European Management Journal 39 (2021) 445–455

Contents lists available at ScienceDirect

European Management Journal

journal homepage: www.elsevier.com/locate/emj

Management Focus

Does task-related conflict mediate the board gender diversity–organizational performance relationship?

Koen van den Oever [a, *], Bart Beerens [b]

[a] Nijmegen School of Management, Radboud University Nijmegen, Institute for Management Research, 6525, AJ Nijmegen, the Netherlands
[b] School of Economics and Management, Tilburg University, P.O. Box 90153, 5000, LE Tilburg, the Netherlands

ARTICLE INFO

Article history:
Received 30 October 2019
Received in revised form
27 July 2020
Accepted 17 September 2020
Available online 17 September 2020

Keywords:
Board of directors
Task-related conflict
Organizational performance
Board diversity

ABSTRACT

The board gender diversity–organizational performance relationship has been criticized for the absence of tests of the underlying mechanisms of this relationship. This study aims to empirically investigate whether task-related conflict – one of the prime theorized mechanisms of board diversity – indeed mediates this relationship. Consistent with the literature, we theorize how board gender diversity affects task-related conflict, and how task-related conflict in turn affects organizational performance. We test our hypotheses in the Dutch water authority sector 2009–2014, where we have access to the detailed board meeting minutes of 27 organizations. Our results find support for a partial mediating effect of task-related conflict in the board gender diversity–organizational performance relationship. We conclude by discussing the implications for the board gender diversity literature.

1. Introduction

Given that it is the highest legal authority in the organization, a large strand of research has focused on the role of the organization's board of directors in explaining performance heterogeneity (Johnson, Schnatterly, & Hill, 2013). This work has mainly argued that the board serves as a critical governance mechanism: its function is primarily and most importantly to monitor top management (Boivie, Bednar, Aguilera, & Andrus, 2016; Campbell, Campbell, Sirmon, Bierman, & Tuggle, 2012; Daily, Dalton, & Cannella, 2003). As such, considerable evidence has been built on what combination of board characteristics will lead to greater organizational outcomes (Boivie et al., 2016).

One particular domain in this research focuses on the role of board diversity in explaining organizational performance (e.g. Forbes & Milliken, 1999; Johnson et al., 2013; Triana, Miller, & Trzebiatowski, 2013; Williams & O'Reilly, 1998), and in particular board gender diversity (Hoobler, Masterson, Nkomo, & Michel, 2018; Post & Byron, 2015). The main claim is that gender diversity implies diverse perspectives (by reference to upper echelon theory's core assumption; Hambrick & Mason, 1984), leading to an optimal pool of knowledge available for the board's monitoring activities (Post & Byron, 2015). However, this literature on the board gender diversity–organizational performance relationship has found ambiguous results (Hoobler et al., 2018; Johnson et al., 2013; Lawrence, 1997; Miller & Triana, 2009).

One particular need in this strand of work is insights into the mechanisms by which board gender diversity impacts organizational performance. That is, "a consistent shortcoming of business case data sets, and thereby a shortcoming of published studies, is the missing test of a mechanism linking women leaders to performance" (Hoobler et al., 2018, p. 2485). Without strong and empirically supported theory, the literature thus remains in a deadlock of work that is primarily data rather than theory driven. In fact, even more than two decades ago, Lawrence (1997: 16) argued that "demographic variables should play no role in organizational studies unless we understand what role they are playing."

Disagreement, however, may be one of the key elements that can explain how board gender diversity impacts organizational performance. One often theorized mechanism by which board diversity impacts organizational performance concerns task-related conflict (e.g. Carpenter & Westphal, 2001; Frijns, Dodd, & Cimerova, 2016; Goodstein, Gautam, & Boeker, 1994). Task-related conflict is conceptualized as disagreement between group members, based on task-related issues, due to different opinions or perspectives (Forbes & Milliken, 1999; Jehn, 1995). This kind of conflict produces the beneficial generation of discussion and the

* Corresponding author.
E-mail addresses: K.vandenOever@fm.ru.nl (K. van den Oever), bartbeerens93@hotmail.com (B. Beerens).

https://doi.org/10.1016/j.emj.2020.09.008
0263-2373/

evaluation of alternatives, especially when it concerns complex tasks (Jehn, Northcraft, & Neale, 1999) as the board's tasks imply (Forbes & Milliken, 1999; Zajac & Westphal, 1996). Increasing gender diversity would thus lead to an optimal pool of knowledge available for the board's monitoring activities (Post & Byron, 2015), that could increase organizational performance *through* task-related conflict.

Conversely, research in social psychology has consistently found that teams that have individuals with different information at their disposal share less information within their team (Lu, Yuan, & McLeod, 2012; Mesmer-Magnus & DeChurch, 2009; Stasser & Titus, 1985). Thus, a social psychology lens would predict that although diversity can potentially increase the total information pool, discussion would be on commonly shared information and not on unique information. This mechanism is called 'biased information sampling' (Stasser & Titus, 1985), which potentially prevents task-related conflict to arise.

In this study, we synthesize the literature streams on board gender diversity and biased information sampling to uncover the role of task-related conflict in the board gender diversity–organizational performance relationship. In essence, we set out to study whether task-related conflict within the board of directors has a mediating effect on the board gender diversity–organizational performance relationship. Therefore, the central question dealt with in this study is: '*What is the influence of board gender diversity on organizational performance and does task-related conflict mediate this relationship*?'. We study this research question in the Dutch water authority sector. This setting allows us to collect a unique set of meeting minutes of 27 organizations over six years to measure task-related conflict objectively and unobtrusively. Also, in other studies, board gender diversity is often an endogenous construct (Johnson et al., 2013). In contrast, board members of water authorities are elected by the residents of the water authority, which allows for a better identification of the effects of board gender diversity.

This study is intended to contribute to the literature in multiple ways. First, we make a theoretical contribution to the board gender diversity literature by showing that task-related conflict is an important mechanism by which gender diversity affects organizational performance. As previously mentioned, studies analysing the board gender diversity-organizational performance relationship have produced ambiguous results. As such, it is important to analyze the intervening variables in the board gender diversity–organizational performance relationship because this relationship may be 'complex and indirect' (Forbes & Milliken, 1999, p. 490). We do so by drawing in the literature on biased information sampling to explicate the role of task-related conflict.

Second, we make an empirical contribution by offering the first test of the mediating role of task-related conflict in the board gender diversity–organizational performance relationship. Third, we make a methodological contribution by developing a content analysis dictionary for task-related conflict. Although partly tailored to the context of our study, this dictionary can be used in future research to analyze the role of task-related conflict in other contexts, but also in its relations to other concepts.

2. Hypothesis development

In general, the literature on board gender diversity has consistently found evidence that it increases firm financial performance (as evidenced by a recent literature review and a meta-analysis: Hoobler et al., 2018; Post & Byron, 2015). Although different theories are used to explain this relationship, such as agency theory, resource dependence theory, and the resource-based view (Hoobler et al., 2018), the main argument used is that men and women have different cognitive frames (Dezsö & Ross, 2012; Hoobler et al., 2018). These cognitive frames influence what issues directors pay attention to, how they interpret these issues, and how they respond to these issues, thus materializing in contributions of perspectives in board meetings (Hambrick, 2007). Since boards encounter ambiguous and complex problems, directors are likely to have different perspectives and opinions on these problems, leading to disagreements (Forbes & Milliken, 1999). The literature refers to this as task-related conflict: the differences in opinions, perspectives or other task-related disagreements among individuals (Forbes & Milliken, 1999). In the complex decision-making tasks boards are confronted with, a consideration of more information and alternatives is expected to improve decision-making, and thus organizational performance (Miller & Triana, 2009).

