GEDII GENDER DIVERSITY IMPACT – Improving research and innovation through gender diversity

WWW.GEDII.EU

Project Deliverable

D1.1 Conceptual Framework

31st of May 2016 – **Version 1.0**

Project Consortium

Universitat Oberta de Catalunya, Spain

Project Coordination

Dr. Jörg Müller
Dr. Milagros Sáinz
Dr. Julio Meneses
Dr. Sergi Fàbregues
Dr. Rosa Borge
Dr. Julià Minguillón

Hochschule Furtwangen, Germany

Prof. Dr. Ulrike Busolt
Sandra Klatt
Wiebke Kronsbein

Örebro University, Sweden

Prof. Dr. Ulf Sandström
Dr. Anne-Charlott Callerstig
Prof. Dr. Liisa Husu

Verband der Elektrotechnik Elektronik Informationstechnik e.V., Germany

Dr. Walter Börmann
Nicole Bienge
Andreas Landwehr
Annika Gereke

Cranfield University, UK

Dr. Anne Laure Humbert

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 665851.

Table of Contents

Introduction

The following document presents the conceptual framework of the GEDII project. As such it reviews the literature roughly pertaining to three quite separate knowledge fields, namely Team Science (Chapter 1), Research Assessment (Chapter 2) and Gender Research (Chapter 3). Since each of these areas is huge in terms of publications and thematic subdivisions, the challenge to articulate their potential overlaps is not an easy task, let alone complete. Overall, we have proceeded to review first the literature on the main factors that affect team performance, followed by the different ways that research is currently assessed and evaluated. The third chapter then aims to flesh out the gender dimensions of team science and research assessment in order to arrive at an understanding how gender diversity specifically affects team performance. The fourth chapter finally spells out the methodological implications for our case study research, the development of the Gender-Diversity-Index and the survey.

In a nutshell, gender diversity influences team performance along two dimensions. First, in a negative sense, gender bias can prevent research groups from utilizing the available expertise in their teams. Status differentials as well as formal power relations, homophily or sexual harassment among group members all potentially undermine optimal sharing of information. Second, gender diversity can also improve group performance by fostering conversational turn-taking and facilitating a favorable team climate. Some research has shown that a more balanced participation of all team members improves team performance. In addition, gender equality and diversity policies facilitate a more inclusive environment which also can strengthen individual capacity, motivation and hence performance. As this basic setting already suggests, we do not expect a simple, straightforward relation between the proportion of women and men in teams and their performance. Rather on the contrary: as the literature review on team science demonstrates, there are many more variables besides gender diversity that complexify our understanding of how research performance can be twisted and tuned. Thus, one of the main challenges of the conceptual framework consists of providing a plausible and encompassing configuration of variables that affect research performance without loosing sight of the effects of gender diversity among all these factors.

This immediately conjures up another challenge: research on teams itself is constantly evolving. For example, there is a wealth of publications regarding the effect of personality traits, motivational factors, the role of conflict, the importance of emotions, trust and psychological safety for the internal workings of teams. Equally, the impact of collaboration networks, team boundary spanning or the role of leadership have been very productive areas of research for decades. However, there is no reason to assume that this list is complete. Unresolved issues due to theoretical but also methodological constraints concern for example the role of time in team processes. Departing from a more or less static understanding of teams, researchers increasingly advocate a more dynamic, "organizing" perspective that is sensitive to the "differential team needs over time or across situations, differential team member capabilities to deal with team situations and therefore fluid set of power, status, interdependencies"[1]. And of course, gender studies as well as research

1 Humphrey SE and Aime F (2014) Team Microdynamics: Toward an Organizing Approach to Teamwork. The Academy of Management Annals, Routledge 8(1): 443–503.

assessment are just as dynamic and evolving fields as the research on teams. How "performance" is measures beyond and with publication counts, citation impacts is a hotly debated question, last but not least from feminist perspectives questioning the "metric tide" of the neo-liberal university. Hence, the question arises which contribution can GEDII possible make to these research fields, let alone all three?

The primary aim of GEDII is not to go deeper into any of these fields and contribute to any of these single topics. The objective cannot be to add substantially to the literature on trust, leadership or any of these existing research programs. Rather, the objective is to systematize the already existing evidence regarding the potential impact of gender diversity on team performance and integrate the available insights from gender studies, team science, and research assessment into a coherent, new methodology. GEDII aims, in short, to provide the means for downstream research to address the relation between gender diversity and research performance more reliably.

