

Krefft, Anja; Marcus, Bernd; Weigelt, Oliver

**Article — Published Version**

## Facets of Constructive Voice Behavior: Construct Clarification, Theoretical Refinement, and Measurement

Journal of Business and Psychology

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Krefft, Anja; Marcus, Bernd; Weigelt, Oliver (2024) : Facets of Constructive Voice Behavior: Construct Clarification, Theoretical Refinement, and Measurement, Journal of Business and Psychology, ISSN 1573-353X, Springer US, New York, NY, Vol. 39, Iss. 5, pp. 1085-1105, <https://doi.org/10.1007/s10869-024-09932-z>

This Version is available at:

<https://hdl.handle.net/10419/315317>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>



# Facets of Constructive Voice Behavior: Construct Clarification, Theoretical Refinement, and Measurement

Anja Krefft<sup>1</sup> · Bernd Marcus<sup>1</sup> · Oliver Weigelt<sup>2,3</sup>

Accepted: 30 December 2023 / Published online: 23 January 2024  
© The Author(s) 2024

## Abstract

Constructive voice encompasses a multitude of voice message content. However, existing measures of voice do not capture these differences in a precise way and contribute to inconclusive evidence. Based on existing conceptualizations of constructive voice, we suggest that employee voice messages differ along three main dimensions: (1) suggesting innovative opportunities versus highlighting harm (functional orientation), (2) offering specific suggestions versus pointing out problems (substantive orientation), and (3) addressing existing versus potential future issues (temporal orientation). Using this three-dimensional approach, we define five distinct prototypical facets of voice with well-defined conceptual boundaries. Through two expert studies ( $N = 12$  and  $N = 10$ ), we identify uncertainties in the content of current voice measures. We then introduce a newly developed measure that addresses these shortcomings. In two consecutive survey studies ( $N = 132$  and  $N = 553$ ), we present initial evidence supporting the reliability, convergent validity, and discriminant validity of this measure. The correlations among the five facets of voice were high. However, we also found evidence for differential associations with some antecedents.

**Keywords** Proactive work behavior · Constructive voice behavior · Promotive voice · Prohibitive voice

Proactive voice behavior (or simply “voice”) is the voluntary expression of constructive-intended and change oriented messages of innovative ideas, adaptive solutions, foreseen concerns or existing difficulties about work-related topics (Liang et al., 2012; Morrison, 2011). Research capturing voice as a unidimensional construct has found that voice is either beneficial, detrimental, or unrelated to employee well-being (e.g., Röllmann et al., 2021; Weiss & Zacher, 2022).

The causes underlying these inconclusive findings are underexplored. We argue that inconsistencies emerge from imprecisions and confounds inherent in the way voice is studied. One plausible explanation is that unidimensional measures of voice are unspecific or ambiguous about voice message content. For instance, van Dyne et al. (1995) introduced a general differentiation of extra-role behaviors into promotive and prohibitive categories, which were intended to clarify the differences and similarities in the antecedents and consequences of the various extra-role behavior constructs.

Building on this, Liang et al. (2012) introduced promotive and prohibitive voice<sup>1</sup> facets. According to Liang et al.

Additional supplementary materials may be found here by searching on article title <https://osf.io/collections/jbp/discover>. Oliver Weigelt's contribution to this work was funded by Volkswagen Foundation (Az. 96 849, “The role of work in the development of civilization diseases”).

✉ Anja Krefft  
anja.krefft2@uni-rostock.de

<sup>1</sup> Organizational and Personnel Psychology, Faculty of Economics and Social Sciences, University of Rostock, Rostock, Germany

<sup>2</sup> Wilhelm Wundt Institute of Psychology, Leipzig University, Leipzig, Germany

<sup>3</sup> Department of Psychology, University of Groningen, Groningen, Netherlands

<sup>1</sup> The present research is concerned with constructive, change focused and organization-promoting definitions in this tradition of dichotomous distinctions, which currently represent the mainstream of voice research within Organizational Behavior. Although we agree with recent calls to broaden the voice concept to cover all forms of voluntary employee communication (Maynes and Podsakoff, 2014), or collective voice in a unionist tradition (e.g., Barry and Wilkinson, 2016; Wilkinson et al., 2020), we believe that such broader approaches may also benefit from clarifying the predominant narrower definitions in the first place. This focus on individual constructive voice also aligns with Morrison's (2023) conceptual arguments in her most recent review of voice research.

(2012), promotive voice refers to the expression of innovative ideas and suggestions for improvement directed towards a future ideal state or what could be. In contrast, they define prohibitive voice as the expression of concern or problems about work practices, incidents, or employee behavior that is or could be harmful to the organization. Research that differentiates between different voice facets is largely restricted to the two-fold promotive-prohibitive distinction (e.g., Köllner et al., 2019; Lin & Johnson, 2015; Song et al., 2019). Empirical evidence based on this distinction is inconclusive (e.g., Chamberlin et al., 2017; Köllner et al., 2019; Liang et al., 2012). In their meta-analysis, Chamberlin et al. (2017) showed that promotive and prohibitive voice did not relate differentially to 12 out of 19 criteria studied (e.g., agreeableness, openness to experience, social support, workplace stressors). These authors suggested that important differences are obscured by common variance of items in existing voice scales and called for new measures allowing for clearer separation of voice content. In other words, measures of promotive and prohibitive voice may capture facets of voice less distinct than desired, because focal items may be either mute or ambiguous about more specific aspects of the voice message.

Drawing on the definition of voice outlined above, we argue in the present paper that voice encompasses a multitude of voice message facets and that existing measures of voice do not capture these differences in a precise way. Our research was therefore set out to (1) scrutinize conceptually whether the simple promotive-prohibitive voice dichotomy provides an accurate description of the range of voice facets, (2) examine empirically how content-validly voice can be captured with existing measures, and (3) investigate the content validity and psychometric properties of a new instrument designed to overcome the limitations of existing measures. Our primary objectives are to foster conceptual clarity and parsimony and to develop an appropriate questionnaire to enable research to address the question of specificity described. Drawing on Liang et al.'s (2012) work, we first seek to identify behavioral dimensions underlying the promotive-prohibitive distinction that may be advanced to arrive at a clearer separation of facets. We then present two expert studies and two employee studies in which we critically examine existing voice measures and validate our conceptualization and proposal for a new measure.

## Conceptualizing Constructive Voice Behavior

As outlined by Liang et al. (2012), universally shared attributes resonate in both promotive and prohibitive voice, encompassing extra-role conduct, constructive intent, and organizational promotion. Evaluating voice as constructive and organization-promoting is contingent upon the

perceptions and evaluations of the voice recipient and thus, beyond the act of voicing itself. However, the presence of constructive intent and organizational promotion are fundamental characteristics inherent in all voice messages, setting the voice construct apart from others (e.g., whistleblowing; e.g., Blenkinsopp & Edwards, 2008). Despite significant frameworks addressing various types of employee communication (e.g., Maynes & Podsakoff, 2014), adopting these frameworks would disregard the essential elements of the widely accepted definition of employee voice. Omitting these defining characteristics would, in accordance with Morrison (2023), diminish rather than enhance the precision of the voice construct's conceptualization.

Besides particular features common to all facets of voice behavior, promotive and prohibitive voice can be distinguished along the *function*, *behavioral content*, and *implications for other organizational members* (Liang et al., 2012). Notably, the latter dimension lies outside the behavioral domain, as, by definition, implications for others do not belong to the sphere of the act itself but to possible effects of behavior (and thus may be measured as a separate variable). Function refers to the (intended) purpose of the message, in which the voice actor may either point to innovative improvement (promotive) or to harmful factors in order to stop or prevent negative consequences for the organization (prohibitive). Behavioral content covers two distinct features of the voice message: (1) expressing innovative ideas or solutions (promotive) vs. existing or impending harmful factors (prohibitive), and (2) temporal orientation of the message content. With regard to the latter feature, Liang et al. (2012) point out that promotive voice (innovative ideas and solutions) is more future oriented but prohibitive voice (problems) may be either future or past oriented. Further, promotive voice includes both novel ideas and adaptive solutions to difficulties (behavioral content), thereby adding elements of ambiguity to their description of the two facets of voice. We also note that the first feature of behavioral content (expression of new ideas or solutions vs. harmful factors) appears closely related to both the same authors' function dimension (pointing out potential improvements vs. harmful factors) and to Morrison's (2011) earlier distinction of suggestion focused vs. problem focused voice. Van Dyne et al. (2003) also explicitly differentiate between communicating *problems* vs. expressing *ideas and suggestions* to improve the organization. Taken together, the distinction between promotive and prohibitive voice covers three descriptive behavioral dimensions for which it is not clear to which degree these features can and should be distinguished conceptually and psychometrically.

Several groups of authors proposed features of constructive voice in addition to those outlined by Liang et al. (2012), Morrison (2011), and van Dyne et al. (2003). Those additional features include the target of the voice message (e.g., manager, team, or employee) (Burris et al., 2022), the

channel through which the voice message is expressed, the relevance/value, urgency, and feasibility of the voiced message (Brykman & Raver, 2021; Burris et al., 2017). However, we argue that all these dimensions should be separated from the measurement of different voice facets because the target and channel are by definition not part of the sphere of the voice message itself but rather may be of interest as correlates of voice behavior. If so, such variables are more appropriately measured separately from voice behavior. Similarly, the substantively same voice message may be directed at different targets and communicated through different channels, which may yield different consequences independent of behavioral content. Further, whether a proposal is important or even feasible and implementable, or whether a problem is urgent, lies in the subjective perception of both the voicer and the voice recipient.

Building on the distinction between promotive and prohibitive voice and extending the dimensional approach elaborated by Liang et al. (2012), we suggest a taxonomy that disentangles specific features of the two-fold distinction in order to avoid conceptual confounds. More specifically, we propose to begin with dimensions that could be defined without ambiguity, and then use those dimensions to develop an unambiguous typology of voice message facets. We adopt the function dimension (innovative improvement vs. harmful factors), but remove confounds inherent in the behavioral content dimension of Liang et al. (2012), resulting in three fundamental dimensions which qualitatively distinguish various voice messages. First, the *functional orientation* dimension of voice serves as a key discriminator for voice messages, differentiating messages of highlighting innovative opportunities versus identifying harm (Liang et al., 2012). It is imperative to emphasize that voice messages highlighting innovative opportunities inherently encompass a component of tackling perceived challenges (potential harm) that gradually emerge. However, aligning with Liang et al. (2012), the focal point primarily revolves around propelling the organization forward through progressive advancement, spearheading novel approaches and pioneering change. Conversely, voice messages functionally pointing out harm are steadfastly committed to the meticulous eradication of obstacles, ensuring the seamless maintenance of overall functionality. Second, the *substantive orientation* dimension characterizes voice messages as either identifying issues without proposing improvements or providing specific suggestions (Whiting et al., 2012). Third, the *temporal orientation* dimension distinguishes voice messages by either addressing the current status quo or contemplating matters that could impact the organization in the future (Liang et al., 2012; van Dyne & Lepine, 1998).

Any specific message of voice may be categorized unequivocally along the dimensions of functional, substantive, and temporal orientation, which may theoretically result in

$2 \times 2 \times 2 = 8$  partially overlapping or mutually exclusive prototypical configurations of respective dimensions and thus facets of voice with potentially different implications. These three dimensions describing prototypical voice message facets may help reconcile the ambiguity of previous findings. While voicing constructive messages is defined as proactive behavior, characterized by voluntary, autonomous (self-initiated), and change oriented behavior (Parker & Collins, 2010), the functional, substantive, and temporal orientation of the message could have differential implications regarding the extent to which the behavior is perceived as positively challenging or rather as stressful and risky obligation (Köllner et al., 2019; Liang et al., 2012), the extent to which voice is self-initiated and intrinsically motivated (Cangiano & Parker, 2016; Strauss & Parker, 2014), the extent to which voice necessitates the investment of personal resources in envisioning, planning, voicing, and implementation, such as time and energy (Bolino et al., 2010), and the extent to which it enhances relationships with others at work or jeopardize them (Burris, 2012; Liang et al., 2012).