The other argument is that women leaders can create gender supportive climates, yet such climate is then again assumed to allow women to bring their unique perspectives to the table (Hoobler et al., 2018). The general assumption for these proposed mechanisms is that these perspectives cause discussion, benefiting the board in their monitoring and advising functions (e.g. Erhardt, Werbel, & Shrader, 2003).

Yet, given that gender is a highly salient demographic, it provides a strong basis for social categorization (Milliken & Martins, 1996). Such social processes include stereotyping and categorization (Pelled, Eisenhardt, & Xin, 1999; Watson, Kumar, & Michaelsen, 1993). Such social categorization implies that board members categorize themselves and others into groups, distinguishing between ingroup (similar) members and outgroup (dissimilar) members (Van Knippenberg & Schippers, 2007). Such categorization implies that people tend to favor the opinions and ideas of ingroup members more than outgroup members (Tajfel & Turner, 1979). Therefore, when stereotyping and categorization within the board exists, directors feel less attracted to the rest of the group, group cohesion deteriorates, which subsequently limits group communication and cooperation (Dahlin, Weingart, & Hinds, 2005; Williams & O'Reilly, 1998). Consequently, cognitive resources may not be utilized to the maximum, directors have more complications in reaching consensus, and organizational performance will eventually decrease (Barkema & Shvyrkov, 2007).

Thus, board gender diversity scholars have argued that board members in gender diverse boards may be less prone to speak up and convey their perspectives (Westphal & Bednar, 2005). And if board members do speak up, the diversity of perspectives has been argued to impede the board's ability to reach consensus (Triana et al., 2013). Nevertheless, research has failed to find evidence that men or women have less influence in more gender diverse boards (Westphal & Milton, 2000). In fact, there may be evidence that gender diversity in boards leads to less social categorization, per the decision-making styles of women. That is, women in upper echelon positions have been found to exhibit communication styles that emphasize inclusion (Rosener, 1995), to be more communal than men (Rosette & Tost, 2010), and to value benevolence more (Adams & Funk, 2012). Hence, although gender could form a basis for social categorization, the literature has not found evidence for it on the board level; if anything, the literature points to greater inclusion of perspectives when more women are on the board, i.e. when there is greater gender diversity.

Notably nonexistent in the board diversity literature are the insights from biased information sampling. The biased information sampling model has shown that teams often do not optimally utilize information when making decisions; intragroup discussion is often revolved around strengthening individual pre-discussion preferences rather than as a forum to share new information (Gigone & Hastie, 1993; Stasser & Titus, 1985). For instance, information that is held by only one individual is rarely discussed

(Stasser, Taylor, & Hanna, 1989). In particular, this is the case when the team needs to make a decision that does not have one unambiguous correct alternative (Stasser & Titus, 1985), as is often the case for board decisions. Hence, although the board gender diversity literature assumes that different information will be shared, a social psychological lens would argue that only a sample of the information will be contributed in a discussion (specifically information that group members already had in common before the discussion). Indeed, the presence of biased information sampling has been consistently found in the literature (for a meta-analysis, see Mesmer-Magnus & DeChurch, 2009; Stasser & Titus, 1985).

To understand how gender could impact the likelihood of biased information sampling, expectation states theory provides guidance. Expectation states theory argues that group members form expectations of the potential usefulness of their own and others' contributions (Berger, Conner, & Fisek, 1974). Individuals who are gauged to have the most potential useful contributions will have increased participation (Dovidio, Brown, Heltman, Ellyson, & Keating, 1988). Thus, when gender is associated with this assessment of usefulness, we would be likely to find biased information sampling. That is, when women or men feel their contributions may not be useful, they will be less prone to speak up and discuss their own contributions. However, when other, more task-relevant cues, are available, group members are likely to use those cues as a basis to assess potential usefulness of contributions (Wittenbaum, 1998). Given that gender may not be that important for the board tasks (the role of other demographics, such as industry background and ideology is more likely to be pronounced, given that these are task-relevant forms of diversity), it is unlikely that gender diversity would spawn biased information sampling.

In all, the literature provides indication that more gender diverse boards will have a larger pool of information and perspectives to their disposal, which in turn is a necessary condition for task-related conflict. After all, as task-related conflict is defined as the disagreement between group members, based on task-related issues, due to different opinions or perspectives (Forbes & Milliken, 1999; Jehn, 1995), different opinions or perspectives need to exist to enable task-related conflict. As social categorization and biased information sampling are deemed less of an issue for gender diverse boards, we hypothesize the following:

Hypothesis 1. The higher the level of board diversity in gender, the higher the level of task-related conflict within the board.

2.1. Task-related conflict and organizational performance

Forbes and Milliken (1999) suggest that task-related conflict is a board process that influences board task performance, and thus organizational performance. Task-related conflict can affect strategic decision-making, because of the consideration of multiple perspectives and a careful evaluation of alternatives (Forbes & Milliken, 1999). The constructive debate and exchange of comments help boards to carry out their intellectual tasks more effectively (Zona & Zattoni, 2007). In particular, the control task of the board is enhanced by this kind of interactive processes, because disagreement and criticism require CEOs and other top-level managers to justify and evaluate strategic decisions (Forbes & Milliken, 1999). Thus, task-related conflict stimulates board members to generate novel concepts and approaches (Jehn, 1995). Consequently, this will induce boards to take unconventional decisions and lead firms to strategic change (Barkema & Shvyrkov, 2007). Summarizing, task-related conflict in the board of directors is expected to have a positive effect on organizational performance, formalized in the following hypothesis:

Hypothesis 2. Task-related conflict has a positive effect on organizational performance.

2.2. Mediation of the board diversity-organizational performance relationship

Task-related conflict has only been theorized to influence the board diversity–organizational performance relationship (e.g. Forbes & Milliken, 1999). Furthermore, Heemskerk, Heemskerk, and Wats (2017) conducted research on the conflict–performance relationship and recommended future studies to include board diversity in their framework. Jehn et al. (1999) found that informational diversity positively influences group performance, mediated by task-related conflict. This indicates that task-related conflict can have a mediating effect. Below, a visual overview is provided in Fig. 1.

Hypothesis 3. Task-related conflict mediates the relationship between board gender diversity and organizational performance.

3. Methodology

3.1. Sample

The sample of this study contains data on all the water management authority boards in The Netherlands from 2009 to 2014. These water management organizations are responsible for territorial safety (e.g. against flooding of rivers), water quantity and quality (for irrigation purposes for local farming), and sewage treatment. Given increasingly erratic weather conditions (e.g. extremely dry summers and occasional intense rainstorms), these water authorities have become crucial actors in safeguarding the Netherlands against the consequences of climate change, such as drought, heat stress, and flooding. The efficient and effective allocation of resources, on which the board of directors provides oversight, is thus paramount to making the necessary countermeasures. As of 2009, when our sampling frame starts, 27 water authorities exist in the Netherlands.

The water authorities represent a separate public governance level—operating alongside the Dutch national government, provincial, and municipal administrations—and can thus levy their own taxes. However, water authorities are considered as functional bodies whose remit is limited to public water management (Water Authority Act, 1991), unlike other more general government authorities—such that they very similar to more general organizations in their focus and structure. Yet, in contrast to general organizations, a large part of the directors in these organizations are selected on the basis of votes by the inhabitants of the organizations' regions—limiting potential concerns about processes that harm director selection in board settings (Johnson et al., 2013). The directors that are not elected are nominated by important stakeholder associations, such as the Dutch Chamber of Commerce, the Agricultural Association, and the Nature Association. As such, although these organizations are similar to many other organizations, their differences (e.g. the exogenous selection of directors and the availability of meeting transcriptions) allow for a unique peek into the effects of diversity in the board room.