By synthesizing the available evidence regarding the impact of gender diversity on teams and molding it into specific research instruments, we aim to go beyond the prevailing conceptualization of gender diversity as counting women and men in research teams. The aims is precisely to provide a more complex but nevertheless fairly easy to use instrument that operationalizes in a more holistic manner the crucial dimensions of gender diversity in the context of research teams.

The conceptual framework feeds informs the work of subsequent workpackages. On the one hand it informs the design and execution of the Case Studies in workpackage 2. As chapter four will document in more detail, the case studies aim for a methodological contribution regarding a more dynamic understanding of teams. Sociometric badges can provide micro-second details of team interaction and non-verbal language among co-workers that bring without our reach the type of data necessary to observe teams as a dynamic system. It is exciting to explore these methodological advancements from the start incorporating a gender perspective.

On the other hand, the conceptual framework also feeds into the development of the Gender-Diversity-Index in WP3. The Index is conceived as a research instrument to measure the available gender diversity within research teams in a more sophisticated manner than simply determining the proportion of women. After having been deployed and tested in WP4, it should provide other researchers with the necessary knowledge and instrumental tools to integrate the gender diversity dimension into their research on teams.

GEDII GENDER DIVERSITY IMPACT – Improving research and innovation through gender diversity

WWW.GEDII.EU

D1.1. Conceptual Framework

Chapter 1 – Team Science

Jörg Müller

31st of May 2016

Table of Contents

1. Team science

An increasing amount of work in science, research and development happens collaboratively. Knowledge production in science has left behind the single author in favor of collaborative networks of co-authors – if we follow the general consensus that joint publications are a valid approximation of collaborative efforts (Katz & Martin, 1997; Wagner & Leydesdorff, 2005; Wuchty, Jones, & Uzzi, 2007). Collaboration is both an advantage and a necessity. Collaboration is first a necessity because scientific knowledge is ever expanding via sub-divisions and specialized knowledge fields while important societal and global challenges require the re-integration of expertise across disciplinary boundaries. Team-science in its US-flavor stresses particularly large scale collaborations involving 50 to 200 investigators where unidisciplinary approaches are overcome in favor of interdisciplinary work and transdisciplinary integration necessary for tackling complex challenges (Stokols, Hall, Taylor, & Moser, 2008).[1] Second, collaboration is seen as an advantage, because the locus of innovation has moved beyond individual persons or organizations towards the respective networks they are embedded in. Independent of disciplinary boundaries, networks between people, teams, organizations or even nations contain diversity that yields innovative results via recombinatory efforts of its elements (Savino, Messeni Petruzzelli, & Albino, 2015). In the case of science, "recombination" can happen in the form of collaborative efforts between scholars that have been shown to affect quantity and quality of outcomes (Figg et al., 2006; Lee & Bozeman, 2005; Rigby & Edler, 2005). There is evidence that high research performance – in terms of high publication outputs – is positively associated with high levels of collaborations (Katz & Martin, 1997). Besides specialization and performance issues, resource sharing, new enabling technologies or new public policies have also been identified as the driving factors behind the "collaboration imperative" (Bozeman & Boardman, 2014).

Although collaboration can happen between single scientists across research organizations, "teams" provide an important intermediate organizational unit that resides in-between the individual and organizational level. Teams provide the structural unit through which many joint efforts happen, offering as such a rich analytic perspective upon a central unit of knowledge production in contemporary science.