For instance, an employee might suggest a radical departure from the conventional marketing strategies that the company has been using (e.g., leverage emerging virtual reality technology to create immersive product experiences for customers). The suggestion is not mainly prompted by any concrete harmful situation but rather draws attention to future trends and possibilities that could shape the company's direction. On the other side, an employee might address recurring difficulties that the team has been facing during project execution (e.g., suggesting modifications to the current processes). Both of these examples are categorized as promotive voice in accordance with Liang et al. (2012), because both messages *substantively* express new ideas or solutions for how to improve the organization. We contend that the *functional* and *temporal* orientations exhibit distinct characteristics. The initial example emphasizes the identification of innovative opportunities in the future, while the subsequent example centers on rectifying problems within the company by proposing a solution. This distinction potentially leads to significant variations in risk considerations, investment of personal resources and behavioral response between the two examples, impacting both the predictors and outcomes of such behavior.

Put differently, the now popular dichotomy of promotive vs. prohibitive voice, to the extent, it is defined by partially ambiguous dimensions, may in some way put the cart before the horse. Instead of beginning with a simple dichotomy and then trying to describe it with a number of distinct dimensions, we propose to start with dimensions that could be defined without ambiguity, and then use those dimensions to develop an unambiguous, yet parsimonious typology of voice facets. We hold that this process must be based on unequivocal measurement at the item level. That is, only if

we have single indicators (i.e., items), each of which unequivocally cover a particular combination of unambiguous dimensions (i.e., if items are content valid), may we use this set of indicators to arrive at a meaningful description of the construct (i.e., the actual structure of voice) and test whether constructs narrower than general voice may be affected by different antecedents and yield different effects, or whether the different facets of voice do not need to be measured differentially because common elements are paramount.

## Overview of Present Research

The present studies are set out to, theoretically and empirically, clarify relations among the three dimensions of the voice message, as outlined above based on previous work (Liang et al., 2012; Morrison, 2011; Whiting et al., 2012) and to provide researchers of voice with a set of indicators covering all meaningful voice facets defined by configurations of those core defining dimensions in an unequivocal manner. We thereby also echo calls to move beyond the promotive-prohibitive dichotomy (Chamberlin et al., 2017; Köllner et al., 2019; Morrison, 2014, 2023) and to establish a more precise and nuanced measure of voice.

Fortunately, we do not have to start from scratch for this endeavor. Rather, we deductively used the established definition of constructive voice and the content of a number of established voice measures as our starting points. In “Study 1,” we first submitted a comprehensive set of existing voice items to a panel of experts of constructive voice and asked them to assign each item to each of our three dimensions. In the (expected) event that “Study 1” failed to yield a sufficiently large and comprehensive set of unambiguous items, new items were developed and again subjected to a sample of experts for rating its content in “Study 2.” The set of items that passed this stage of expert judgment were administered in a self-report format to two larger samples of employees (“Study 3” and “Study 4”) for further refinement and for validating the scale’s psychometric properties.

## Study 1

### Sample, Procedure, and Materials

First, we reviewed the content of several constructive voice scales (Farh et al., 2007; Liang et al., 2012; Parker & Collins, 2010; Premeaux & Bedeian, 2003; van Dyne & Lepine, 1998; van Dyne et al., 2003; Zhou & George, 2001) to select a set of comprehensive yet non-redundant items. We included items that capture voice in terms of the outlined established construct definition and are frequently used in

research practice. Forty items were finally selected for expert review. We presented these items to the experts and asked them to rate the content of each item on dichotomous options for each of the three dimensions (*functional*: focus on innovative opportunities vs. focus on harm; *substantive*: proposal vs. no proposal; *temporal*: status quo focus vs. future focus). Our goal was to procure items that clearly exemplify conceptual distinctions. A “Not clearly attributable” option was also included, along with a field for free expert comments. Hence, rather than gathering subjective ratings of the overall degree of overlap with the target construct, as sometimes suggested for content validation (Hinkin & Tracey, 1999; Polit et al., 2007), the task for experts in this study was simply to assign items on the three predefined dimensions. We considered this less cognitively demanding and judgment prone than other traditional approaches.

A total of 48 experts of constructive voice were contacted by email in the first round. The prerequisite was a research contribution with naming as first author on constructive voice after 2010. We implemented precautions to ensure that the panel of experts possessed a clearly outlined expectation and an all-encompassing comprehension of the task in question. In addition to furnishing a comprehensive description of our undertaking and the three dimensions right from the outset, we prominently displayed the definition of each dimension alongside each item under assessment. With only six experts having completed the questionnaire after two weeks, the link of “Study 1” was posted in the *Academy of Management Forum*. As this meant giving up control of expert status, an item was included that captured expertise. As a result of this, we were able to engage additional six experts who successfully completed the questionnaire and affirmed their well-founded expertise in the domain of voice. Consequently, we attained a total of 12 experts who diligently responded to the questionnaire in its entirety. Hence, our sample size is above the median of experts participating in content validation studies within the field of Organizational Behavior research (Colquitt et al., 2019). Items were considered clearly classified if at least 75% of the experts agreed in their assignment of each dimension, which most closely resembles the requirement of 0.78 agreement of the overall degree proposed by Polit et al. (2007) for our sample size of 12.

## Results

Of the 40 items, a total of seven items were clearly classified, i.e., mutually exclusive on *all three dimensions*. Within this set of seven clearly classified items, only two out of eight possible voice facets were covered. The coverage of dimensions in “Study 1” is shown in Table 1. Four items covered the configuration of (1) focus on innovative opportunities (functional orientation), (2) proposal (substantive orientation), and (3) future focus (temporal orientation) (e.g., “This



**Table 1** Item coverage of voice facets in “Study 1”

Item coverage of voice message facets									
Items classified to none of the dimension				8 items					
Functional orientation		Focus on innovative opportunity				Focus on harm			
Items classified only to the functional dimension		Not covered				1 item			
Substantive orientation		Proposal		No proposal		Proposal		No proposal	
Items classified only to the substantive dimension		2 items		Not covered					
Items classified only to the functional and substantive dimensions		16 items		Not covered		1 item		5 items	
Temporal orientation		Future focus	Status quo focus	Future focus	Status quo focus	Future focus	Status quo focus	Future focus	Status quo focus
Items classified only to the temporal dimension		Not covered							
Items classified to all three functional, substantive, and temporal dimensions		4 items	Not covered	Not covered	Not covered	Not covered	Not covered	Not covered	3 items

employee speaks up with ideas for new projects that might benefit the organization”; van Dyne et al., 2003). Three items comprised the configuration of (1) focus on harm (functional orientation), (2) no proposal (substantive orientation), and (3) status quo focus (temporal orientation) (e.g., “This employee expresses his/her concerns about current work practices to alarm undetected problems.”; Liang et al., 2012). Noticeably, no item was assigned to both the configuration of innovative opportunity and no proposal and to the configuration of innovative opportunity and status quo focus. Also, there is no item that is clearly assigned to innovative opportunity, but remains unclear on the substantive orientation. All items assigned to innovative opportunity have a proposal assignment. However, there are clearly functionally identified items for harm focus, which remain unclear on the substantive orientation. The expert judgment on the 40 items and original materials may be found in an online supplement for “Study 1”: [https://osf.io/9ubh6/?view\\_only=e188526419c643f19d479e4bd2b79946](https://osf.io/9ubh6/?view_only=e188526419c643f19d479e4bd2b79946).

## Discussion

The most notable findings of “Study 1” are (1) that the vast majority of existing voice items cannot be unequivocally assigned to the three dimensions describing promotive vs. prohibitive voice, (2) that only two out of eight possible configurations of those dimensions seem to be represented at all in the prominent scales, and (3) that all items assigned to innovative opportunity also have a proposal and a future assignment. Hence, the present item pool seems deficient for measuring promotive and prohibitive voice with regard to both the sheer number of unequivocal items and the substantive coverage of the potential construct space. However, as most items (with the exception of the items of Liang et al., 2012) were not developed

to measure different facets of voice and to cover the central dimensions, it is difficult to know which of the apparent gaps point to actual deficiencies. An alternative explanation for these findings could stem from our approach of measuring the expressions of the dimensions as discrete categories, rather than along a continuous spectrum. Within this spectrum, a wide range of voice messages exists, for instance from those displaying highly innovative ideas to those showcasing only minimal deviations from the established framework. Some messages may simply highlight seemingly minor challenges, while others may draw attention to problems of substantial consequence. Nonetheless, our primary emphasis remains on unequivocally delineating the two qualitative expressions (e.g., innovation vs. harm) as separate categories, and further, to render them quantifiable as prototypes with well-defined conceptual boundaries. This approach enables us to distinctly characterize these voice facets and make them unambiguously amenable to measurement.

We therefore proceeded with creating new items designed to measure the potentially overlooked and meaningful configurations unequivocally and thus, a measure that captures the content domain of the construct. “Study 2” was set out to develop an item pool satisfying these requirements and to subject this pool to another panel of experts for initial content validation in similar vein as “Study 1”.

## Study 2

### Development of an Extended Item Pool

“Study 1” indicated that not all eight voice facets emerging from a crossing of the three main dimensions shown in Table 2, may be equally logically possible. Specifically, the

**Table 2** Meaningful configurations of the three dimensions

Functional orientation	Substantive orientation	Temporal orientation	Voice example quote
Is the voice message about innovation or harm?	Is the voice message confined to a description of harm or does it make a specific suggestion for improvement?	Does the voice message refer to the status quo (existing issues) or issues that may be relevant in the future?	
Focus on innovative opportunities	Proposal	Future	“I heard about idea X. We haven't tried this opportunity yet. It could give us advantages Y.”
	No proposal	Status quo	
		Future Status quo	
Focus on harm	Proposal	Future	“We might get problem X next month. Maybe we'll do Y to prevent that.”
		Status quo	“Last week we had problem X. Maybe next time we'll do Y.”
	No proposal	Future	“Next month, problem X may arise. We should do something to prevent this.”
		Status quo	“I noticed a problem X that we ought to solve.”

functional dimension appears to restrict the set of meaningful combinations. Functional orientation distinguishes voice messages directed at innovative opportunities from those focused on avoiding and stopping harm. A functional focus on innovative opportunities seems future focused by definition, as it points to a novel future state and it appears to require at least some concrete substantive idea as to where such opportunity may be sought (i.e., at least some kind of vague proposal). This notion is consistent with our findings of “Study 1,” as no item assigned to focus on innovative opportunities had also been assigned to either lack of substantive proposal or to status quo focus temporally. By contrast, a harm focused functional orientation does not restrict the other dimensions logically. Employees may either simply point to harm or additionally propose a solution (substantive orientation), and they may either point out existing problems or forecast problems that may arise in the future (temporal orientation). Notably, only one out of these four possible configurations has been identified unequivocally in existing items in “Study 1.” As all four combinations appear logically possible and potentially meaningful, we consider existing measures deficient in this respect. Based on this reasoning, we developed items covering the five potentially meaningful facets of voice listed in Table 2.