The board is particularly involved in all the activities concerning water environmental issues in their appointed region. They determine the budget, taxes and the annual report of the organization. The control task, which is the largest responsibility of the board, includes the monitoring of the executed strategy by the top management team. The board characteristics data consists of gender, educational level, industry background, age, political group and board size. This data was collected via the organization Unie van

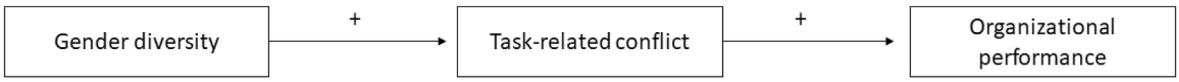

Fig. 1. Proposed conceptual framework.

Waterschappen, which is the association of water management authorities, and via publicly available sources as websites of the water management authorities. Moreover, the notes of meetings and the annual costs per water management authority were gathered. In total, the sample size included 79 organization-year observations.

3.2. Dependent variable

The dependent variable organizational performance is operationalized as accounting-based performance, using costs divided by the number of regional inhabitants and subsequently negatively scaled. This variable is suitable, since the water management authority's performance is monitored by annual costs. The costs of the water management authority should be fully covered by the taxes paid by the local citizens. Therefore, it is important for the board to lower the costs, since subsequently, the taxes can decrease. When the board succeeds in decreasing taxes, the citizens will be more satisfied with the board and the probability of being re-elected increases. Other studies have primarily focused on the firm revenues, profit or stock price. The incentive for directors is to minimize the annual costs, while achieving an efficient way of operating, as others have also conceptualized organizational performance in related settings (e.g., Opstrup & Villadsen, 2014).

The importance of cost control at the water authorities is expressed in various ways. For instance, a business comparison between the water authorities is made every two year (Unie van Waterschappen, 2018). This comparison pays much attention to the cost development of the different water authorities, for instance by describing ways to save costs. "With innovative solutions, cooperation with other parties and efficiency measures, the boards keep the increment of costs, and therefore the increment of the burden for citizens and businesses, as low as possible" (Unie van Waterschappen, 2018: 58). The comparison also considers future developments by describing the agreements that have been made in order to keep the increment of taxes limited. "In 2011, the water boards made agreements (…) to ensure that the costly investments required for safety against high water, good water quality and not too much and not too little water can be carried out without leading to a sharp rise in taxes for households and businesses" (Unie van Waterschappen, 2018: 65). In addition to this, the water authorities focus on costs in the annual reports and the preparation of the budget.

3.3. Independent variables

Board gender diversity. To calculate the diversity of a group based on a categorical variable such as gender, experts in the field of diversity have used and recommended Blau's heterogeneity index (1977) (Bantel & Jackson, 1989; Harrison & Klein, 2007; Miller & Triana, 2009). Since this research theorizes diversity as a variation among directors, an operationalization of diversity representing variability is suitable (Harrison & Klein, 2007). Moreover, the four criteria for good measurement of diversity are met: a higher index indicates a higher level of diversity, the index does not allow negative values, the index has a zero point to represent perfect homogeneity, and the index is not unbounded (Miller & Triana, 2009). Hence, Blau's index (1977) is utilized to capture an objective, relative measure of diversity (Triana et al., 2013) and is calculated as follows:

$$\text{Heterogeneity} = (1\text{-}\sum p_i^2)$$

where p equals the proportion of group members in a category and i represents the number of different categories that are present in the board. The range of the index is dependent on the number of categories for a characteristic. The range theoretically starts at 0, meaning perfect homogeneity in the board. The upper level is calculated by (i - 1)/i (Miller & Triana, 2009). As a result, the categorical variables are transformed into continuous variables.

3.4. Mediator variable

Task-related conflict. Notes of board meetings in the period 2009–2014 were analyzed to investigate task-related conflict. An objective measurement of task-related conflict could be obtained in the notes of board meetings, due to the unbiased nature of notes. Furthermore, Jehn (1997) found that her objective measure (i.e., categorical tree diagram) indicated task-related conflict in 95% of the cases, which was the most accurate score of the study's measures. Therefore, an objective measure for task-related conflict is preferred. Content analysis is suitable to examine minutes, since the concept of task-related conflict could be captured in dictionary keywords (Abrahamson & Eisenman, 2008; Wade, Porac, & Pollock, 1997). Considering the content analysis, this study made use of a computer-automated analysis of keywords. The content analysis was conducted in a rather similar way as in a previous study of Wade et al. (1997), which is explained hereafter.

First, to collect the data from these minutes, dimensions of task-related conflict were constructed, based on existing literature (e.g. Jehn, 1995; Jehn, 1997). Next, a dictionary with keywords for every dimension was developed and tested to assess content validity. By manually testing the keywords in a sample of 0.5% of all notes, the number of hits and the false hit rates were computed. After the manual test, the number of dimensions was reduced from four to three, since the dimension 'disagreeing on goals and objectives' was highly similar to dimension A. Below, the three dimensions for task-related conflict are displayed in Table 1.

Moreover, the most appropriate keywords were preserved, hence the keywords with the highest number of hits and the lowest false hit rates (i.e., below 0.15) (Wade et al., 1997). The final dictionary is displayed in Table 2. Then, content analysis on the frequency of these task-related conflict keywords was applied with the program LIWC 2015. LIWC 2015 is an automated content analysis program that can search and count the number of dictionary words within a text. LIWC 2015 returns the number of dictionary words in a percentage of the total number of words in a text. In this way, the results of a text with a larger number of words, and thus a higher probability of having more dictionary words, is placed in perspective. Also, in a study of Abrahamson and Eisenman (2008) a ratio was applied for content analysis. Hence, the measure for task-related conflict is the weight of the number of dictionary words in the total number of words in a text. In this study, the context of task-related conflict keywords was taken into account, since it is recommended to look at context, when investigating a smaller sample (Abrahamson & Eisenman, 2008). In Table 3,

Table 1
Dimensions of task-related conflict.

Dimension	Measurement	Author(s)
A	*Disagreeing on content and outcomes of the tasks being performed*	Amason (1996); De Dreu and Weingart (2003); De Wit, Greer, and Jehn (2012); Jehn (1995); Jehn (1997); Jehn and Bendersky (2003); Pelled (1996); Pelled et al. (1999)
B	*Evaluating different perspectives/ alternatives*	Amason (1996); De Wit et al. (2012); Jehn (1995); Jehn and Bendersky (2003); Pelled et al. (1999)
C	*Inducing constructive debate*	De Wit et al. (2012); Gibson and Vermeulen (2003); Wall and Nolan (1986)

Table 2
Corresponding words and (false) hit rates.

Dimension	Corresponding Words	Words in Dutch	Hit Rate	False Hit Rate
A	Objective(s)	Doelstelling (en)	15	0.13
	To not support	Niet steunen/steunt	5	0
	Task(s)	Taak/Taken	17	0.06
	Responsibility	Verantwoordelijkheid/Verantwoording	9	0.11
	To question	Vraagt zich af/Vraag mij af/Vragen ons af/Afvragen	8	0
B	Consequences	Consequenties	13	0
	Possibilities	Mogelijkheden	16	0.125
C	(Request) More information	(Vraagt/vragen om) Meer/Nadere informatie	3	0

Table 3
Examples of keywords in the context of task-related conflict.