1.1. What is a "team"?

As a minimal working definition, a team is conceived as a set of interdependent individuals that have a shared goal (Guzzo & Dickson, 1996; Sundstrom, de Meuse, & Futrell, 1990). That is to say that team members have more or less differentiated roles and rely on the contribution of other team members for accomplishing a common mission. Teams come in all types and shapes: their size, boundaries and life-span is flexible, ranging from relatively loose, ad-hoc humanitarian aid teams to more stable teams in routinized and complex work processes such as pilots or nurses. Basic taxonomies of team types draw a distinction between "action" oriented teams such as sports, assembly line or military combat teams

1 For the purpose of the current text, team-science however refers not just to this more recent US approach to but research on teams in general with important traditions in the management literature, the social psychological research on small groups and organization studies among many others.

which integrate physical joint action versus “decision” making teams that integrate primarily knowledge and information such as management teams (Sundstrom et al., 1990). However, there exists a “perplexing overabundance of conflicting taxonomies for classifying teams in discrete and categorical ways” (Lee, Koopman, Hollenbeck, Wang, & Lanaj, 2015, p. 5). The distinction between “action oriented” vs. “decision oriented” teams is just one among many other taxonomies used in research. Hollenbeck, Beersma, & Schouten (2012) in fact review as many as 42 different classifications, highlighting the difficulties this implies in terms of building a cumulative knowledge base since idiosyncratic definitions prevent the aggregation of insights across studies. In order to remedy this situation, the authors extract three basic descriptive dimensions that underlie most of the reviewed classification systems and that situate teams in a continuous three dimensional space according to their (1) skill differentiation, (2) authority differentiation, and (3) temporal stability (ibid.). As we will see, although some characteristics of research teams and their processes can be sufficiently captured with these dimensions, they need to be complemented in some other aspects, outlined in the following paragraphs.

Looking at science and research teams more specifically, one can observe certain distinct features. First, research teams tend to operate under high task complexity and high degree of uncertainty. Considering the nature of the task involved is especially important for assessing the benefits of diversity for scientific teams.

> *“That is, group performance should benefit primarily from diversity on complex and nonroutine information-processing and decision-making tasks where performance is primarily defined in terms of the quality of the group’s product, such as those facing research and development teams.” (van Knippenberg et al 2004, p. 1012)*

Where problems are by definition ill-defined, the search for solutions is an open, non-deterministic process that benefits from multiplicity of perspectives, approaches and ideas. Diversity in whatever form and shape is indeed a much sought after resource in the case of scientific- or innovation work. It is under these circumstances of high task complexity that the potential benefits of (gender) diversity can become apparent (Bowers, Pharmer, & Salas, 2000; Jehn, Northcraft, & Neale, 1999). Task complexity evokes in this context the “skill differentiation” advocated by Hollenbeck et al. (ibid.) since “skill” is conceived very widely, i.e. including any unique perspective that team members bring to the group based upon their differences in terms of education, experience, culture, or gender. Different views and diverse skills can shine when put to work on the resolution of complex tasks, an argument that will be further developed in the context of a structural perspective on diverse networks. However, it also is true that diversity is not always beneficial to team work: in situations where task complexity is low and “successful” performance of a team is dependent upon following well established work routines and procedures, diversity undermines rather than drives team performance (Joshi & Roh, 2007, 2009). Having to manage multiple viewpoints that question well established work procedures subverts rather than supports the accomplishment of routine tasks. In this sense, research teams are expected to score high on skill differentiation which is expected to produce a positive effect when working on complex tasks.

A second distinct characteristics of research teams concerns their relative high level of autonomy and self-governance especially in the public sector (Wang & Hicks, 2015). Hollenbeck and co-authors capture this crucial dimension of teams with their second dimension, namely authority differentiation. Research teams in public organizations usually include a team leader who has some level of formal authority over team members but otherwise are organized in a relatively flat hierarchical structure, often using democratic decision making among its members. Although less and less true, research was (is) relatively protected from the constraints of the market economy, politics or other spheres of society (Ball, 2012; Slaughter & Rhoades, 2000). Teams and team members enjoy in principal a high autonomy to develop ideas and carry out their research that makes long, hierarchical chain of command structures unnecessary. Herein might lie an important distinguishing factor to research teams in the private sector where the vertical hierarchy can be more pronounced. At the same time, however, relations of authority and power are not just characteristic of research teams but also important from a gender perspective. As the following paragraphs and chapters will show, differences in access to power between women and men introduces an important bias what information gets shared in teams that has consequences for team performance.

A third important dimension of teams concerns their temporal stability. Especially in research teams new members such as Phd students or visiting scholars are continuously entering or leaving. The relative stability of group membership is important because it correlates with central conditioning factors of team performance such as team identification, the development of a shared vision, norms and motivation (Meeussen & van Dijk, 2015). Temporal stability of course gives also important insights regarding the influx of new ideas as when new members join the group or the availability or loss of knowledge and expertise when members leave it.