As we were able to build on developed theoretical assumptions and accumulated knowledge to offer precise conceptual definitions of our target constructs, we opted for a deductive approach to item generation, which is most appropriate for fostering content validity under these conditions (Hinkin, 1995,

1998). Specifically, following best practice recommendations for developing measurement instruments (Boateng et al., 2018), we had specified the purpose of voice based on a thorough literature review, defined the target domain of constructive voice, provided a preliminary conceptual definition, and a priori specification of the dimensions of constructive voice. We had then confirmed in “Study 1” that there are no existing instruments that adequately serve the same purpose based on content analysis of literature and extant scales. Thereby, we justified why the development of a new instrument is appropriate and how it should differ from existing instruments. These efforts resulted in final conceptual definitions for each facet of the voice construct.

Based on considerations outlined above, we developed 49 new and partially revised items designed to unequivocally cover all five meaningful combinations of the three defining dimensions. We varied the amount of information and exact wording to some extent in order to calibrate item content necessary for unambiguous classification on all three dimensions (see Clark & Watson, 1995). Some of the new items used the established items as a basis, which were clearly assigned to two out of three dimensions in “Study 1”. The first author drafted the initial set of items and revised items across several cycles until consensus across all authors was reached.

### Sample, Procedure, and Materials

The new and partially revised set of items (49 items) and the seven items, which were clearly assigned to each of the

three dimensions in “Study 1,” were examined in a second expert survey ( $N=10$ ) in the same manner as in “Study 1” to obtain evidence about the content validity of the new items in relation to the three defining dimensions and to determine the final item selection for “Study 3.” A total of 63 experts were contacted by email to investigate the content validity of the new items. Prerequisite was a research contribution with naming as first author on constructive voice behavior after 2005. This resulted in a response of 10 experts, who completed the questionnaire in full. As participation was anonymous, we are unaware of potential sample overlap across the two expert studies, yet the substantial number of experts invited, along with the added factor of reaching out to different experts in “Study 1” and “Study 2” (i.e., in “Study 1,” the link was shared in the Academy of Management Forum), does introduce the potential for variations in perspectives and responses.

## Results

Of the 56 items, a total of 32 items were clearly classified to functional, substantive, and temporal orientation. Unexpectedly, five of the seven clearly assigned items from “Study 1” were no longer clearly assigned in “Study 2.” A possible explanation for this could be that anchor effects occurred due to the new items that were explicitly developed to reflect the three dimensions and contain more information about the functional, substantive and temporal orientation of the voice message. The two items that continued to be clearly assigned were slightly revised. Of the 32 items, four additional items were also sorted out by expert comments that criticized substantive components of these items. Classification of the remaining 28 items resulted in five item groups that corresponded to the postulated five configurations of the three dimensions. At least five items for each postulated facet were clearly classified, so that no revision and thus no further extension of our initial item pool was necessary. As expected, no item rated as functionally innovative opportunity focused had been classified as either focusing status quo temporally or as covering no proposal substantively. The expert judgment on the 56 items and original materials may be found in the online supplement for “Study 2.”

## Discussion

The most notable finding of “Study 2” appears to be that a sufficiently large number of voice items can be unequivocally assigned to each of the three dimensions. Of the three dimensions we adopted for describing voice message facets, functional orientation appears to be the key feature for developing a typology that satisfies the criteria of both parsimony and comprehensiveness. If the functional focus is merely on pointing out innovative opportunity, this appears to imply substantive orientation towards suggesting an innovative proposal and temporal orientation towards the future, whereas no such implications arise from a functional orientation towards pointing out harm. Hence, values on the latter two dimensions partially depend on the value of the functional dimension. These differences do not seem to be fully captured by the simple dichotomy of promotive vs. prohibitive voice.

Items that were clearly assigned to the configuration of innovative opportunity, proposal and future focus, consistently referred to novel and innovative proposals of how to do things better in the future without explicitly pointing to harm. Hence, we suggest to label this facet innovative voice. Items that were clearly assigned to harm focus, proposal, and either status quo or future focus, referred more to proposals for change in a known framework to prevent (solve) harmful situations and conditions. Therefore, we suggest to collectively describe these facets as adaptive voice as a solution to existing or potential harm. Thus, we account for two different types of proposals within the substantive orientation: *innovative* proposal vs. *adaptive* proposal vs. *no* proposal (i.e., mere problem focus). Already van Dyne and Lepine (1998, p. 109) stated earlier that “Voice is making innovative suggestions for change and recommending modifications to standard procedures [...]” Based on these considerations, we propose an integrated and extended five-facet typology of voice messages, which is shown in Table 3. With the differentiation between the substantive orientation of the proposal, the dimension of functional orientation becomes negligible for the typology, but not for the definition of the five facets of prototypical constructive voice behavior.

We define *innovative voice* (IV) as the communication of novel ideas and creative possibilities for advancing the organization or work group in the future by leaving the

**Table 3** Five facets of voice—integrated and extended typology

Functional orientation	Substantive orientation	Configuration of functional and substantive orientation	Temporal orientation	Configuration of functional, substantive, and temporal orientation
Focus on innovative opportunities	Innovative proposal	Innovative voice	Future	Innovative voice (IV)
Focus on harm	Adaptive proposal	Adaptive voice	Future	Adaptive and future focused voice (AFV)
			Status quo	Adaptive and status quo focused voice (ASV)
	No proposal	Problem focused voice	Future	Problem and future focused voice (PFV)
			Status quo	Problem and status quo focused voice (PSV)



given framework and doing something completely different and new for the organization or work group. It calls the organizations or work group's attention to key trends and future developments in the absence of a specific harmful background (configuration: function of pointing out innovative opportunity, substantive innovative proposal, future temporal orientation).

*Adaptive and status quo focused voice* (ASV) is the expression of adaptive proposals to solve existing problems or difficulties in a given known framework with the aim of remedying them (configuration: function of pointing out harm, substantive solution proposal, status quo temporal orientation).

*Adaptive and future focused voice* (AFV) is the expression of adaptive proposals in a given known framework to prevent anticipated problematic conditions in the future and to avoid possible negative consequences for the organization or work group by adapting future actions (configuration: function of pointing out harm, substantive solution proposal, future temporal orientation).

*Problem and status quo focused voice* (PSV) is the expression of existing problems in order to stop negative consequences for the organization or work group without proposing a solution (configuration: function of pointing out harm, no substantive solution proposal, status quo temporal orientation).

*Problem and future focused voice* (PFV) is the expression of looming problems in the future in order to protect against possible negative consequences for the organization or work group through prevention without proposing a solution (configuration: function of pointing out harm, no substantive solution proposal, future temporal orientation).

In summary, the results from “[Study 2](#)” provided support for the content validity of the 28-item questionnaire, which we label *Five-Facet Constructive Voice Questionnaire* (5F-CVQ). In “[Study 3](#)”, we next test the factorial structure of the new 5F-CVQ.

### Study 3

In this study, the primary aim was to determine if any of the 28 content-valid items require revision before conducting a larger construct validation and to check whether the data fit the theoretically expected structure. We examine the proposed five-facet structure of voice by comparing the fit of the focal 5-factor model to alternative measurement models.

### Sample

The sample was composed of 132 participants from Germany. All participants held regular employment, but 76

(57.6%) of them were also enrolled in an open university. Participants included 39 men (29.5%) and 91 women (68.9%). One participant stated to be non-binary and one participant did not provide any information. Ages ranged from 18–20 years (4, 3%), 21–29 years (94, 71.2%), 30–39 years (21, 15.9%), 40–49 years (1, 0.8%), 50–59 years (11, 8.3%) to 60 years or more (1, 0.8%). Job tenure ranged from less than a year (44, 33.3%), 1–2 years (31, 23.5%), 2–5 years (32, 24.2%) to 10 or more years (11, 8.3%). In terms of educational level, 2 (1.5%) finished general secondary school, 5 (3.8%) finished intermediate secondary school, 32 (24.2%) had obtained a high-school degree, 84 (63.6%) held a university degree, 6 (4.5%) went to university without graduating, and 3 (2.3%) received a doctoral level degree. One hundred eight (81.8%) participants had no management position, 16 (12.1%) had a position as a team leader at the operational level, 4 (3%) had a management position at the middle level, and 3 (2.3%) had a management position at the higher level. One participant did not provide any information.

### Measures and Analytic Procedure

Following translation/back-translation procedures, we had all the English items translated into German. Efforts were directed towards ensuring quality in the translation itself by a team approach. The work group for translation was composed of the three authors, who are fluent in both German and English and knowledgeable about the content of the scale and culture, and an additional person who is fluent in both German and English and did not have access to the original English version of the 5F-CVQ. Thus, the final German version of the 5F-CVQ met established criteria for translation (Geisinger, 1994; Wild et al., 2005). We applied a cross-sectional survey study and asked participants to reflect on their voice behavior in the past twelve months. We measured voice with our new 5F-CVQ. Response options ranged from 1 (strongly disagree) to 7 (strongly agree). Correlations among all manifest study variables and correlations among all latent factors are provided in the online supplement for “[Study 3](#)” in Table SM1 and Table SM2. Internal consistency reliability was 0.97 for IV, 0.97 for ASV, 0.95 for AFV, 0.92 for PSV, and 0.96 for PFV.

We conducted CFA using Mplus 8.8. We modeled each item as a manifest variable and did not combine items into parcels. The distributions tended to be quite skewed, violating the assumptions of the normal theory-based maximum likelihood estimation. The robust maximum likelihood estimator in Mplus was thus used as it provides correct standard errors and handles potential non-normal data (Muthén & Muthén, 2017).

We compared the theorized five-factor model to plausible alternatives, namely a model with (1) only one factor ignoring the dimensionality, (2) the two prominent promotive (innovative and adaptive ideas and solutions [IV + ASV + AFV]) and prohibitive factors (concerns [PSV + PFV]), (3) two functional factors (innovative opportunity focused voice [IV] and harm focused voice [ASV + AFV + PSV + PFV]), (4) three substantive factors (innovative [IV], adaptive [ASV + AFV], and problem focused [PSV + PFV] voice), (5) two temporal factors (status quo [ASV + PSV] and future focused voice [IV + AFV + PFV]), and (6) four factors in which the items from the two most highly intercorrelated facets ((6a) IV + ASV and (6b) AFV + ASV) load on the same factor (see Table SM2 in the online supplement for “Study 3”). In view of the fact that the alternative models and the focal five factor model are conceptually nested, we focus on Satorra-Bentler scaled  $\chi^2$ -difference tests for robust maximum likelihood estimation in the comparative interpretation to determine the best fitting model. Acceptable fit of individual models was tested using conventional standards for comparative fit index (CFI,  $\geq 0.9$ ), Tucker-Lewis Index (TLI,  $\geq 0.9$ ), root mean square error of approximation (RMSEA,  $< 0.08$ ), and standardized root mean square residual (SRMR,  $< 0.08$ ) (Marsh et al., 2005).

## Results and Discussion

The model fit results are shown in Table 4, with the focal model numbered #7. Our theorized model, which specified five facets of voice (Model 7), had an acceptable fit on all indices according to conventional standards (Marsh et al., 2005), and it fit the data significantly better than any of the alternative models. The  $\chi^2$  difference tests of the alternative Models (1–6) each indicated considerable and significant

increase in misfit of the models 1–6 compared to the measurement Model 7. A one-factor model (Model 1) fit the data worst according to all indices.

Further evidence for the hypothesized five-factor structure was provided by the standardized factor loadings, which ranged from 0.75 to 0.92, meaning that each item had a high and significant loading ( $p < 0.001$ ) on its respective voice factor (see Table SM3 in the online supplement for “Study 3”). However, the results indicated that correlations between the factors are substantial (range of latent correlations: 0.75 to 0.94; see Table SM2 in the online supplement for “Study 3”). The high correlations between the five factors are not entirely unexpected (c.f., Liang et al., 2012). Nonetheless, with regard to discriminant validity at the item level, it cannot be shown that the correlation between items representing the same factor is consistently stronger than the correlation between items assigned to different factors.