Keyword in Dutch	Keyword in English	Conflict Example English
Afvragen	To question	And then I keep questioning myself: is a water authority treating its citizens like this?
Consequenties	Consequences	And you say, well, maybe that can have consequences for home visits. Those home visits should just take place, I think (..)
Doelstelling	Objective	(..) a problem is a problem, when it hinders the goals of the organization. However, it is also about the outside world and why is this not taken into account.
Mogelijkheden	Possibilities	His request is critically looked at again to investigate the possibilities of economizing.
Niet steunen	To not support	The party thinks that resolution 2 is completely out of order and that this belongs to the budget estimation of 2013. WN will not support the resolution.
Taak	Task	It is a major project of a different authority, but the general board has the task to consider it.
Verantwoordelijk-heid	Responsibility	The party calls the coalition parties to shed a light on the agreement again, keeping the international financial crisis in mind and tie consequences to this by decreasing expenses, to take responsibility in this way.

examples of sentences containing these keywords are gathered.

3.5. Control variables

As suggested by an extensive review of the literature and a logical reasoning on the research environment, the following control variables are included in the model: educational level diversity, industry background diversity, age diversity, ideological diversity, board size, organization size, slack resources, attainment discrepancy and average meeting duration.[1]

Educational level diversity. For this variable, three categories are selected: WO, HBO and 'No Title'. The categories stand for the highest level of Dutch education that a director has completed, with WO being a graduate university degree and the highest category. HBO is an undergraduate degree from a university of applied sciences. We use Blau's heterogeneity index to measure this type of diversity.

Industry background diversity. In this research, the data is compiled into seven categories: no paid work, self-employed, employee in business, employee in health care, employee in education, employee in the government, and employee in other type of industry. We again use Blau's heterogeneity index to measure this type of diversity.

Ideological diversity. We control for ideological diversity by signaling which political party each specific board member is affiliated with. Again, we use Blau's heterogeneity index to measure this type of diversity.

Age diversity. Age diversity is examined by other scholars (Milliken & Martins, 1996) and could influence the formation of subgroups and organizational performance. Therefore, it is included as a control variable in this study. The directors' age is divided into five categories: ≤ 30, $30 \leq 40$, $40 \leq 50$, $50 \leq 60$, $60 >$. Here, we use Blau's heterogeneity index to measure the diversity.

Board size. The size of the board is acknowledged to have an impact on group dynamics (Li & Hambrick, 2005; Pelled et al., 1999). Hence, according to other literature, studies on diversity should include this control variable. The board size is measured by the number of directors.

Organization size. Previous literature on board diversity and organizational performance included organization size as a control variable (Miller & Triana, 2009). Organization size is measured as the logarithm of the regional number of inhabitants per water management authority, since the number of employees is not available and the number of inhabitants indicates the size of the operations.

Slack resources. Given that slack resources can buffer intra-organizational conflict (Cyert & March 1992), we controlled for slack resources, the level of cash reserves of water authority i at time t. Cash is the most easily deployed resource and is, as such, an ideal measure of slack resources. Given the non-normal distribution of this variable, we took its cube root to correct for skew.

[1] Although it would have been useful to control for faultline strength, the data is on the team-level, hence we do not know per board member each demographic, and are thus unable to calculate faultline strength.

Attainment discrepancy. In this context, attainment discrepancy could best be described as the difference between the aspired (i.e., budgeted) costs and the realized costs (Lant & Montgomery, 1987). Performance that is below aspired performance is considered to be unsatisfactory performance (Cyert & March 1963), which could lead to an increase in conflict. This could influence task-related conflict, since discussions can emerge when the budgeted costs are not reached. On the other hand, a positive result can lead to less conflict, due to a good fulfillment of the task. Therefore, and to preserve as much observations as possible, the variable is distinguished in two separate variables: positive and negative attainment discrepancy (i.e., PAD and NAD).

Average meeting duration. Since the duration of the meeting can determine how much task-related conflict can occur, we controlled for it by subtracting the closing time of the meeting by the opening time of the meeting. This data was included in the meeting minutes. This length of the meeting was measured in minutes. We then averaged the meeting duration per organization *i* at year *t*.

Year dummies. To control for time-specific effects, year dummies are included in the model.

3.6. Analysis

Hypotheses 1 and 2 were tested by using random-effects generalized least squares (GLS) regression, since the variables consist of panel data. Panel data are likely to have a certain degree of serial correlation within the error term of the regression, due to unobserved individual−specific time-invariant effects (Ahn, Lee, & Schmidt, 2013). Therefore, the panel data were analyzed with the random-effects model, since this controls for the before mentioned effects. Given that the gender diversity measure did not deviate per year within the organization, we chose to use random-effects over fixed-effects.

Hypothesis 3 was analyzed by using seemingly unrelated regression with bootstrapping (Preacher & Hayes, 2008). Bootstrapping is a random sampling method with replacement, which does not demand the assumption of a normally distributed sample. By resampling 5000 times, an approximation of the indirect effect's confidence interval can be constructed. While smaller samples are often not normally distributed, bootstrapping is a method that can help to investigate the mediating effect in smaller samples. Therefore, bootstrapping is the preferred method in this study. First, a regression of task-related conflict on the independent and control variables was executed using bootstrapping. After that, organizational performance was regressed on all independent and control variables with bootstrapping.

4. Results

The descriptive statistics of the included variables are shown in Table 4. Organizational performance was scaled negatively for ease of interpretation, since performance is measured as costs divided by the number of regional inhabitants. The mean is −134.748 and the range of organizational performance is rather wide. Hence, a high level of variation in the dependent variable is observed. In Table 5, we display the correlation table.

Hypothesis 1 predicted that gender diversity is positively related to task-related conflict. In Table 6, the results are shown of regressing task-related conflict on these variables and all control variables. Model 1 excludes gender diversity to test the effect of the control variables on task-related conflict. In model 2, we report the results when board gender diversity and all control variables are included. Compared to model 1, the R^2 in model 2 increases from 0.351 to 0.592, which suggests that the explanatory variables are adding value. Hence, the hypothesis will be tested in the full model, unless stated otherwise.

Model 2 shows that Hypothesis 1 is supported, since gender diversity has a positive effect of 0.007 on task-related conflict with significance level 0.004. For every standard deviation increase (0.077) in gender diversity, the percentage points of words in minutes related to task conflict increases with 0.000539 approximately. Considering the mean of task-related conflict (0.002), the magnitude of this effect is moderately low.

In Table 7, the regression models for testing Hypothesis 2 are shown, which suggests that task-related conflict is positively related to organizational performance. First, Model 3 shows the effect of all control variables on organizational performance. Next, Model 4 includes the explanatory variables in the regression. Gender diversity is positively and significantly related to organizational performance (167.689, $p > 0.000$). Task-related conflict shows a positive significant relationship with organizational performance (4887.199, $p = 0.002$). Therefore, Hypothesis 2 is supported. For every standard deviation increase (0.001) in task-related conflict, the costs per inhabitant decreases with 4.89 euros approximately. Given that the mean cost per inhabitant is 134.75 euro, this is an economically significant effect.

Testing Hypothesis 3 required mediation analyses. The output of these analyses is displayed in Table 8. We started with a simple Sobel test (Sobel, 1982) for each of the diversity measures (gender, educational level, and industry). For the gender diversity−organizational performance relationship, we find that task-related conflict plays a mediating role ($z = 2.380$, $p = 0.017$).

To ensure the robustness of our findings to alternative tests of mediation, we also followed the stepwise approach of Blau (1977). According to this approach, the presence of a mediation effect is indicated if; (1) the independent variable significantly predicts the mediating variable, (2) the independent variable significantly predicts the dependent variable, and (3) the mediating variable significantly predicts the dependent variable while controlling for the effect of the independent variable, (4) full mediation is indicated when the effect of the independent variable on the dependent variable is 0 when the mediator is included in the same model.