Although these three dimensions capture important features of research teams and teams more broadly, it does not address the issue of team membership itself. Where and when to draw the line between team members and nonmembers is neither an easy nor a trivial question especially in relation to assessing team performance! Depending on the formal criteria of team membership used, one achieves smaller or larger networks of people with possibly drastic consequences for resulting performance scores. What more, the preoccupation with team boundaries actually runs much deeper. As Mortensen (2012) argues, research on teams has lost to a large degree touch with the fluid and changing boundaries of real world teams that seldom correspond to the neat, clearly defined and stable work groups that most theories presuppose. Team boundaries might be *changing* with members arriving or leaving, they might be *overlapping* with team members working in multiple teams simultaneously, or team boundaries even might be *disagreed upon*, when team members do not coincide whom to count in- or out.

As Mortensen recognizes, taking these shifting boundaries into account in empirical research can be quite challenging and costly (ibid.). In addressing this issue it is helpful to consider one of the principal objective of GEDII, namely to better understand the impact of (gender) diversity on research performance. As the literature review suggests, one can distinguish between two sources of diversity that affect team performance: team internal diversity and diversity inherent in the network of external collaborations. However, this simply suggests to

define team membership along similar lines, namely to distinguish between work carried out among members of the core team that are embedded in the same organization vs. more informal and fluid collaborations beyond the organizational boundary of the team. No matter if the team operates within a private company or a public research institute, the wider organizational context provides a more binding context to team members than simple collaborations. In contrast to external collaborations, team membership does not simply depend upon a common interest but some form of organizational anchorage, such as labor contracts, shared funding and hence responsibilities, or formal recognition by authorities of team activity. Research collaborations, insofar they are distinct from the collaboration within teams, are usually less formalized, depending rather on the voluntary exchange between the involved parties. In other words, persons form part of a specific team if they have some form of organizational binding such as labor contract, Phd stipend, fellowship. As a consequence of this clear membership criteria, external, *less formal and more voluntary and flexible forms of "collaboration"* activities are clearly defined as well: everybody that does not have an organizational anchor in a team. Thus, a more precise definition of teams that integrates an organizational perspective emerges. A team comprises:

> "(a) two or more individuals who (b) socially interact (face-to-face or, increasingly virtually) (c) possess one or more common goals; (d) are brought together to perform organizationally relevant tasks; (e) exhibit interdependencies with respect to workflow, goals, and outcomes; (f) have different roles and responsibilities; and (g) are together embedded in an encompassing organizational system, with boundaries and linkages to the broader system context and task environment" (Kozlowski & Ilgen, 2006)

This emphasis on the organizational binding of team membership makes especially sense in the context of the present project out of two reasons: first, in order to capture "invisible" contributions of group members not listed as author of publications and second, to take into account the effect of organizational gender equality and diversity policies on team work.

First, an organizational team membership criteria allows us to go beyond the bias of many bibliometric studies that define membership through co-authorship. As Katz & Martin (1997, p. 3) suggest, "bibliometric analysis of multiple-author papers can only be used as partial indicator of collaborative activity". A recent study by Conti & Liu (2015) confirms this point by showing that inferring scientific output from publications author list leads to severe bias. In their analysis of the personnel composition of the MIT Biology Department over the period from 1966-2000 they show that experience and external funding are critical determinants of labs productivity. Focusing on high-profile / breakthrough publications, several interesting results emerge that emphasize the contribution of technicians and graduate students:

> *"First, graduate students, who make only nominal contributes to overall contribution counts, contribute as much to breakthroughs as postdocs with external funding. Second, postdocs without fellowships have no observable impact on breakthrough publications. In a final intriguing finding, technicians are instrumental to high-profile publications, but have no observable impact on lower impact publication output" (Conti & Liu, 2015)*

“Invisible” persons working in the lab may be key for production high impact publications, although not listed. Unfortunately, the paper does not examine the sex distribution of the staff nor possible differences in scientific output.[2] However, it makes the important point that staff categories and even “invisible” workers in the laboratory are important determinants of scientific production that have to be taken into account – especially from a gender perspective which is sensitive to the vertical gender segregation across staff categories! By defining team membership through an organizational criteria, we will be able to capture these important but often invisible contributions to scientific work.