Regarding limitations, it is important to acknowledge that the sample used for “Study 3” primarily comprised university students who were concurrently employed. This demographic composition may have imposed constraints on personal, intellectual, or demographic characteristics that differ from the broader working population. Furthermore, it is noteworthy that the sample size in this study was relatively small, which can introduce biases in parameter estimates and estimated standard errors (Muthén & Muthén, 2002).

In summary, “Study 3” provided initial support for our measure of constructive voice in terms of fit of the intended five-facet structure, but also did show considerable covariation across factors. Therefore, we slightly resharpened the formulation of 12 items to more clearly delineate the items between facets. For adaptive voice items, we have now consistently introduced the word “adaptive” preceding the word “suggestions” or “proposals” (e.g., “I speak up with *adaptive* suggestions to existing problematic procedures or

**Table 4** “Study 3” confirmatory factor analysis model fit for the 5F-CVQ

Measurement model	$\chi^2$	DiffTest	df	AIC	BIC	TLI	CFI	SRMR	RMSEA
General factor model 1	936.92	325.54***	350	11272.32	11514.47	.79	.80	.064	.113
Promotive-prohibitive two-factor model 2	738.57	158.00***	349	11034.35	11279.39	.86	.87	.055	.092
Functional two-factor model 3	858.55	287.59***	349	11175.21	11420.25	.81	.83	.062	.105
Substantive three-factor model 4	711.32	150.67***	347	11001.31	11252.11	.87	.88	.052	.089
Temporal two-factor model 5	860.61	245.41***	349	11183.65	11428.69	.81	.83	.062	.105
Four-facet factor model 6a (IV + ASV loading on one factor)	563.51	23.33***	344	10827.34	11086.80	.91	.92	.046	.070
Four-facet factor model 6b (AFV + PFV loading on one factor)	623.59	77.43***	344	10899.69	11159.15	.90	.91	.049	.078
Five-facet factor model 7	537.66		340	10803.01	11073.99	.93	.93	.043	.066

$N = 132$ . Models allowed the factors to correlate.  $\chi^2$  = chi-squared-value; DiffTest = Satorra-Bentler scaled chi-square difference test for robust maximum likelihood estimation comparing each model to the hypothesized model 1;  $df$  degrees of freedom. *AIC* Akaike’s information criterion, *BIC* Bayesian information criterion, *TLI* Tucker-Lewis Index, *CFI* comparative fit index, *SRMR* standardized root mean square residual, *RMSEA* root mean square error of approximation. The following cut-off values to determine goodness of fit were applied: TLI values  $\geq .9$ , CFI values  $\geq .9$ , SRMR  $\leq .08$ , and RMSEA  $\leq .08$  (Marsh et al., 2005)

\*\*\* $p < .001$

processes.”). For the future focused voice items, we have included formulations as “in the future” (e.g., “I speak up with *impending* problems that could have a negative impact on the organization/my work group *in the near future*.”). Any changes to the wording of the items may be found in the online supplement for “Study 3” in Table SM4.

## Study 4

In “Study 4” we consider two types of information (Borsboom et al., 2004; Newton & Shaw, 2013), namely psychometric properties and links of the voice facets within a nomological net. In this preregistered<sup>2</sup> study, our primary aims were (1) to replicate the five-factorial structure of our final 5F-CVQ in an independent larger sample, and to assess whether the five facets of voice (2) show differential relationships to predictor trait variables, and (3) converge with other measures of constructive and other types of voice whereas still being distinguishable from those measures. The first objective listed relates to psychometric properties. Analogous to “Study 3,” we compare the fit of the hypothesized five-factor structure of the revised set of voice items to the fit of the other plausible measurement models. The second and third objective refer to the nomological network of the 5F-CVQ. Below, we first discuss how the five facets of voice may relate differentially to predictor variables, aside from relations expected to be common across voice facets. We then outline expected empirical associations between the five facets of voice and established measures of constructive and other types of voice. In doing so, we also aim at calibrating the appropriate level of specificity vs. parsimony in measuring voice in person focused approaches.

## Nomological Network

### Antecedents of Voice

We identified three antecedents closely tied to the three dimensions of temporal, functional, and substantive, orientation. Specifically, (1) *temporal focus* corresponds to temporal orientation of the voice message. Temporal focus characterizes an individual’s innate inclination towards the past, present, or future (Bluedorn, 2001; Shipp & Aeon, 2019; Shipp et al., 2009), guiding the incorporation of perceptions from past events, current circumstances, and future expectations into behaviors. We hypothesize that voice facets focused towards the status quo will exhibit a stronger correlation with present orientation than those focused towards the future. A

future oriented temporal focus is likely to divert employees’ attention away from existing issues and challenges. Furthermore, we postulate that a focus on the past may exhibit a negative correlation with all five facets of voice. Given that all voice facets involve efforts to alter existing situations, a pronounced focus on the past could potentially clash with this proactive orientation. (2) *Regulatory focus* distinguishes between promotion and prevention focus, which were shown to have differential relationships to specific behaviors (Gamache et al., 2015; Higgins, 1997; Johnson et al., 2010; Lin & Johnson, 2015). This distinction pertains to functional voice behavior. Consistent with prior empirical research relating promotion and prevention focus to promotive and prohibitive voice (Lin & Johnson, 2015), we expect promotion focus to relate primarily to innovative voice, as both constructs share an orientation towards ideal states and thus imply compatible goals and strategies. By contrast, prevention focus and functionally harm focused voice share an orientation towards eliminating or avoiding undesirable states and thus should be positively associated. (3) *Innovativeness* (innovative and adaptive attitudes) is expected to correspond to the substantive voice dimension. According to Kirton (1976), employees with adaptive attitudes seek minor improvements that are close to existing organizational practices, and push the boundaries incrementally, whereas innovators give free reign to their creativity and tend to do things a previously unknown way. These two mindsets correspond to adaptive and to innovative forms of voice behavior, respectively. As an additional antecedent, we selected (4) *psychological safety* as this construct has generated conflicting findings in previous voice research. While Liang et al. (2012) emphasized that psychological safety was only significantly associated with change in prohibitive voice ( $\gamma = 0.19$ ), the findings from Köllner et al. (2019) indicated a lack of significant differentiation between the two promotive-prohibitive voice facets. We posit that voice emphasizing harm focused functions (i.e., problem focused voice and also adaptive voice as a subset of promotive voice) may involve increased perceptions of personal risk. Therefore, we hypothesize that facets of voice functionally centered on harm will demonstrate a more robust correlation with psychological safety compared to voice functionally focused on innovative opportunities.

### Constructive and Non-Constructive Voice

Differential relationships of our newly proposed five voice facets were further explored with regard to established measures of constructive voice (Liang et al., 2012; Maynes & Podsakoff, 2014) and to other types of employee communication (Maynes & Podsakoff, 2014). The *promotive*

<sup>2</sup> For reasons of common scope, we limit this manuscript to a subset of the preregistered research questions as we consider them essential.

*voice scale* (Liang et al., 2012) revolves around voicing ideas and solutions that drive constructive enhancements. As a result, we anticipate a more pronounced correlation between this particular subscale and the innovative and adaptive voice scales, rather than the problem focused voice scales. In contrast, we predict a stronger association between the *prohibitive voice scale* and our problem focused voice scales, as opposed to the innovative and adaptive ones. Among the subscales of Maynes and Podsakoff's (2014) measure, *constructive voice* is hypothesized to exhibit the most robust correlation with our five facets, because the three remaining scales—*supportive*, *destructive*, and *defensive*—were all outlined to refer to non-constructive forms of voice distal to the present operationalization.

### Sample and Analytic Procedure

Participants were contacted by the survey company *Respondi*, an ISO-certified panel provider of digital online data that supports a wide range of research projects (e.g., Bach et al., 2021; Munzert et al., 2021; Sandner et al., 2021). Potential participants were first screened through some questions that ensure that the final sample is nationally representative on gender, age, and income. Additionally, participants who indicated in a screening question that they were not employed or had an employment relationship of less than 12 months or/and worked less than 30 h per week were excluded from the sample. Participants were paid for full completion of the survey. Twenty-eight cases were deleted due to extremely short participation times of less than two minutes combined with systematic response patterns across multiple items or large portions of missing data. There were very little missing data in the remaining sample (only one item was missed by more than one participant in total). The final sample size was  $N=553$ . Demographic variables were distributed as follows: gender: men (46.8%), women (53.0%) (one person did not indicate gender); age: 21–29 years (10.1%), 30–39 years (27.8%), 40–49 years (20.8%), 50–59 years (30.6%), 60 years or older (10.7%); educational level: basic secondary school (5.4%), intermediate secondary school (26.2%), advanced secondary school (27.5%), university degree (23.3%), doctoral degree (10.7%) (6.9% did not provide information on education level). Work-related variables were distributed as follows: job tenure: 1–2 years (7.8%), 2–5 years (22.2%), 5–10 years (23.5%), 10 years or more (46.5%); hours per week: 30–39 h per week (43.0%), 40 or more (57.0%); primary place of work: on-site (73.8%), telework (24.2%) (2.0% did not provide that information); and hierarchical position: higher-level management (15.2%), mid-level management (10.7%), team leader (19.7%), and entry level (54.4%).

### Measures and Analytic Procedure

We asked participants to reflect on their work-related behavior and their attitudes and perceptions in the past twelve months. The responses ranged from 1 (strongly disagree) to 7 (strongly agree) for all variables. All scales for which there was no German language version available (i.e., alternative voice scales, innovativeness, psychological safety) were translated following translation/back-translation procedures. The procedure was the same as for the translation of the 5F-CVQ (the German versions of the scales are available in the online supplement). Latent correlations among all study variables (factors) (see Table SM1 in the online supplement for “Study 4”) and among all manifest items are provided in the online supplement. Each of the variables is described below.

#### Constructive and Non-constructive Voice

We measured voice behavior with our final 5F-CVQ. The final survey instrument 5F-CVQ contains 28 items and is shown in the Appendix (the German original is available in the online supplement). Internal consistency reliability was 0.96 for IV, 0.96 for ASV, 0.96 for AFV, 0.92 for PSV, and 0.96 for PFV. Furthermore, for testing discriminant and convergent validity we measured voice behavior with the 5-item promotive and 5-item prohibitive voice scales by Liang et al. (2012), and the four facets of voice by Maynes and Podsakoff (2014) with five items for each subscale. Internal consistency reliability for the voice scales by Maynes and Podsakoff were 0.96 for constructive voice, 0.97 for supportive voice, 0.96 for destructive voice, 0.97 for defensive voice. Internal consistency reliability for the voice scales from Liang et al. was 0.95 for promotive voice, and 0.92 for prohibitive voice.

#### Temporal Focus

Temporal focus was measured with 13 items (four items for past orientation, four items for present orientation and five items for future orientation) by a German version (Geiger et al., 2018) of the Temporal Focus Scale (TFS) by Shipp (2009). Internal consistency reliability was 0.92 for past focus, 0.91 for present focus, and 0.92 for future focus.

#### Regulatory Focus

Regulatory Focus was assessed with 24 items from the German Regulatory Focus Scale (Büttner, 2012), with 12 items each assessing promotive focus and prevention focus. Internal consistency reliability was 0.89 for prevention focus, and 0.88 for promotion focus.



## Innovativeness

Innovativeness was assessed with a shorter version of the Kirton-Adaptation-Innovation-Inventory (KAI) (Bobic et al., 1999) with nine items for innovation orientation and nine items for adaptation orientation. Internal consistency reliability was 0.86 for innovation orientation, and 0.87 for adaptation orientation.

## Psychological Safety

Psychological safety was measured with four items (Liang et al., 2012) as the extent to which an individual perceived it to be safe to express himself or herself at work. Internal consistency reliability was 0.92.