Results of the stepwise test confirm that task-related conflict significantly mediates the relationship between one type of diversity (gender) and organizational performance. This is shown by; (1) the significant coefficient of gender diversity (coefficient = 0.007, $t = 2.91$, $p = 0.004$) when explaining task-related conflict, (2) the significant coefficients of gender diversity (coefficient = 167.689, $t = 5.73$, $p < 0.001$) when explaining organizational performance, and (3) the significant effect of task-related conflict on organizational performance when controlling for gender diversity (coefficient = 4887.199, $t = 3.15$, $p = 0.002$), (4) yet the effect is not fully mediated as gender diversity remains to significantly affect organizational performance even when controlling for task-related conflict (coefficient = 167.689, $t = 5.73$, $p < 0.001$). For ease of reading, we report the coefficient estimates meaningful to our analyses (from Tables 6 and 7), together with the results of the tests, in Table 8.

Additionally, we performed a nonparametric (bootstrap) test. We used the bootstrap approach developed by Preacher and Hayes (2004): we estimated the indirect effects using coefficients from the full model (i.e., those of model 4) and utilized bootstrapping procedures with 5000 resamples to place bias-corrected and accelerated 95% confidence intervals (CIs) around the estimates of the indirect effects. An indirect effect is significant at the 0.05 level when the 95% CI does not include zero (Shrout & Bolger, 2002). Table 8 shows that the indirect effect of gender diversity on organizational performance was significant through task-related conflict (coefficient = 34.210, 95% CI = 3.344, 109.626). In conclusion, our empirical evidence provides support for H3, which holds that

Table 4
Descriptive statistics.

Variable	Mean	S.D.	Min	Max
1. Organizational Performance	−134.748	31.390	−262.083	−82.110
2. Task Conflict	0.002	0.001	0.0004	0.009
3. Gender Diversity	0.313	0.077	0.147	0.420
4. Educational Diversity	0.591	0.083	0.444	0.714
5. Industry Diversity	0.739	0.041	0.656	0.812
6. Ideological Diversity	0.783	0.126	0.477	0.873
7. Age Diversity	0.463	0.066	0.331	0.561
8. Board Size	28.329	2.673	23	30
9. Organizational Size	5.851	0.193	5.342	6.130
10. Slack Resources	72.945	69.230	0	320.739
11. Positive Attainment Discrepancy	3.336	4.916	0	38.111
12. Negative Attainment Discrepancy	1.267	3.200	0	16.223
13. Average Meeting Duration	145.682	41.752	62.857	254.625

Table 5
Correlation table.

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Organizational performance	1												
2	Task Conflict	0.12	1											
3	Gender Diversity	0.50[b]	0.29[b]	1										
4	Educational Diversity	0.19	−0.06	−0.07	1									
5	Industry Diversity	−0.02	0.18	0.38[b]	−0.07	1								
6	Ideological Diversity	−0.33[b]	−0.05	0.17	−0.28[a]	0.21	1							
7	Age Diversity	−0.20	−0.05	0.03	−0.35[b]	−0.39[b]	0.09	1						
8	Board Size	0.01	−0.33[b]	0.00	0.11	−0.11	0.62[b]	−0.13	1					
9	Organization Size	0.39[b]	−0.22	0.19	0.26[a]	−0.11	0.27[a]	−0.16	0.70[b]	1				
10	Slack Resources	−0.15	0.26[a]	0.07	−0.33[b]	−0.07	0.26[a]	0.06	−0.03	0.04	1			
11	Positive Attainment Discrepancy	0.31[b]	−0.10	0.08	−0.01	0.01	−0.03	−0.08	0.01	−0.03	−0.12	1		
12	Negative Attainment Discrepancy	−0.42[b]	0.03	−0.23[a]	−0.02	−0.08	0.05	−0.06	0.10	−0.10	0.13	−0.27[a]	1	
13	Average Meeting Duration	0.15	−0.14	0.06	−0.19	−0.33[b]	0.32[b]	0.09	0.29[b]	0.19	0.39[b]	0.21	−0.01	1

[a] p-value < 0.05.
[b] p-value < 0.01.

task-related conflict mediates the relationship between the gender diversity and organizational performance.

Further notable is that in model 4 of Table 7, gender diversity remains to have a significant and positive effect on performance, even when task-related conflict is included. Thus, although we can conclude that task-related conflict does fully mediate this relationship, a separate, direct effect of gender diversity on organizational performance remains. This is also evident when considering the direct effect to total effect ratio, which is 0.83 (167.689/201.899). This ratio indicates that only 17% of the effect of gender diversity on organizational performance is mediated by task-related conflict.

4.1. Robustness checks

Since achieving complete accuracy in automated content analysis is complicated (Wade et al., 1997), the results were adapted with the support of the tested false hit rates, to obtain the most reliable data. Hence, the results of content analysis were multiplied by (1-FHR), where FHR is the false hit rate of the keyword. This is different from previous research (e.g. Abrahamson & Eisenman, 2008; Wade et al., 1997), where false hits were rather accepted than addressed. Furthermore, the different dictionaries with their respective false hit rate adjustments, were analyzed separately to investigate dissimilarities between the analyses. Table 9 reports the result of this check for the seemingly unrelated regression using bootstrapping with 5000 resamples (replication results for the other models are available upon request). The results are more conservative with regard to our original results given that we account for the false hit rate, and these results remain consistent with our original results.

Moreover, reverse causality regarding the task-related conflict – organizational performance relationship could be present. As such, in model 7 we replicate the random effects GLS regression where organizational performance was the dependent variable (model 4), but leading organizational performance by one year. Our results are robust to this specification, further validating the effect of task-related conflict on performance.[2]

5. Discussion

In this study, we set out to study whether task-related conflict is indeed a core mechanism to explain the relationship between board gender diversity and organizational performance as was assumed in the literature (Forbes & Milliken, 1999). Past board of director "studies frequently assume that certain compositional attributes (independence, director experience, social ties) will lead to certain conduct in the board room, leading to firm-level outcomes" (Johnson et al., 2013, p. 253). In fact, Lawrence (1997) even claimed that "Demographic variables should play no role in organizational studies unless we understand what role they are playing" (Lawrence, 1997, p. 16).

We find that task-related conflict is indeed a powerful

[2] We did not lead the task-related conflict variable for other robustness checks, since diversity is in place before the meetings start and task-related conflict occurs as these meetings take place. As such, diversity naturally precedes task-related conflict. In fact, leading the task-related conflict measure could lead to spurious results as the amount of task-related conflict in a year is contingent on the diversity of the board of that year, not on the diversity of the board in previous years.

Table 6
Random-effects GLS regression results independent variables on mediating variable.

