaking up: Managerial Responses to Employee Voice. *Academy of Management Journal*, 55(4): 851–875.

Chiaburu, D. S., Marinova, S. V., & Dyne, L. V. (2008). Should I Do It or Not? An Initial Model of Cognitive Processes Predicting Voice Behaviors. Kane and M. R. Poweller (Eds.), *Citizenship in the 21st Century*: 127-153. Nova Science Publishers, Inc.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30(2): 163–183.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1): 42–45.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, 24(1): 154– 176.

Crant, J. M., Kim, T.-Y., & Wang, J. (2011). Dispositional Antecedents of Demonstration and Usefulness of Voice Behavior. *Journal of Business and Psychology*, 26(3): 285–297.

Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L., & Bunge, S. A. (2006). Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences*, 103(24): 9315–9320.

De Dreu, C. K. W. (2002). Team innovation and team effectiveness: The importance of minority dissent and reflexivity. *European Journal of Work and Organizational Psychology*, *11*(3): 285–298.

De Dreu, C. K. W., Baas, M., & Nijstad, B. A. (2008). Hedonic tone and activation level in the mood-creativity link: Toward a dual pathway to creativity model. *Journal of Personality and Social Psychology*, 94(5): 739–756.

De Dreu, C. K. W., & West, M. A. (2001). Minority Dissent and Team Innovation: The Importance of Participation in Decision Making. *Journal of Applied Psychology December 2001*, 86(6): 1191–1201.

De Dreu, C. K. W., Nijstad, B. A., & Baas, M. (2011). Behavioral Activation Links to Creativity Because of Increased Cognitive Flexibility. *Social Psychological and Personality Science*, 2(1): 72–80.

De Dreu, Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working Memory Benefits Creative Insight, Musical Improvisation, and Original Ideation Through Maintained Task-Focused Attention. *Personality and Social Psychology Bulletin*, 38(5): 656–669.

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. Penguin.

Detert, J. R., & Burris, E. R. (2007). Leadership behavior and employee voice: is the door really open? *Academy of Management Journal*, 50(4): 869–884.

Detert, J. R., Burris, E. R., Harrison, D. A., & Martin, S. R. (2013). Voice Flows to and around Leaders:

Understanding When Units Are Helped or Hurt by Employee Voice. *Administrative Science Quarterly*, 58(4): 624–668.

Detert, J. R., & Edmondson, A. C. (2011). Implicit voice theories: Taken-for-granted rules of self- censorship at work. *Academy of Management Journal*, 54(3): 461–488.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & Baaren, R. B. van. (2006). On Making the Right Choice: The Deliberation-Without-Attention Effect. *Science*, 311(5763): 1005–1007.

Dutton, J. E., Ashford, S. J., O'neill, R. M., Hayes, E., & Wierba, E. E. (1997). Reading the wind: How middle managers assess the context for selling issues to top managers. *Strategic management journal*, 18 (5): 407-423.

Edmondson, A. C. (1996). Learning from Mistakes is Easier Said Than Done: Group and Organizational Influences on the Detection and Correction of Human Error. *The Journal of Applied Behavioral Science*, 32(1): 5–28.

Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(1), 255–278.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2(4): 290–309.

Fink, A., Benedek, M., Grabner, R. H., Staudt, B., & Neubauer, A. C. (2007). Creativity meets neuroscience: Experimental tasks for the neuroscientific study of creative thinking. *Methods*, *42*(1), 68–76.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, 13(1): 1–17.

Frese, M., Kring, W., Soose, A., & Zempel, J. (1996). Personal Initiative at Work: Differences between East and West Germany. *The Academy of Management Journal*, 39(1) 37–63.

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, 17(5): 673–679.

Grant, A. M., Parker, S., & Collins, C. (2009). Getting Credit for Proactive Behavior:supervisor Reactions Depend on What You Value and How You Feel. *Personnel Psychology*, 62(1): 31–55.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: a new hypothesis. *Trends in Cognitive Sciences*, 11(6): 236–242.

Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., … Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology*, 27(2): 220–229.

Jolles, D. D., Kleibeuker, S. W., Rombouts, S. A. R. B., & Crone, E. A. (2011). Developmental differences in prefrontal activation during working memory maintenance and manipulation for different memory loads. *Developmental Science*, 14(4): 713–724.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2): 263–291.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130(2): 169-183.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, 9(4): 637–671.

Kane, M. J., & Engle, R. W. (2003). Working-Memory Capacity and the Control of Attention: The Contributions of Goal Neglect, Response Competition, and Task Set to Stroop Interference. *Journal of Experimental Psychology*, 132(1): 47–70.

Kish-Gephart, J. J., Detert, J. R., Treviño, L. K., & Edmondson, A. C. (2009). Silenced by fear: The nature, sources, and consequences of fear at work. *Research in Organizational Behavior*, *29*, 163–193.

LePine, J. A., & Van Dyne, L. (1998). Predicting voice behavior in work groups. *Journal of Applied Psychology*, 83(6): 853-868.

LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: evidence of differential relationships with big five personality characteristics and cognitive ability. *Journal of Applied Psychology*, 86(2): 326–336.

Liang, J., C. Farh, C. I., & Farh, J.-L. (2012). Psychological Antecedents of Promotive and Prohibitive Voice: A Two-Wave Examination. *Academy of Management Journal*, *55*(1), 71–73.

Limb, C. J., & Braun, A. R. (2008). Neural Substrates of Spontaneous Musical Performance: An fMRI Study of Jazz Improvisation. *PLoS ONE*, 3(2).

Liu, W., Zhu, R., & Yang, Y. (2010). I warn you because I like you: Voice behavior, employee identifications, and transformational leadership. *The Leadership Quarterly*, *21*(1), 189–202. Loewenstein, G. F.,

Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic*

Bulletin & Review, 1(4): 476–490.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657): 279–281.

Maynes, T. D., & Podsakoff, P. M. (2014). Speaking more broadly: An examination of the nature, antecedents, and consequences of an expanded set of employee voice behaviors. *Journal of Applied Psychology*, 99(1): 87-112.

Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An Exploratory Study of Employee Silence: Issues that Employees Don't Communicate Upward and Why*. *Journal of Management Studies*, 40(6): 1453–1476.

Mishra, S. (2014). Decision-Making Under Risk Integrating Perspectives From Biology, Economics, and Psychology. *Personality and Social Psychology Review*, 18(3): 280–307.

Morrison, E. W. (2011). Employee Voice Behavior: Integration and Directions for Future Research. *The Academy of Management Annals*, 5(1): 373–412.

Morrison, E. W. (2014). Employee voice and silence. *Annu. Rev. Organ. Psychol. Organ. Behav.*, 1(1): 173–197.

Morrison, E. W., & Milliken, F. J. (2000). Organizational Silence: A Barrier to Change and Development in a Pluralistic World. *The Academy of Management Review*, *25*(4): 706–725.

Morrison, E. W., & Phelps, C. C. (1999). Taking Charge at Work: Extrarole Efforts to Initiate Workplace Change. *The Academy of Management Journal*, 42(4): 403–419.

Morrison, E. W., Wheeler-Smith, S. L., & Kamdar, D. (2011). Speaking up in groups: A cross-level study of group voice climate and voice. *Journal of Applied Psychology*, 96(1): 183–191.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology*: 37–72. New York, NY, US: Guilford Press.

Runco, M. A., & Charles, R. E. (1993). Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences*, 15(5): 537–546.

Sawyer, K. (2008). *Group Genius: The Creative Power of Collaboration*. Basic Books.

Sawyer, K. (2011). The Cognitive Neuroscience of Creativity: A Critical Review. *Creativity Research Journal*, 23(2): 137–154.

Scott, G., Leritz, L. E., & Mumford, M. D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal*, 16(4): 361–388.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality. *Risk Analysis*, 24(2): 311–322.

Slovic, P., Peters, E., Finucane, M. L., & MacGregor, D. G. (2005). Affect, Risk, and Decision Making. *Health Psychology Supplemental Issue: Basic and Applied Decision Making in Cancer Control*, 24(4): 35–40.

Stasser, G., & Titus, W. (1985). Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussion. [Miscellaneous Article]. *Journal of Personality*, 48(6): 1467–1478.

Tangirala, S., & Ramanujam, R. (2008). Exploring Nonlinearity in Employee Voice: The Effects of Personal Control and Organizational Identification. *The Academy of Management Journal*, 51(6): 1189–1203.

Tornau, K., & Frese, M. (2013). Construct Clean-Up in Proactivity Research: A Meta-Analysis on the Nomological Net of Work-Related Proactivity Concepts and their Incremental Validities. *Applied Psychology*, 62(1): 44–96.

Unsworth, K. (2001). Unpacking Creativity. *The Academy of Management Review*, *26*(2), 289–297.

Van Dyne, L., Ang, S., & Botero, I. C. (2003). Conceptualizing Employee Silence and Employee Voice as Multidimensional Constructs*. *Journal of Management Studies*, 40(6): 1359–1392.

Van Dyne, L., & LePine, J. A. (1998). Helping and Voice Extra-Role Behaviors: Evidence of Construct and Predictive Validity. *The Academy of Management Journal*, 41(1): 108–119.

Van Dyne, L., & Saavedra, R. (1996). A naturalistic minority influence experiment: Effects on divergent thinking, conflict and originality in work-groups. *British Journal of Social Psychology*, 35(1): 151–167.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984): 748–751.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438(7067): 500–503.

Woodman, R. W., Sawyer, J. E., & Griffin, R. W. (1993). Toward a Theory of Organizational Creativity. *The Academy of Management Review*, 18(2): 293–321.

Zhou, J., & George, J. M. (2001). When Job Dissatisfaction Leads to Creativity: Encouraging the Expression of Voice. *The Academy of Management Journal*, 44(4): 682–696.

Zhou, J., & Shalley, C. E. (2003). Research on employee creativity: A critical review and directions for future research. *Research in Personnel and Human Resources Management*, 22: 165–218.