We applied the same analytic strategy as in the “Study 3” to examine the factorial structure of the 5F-CVQ. We focused on correlations among latent factor scores to examine differential associations between 5F-CVQ facets and the other variables. We tested the difference of two correlation coefficients that share one variable in common using three different approaches (Diedenhofen & Musch, 2015). We tested the difference according to Steiger (1980), Eid et al. (2011) and Zou (2007). Thus, we used tests based on significance testing and based on the computation of confidence intervals in comparing correlations.

## Results

### Psychometric Evaluation

We compared the six alternative models presented earlier with the hypothesized five-factor model using the same fit indices as in “Study 3.” The model fit results are shown in Table 5. Our theorized model, which specified five facets of voice

(Model 7), had again an acceptable fit to the data according to all indices, and a significantly better fit than all alternative Models (1–6). Further supporting the hypothesized five-factor structure, each item had a large standardized factor loading on its respective factor ranging from 0.76 to 0.94 (see Table SM3 in the online supplement for “Study 4”). The correlations among the 5F-CVQ-factors were not as high as in “Study 3,” although still substantial (range of  $r = 0.66$ – $0.87$ ; see Table SM1 in the online supplement for “Study 4”). Similar results were observed for the alternative promotive/prohibitive voice scale (Liang et al., 2012) ( $r = 0.78$ , see Table SM1 in the online supplement for “Study 4”). With regard to discriminant validity at the item level, correlations between items representing the same factor are stronger than the correlations between items assigned to different factors. The discriminant validity between the voice facets at the construct level can be determined according to the Fornell-Larcker (1981) criterion. In line with this criterion, each construct facet average variance extracted (AVE) is greater than the squared inter-factor correlations for all other facets (see Table SM4 in the online supplement for “Study 4”).

### Nomological Network

The relationships between personality and perception variables and voice facets were assessed by latent factor zero-order correlations. The goal was to examine nomological validity and to explore differences in the pattern of correlations across the five facets of voice. The respective coefficients are shown in Table 6. As the 5F-CVQ is set out as a refinement of Liang et al.’s (2012) model and measures, we also report comparable results for the latter scales in Table 7.

**Antecedents of Voice** Regarding the *temporal focus* variables, we found little support for expected differential

**Table 5** “Study 4” confirmatory factor analysis model fit for the 5F-CVQ

Model	$\chi^2$	DiffTest	<i>df</i>	AIC	BIC	TLI	CFI	SRMR	RMSEA
General factor model 1	2689.83	1104.19***	350	44739.69	45102.18	.75	.77	.071	.110
Promotive-prohibitive two-factor model 2	1815.95	655.87***	349	43129.26	43496.06	.84	.85	.049	.087
Functional two-factor model 3	2098.235	956.04***	349	43612.99	43979.81	.81	.83	.061	.095
Substantive three-factor model 4	1309.17	467.22***	347	42178.09	42553.53	.89	.90	.040	.071
Temporal two-factor model 5	2451.56	985.59***	349	44291.11	44657.91	.77	.79	.075	.104
Four-facet factor model 6 (AFV + ASV loading on one factor)	1040.79	308.01***	344	41688.77	42077.16	.92	.93	.034	.061
Five-facet factor model 7	577.20		340	40858.92	41264.56	.97	.97	.028	.036

$N = 553$ . Models allowed the factors to correlate.  $\chi^2$  = chi-squared-value; DiffTest = Satorra-Bentler scaled chi-square difference test for robust maximum likelihood estimation comparing each model to the hypothesized model 6; *df* degrees of freedom, *AIC* Akaike’s information criterion, *BIC* Bayesian information criterion, *TLI* Tucker-Lewis Index, *CFI* comparative fit index, *SRMR* standardized root mean square residual, *RMSEA* root mean square error of approximation. The following cut-off values to determine goodness of fit were applied: TLI values  $\geq .9$ , CFI values  $\geq .9$ , SRMR  $\leq .08$ , and RMSEA  $\leq .08$  (Marsh et al., 2005)

\*\*\* $p < .001$



**Table 6** Correlations between nomological network variables and five facets of voice

	IV	ASV	AFV	PSV	PFV
<b>Temporal focus</b>					
Employee's present focus	.34 <sup>a</sup>	.41 <sup>b</sup>	.32 <sup>a</sup>	.32 <sup>a</sup>	.32 <sup>a</sup>
Employee's future focus	.32 <sup>a</sup>	.37 <sup>b</sup>	.36 <sup>a,b</sup>	.31 <sup>a,b</sup>	.31 <sup>a</sup>
Employee's past focus	.16 <sup>a</sup>	.19 <sup>a</sup>	.19 <sup>a</sup>	.19 <sup>a</sup>	.20 <sup>a</sup>
<b>Regulatory focus</b>					
Employee's promotion focus	.53 <sup>c</sup>	.48 <sup>b</sup>	.52 <sup>c</sup>	.46 <sup>a,b</sup>	.43 <sup>a</sup>
Employee's prevention focus	.30 <sup>a</sup>	.35 <sup>b</sup>	.35 <sup>b</sup>	.41 <sup>c</sup>	.37 <sup>b,c</sup>
<b>Innovative-/adaptive-attitude</b>					
Innovative employee attitude	.46 <sup>a</sup>	.43 <sup>a</sup>	.46 <sup>a</sup>	.41 <sup>a</sup>	.41 <sup>a</sup>
Adaptive employee attitude	.41 <sup>a</sup>	.49 <sup>b</sup>	.47 <sup>b</sup>	.49 <sup>b</sup>	.51 <sup>b</sup>
Psychological safety	.42 <sup>a</sup>	.47 <sup>b</sup>	.46 <sup>b</sup>	.46 <sup>b</sup>	.47 <sup>b</sup>
<b>Other voice scales</b>					
Promotive Voice	.82 <sup>b</sup>	.83 <sup>b</sup>	.84 <sup>b</sup>	.73 <sup>a</sup>	.81 <sup>b</sup>
Prohibitive Voice	.67 <sup>a</sup>	.67 <sup>a</sup>	.70 <sup>a</sup>	.81 <sup>c</sup>	.75 <sup>b</sup>
Supportive voice	.69 <sup>a,b</sup>	.74 <sup>c</sup>	.70 <sup>b</sup>	.65 <sup>a</sup>	.70 <sup>a,b</sup>
Constructive voice	.79 <sup>a,b</sup>	.82 <sup>c</sup>	.81 <sup>b,c</sup>	.76 <sup>a</sup>	.79 <sup>b</sup>
Defensive voice	.07 <sup>b</sup>	-.00 <sup>a</sup>	.12 <sup>c,d</sup>	.15 <sup>d</sup>	.10 <sup>b,c</sup>
Destructive voice	.04 <sup>b</sup>	-.03 <sup>a</sup>	.06 <sup>b</sup>	.10 <sup>b</sup>	.05 <sup>b</sup>

$N=553$ . *IV* innovation focused voice, *ASV* adaptive and status quo focused voice, *AFV* adaptive and future focused voice, *PSV* problem and status quo focused voice, *PFV* problem and future focused voice. Latent factor zero-order correlations within a row sharing a superscript are not significantly different; coefficients within a row with different superscripts are significantly different ( $p < .05$ ) (calculation according to Steiger, 1980; Eid et al., 2011, and Zou, 2007; results do not differ)

Correlations with absolute values  $> .12$  are significant at  $p < .001$

Correlations with absolute values  $> .09$  are significant at  $p < .05$

relations to 5F-CVQ facets (except that ASV correlated higher with present focus than remaining voice facets). All five facets of voice demonstrated significant and positive associations with an employee's present focus, with correlations ranging from  $r = 0.32$  to  $r = 0.42$ . Similarly, these facets displayed positive relationships with an employee's future focus, ranging from  $r = 0.31$  to  $r = 0.37$ . Although we did not find the expected negative correlations of voice with temporal past focus, at least all five voice facets exhibited weaker positive relationships ( $r \leq 0.20$ ) compared to the associations observed with both employee's temporal present and future focus. By contrast, whereas Liang et al.'s promotive and prohibitive subscales showed statistically indifferent correlations with present focus ( $r = 0.34/0.38$ ), the former scale correlated more highly with both future ( $r = 0.33$  vs.  $0.25$ ) and—surprisingly—with past ( $r = 0.24$  vs.  $0.15$ ) orientation.

Regarding *regulatory focus*, again all five facets of voice demonstrated significant and positive correlations to an employee's promotion focus, with correlations spanning from

**Table 7** Correlations between nomological network variables and promotive and prohibitive voice

	Promotive voice	Prohibitive voice
<b>Temporal focus</b>		
Employee's present focus	.34 <sup>a</sup>	.38 <sup>a</sup>
Employee's future focus	.33 <sup>b</sup>	.25 <sup>a</sup>
Employee's past focus	.24 <sup>b</sup>	.15 <sup>a</sup>
<b>Regulatory focus</b>		
Employee's promotion focus	.54 <sup>a</sup>	.53 <sup>a</sup>
Employee's prevention focus	.36 <sup>a</sup>	.31 <sup>a</sup>
<b>Innovative-/adaptive-attitude</b>		
Innovative employee attitude	.50 <sup>a</sup>	.52 <sup>a</sup>
Adaptive employee attitude	.47 <sup>a</sup>	.45 <sup>a</sup>
Psychological safety	.46 <sup>a</sup>	.51 <sup>b</sup>
<b>Other voice scales</b>		
Promotive voice		.78
Prohibitive voice	.78	
Supportive voice	.76 <sup>a</sup>	.74 <sup>a</sup>
Constructive voice	.88 <sup>b</sup>	.76 <sup>a</sup>
Defensive voice	.13 <sup>a</sup>	.19 <sup>b</sup>
Destructive voice	.08 <sup>a</sup>	.15 <sup>b</sup>

$N=553$ . Latent factor zero-order correlations within a row sharing a superscript are not significantly different; coefficients within a row with different superscripts are significantly different ( $p < .05$ ) (calculation according to Steiger, 1980; Eid et al., 2011, and Zou, 2007; results do not differ)

Correlations with absolute values  $> .12$  are significant at  $p < .001$

Correlations with absolute values  $> .09$  are significant at  $p < .05$

$r = 0.43$  to  $r = 0.53$ . Similarly, these facets exhibited positive relationships with prevention focus, ranging from  $r = 0.30$  to  $r = 0.41$ . While we initially hypothesized that functionally innovative voice would display a stronger correlation with promotion focus compared to functionally harm focused voice, our findings did not entirely support this expectation. Instead, our analysis did highlight the significance of temporal orientation towards the future in shaping these distinct associations. Specifically, the facets IV (functionally innovative) and AFV (functionally harm focused), which both are future focused, displayed the most substantial correlations with an employee's promotion focus ( $r = 0.53$  and  $r = 0.52$ , respectively), significantly surpassing the strength of correlations observed for all status quo focused voice facets. Largely in line with expectations, problem focused voice exhibited stronger ( $r = 0.37$ – $0.41$ ) correlations with prevention focus than innovative voice ( $r = 0.30$ ), whereas adaptive facets lay somewhat in between ( $r = 0.35$ ). In contrast to those nuanced findings on the 5F-CVQ, Liang et al.'s (2012) subscales did not show any differential relations to regulatory foci.

Turning to *innovativeness*, both innovative and adaptive employee attitudes displayed significant and positive correlations with all facets of voice messages, with correlations ranging from  $r=0.41$  to  $r=0.46$  for innovative attitude and from  $r=0.41$  to  $r=0.51$  for adaptive attitude. Contrary to our assumptions, there were no significant differences in the correlation patterns between the five facets of voice and innovative attitude. By contrast, adaptive employee attitude exhibited the expected significantly weaker relationship with innovative voice ( $r=0.41$ ) in comparison to the correlations observed with the other voice facets. By comparison, we found no statistically significant differences at all for the Liang et al. (2012) subscales in relation to either innovative or adaptive attitudes.