Model	DV = Task-related conflict			
	1		2	
	Coef.	p-value	Coef.	p-value
Gender Diversity (H1a +)			0.007	0.004
			(0.002)	
Educational Diversity (H1b -)	0.001	0.899	0.001	0.532
	(0.007)		(0.002)	
Industry Diversity (H1c -)	−0.002	0.910	−0.004	0.445
	(0.015)		(0.006)	
Ideological Diversity	0.005	0.457	0.002	0.405
	(0.006)		(0.002)	
Age Diversity	−0.004	0.663	−0.003	0.250
	(0.010)		(0.003)	
Board Size	−3.70e-04	0.259	−3.56e-04	0.171
	(3.26e-04)		(2.38e-04)	
Organization Size	0.001	0.672	−0.001	0.446
	(0.003)		(0.001)	
Slack Resources	1.68e-06	0.511	2.10e-06	0.024
	(2.56e-06)		(2.53e-06)	
Positive Attainment Discrepancy	8.22e-06	0.749	7.34e-06	0.689
	(2.57e-05)		(2.62e-05)	
Negative Attainment Discrepancy	4.41e-05	0.265	4.40e-05	0.547
	(3.96e-05)		(4.03e-05)	
Average Meeting Duration	−6.76e-06	0.170	−6.50e-06	0.231
	(4.92e-06)		(4.87e-06)	
Constant	0.004	0.861	0.013	0.106
	(0.022)		(0.008)	
Observations	79		79	
R^2 (between)	0.351		0.592	
Root Mean Square Error	0.001		0.001	
Wald Chi^2	8.72		32.92	
P-value	0.892		0.008	
Random effects and time-fixed effects are included in all models.				

Table 7
Random-effects GLS regression results independent variables on dependent variable.

Model	DV = Organizational performance			
	3		4	
	Coef.	p-value	Coef.	p-value
Task-related conflict (H2 +)			4887.199	0.002
			(1550.808)	
Gender Diversity	199.599	0.000	167.689	0.000
	(29.342)		(29.241)	
Educational Diversity	−30.797	0.265	−37.252	0.150
	(27.623)		(25.907)	
Industry Diversity	−100.013	0.169	−79.200	0.246
	(72.752)		(68.337)	
Ideological Diversity	−123.049	0.000	−131.550	0.000
	(27.270)		(25.638)	
Age Diversity	−91.826	0.017	−75.302	0.038
	(38.367)		(36.251)	
Board Size	0.489	0.756	1.296	0.386
	(1.575)		(1.495)	
Organization Size	57.096	0.000	61.737	0.000
	(16.277)		(15.288)	
Slack Resources	−0.064	0.068	−0.093	0.006
	(0.035)		(0.034)	
Positive Attainment Discrepancy	1.033	0.019	1.099	0.008
	(0.440)		(0.412)	
Negative Attainment Discrepancy	−2.285	0.000	−2.432	0.000
	(0.652)		(0.612)	
Average Meeting Duration	0.093	0.141	0.121	0.042
	(0.063)		(0.060)	
Constant	−314.406	0.002	−376.357	0.000
	(102.394)		(97.729)	
Observations	79		79	
R^2 (between)	0.851		0.890	
Root Mean Square Error	16.174		15.122	
Wald Chi^2	231.80		275.12	
P-value	>0.001		>0.001	
Random effects and time-fixed effects are included in all models.				

mechanism by which we can partly explain the effects of board gender diversity. Although we find support for the partial mediating effect of task-related conflict in the board gender diversity–organizational performance relationship, one particular striking result is that gender diversity remains to positively affect organizational performance even when accounting for task-related conflict (83% of the total effect of gender diversity on organizational performance is direct, thus not mediated). As such, our paper shows that task-related conflict is an important mechanism by which gender diversity affects organizational performance, yet it is by far the only mechanism through which gender diversity affects organizational performance.

One of these possible mechanisms may be that female and male directors voice unique ideas and perspectives (Post & Byron, 2015), but also that female leaders can create a gender supportive climate (Hoobler et al., 2018). Other possible reasons that have been put forward are that of an increased pool of talented directors when both male and female directors are considered and that female directors may be motivated to outperform men because of their perceived discrimination at the work floor (Erhardt et al., 2003). Future work empirically testing these mechanisms is warranted, so that we are able to build theory on *why* we find consistent positive effects of board gender diversity on organizational performance.

Other remarkable results concern our study's non-findings for the other diversity measures. This study found no evidence for the relationship between educational level diversity and task-related conflict. Apparently, this type of diversity in this context does not have a significant impact on the level of task-related conflict in the board of directors. An explanation might be, that this type of diversity does not provide different perspectives on the tasks to be performed. For example, a different level of education may provide a different knowledge base, but is not necessarily an assurance for varying views and opinions on the task. However, a distinction in type of education (e.g., business, law, biology, etc.) might show different results, since varying areas of expertise could imply a diversity in perspectives (Johnson et al., 2013).

Furthermore, we found no evidence that industry background diversity induces task-related conflict. In previous board diversity literature, industry background diversity has been found to influence organizational performance (e.g. Kor & Sundaramurthy, 2009), yet it has not been claimed to have an impact on task-related conflict. This study also did not find evidence for a direct influence of industry background diversity on organizational performance. Coupled with the positive findings of Kor and Sundaramurthy (2009), we surmise that different – and in our data, countervailing – mechanisms influence how industry background diversity affects organizational performance. For instance, from an upper echelon theory standpoint, industry background diversity could affect performance as it enables the board to tap onto a diversity of perspectives. Yet, simultaneously, high industry background diversity implies a dispersion of interests. According to Kor (2006), reaching consensus is difficult when interests are diverse and conflicting. Particularly, with the allocation of means, this could affect the firm performance negatively (Bourgeois, 1980).

Another reason why both educational level diversity and industry background diversity non-significantly impact task-related conflict and organizational performance may be that the biased information sampling model of Stasser and Titus (1985) is in fact very relevant for both these forms of diversity on the board level.

Table 8
Mediation test results.

	Sobel Test									
	z		P							
Gender diversity	2.380		0.017			*Mediation supported*				
	Baron and Kenny's stepwise approach									
	Effect of X on M			Effect of X on Y			Effect of M on Y			
	Coef.	T	P	Coef.	T	P	Coef.	T	P	
Gender diversity	0.007	2.91	0.004	167.689	5.73	0.000	4887.199	3.15	0.002	*Mediation supported*
	Bootstrapped estimate									
	Direct effects		Indirect effects		95% bias-corrected accelerated CI			Total effects		
	B_{YX}		(B_{YM} * B_{MX})		LLCI - ULCI			B_{YX} + (B_{YM} * B_{MX})		
Gender diversity	167.689		34.210		(3.344–109.626)			201.899		*Mediation supported*

Table 9
Robustness checks.

Model	Accounting for the false hit rates in the seemingly unrelated regression models				Leading the DV in a random-effects GLS regression	
	DV = Task-related conflict		DV = Organizational performance		DV = Organizational performance	
	5		6		7	
	Coef.	p-value	Coef.	p-value	Coef.	p-value
Task-related conflict			4887.199	0.005	5435.909	0.019
			(1760.301)		(2316.627)	
Gender Diversity	0.009	0.000	167.689	0.000	192.590	0.000
	(0.002)		(33.215)		(38.955)	
Educational Diversity	−0.001	0.688	−37.252	0.238	−44.652	0.201
	(0.002)		(31.597)		(34.935)	
Industry Diversity	−0.005	0.203	−79.200	0.343	−78.131	0.403
	(0.004)		(83.524)		(93.406)	
Ideological Diversity	−0.001	0.548	−131.550	0.000	−102.467	0.005
	(0.002)		(29.198)		(36.425)	
Age Diversity	−0.003	0.205	−75.302	0.080	−88.400	0.066
	(0.003)		(42.969)		(48.008)	
Board Size	−4.12e-06	0.979	1.296	0.465	−1.183	0.576
	(1.57e-04)		(1.774)		(2.117)	
Organization Size	−0.002	0.105	61.737	0.002	84.498	0.000
	(0.001)		(20.096)		(22.082)	
Slack Resources	7.18e-06	0.011	−0.093	0.009	−0.184	0.000
	(2.82e-6)		(0.035)		(0.052)	
Positive Attainment Discrepancy	−1.69e-05	0.648	1.099	0.023	−0.141	0.785
	(3.7e-05)		(0.483)		(0.518)	
Negative Attainment Discrepancy	2.62e-05	0.482	−2.432	0.019	−0.107	0.888
	(3.72e-05)		(1.033)		(0.762)	
Average Meeting Duration	−1.33e-05	0.003	0.121	0.043	0.221	0.008
	(4.49e-06)		(0.060)		(0.084)	
Constant	0.020	0.007	−376.357	0.001	−471.984	0.000
	(0.007)		(117.897)		(133.541)	
Observations	79		79		64	
R^2 (between)	0.498		0.819		0.897	
RMSE	0.001		13.288		18.071	
(Wald) Chi^2	41.94		356.30		135.68	
P-value	>0.000		>0.000		>0.000	
Random effects and time-fixed effects are included in all models.						