The second reason to base team membership on an organizational criteria side by side with wider and informal collaborations has to do with the importance of the organizational context for team work. Team membership binds actors together in a more formal way where team members are “stuck” with each other, their superiors or the institution – which might imply quite divergent emotional investments, political- and power struggles and social dynamics than engaging in more informal collaborations. Actors are bound together by the organizational context whereby this context at the same time can become an important driving factor for shaping team work itself. Gender equality- and diversity policies constitute key elements for recognizing and valuing the contributions of any type of minority within an organization. These policies thus influence team work and provide an additional layer to the social dynamics and interactions that happen among team members but not necessarily wider research collaborations. Overall, we conceive of “teams” as a form of research collaboration that constitutes an intermediate level where organizational factors mold research collaborations between individuals.

In a final step it also has to be noted, that using an organizational criteria for team membership is compatible with the current shifting forms of collaborative work where co-workers do not necessarily share the same physical location. Although it seems plausible to assume that external, informal and more sporadic collaborations happen more through electronic communication than exchanges among core team members in face-to-face situations, this is not necessarily a good indicator to distinguish “virtual” from “real” teams. As Foster et al. (2015) argue, electronic communication is used in teams whose members might be working door by door or in the same lab. A more promising – in fact a fourth dimensions to Hollenbecks (see above) proposal on (1) skill differentiation, (2) authority differentiation and (3) temporal stability - is (4) “virtuality”, now simply measuring the physical or geographic distance between team members at work. The “virtual” dimension emphasizes the importance of the organizational context in the sense that even though team members are located across regions and even countries, it is the same organizational environment within which their cooperation with others develops. In other words, team membership does not depend upon co-location but physical distance nevertheless is an important characteristic to bear in mind when describing teams.

2 Unfortunately, no analysis was done regarding sex composition of the laboratory, as confirmed by email exchange with Annamaria Conti, first author of the article.

1.2. Conceptualizing the “diversity – performance” link

Scientists themselves see collaboration as an essential ingredient of their success (Leahey & Cain, 2013). Research on diversity in work groups has produced over the years a wealth of explanatory factors that detail why collaboration and team work excels – or not. Very schematically, the available accounts have focused either on the different structural configurations of diversity and their impact on performance or the social psychological factors that condition the processing of diverse resources within work groups.

In part, these explanatory accounts have gathered momentum by the insight that any straightforward explorations regarding “direct diversity-performance relationships seems to have been laid to rest” (Haas, 2010). We no longer expect a simple linear relationship between certain diversity measures and performance indicators but conceive it as a more complex relationship where a variety of contextual and moderating variable intervene (Richard & Miller, 2013). For example while the meta-analysis conducted by Stewart (2006) or Webber & Donahue (2001) neither could detect a relationship between work group diversity and “performance”, Hülsheger et al. (2009) demonstrate a moderate effect of functional diversity on “innovation”. Likewise, while some studies (Catalyst, 2010; Erhardt, Werbel, & Shrader, 2003; McKinsey & Company, 2007; Shrader, Blackburn, & Iles, 1997) found a significant positive effect between diversity and firm level performance, several meta-analytic studies coincide that there is no or only a very moderate effect while highlighting in general the inconsistency of the results (Certo, Lester, Dalton, & Dalton, 2006; Homberg, 2010; Pletzer, Nikolova, Kedzior, & Voelpel, 2015). A similar situation holds for R&D teams: although it has been argued that gender-heterogeneous working groups produce higher quality science (Campbell, Mehtani, Dozier, & Rinehart, 2013) and have a higher innovation potential (Østergaard, Timmermans, & Kristinsson, 2011; The Lehman Brothers Centre for Women in Business, 2007), these findings are not consistent (Bowers et al., 2000; Mannix & Sauer, 2006; Shore et al., 2009; van Knippenberg & Schippers, 2007; Webber & Donahue, 2001; Williams & O’Reilly, 1998), resulting in a need to develop more sophisticated and systematic accounts that explain the elements and processes by which diversity can be harnessed in teams.