Finally, all voice facets demonstrated significant and positive associations with *psychological safety*. Yet, differential relations offered some support for our expectations. The facet functionally characterized by innovation showcased the weakest correlation with psychological safety ( $r=0.42$ ), whereas correlations with functionally harm focused voice facets were consistently higher ( $r=0.46$ – $0.47$ ). The latter finding aligns with the observation that Liang et al.'s (2012) prohibitive scale correlates slightly more strongly with psychological safety than their promotive subscale ( $r=0.51$  vs.  $0.46$ ).

**Constructive and Non-constructive Voice** As expected, the association of *prohibitive voice* (Liang et al., 2012) with PSV was notably stronger ( $r=0.81$ ) compared to its correlations with the substantively proposal focused facets IV, ASV, and AFV (range from  $r=0.67$  to  $r=0.70$ ). However, we did not find expected differences in correlations between substantively proposal focused facets and problem and future focused voice. In line with our hypothesis, *promotive voice* (Liang et al., 2012) displayed a significantly weaker relationship with problem and status quo focused voice ( $r=0.73$ ) compared to its correlations with innovative voice facet and adaptive voice facets (range from  $r=0.82$  to  $r=0.84$ ).

Shifting focus to the voice scales introduced by Maynes and Podsakoff (2014), all facets of the 5F-CVQ displayed their highest correlations (ranging from  $r=0.76$  to  $0.82$ ) with the *constructive voice subscale*, surpassing the other three subscales of *defensive*, *destructive*, and *supportive voice* as defined by Maynes and Podsakoff (2014). These results were in line with our expectations. Furthermore, Maynes and Podsakoff's constructive voice scale showed the most pronounced associations with the proposal focused voice facets (IV, AFV, and ASV) of the 5F-CVQ. Correlations with defensive and destructive voice were weakest and practically zero for ASV, whereas relations of remaining 5F-CVQ facets to those two non-constructive forms of voice hovered around  $r=0.10$ . Of Liang et al.'s (2012) subscales,

promotive voice correlated extremely high and more strongly than prohibitive voice with Maynes and Podsakoff's (2014) constructive voice ( $r=0.88$  vs.  $0.76$ ), whereas prohibitive voice displayed stronger correlations than promotive voice with the defensive ( $r=0.19$  vs.  $0.13$ ) and destructive ( $r=0.15$  vs.  $0.08$ ) facets of the latter instrument.

## Discussion

The present “Study 4” was set out to examine the internal structure of the 5F-CVQ, thereby replicating “Study 3” with a slightly revised version of the instrument, and to test its convergent and discriminant validity within a nomological net of selected antecedents and alternative measures of voice behavior. Below, we discuss our findings in separate subsections referring to those objectives.

### Internal Structure of the 5F-CVQ

Replicating “Study 3,” CFA results supported the validity of the intended multi-dimensional structure. In fact, differences in fit between our theoretical model and plausible alternatives were even more pronounced than in the previous study. Moreover, inter-scale correlations were no longer as excessively high, and even slightly lower than those between constructive subscales of alternative voice measures (see Table SM1 in the online supplement for “Study 4” and Table 7). Individual items now fit target subscales better than in the previous sample. Overall, the revised 5F-CVQ's psychometric properties in terms of reliability and structural fit appear adequate at both the subscale and the item level.

### Antecedents of Facets of Voice

Overall, the results showed that the differences in correlations in “Study 4” between the facets of voice and predictor variables are rather small, although in some cases significant differences emerge as expected. Recall that all theoretical assumptions were pre-registered, ensuring a balanced evaluation of support. Given the substantial inter-correlations between our, as well as alternative, facet measures of constructive voice, it is not surprising that observed communalities are more pronounced than differences within the nomological net. The inherent proactive nature shared by all constructive voice facets (Parker & Collins, 2010) might have contributed to the lack of large differences in their relationships.

Notably, differential relations of 5F-CVQ facets to outside variables tended to be more prevalent than for Liang et al.'s (2012) subscales operationalizing their popular dichotomous distinction. For example, the latter subscales did not differentiate at all with regard to regulatory foci or innovativeness,

whereas 5F-CVQ facets displayed limited, but at least partial evidence of discriminant validity in relation to those variables. Where promotive and prohibitive voice showed differential relations, these were either similar to those observed for 5F-CVQ facets (psychological safety) or hard to make sense of (the analogous pattern observed for both past and future temporal focus). Moreover, relations of the 5F-CVQ to Maynes and Podsakoff's (2014) measure of constructive and non-constructive voice facets tended to be somewhat clearer than those of Liang et al.'s scales. Although differences between 5F-CVQ facets and the more established Liang et al. scales were generally small, when observed, they consistently pointed to superior construct validity of our new measure.

Turning to findings on the 5F-CVQ beyond the facet level, most observed correlations were in line with expectations. As expected, all 5F-CVQ subscales correlated positively with temporal foci on presence and future, with regulatory foci on promotion and prevention, with positive attitudes towards innovation and adaptation, and with psychological safety. These relations support the assumption that all facets of the 5F-CVQ share the feature of constructiveness. Unexpectedly, though, we also found moderately positive relations of all voice facets to temporal past focus. One explanation for these positive correlations may be that voice shares with the content of the temporal past focus scale (e.g., "I reflect upon what has happened in my life") a tendency to ruminate about significant issues in one's situation, which may have compensated for the lack of proactivity in the latter measure.

At the facet level of the 5F-CVQ, we found that these voice facets hardly differ with respect to temporal focus. Morrison's (2014) voice framework suggests that a core set of variables serve to foster a general drive to speak up in various constructive ways. This general idea appears to be supported by our findings. One explanation could also be that temporal focus is a disposition too generic to affect behaviors that respond to specific job characteristics and events in the job context.

More support was found for differential relations of 5F-CVQ facets to a number of other common correlates of constructive voice. First, correlations with the regulatory foci underscore the significance of our temporal voice dimension for promotion focus, and of the functional dimension prevention focus. The simple dichotomy of promotive/prohibitive voice may obscure such differences by confounding underlying dimensions, as is apparent from 5F-CVQ patterns with Liang et al.'s (2012) scales and the lack of discriminatory power for the latter instrument. Second, correlations of adaptive, but not of innovative, attitudes with innovative and adaptive facets of the 5F-CVQ followed the expected pattern. We have no straightforward explanation for the lack of support for expected differences to innovative

attitudes. The slightly, though insignificantly, higher correlations with future focused voice observed may indicate that the effect is simply too small to be detected with present sample size. With regard to psychological safety, we found support for the expected effect of our functional dimension, whereas no significant disparities emerged between problem focused facets and adaptive facets (both are functionally harm focused voice). Liang et al. (2012) emphasized the impact of psychological safety exclusively on prohibitive voice, whereas Köllner et al.'s (2019) findings indicated a lack of differential effects on promotive and prohibitive voice. The present findings may help to explain this apparent inconsistency, as the innovative facet we found to stand out shares its functional focus with much of Liang et al.'s (2012) promotive scale, whereas Köllner et al.'s (2019) promotive voice scenario emphasized functional harm mitigation. Our findings point to the conclusion that the functional dimension distinguishing IV from remaining facets may explain observed differences. This way, the clearer distinction of three dimensions refined model again adds to the extant literature.

### Constructive Voice and Other Voice Measures

The findings regarding the relationships between the 5F-CVQ scales and other voice-related measurements largely aligned with our theoretical expectations. Correlations with Liang et al.'s (2012) promotive and prohibitive voice measures, as well as with the constructive scale introduced by Maynes and Podsakoff (2014), were all positive and substantial. Still, compared to the correlations among those previous measures ( $r_s = 0.76$  to  $0.88$ ), especially the differences observed for the substantive dimension of the 5F-CVQ (proposal vs. no proposal) in distinguishing promotive from prohibitive voice (all 5F-CVQ facets including a proposal correlate at  $\geq 0.82$  with promotive, and  $\leq 0.70$  with prohibitive voice) at the same time point to a greater potential of discriminant validity for the 5F-CVQ facets. Conversely, relations of 5F-CVQ facets to non-constructive voice were consistently insubstantial and especially so when including a proposal (5 out of 6  $r_s < 0.10$ ).

### General Discussion

Drawing on a taxonomy describing constructive voice messages along three dimensions (functional, substantive, and temporal) that lead to specific facets of voice, we set out to clarify the conceptualization and remove confounds in the measurement of constructive voice. "Study 1" findings revealed that established items often lacked clarity concerning their defining dimensions, and that existing voice scales failed to capture potentially crucial combinations of these

dimensions. Accordingly, our primary focus in the subsequent studies was to develop a comprehensive self-report measure of voice that unequivocally encompassed all meaningful configurations of the defining dimensions. “Study 2” lent support to the content validity of the measure, while “Study 3” and “Study 4” offered preliminary evidence of psychometric properties and of the construct validity of the 5F-CVQ. Taken together, our findings provide initial support for our model describing constructive voice in terms of five facets based on meaningful configurations of three dimensions underlying each voice message: its functional, substantive, and temporal orientation. The items within the 5F-CVQ have been carefully crafted to assure both theoretical and practical significance of their content.

Turning to less supportive results, findings especially from “Study 4” certainly also revealed that the discriminatory power of resultant scales remains somewhat limited. Intersubscale correlations are substantial, and patterns of correlations with outside variables tend to be similar across subscales. Some additional evidence addressing this issue comes from a longitudinal study with six measurement waves using a shortened version of the 5F-CVQ. Main objectives of that study lie beyond the scope of the present paper, but we may report in a cursory fashion some results on discriminant validity (see online supplement for the additional study for more details). The latent correlations within individuals for the five voice facets ranged from  $\rho = 0.41$  to  $\rho = 0.88$  (mean  $\rho = 0.63$ ), indicating considerably better discrimination within than between persons. By far the single highest correlation was observed between adaptive facets (ASV and AFV), pointing to a need to further scrutinize the theoretical value specifically of the temporal dimension within our model. Much lower correlations were observed between facets that diverge in their functional (IV vs. all others) and substantive (PSV/PFV vs. IV/ASV/AFV) orientations (all  $\rho$ s < .70). Taken together, our results converge in the conclusion that the 5F-CVQ may reveal meaningful differences that are inevitably confounded with a simple distinction of promotive vs. prohibitive voice.

From a pragmatic perspective, findings on discriminant validity suggest that, although a general factor model was rejected, measuring the common core of constructive voice may be a defensible strategy, as long as a more nuanced understanding is not the primary focus of research (e.g., in studies measuring voice as moderator or control variable). The 5F-CVQ may unfold its strengths particularly through its ability to provide a more nuanced separation of underlying dimensions than is currently possible, as is most evident in comparison to Liang et al.’s (2012) widely used measure. This becomes especially relevant in situation focused approaches aimed at comprehending intricacies unique to certain voice facets, thereby contributing to a more comprehensive understanding of the phenomenon.

Differential results shed light on several distinct relationships. One such instance involves the differentiation between innovative and adaptive voice, both of which encompass the promotive aspect of voice (Liang et al., 2012). In the present model, respective facets share the same substantive orientation but differ on the functional dimension. This differentiation proved meaningful in relation to several antecedents. Compared to the harm focused adaptive facets, IV, which refers to a functional orientation towards taking innovative opportunities, distinguished more clearly between promotion and prevention regulatory foci, showed the opposite pattern of correlations with innovative and adaptive attitudes, and weaker association with psychological safety. All these nuanced differences appear to make intuitive sense. The simple dichotomy of promotive vs. prohibitive would have concealed these differences, as it relates primarily to the substantive rather than the functional dimension according to our findings. Differential patterns distinguishing 5F-CVQ facets on substantive and temporal dimensions were less consistent and easily interpretable, which calls for additional research scrutinizing the relative merits of dimensions underlying constructive voice (see section on future research below).