The board diversity literature has mainly assumed that information will be shared (although there will be social categorization), but has not engaged with the work on biased information sampling. In fact, although diversity would ensure different perspectives will be present, understanding the conditions under which these perspectives will be shared is essential. This is generally conceptualized as the openness of information sharing and is indeed important for information sharing (Mesmer-Magnus & DeChurch, 2009). Studying both uniqueness and openness in information sharing is thus warranted (Mesmer-Magnus & DeChurch, 2009).

As suggested by Dahlin et al. (2005), diversity is inconsistent in its effect on organizational performance. Some types of diversity affected organizational performance positively, while others had no effect or even a negative influence. For instance, age diversity showed a negative relationship with both task-related conflict and organizational performance. Strong disagreement could lead the board to never reach consensus on issues, which is negative for performance in this type of stable environment (Priem, 1990). However, this study has not found evidence for a mediating relationship of task-related conflict. Age diversity might have a negative influence on social cohesion within the board, leading to decreased performance.

According to this study, age diversity negatively affects task-related conflict. The explanation for this relationship could reside in the fact that the average director age in this context is very high (the maximum number of directors in a board with an age of less than 40 years is three). Johnson et al. (2013) previously mentioned that age cohorts are probable to have varying opinions and values, because the cohorts' values have been shaped by different environments and these values also develop by maturation.

This study contributes to the literature in two ways. First, the mediating role of conflict in the board gender diversity–organizational performance relationship has not been empirically investigated before. We found task-related conflict is a powerful mechanism by which board gender diversity impacts organizational performance, our study also shows that this is not the only mechanism. As others have vouched for before: it remains imperative to understand the mechanisms by which board gender diversity impacts firm outcomes (Hoobler et al., 2018; Johnson et al., 2013; Lawrence, 1997).

Second, the method of measuring task-related conflict was novel and different from previous literature on conflict. Previously, scholars measured conflict in boards of directors with questionnaires among directors. However, this type of measurement is sensitive for biased results. Content analysis offers an unobtrusive solution to this complication. The developed dictionary is applicable in other settings as well, offering a new way to measure task-related conflict.

5.1. Limitations and recommendations for future research and practice

The study has limitations as well. First, the most important limitation is the size of the panel data. Second, the context of this study is specific. Therefore, the generalizability to other industries remains an empirical question. Future researchers should further investigate the mechanisms by which board diversity affects organizational performance. First, the 'black box' with the underlying mechanisms between board diversity and organizational performance needs to be explored more and in different ways to understand these mechanisms. Second, national culture also plays a role. National culture could have an effect in the underlying mechanisms between board diversity and organizational performance. For instance, in a culture that is very open to discussion, task-related conflict will rarely be associated with negative emotions. However, in other countries the level of task-related conflict could have a higher impact on negative emotions between directors, resulting in relationship conflict. Eventually, this could decrease group cohesion and organizational performance (Westphal & Bednar, 2005).

This research also offers important implications for practice. Possibly, the most important lessons learned, are that gender diversity is positively related with organizational performance through increased task-related conflict. Therefore, the selection of directors can also be seen as strategically important to firms. Since gender diversity is positively related to organizational performance, it could be recommended to ensure a certain level of female representation in the board. This could be accomplished by stimulating women to participate in the board. In the case of water management authorities, the government or the authorities could focus more on the representation of women in the board and approaching them to consider a board position. Our results also provide further evidence that gender enforced policies (e.g. Senden & Kruisinga, 2018) might indeed be advantageous for businesses and society as it allows for task-related conflict, which allows the organization to attain superior performance.

6. Conclusion

This study contributed to the literature in several ways, yet the most important contribution is that task-related conflict is indeed a mechanism by which board gender diversity affects organizational performance. Relatedly, we found that task-related conflict is not a mechanism for other types of diversity. These findings help us redirect the literature in suggesting that more mechanisms affect the board gender diversity–organizational performance relationship and that other types of diversity may affect organizational performance through different mechanisms than task-related conflict. Future research is warranted to include different mechanisms to examine the type of relationship to further explore the 'black box' of the board diversity–organizational performance relationship.

Acknowledgements

This research was supported by a Netherlands Organization for Scientific Research (NWO) Research Talent grant, #406-12-122, and by the Tilburg School of Economics and Management Alumni Fund.

References

Abrahamson, E., & Eisenman, M. (2008). Employee-management techniques: Transient fads or trending fashions? *Administrative Science Quarterly, 53*(4), 719–744.
Adams, R. B., & Funk, P. (2012). Beyond the glass ceiling: Does gender matter? *Management Science, 58*(2), 219–235.
Ahn, S. C., Lee, Y. H., & Schmidt, P. (2013). Panel data models with multiple time-varying individual effects. *Journal of Econometrics, 174*(1), 1–14.
Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of Management Journal, 39*(1), 123–148.
Bantel, K. A., & Jackson, S. E. (1989). Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management Journal, 10*(S1), 107–124.
Barkema, H. G., & Shvyrkov, O. (2007). Does top management team diversity promote or hamper foreign expansion? *Strategic Management Journal, 28*(7), 663–680.
Berger, J., Conner, T., & Fisek, M. H. (1974). *Expectation states theory: A theoretical research program*. Cambridge, MA: Winthrop.
Blau, P. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York: Free Press.
Boivie, S., Bednar, M. K., Aguilera, R. V., & Andrus, J. L. (2016). Are boards designed to fail? The implausibility of effective board monitoring. *The Academy of Management Annals, 10*(1), 319–407.
Bourgeois, L. J. (1980). Performance and consensus. *Strategic Management Journal, 1*(3), 227–248.
Campbell, J. T., Campbell, T. C., Sirmon, D. G., Bierman, L., & Tuggle, C. S. (2012). Shareholder influence over director nomination via proxy access: Implications for agency conflict and stakeholder value. *Strategic Management Journal, 33*(12), 1431–1451.
Carpenter, M. A., & Westphal, J. D. (2001). The strategic context of external network ties: Examining the impact of director appointments on board involvement in strategic decision making. *Academy of Management Journal, 4*(4), 639–660.
Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice Hall.
Cyert, R. M., & March, J. G. (1992). *A behavioral theory of the firm* (2nd ed.). Cambridge, MA: Blackwell.
Dahlin, K. B., Weingart, L. R., & Hinds, P. J. (2005). Team diversity and information use. *Academy of Management Journal, 48*(6), 1107–1123.
Daily, C. M., Dalton, D. R., & Cannella, A. A. (2003). Corporate governance: Decades of dialogue and data. *Academy of Management Review, 28*(3), 371–382.
De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology, 88*(4), 741–749.
De Wit, F. R., Greer, L. L., & Jehn, K. A. (2012). The paradox of intragroup conflict: A meta-analysis. *Journal of Applied Psychology, 97*(2), 360–390.
Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal, 33*(9), 1072–1089.
Dovidio, J. F., Brown, C. E., Heltman, K., Ellyson, S. L., & Keating, C. F. (1988). Power displays between women and men in discussions of gender-linked tasks: A multichannel study. *Journal of Personality and Social Psychology, 55*(4), 580–587.
Erhardt, N. L., Werbel, J. D., & Shrader, C. B. (2003). Board of director diversity and firm financial performance. *Corporate Governance: An International Review,*