The basic dialectic that unfolds the “diversity – performance” relation oscillates between the perspective of social categorization (detectable attributes such as age, sex and racio-ethnicity) vs. informational/functional diversity (personality, knowledge, values) perspectives (van Knippenberg & Schippers, 2007). Whereas the latter perspective emphasizes diversity of surface (age, sex, ethnicity) and functional (tenure, educational background, expertise) attributes increases performance, the former social categorization perspective contends that diversity increases the cost of group management due to the insider/outside dynamics of social groups (Tajfel, 1978). Put bluntly, diverse groups might hamper performance due to increased costs of coordination and negotiation between highly different members, while heterogeneous groups outperform their more homogeneous counterparts due to the more varied resources and better quality decision making available (see van Knippenberg & Schippers 2007).

Contextualizing these contrasting viewpoints on the potential benefits and drawbacks of

diversity with the wider debate on social capital, one can unfold a more systematic and integrative conceptual framework that combines the examination of a) a structural perspective on diversity with an exploration of b) its social psychological and c) cognitive dimensions (Nahapiet & Ghoshal, 1998). That is to say that the antagonistic viewpoints on the merits and problems of heterogeneous and homogeneous groups actually can be reformulated in a more network analytic terms that emphasizes the importance of both "bonding" and "bridging" social capital (Reagans & Zuckerman, 2001). In order for research teams to perform "well", they have to dispose of access to diverse resources (previously discussed as advantage of heterogeneous groups) and be able to integrate and process those resources (previously discussed as advantage of more homogeneous groups) within the existing confinements of their team and organization.

1.3. Structural Perspectives on Diversity

The following section summarizes research on the structural effects of diversity on research performance. The performance of a team is conceived as a structural / topological feature of the network it forms. Topological features of network have consequences in terms of conditioning access to resources, knowledge, funding or the speed of information dissemination (Cummings, Kiesler, Bosagh Zadeh, & Balakrishnan, 2013; Cummings, 2004). What constitutes innovative and high performing teams depends upon the position of individual actors (be that individual scientists or teams) within certain networks. Actor centrality, density of links, or nodes bridging otherwise disconnected clusters have provided valuable insights how these network parameters condition "innovation capacity", "creativity", "research quality" among others (Zheng, 2008). Jones et al. note for example, that the production of high-impact papers is increasingly dependent upon forming part of a cluster of elite universities, implying a growing stratification of our knowledge production (Jones, Wuchty, & Uzzi, 2008). The literature on the structural dimension of social capital complements here the social psychology approaches (see next section). Although the latter has identified the influence of team size and composition as important variable that affect the performance of teams (Hülsheger et al., 2009), it does seldom contemplate explicitly how the position between actors within a network affects performance measures.

The added value of the network perspective can easily be appreciated by considering the early works on the relationship between ego-centered networks and career advancements (Brass, 1985; Lin, Vaughn, & Ensel, 1981). In his groundbreaking work Granovetter argued for the "strength of weak ties", showing that white-collar workers who dispose of weak social relationships can access very valued, unique information outside of their closed peer groups helping them to find better jobs faster (Granovetter, 1973). Similar, Burt (2005) argued that employees that bridge "structural holes" in their organization, i.e. that connect otherwise disconnected sub-groups, access information faster and exercise control through their brokerage relation between disconnected parties. Over time, benefits accrue from these structural positions as "managers" in this particular case performed "better" in terms of receiving higher personal- and team-level evaluations. Consistent empirical results point to the fact that "[s]enior managers with networks richer in structural holes are more likely to get promoted early, receive more positive job evaluations, and take home higher compensation" (Burt, 1998, p. 12). Subsequent studies have refined these approaches by

arguing for the importance of considering contextual factors such as the content of the network ties (Podolny & Baron, 1997; Scheidegger, 2010). However, independent of these methodological advancements, these early studies already alerted to the important gender differences within networks due to "homophily" on the one hand and differential constraints between women and men to convert network position into advantages on the other.

1.3.1. Network Size (Ego, Team)

From a structural perspective, among the most evident variables to explore is the size of the network. Depending on the unit of analysis, ego network size can refer to individuals, teams, organizations. Overall, the literature on "innovation" has found a positive relation between larger networks and innovation. This makes intuitively sense, since larger networks potentially expose actors to more diverse viewpoints and richer information, knowledge, skills and ideas. In their meta-analytic reviews on the team science literature, both Stewart (2006) and Hülsheger et al. (2009) suggest a positive relation between team size and innovation.