## Strengths and Limitations

Our overall strategy of successively building on conceptual arguments, independent expert ratings based on those arguments, and quantitative tests of the suggested typology and pre-registration, may be considered major strengths of the present research. Furthermore, both the initial expert and employee studies were independently replicated.

One limitation of our study lies in our inability to demonstrate evidence of criterion-related validity. The present focus was on providing evidence of content validity, of the intended internal structure, and of construct validity in terms of a nomological net of alternative measures and antecedents of voice behavior. We consider this a reasonable package for an initial validation effort. Criterion-related validity, in terms of relations to job performance criteria, would not come without interpretational difficulty, as voice itself may be considered a facet of job performance (cf. Chamberlin et al., 2017). In the additional longitudinal study mentioned, we also included various types of perceived recipient reactions. Problem focused facets of voice yielded somewhat less positive (perceived appreciation and impact) and more negative (perceived rejection) reactions than innovative and adaptive facets in this study (see online supplement for the additional study), thereby lending initial support to the criterion-related validity of the 5F-CVQ, as related to reactions often investigated in extant voice literature (Liang et al., 2012; Morrison, 2014; Weiss & Zacher, 2022).



Addressing limitations of “Study 3,” the sample for “Study 3” consisted primarily of university students holding employment, which could have constrained personal, intellectual, or demographic characteristics in relation to the general working population. In addition, the sample size was small, which may lead to biases in parameter estimates and estimated standard errors (cf. Muthén & Muthén, 2002). These concerns are lessened by the fact that participants in “Study 4” were much more diverse demographically and that results replicated across studies. In addition, the CFAs conducted for each of the six time points in the additional longitudinal study demonstrated also a favorable model fit (see online supplement for the additional study). A particular strength of “Study 4” was that it used a broad probability sample of employed adults, which was also large enough to provide adequate statistical power to detect effects and accurately estimate effect sizes (Ioannidis, 2008; Muthén & Muthén, 2002). A limitation of both employee studies was that we collected data from a single source in a cross-sectional design. The former issue may lead to common method variance and thus inflated correlations, yet these inflations should have led to overrating commonality in voice at the expense of specificity, which implies that our conclusions on differential relations tend to be conservative rather than overly liberal. The differential pattern of observed correlations within the nomological net is inconsistent with appreciable method bias. The latter issue prevents us from drawing causal conclusions. However, the focus of our studies was on exploring the structure and nomological net of the new instrument, not on testing causal models, which renders our design adequate for the purpose at hand.

### Avenues for Future Research

Future research should delve deeper into determining the circumstances under which it becomes imperative to distinguish between various facets of voice. In a related vein, there is a need to enhance our understanding of the construct's underlying structure. The nature of the general constructive voice construct, as a coherent latent construct or as an umbrella term denoting a composite of different behavioral domains, is unclear. To the best of our knowledge no research paper on voice explicitly specifies whether they assume a formative or a reflective overall construct. The high correlations between the five facets of voice would correspond to an understanding of voice as reflective construct in which some general latent factor underlies all acts of voice. By contrast, the view of general voice as a formative construct (i.e., a collection of voice facets) would propose that its constituent facets should be studied independently, whereas the position that all acts of voice are, at least partially, driven by a common cause would imply we should focus primarily on the common element in both research

and practice aimed at constructive voice. As we did not distinguish between different settings in our research, one avenue for future studies is examining to what extent features of the work situation affect voice facets differently. These approaches appear particularly useful for explaining what is common or unique across voice facets. With a questionnaire that covers the whole construct sphere in a content-valid way with five facet scales, a whole range of structural models above and beyond the measurement model could be tested and related to outside variables.

Employees may engage in various facets of voice sequentially over time. Future research may thus apply the 5F-CVQ to capture voice in shortitudinal studies with time lags of few weeks or even a few days to capture specific voice episodes as they occur over time. Extending this argument for repeated measurement to longer time frames, longitudinal research might capture voice episodes from the cognitive appearance of the information, to voicing the information, to consequences of behavior, thereby permitting clearer causal inferences.

In a similar vein, the ambiguity persists regarding why and under what circumstances employee voice yields either positive or negative repercussions on employee well-being. We posit that these disparities arise due to inherent imprecisions and confounding factors inherent in the methodologies employed to investigate voice dynamics, with a specific emphasis on the content of the voice messages. Consequently, we advocate for future research endeavors to thoroughly explore distinct facets of voice and their divergent implications, given the paramount importance of employee well-being for organizational success. In this regard, we want to point out that differential links of the different voice messages to individual well-being related outcomes are completely lacking. Rather than the broad distinction between promotive and prohibitive voice, the three-dimensional classification of the message content without conceptual confounds could help to shed light on the theoretically double-edged effects of voice on individual well-being (Cangiano & Parker, 2016).

Further, we encourage future research to empirically test our theoretical assumption of excluding three of eight possible configurations of the three dimensions. Following the logic of the Critical Incident Technique (Flanagan, 1954), participants may first be asked to recall and describe a situation where they voiced in the last month. Such inductive approach may not just reveal whether configurations we deemed theoretically irrelevant may actually occur in practice, but might even lead to the discovery of forms of voice that had been overlooked in the hitherto primarily theory-driven voice literature. On a related note, other variations of such a “back to the roots” approach may explore the potential of previously overlooked theories (e.g., the distinction of sender, channel, and receiver, in general communication



models) for understanding the causes and consequences of voice behavior.<sup>3</sup>

Given our evaluation of prior measures of voice, it is possible that measurement contamination in prior measures may have affected past research involving differentiated voice and the relations to predictor and outcome variables. Morrison (2014) argued in her review that our understanding of voice could be deepened by considering characteristics of the message beyond just the promotive-prohibitive distinction, such as message urgency. Different configurations of functional, substantive, and temporal orientation may be crucial in that respect. For example, Burris et al. (2017) stressed the significance of “initiating change” as a critical facet of voice value. Change might be accentuated differently when a situation involves either existing harm or some future potential. In a related vein, Brykman and Raver (2021) introduced novelty as a dimension of voice quality. This dimension could align with our functional orientation towards innovative opportunities, our temporal dimension of future, or possibly both. As a final example, in innovation research, the expression of “voicing innovative ideas” stands out as a pivotal factor (e.g., Pundt et al., 2010). Whereas existing measures of constructive voice lack the means to distinguish proposed ideas along functional and temporal dimensions, facets of the 5F-CVQ allow for such distinctions that may prove meaningful for the production and communication of innovative ideas.

Finally, we encourage researchers to pay attention to the impact of formal and collective voice mechanisms (Wilkinson et al., 2020) as well as the contextual framing and timing of the voice message (Whiting et al., 2012) on the five facets of constructive voice behavior. The five prototypical facets of voice message may occur formally or informally to a variety of targets and through different channels.

## Conclusion

Despite the fact that interest in voice dates back several decades, existing conceptualizations and measures suffer from several deficiencies and confounds. In the present studies, we deducted a system of voice message dimensions which we used to develop an unambiguous, yet parsimonious typology of voice facets and a multidimensional measure, and we presented considerable initial evidence of construct validity. We anticipate that the Five-Facet Constructive Voice Questionnaire (5F-CVQ) will serve as a valuable instrument for advancing future research on constructive voice behavior, enabling researchers to obtain robust evidence, achieve conceptual clarity, and maintain parsimony in their investigations.

<sup>3</sup> We thank an anonymous reviewer for pointing out these potential oversights in the extant voice literature.

## Appendix

### Five-Facet Constructive Voice Questionnaire (5F-CVQ) in English

The statements are about voicing harmful factors or issues and about making suggestions for improvement at work. Please indicate the extent to which the following statements describe your behavior.

The first statements are about voicing novel ideas and creative ways to move the organization or work group forward by stepping outside the box and doing something completely new.

1. I bring forward innovative and creative ideas that may help the organization/workgroup operate more efficiently or effectively in the near future.
2. I suggest creative and innovative projects which could be beneficial to the organization/my work group in the future.
3. I am a good source of creative and innovative ideas that could improve the organization/my work group in the future.
4. I suggest innovative and creative ways to achieve future goals or objectives.
5. I come up with innovative and creative ideas to improve team performance in the near future.
6. I speak up with new and creative ideas to improve the overall functioning of the organization/ my work group in the near future.

The following statements are about voicing adaptive solutions/suggestions to move the organization or work group forward by improving something within the given known framework.

7. I make adaptive suggestions to solve problematic events at work.
8. I speak up with adaptive suggestions to existing problematic procedures or processes.
9. I express adaptive proposals to solve difficult situations at work.
10. I provide adaptive suggestions to solve existing problems with the goal of advancing the organization/my work group.
11. I propose adaptive and change focused recommendations to solve existing difficulties affecting the organization/my work group.
12. I speak up with practical solutions to current problems/difficulties in the workplace.
13. I express suggestions for adaptation to future challenges in order to prevent emerging problems that could harm the organization/my work group in the future.

14. I try to prevent difficulties that may arise in the future by expressing ways to adapt to the changes.
15. I make suggestions for adaptation to circumvent impending difficulties that might otherwise negatively affect the company/my work group in the near future.
16. I make adaptive suggestions to prevent looming difficulties that could otherwise seriously threaten the success of the organization/my work group.
17. I express adaptive recommendations to prevent looming difficulties in the future that might negatively affect the organization/my work group soon.

The following statements are about drawing attention to harmful factors, problems and difficulties in order to stop or preventing them.

18. I report problematic behavior in the workplace that is inconsistent with the organization's/work group goals.
19. I call attention to problematic work conditions that interfere with job performance.
20. I speak up with problematic events in the organization/my work group that cause difficulties and are detrimental to success.
21. I draw attention to problematic work practices or happenings that impair work performance, are not in conformity with values of the organization or disrupt cooperation in my work group.
22. I report existing coordination and communication problems in the workplace to improve team work.
23. I express my concerns about developments in the organization/my work group that I believe could lead to problems in the near future.
24. I point out looming problems that could jeopardize the success of the organization/my work group.
25. I speak up with impending problems that could have a negative impact on the organization/my work group in the near future.
26. I draw attention to impending problems and difficulties that might cause serious harm to the organization/my work group in the future.
27. I speak up if I feel a plan or project won't work in the future and will lead to difficulties.
28. I call attention to potential problems that I see coming in the future.

Innovative and future focused voice (IV) = Item 1–6.

Adaptive and status quo focused voice (ASV) = Item 7–12.

Adaptive and future focused voice (AFV) = Item 13–17.

Problem and status quo focused voice (PSV) = Item 18–22.

Problem and future focused voice (PFV) = Item 23–28.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data Availability** The inter-item correlations of Study 3 and Study 4, along with the latent within-person correlations of the five facets of voice and the correlations between voice reaction variables and five facets of voice from the additional longitudinal study, are available in the OSF repository. Access the data at: [https://osf.io/9ubh6/?view\\_only=e188526419c643f19d479e4bd2b79946](https://osf.io/9ubh6/?view_only=e188526419c643f19d479e4bd2b79946).