11(2), 102–111.
Forbes, D. P., & Milliken, F. J. (1999). Cognition and corporate governance: Understanding boards of directors as strategic decision-making groups. *Academy of Management Review, 24*(3), 489–505.
Frijns, B., Dodd, O., & Cimerova, H. (2016). The impact of cultural diversity in corporate boards on firm performance. *Journal of Corporate Finance, 41*, 521–541.
Gibson, C., & Vermeulen, F. (2003). A healthy divide: Subgroups as a stimulus for team learning behavior. *Administrative Science Quarterly, 48*(2), 202–239.
Gigone, D., & Hastie, R. (1993). The common knowledge effect: Information sampling and group judgment. *Journal of Personality and Social Psychology, 65*(5), 959–974.
Goodstein, J., Gautam, K., & Boeker, W. (1994). The effects of board size and diversity on strategic change. *Strategic Management Journal, 15*(3), 241–250.
Hambrick, D. C. (2007). Upper echelon theory: An update. *Academy of Management Review, 32*(2), 334–343.
Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review, 9*(2), 193–206.
Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review, 32*(4), 1199–1228.
Heemskerk, E. M., Heemskerk, K., & Wats, M. M. (2017). Conflict in the boardroom: A participant observation study of supervisory board dynamics. *Journal of Management & Governance, 21*(1), 233–263.
Hoobler, J. M., Masterson, C. R., Nkomo, S. M., & Michel, E. J. (2018). The business case for women leaders: Meta-analysis, research critique, and path forward. *Journal of Management, 44*(6), 2473–2499.
Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly, 40*(2), 256–282.
Jehn, K. A. (1997). A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative Science Quarterly, 42*(3), 530–557.
Jehn, K. A., & Bendersky, C. (2003). Intragroup conflict in organizations: A contingency perspective on the conflict-outcome relationship. *Research in Organizational Behavior, 25*(1), 187–242.
Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field stury of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly, 44*(4), 741–763.
Johnson, S. G., Schnatterly, K., & Hill, A. D. (2013). Board composition beyond independence: Social capital, human capital, and demographics. *Journal of Management, 39*(1), 232–262.
Kor, Y. Y. (2006). Direct and interaction effects of top management team and board compositions on R&D investment strategy. *Strategic Management Journal, 27*(11), 1081–1099.
Kor, Y. Y., & Sundaramurthy, C. (2009). Experience-based human capital and social capital of outside directors. *Journal of Management, 35*(4), 981–1006.
Lant, T. K., & Montgomery, D. B. (1987). Learning from strategic success and failure. *Journal of Business Research, 15*(6), 503–517.
Lawrence, B. S. (1997). Perspective - the black box of organizational demography. *Organization Science, 8*(1), 1–22.
Li, J., & Hambrick, D. C. (2005). Factional groups: A new vantage on demographic faultlines, conflict, and disintegration in work teams. *Academy of Management Journal, 48*(5), 794–813.
Lu, L., Yuan, Y. C., & McLeod, P. L. (2012). Twenty-five years of hidden profiles in group decision making: A meta-analysis. *Personality and Social Psychology Review, 16*(1), 54–75.
Mesmer-Magnus, J. R., & DeChurch, L. A. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology, 94*(2), 535–546.
Miller, T., & Triana, M.d. C. (2009). Demographic diversity in the boardroom: Mediators of the board diversity–firm performance relationship. *Journal of Management Studies, 46*(5), 755–786.
Milliken, F. J., & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of Management Review, 21*(2), 402–433.
Opstrup, N., & Villadsen, A. R. (2014). The right mix? Gender diversity in top management teams and financial performance. *Public Administration Review, 75*(2), 291–301.
Pelled, L. H. (1996). Demographic diversity, conflict, and work group outcomes: An intervening process theory. *Organization Science, 7*(6), 615–631.
Pelled, L. H., Eisenhardt, K. M., & Xin, K. (1999). Exploring the black box: An analysis of work group diversity, conflict and performance. *Administrative Science Quarterly, 44*(1), 1–28.
Post, C., & Byron, K. (2015). Women on boards and firm financial performance: A meta-analysis. *Academy of Management Journal, 58*(5), 1546–1571.
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.
Priem, R. L. (1990). Top management team group factors, consensus, and firm performance. *Strategic Management Journal, 11*(6), 469–478.
Rosener, J. B. (1995). *America's Competitive Secret: Utilizing women as a management strategy*. New York: Oxford University Press.
Rosette, A. S., & Tost, L. P. (2010). Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders. *Journal of Applied Psychology, 95*(2), 221–235.
Senden, L., & Kruisinga, S. (2018). *Gender-balanced company boards in Europe*. Retrieved from.
Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422–445.
Stasser, G., Taylor, L. A., & Hanna, C. (1989). Information sampling in structured and unstructured discussions of three- and six-person groups. *Journal of Personality and Social Psychology, 57*(1), 67–78.
Stasser, G., & Titus, W. (1985). *Journal of Personality and Social Psychology, 48*(6), 1467–1478, 1985.
Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin, & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
Triana, M.d. C., Miller, T. L., & Trzebiatowski, T. M. (2013). The double-edged nature of board gender diversity: Diversity, firm performance, and the power of women directors as predictors of strategic change. *Organization Science, 25*(2), 609–632.
Van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology, 58*, 515–541.
Wade, J. B., Porac, J. F., & Pollock, T. G. (1997). Worth, words, and the justification of executive pay. *Journal of Organizational Behavior, 18*(S1), 641–664.
Wall, V. D., & Nolan, L. L. (1986). Perceptions of inequity, satisfaction, and conflict in task-oriented groups. *Human Relations, 39*(11), 1033–1051.
Water Authority Act. (1991). Retrieved from the Dutch law site. https://wetten.overheid.nl/BWBR0005108/2018-07-01.
van Waterschappen, U. (2018). Waterschapspeil 2018: Trends en ontwikkelingen in het regionale waterbeheer. Retrieved from 26th July 2019 https://www.waterschapsspiegel.nl/bedrijfsvergelijkingen/.
Watson, W. E., Kumar, K., & Michaelsen, L. K. (1993). Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups. *Academy of Management Journal, 36*(3), 590–602.
Westphal, J. D., & Bednar, M. K. (2005). Pluralistic ignorance in corporate boards and firms' strategic persistence in response to low firm performance. *Administrative Science Quarterly, 50*(2), 262–298.
Westphal, J. D., & Milton, L. P. (2000). How experience and network ties affect the influence of demographic minorities on corporate boards. *Administrative Science Quarterly, 45*(2), 366–398.
Williams, K. Y., & O'Reilly, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. *Research in Organizational Behavior, 20*, 77–140.
Wittenbaum, G. M. (1998). Information sampling in decision-making groups. *Small Group Research, 29*(1), 57–84.
Zajac, E. J., & Westphal, J. D. (1996). Director reputation, CEO-board power, and the dynamics of board interlocks. *Administrative Science Quarterly, 41*(3), 507–529.
Zona, F., & Zattoni, A. (2007). Beyond the black box of demography: Board processes and task effectiveness within Italian firms. *Corporate Governance: An International Review, 15*(5), 852–864.