## Declarations

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bach, R. L., Kern, C., Amaya, A., Keusch, F., Kreuter, F., Hecht, J., & Heinemann, J. (2021). Predicting voting behavior using digital trace data. *Social Science Computer Review*, 39(5), 862–883. <https://doi.org/10.1177/0894439319882896>
- Barry, M., & Wilkinson, A. (2016). Pro-social or pro-management? A critique of the conception of employee voice as a pro-social behaviour within organizational behaviour. *British Journal of Industrial Relations*, 54(2), 261–284. <https://doi.org/10.1111/bjir.12114>
- Blenkinsopp, J., & Edwards, M. S. (2008). Chapter 8 on not blowing the whistle: quiescent silence as an emotion episode. In *Research on Emotion in Organizations. Emotions, Ethics and Decision-Making* 4, 181–206. Emerald (MCB UP ). [https://doi.org/10.1016/S1746-9791\(08\)04008-X](https://doi.org/10.1016/S1746-9791(08)04008-X)
- Bluedorn, A. C. (2002). *The human organization of time: Temporal realities and experience*. Stanford University Press.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Bobic, M., Davis, E., & Cunningham, R. (1999). The Kirton adaptation-innovation inventory. *Review of Public Personnel Administration*, 19(2), 18–31. <https://doi.org/10.1177/0734371X9901900204>
- Bolino, M., Valcea, S., & Harvey, J. (2010). Employee, manage thyself: The potentially negative implications of expecting employees to behave proactively. *Journal of Occupational and Organizational Psychology*, 83(2), 325–345. <https://doi.org/10.1348/096317910X493134>
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061–1071. <https://doi.org/10.1037/0033-295X.111.4.1061>
- Brykman, K. M., & Raver, J. L. (2021). To speak up effectively or often? The effects of voice quality and voice frequency on peers' and managers' evaluations. *Journal of Organizational Behavior*, 42(4), 504–526. <https://doi.org/10.1002/job.2509>

- Burris, E. R. (2012). The risks and rewards of speaking up: Managerial responses to employee voice. *Academy of Management Journal*, 55(4), 851–875. <https://doi.org/10.5465/amj.2010.0562>
- Burris, E. R., Rockmann, K. W., & Kimmons, Y. S. (2017). The value of voice to managers: Employee identification and the content of voice. *Academy of Management Journal*, 60(6), 2099–2125. <https://doi.org/10.5465/amj.2014.0320>
- Burris, E. R., McClean, E. J., Detert, J. R., & Quigley, T. J. (2022). The agency to implement voice: How target hierarchical position and competence changes the relationship between voice and individual performance. *Organization Science*, Article orsc.2022.1640. Advance online publication. <https://doi.org/10.1287/orsc.2022.1640>
- Büttner, M. (2012). *Marke mit System: Die Relevanz des Regulatorischen Fokus für die Markenwahl*. SpringerLink Bücher. Gabler Verlag. <http://swbplus.bsz-bw.de/bsz355113392cov.htm> <https://doi.org/10.1007/978-3-8349-7107-4>
- Cangiano, F., & Parker, S. K. (2016). Proactivity for mental health and well-being. In S. Clarke, T. M. Probst, F. Guldenmund, & J. Passmore (Eds.), *The Wiley Blackwell handbook of the psychology of occupational safety and workplace health* (pp. 228–250). Wiley Blackwell. <https://doi.org/10.1002/9781118979013.ch11>
- Chamberlin, M., Newton, D. W., & LePine, J. A. (2017). A meta-analysis of voice and its promotive and prohibitive forms: Identification of key associations, distinctions, and future research directions. *Personnel Psychology*, 70(1), 11–71. <https://doi.org/10.1111/peps.12185>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Colquitt, J. A., Sabey, T. B., Rodell, J. B., & Hill, E. T. (2019). Content validation guidelines: Evaluation criteria for definitional correspondence and definitional distinctiveness. *The Journal of Applied Psychology*, 104(10), 1243–1265. <https://doi.org/10.1037/apl0000406>
- Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10(3), e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden*. Beltz Verlag. <https://doi.org/10.17877/DE290R-12739>
- Farh, J. L., Hackett, R. D., & Liang, J. (2007). Individual-level cultural values as moderators of perceived organizational support–employee outcome relationships in china: Comparing the effects of power distance and traditionality. *Academy of Management Journal*, 50(3), 715–729. <https://doi.org/10.5465/AMJ.2007.25530866>
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327–358. <https://doi.org/10.1037/h0061470>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39. <https://doi.org/10.2307/3151312>
- Gamache, D. L., McNamara, G., Mannor, M. J., & Johnson, R. E. (2015). Motivated to Acquire? The Impact of CEO regulatory focus on firm acquisitions. *Academy of Management Journal*, 58(4), 1261–1282. <https://doi.org/10.5465/amj.2013.0377>
- Geiger, S. M., Domenech, F., & van der Meer, E. (2018). *Zeitlicher Fokus-Skala (ZFS)*. <https://doi.org/10.6102/zis259>
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304–312. <https://doi.org/10.1037/1040-3590.6.4.304>
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52(12), 1280–1300. <https://doi.org/10.1037/0003-066x.52.12.1280>
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967–988. <https://doi.org/10.1177/014920639502100509>
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1(1), 104–121. <https://doi.org/10.1177/109442819800100106>
- Hinkin, T. R., & Tracey, J. B. (1999). An analysis of variance approach to content validation. *Organizational Research Methods*, 2(2), 175–186. <https://doi.org/10.1177/109442819922004>
- Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, 19(5), 640–648. <https://doi.org/10.1097/EDE.0b013e31818131e7>
- Johnson, R. E., Chang, C.-H., & Yang, L.-Q. (2010). Commitment and motivation at work: The relevance of employee identity and regulatory focus. *The Academy of Management Review*, 35(2), 226–245. <https://doi.org/10.5465/AMR.2010.48463332>
- Kirton, M. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, 61(5), 622–629. <https://doi.org/10.1037/0021-9010.61.5.622>
- Köllner, B., Ruhle, S., & Süß, S. (2019). The moderating role of message content in the formation of employee voice. *German Journal of Human Resource Management: Zeitschrift Für Personalforschung*, 33(1), 56–75. <https://doi.org/10.1177/2397002218791855>
- Liang, J., Farh, C. I. C., & Farh, J.-L. (2012). Psychological antecedents of promotive and prohibitive voice: A two-wave examination. *Academy of Management Journal*, 55(1), 71–92. <https://doi.org/10.5465/amj.2010.0176>
- Lin, S.-H.J., & Johnson, R. E. (2015). A suggestion to improve a day keeps your depletion away: Examining promotive and prohibitive voice behaviors within a regulatory focus and ego depletion framework. *The Journal of Applied Psychology*, 100(5), 1381–1397. <https://doi.org/10.1037/apl0000018>
- Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275–340). Lawrence Erlbaum Associates Publishers.
- 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. *The Journal of Applied Psychology*, 99(1), 87–112. <https://doi.org/10.1037/a0034284>
- Morrison, E. W. (2011). Employee voice behavior: Integration and directions for future research. *The Academy of Management Annals*, 5(1), 373–412. <https://doi.org/10.1080/19416520.2011.574506>
- Morrison, E. W. (2014). Employee voice and silence. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 173–197. <https://doi.org/10.1146/annurev-orgpsych-031413-091328>
- Morrison, E. W. (2023). Employee voice and silence: Taking stock a decade later. *Annual Review of Organizational Psychology and Organizational Behavior*, 10(1), Article annurev-orgpsych-120920-054654. <https://doi.org/10.1146/annurev-orgpsych-120920-054654>
- Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F., & Lowe, W. (2021). Tracking and promoting the usage of a COVID-19 contact tracing app. *Nature Human Behaviour*, 5(2), 247–255. <https://doi.org/10.1038/s41562-020-01044-x>
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599–620. [https://doi.org/10.1207/S15328007SEM0904\\_8](https://doi.org/10.1207/S15328007SEM0904_8)
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

- Newton, P. E., & Shaw, S. D. (2013). Standards for talking and thinking about validity. *Psychological Methods*, 18(3), 301–319. <https://doi.org/10.1037/a0032969>
- Parker, S. K., & Collins, C. G. (2010). Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management*, 36(3), 633–662. <https://doi.org/10.1177/0149206308321554>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467. <https://doi.org/10.1002/nur.20199>
- Premeaux, S. F., & Bedeian, A. G. (2003). Breaking the silence: The moderating effects of self-monitoring in predicting speaking up in the workplace. *Journal of Management Studies*, 40(6), 1537–1562. <https://doi.org/10.1111/1467-6486.00390>
- Pundt, A., Martins, E., & Nerdinger, F. W. (2010). Innovative behavior and the reciprocal exchange between employees and organizations. *German Journal of Human Resource Management: Zeitschrift Für Personalforschung*, 24(2), 173–193. <https://doi.org/10.1177/239700221002400205>
- Röllmann, L. F., Weiss, M., & Zacher, H. (2021). Does voice benefit or harm occupational well-being? The Role of Job Insecurity. *British Journal of Management*, 32(3), 708–724. <https://doi.org/10.1111/1467-8551.12471>
- Sandner, S., Merz, E.-M., van den Hurk, K., van Kraaij, M., Mikkelsen, C., Ullum, H., & Clement, M. (2021). Validation of a standardized donor health questionnaire across substances of human origin. *Vox Sanguinis*, 116(6), 645–655. <https://doi.org/10.1111/vox.13054>
- Shipp, A. J., & Aeon, B. (2019). Temporal focus: Thinking about the past, present, and future. *Current Opinion in Psychology*, 26, 37–43. <https://doi.org/10.1016/j.copsyc.2018.04.005>
- Shipp, A. J., Edwards, J. R., & Lambert, L. S. (2009). Conceptualization and measurement of temporal focus: The subjective experience of the past, present, and future. *Organizational Behavior and Human Decision Processes*, 110(1), 1–22. <https://doi.org/10.1016/j.obhdp.2009.05.001>
- Song, J., Gu, J., Wu, J., & Xu, S. (2019). Differential promotive voice–prohibitive voice relationships with employee performance: Power distance orientation as a moderator. *Asia Pacific Journal of Management*, 36(4), 1053–1077. <https://doi.org/10.1007/s10490-019-09644-6>
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245–251. <https://doi.org/10.1037/0033-2909.87.2.245>
- Strauss, K., & Parker, S. K. (2014). *Effective and sustained proactivity in the workplace: A self-determination theory perspective*. <https://doi.org/10.13140/2.1.2809.1845>
- van Dyne, L., & Lepine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal*, 41(1), 108–119. <https://doi.org/10.2307/256902>
- van Dyne, L., Cummings, L. L., & McLean Parks, J. (1995). Extra-role behaviors: in pursuit of construct and definitional clarity. *Research in Organizational Behavior*, 17, 215–285.
- 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. <https://doi.org/10.1111/1467-6486.00384>
- Weiss, M., & Zacher, H. (2022). Why and when does voice lead to increased job engagement? The role of perceived voice appreciation and emotional stability. *Journal of Vocational Behavior*, 132, 103662. <https://doi.org/10.1016/j.jvb.2021.103662>
- Whiting, S. W., Maynes, T. D., Podsakoff, N. P., & Podsakoff, P. M. (2012). Effects of message, source, and context on evaluations of employee voice behavior. *The Journal of Applied Psychology*, 97(1), 159–182. <https://doi.org/10.1037/a0024871>
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 8(2), 94–104. <https://doi.org/10.1111/j.1524-4733.2005.04054.x>
- Wilkinson, A., Dundon, T., Donaghey, J., & Freeman, R. B. (2020). Employee voice: Bridging new terrains and disciplinary boundaries. In A. Wilkinson, J. Donaghey, T. Dundon, & R. Freeman (Eds.), *Handbook of Research on Employee Voice* (pp. 2–18). Edward Elgar Publishing. <https://doi.org/10.4337/9781788971188.00007>
- Zhou, J., & George, J. M. (2001). When job dissatisfaction leads to creativity: Encouraging the expression of voice. *Academy of Management Journal*, 44(4), 682–696. <https://doi.org/10.2307/3069410>
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, 12(4), 399–413. <https://doi.org/10.1037/1082-989X.12.4.399>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.