How does the size of a team influence its performance? Several publications suggest that size and performance relation follows an inverted u-shape: with increasing size of a team, diversity assets increase. However, at the same time, larger teams also face higher coordination costs especially when diverse viewpoints, cultural backgrounds and experiences have to be integrated. Research has shown for example that individuals perform worse in large teams due to the perceived "relational loss" (Cummings et al., 2013). The extent to which individuals perceive that social support is available within the team diminishes the larger the group, that is, the higher the probability to maintain superficial social relations. Team members are less motivated to overcome differences and identify less with the overall team. Psychological safety has been found to moderate this relationship: larger teams with higher participant safety are more innovative (Mueller, 2012). In general, the rising costs for managing larger teams may run out of balance and ultimately fail to justify the potential efficiency gains; it also seems that the cost of these increasing frictions of large teams is likely to be underestimated (Peltokorpi & Hasu, 2013).

The fact that network size and innovation capacity follow an inverted u-shaped relation suggests a a "bliss" point between new and repeating members in scientific work (Cheruvelil et al., 2014; Whitfield, 2008; Wuchty et al., 2007). As a study in bio-medical research for example showed, once the network of co-authors exceeds 93, the research performance returns diminish (McFadyen & Cannella, 2004). However, as the meta-analysis by Baer, Evans, Oldham, & Boasso, (2015) argues, brokerage, closure and diversity fully mediate the effect of network size on innovation, since it is not the size of the network per se that matters but access to non-redundant information. As the following paragraphs will argue, centrality measures and tie strength of social networks throw further light onto these diversity configurations that need to be considered.

1.3.2. Topology of Diversity

Most literature on social networks and research performance discusses "bonding" and "bridging" social capital. "Bonding" refers to the strength of ties indicating the quality of the relation between two actors given by the frequency of project collaborations or shared

publications. “Bridging” ties indicate how central or peripheral an actors is within a given network; it thus concerns the the overall topology of the network.

Several authors have examined the relationship between centrality measures of scientists within their authorship networks and the impact of the produced papers. The more centrally connected a given author is, the more productive s/he is and the higher the quality of the produced papers (Abbasi, Altmann, & Hossain, 2011; Abbasi, Chung, & Hossain, 2012; Liao, 2010). Teams at different labs at the US National High Magnetic Field Laboratory whose members have central structural positions perform better than other teams (Conti & Liu, 2015, p. 1634). The insight that scholars who are connected to many distinct scholars fare better in citation based performance indicators makes intuitively sense because they can access more varied, non-redundant information which furthermore fosters innovative results. However, the quality (i.e. higher citation count) depends not only on the diversity of ties but also on the *strength* of the relationships between co-authors: scholars which repeat co-authorship (“bonding” ties) have better research performance that those with weak ties, i.e. publishing with many different authors. In terms of efficiency, Abbasi et al. suggest that strong relationships with one co-author of a cluster of collaborating authors is more beneficial than maintaining redundant ties with all members of the co-author network (ibid.).

Similar findings are available regarding the role of centrality vs. peripheral network positions and ratings of individual creativity in research teams (Perry-Smith, 2006). Weak ties among actors within the research lab facilitated creativity by exposing actors to diverse approaches and perspective (Perry-Smith & Shalley, 2003). Strong ties on the other hand were seen as constraining creativity by social pressure to conform and “group think”. Interestingly, the relationship between centrality and creativity depended on the number ties outside the research team: peripheral scientists were more creative to the degree that they accumulated more external ties whereas outside ties had a less positive impact on more central scientists (ibid.). More recently, Perry-Smith and Shalley argue that team creativity increases when teams are composed of members with nationally diverse outside ties (Perry-Smith & Shalley, 2014). The more team members are exposed to unique perspectives and diverse viewpoints, the more flexible their cognitive schemas will be, enhancing their capacity to think creatively.

Peter Gloor and colleagues (Gloor et al., 2011, 2012) have explored more closely network measures in relation to trust and its impact on creativity in teams. As their research shows, higher in-between centrality measures of face-to-face interactions is positively related with higher levels of trust. Higher levels of trust between team members influence in turn the creativity of the group.

These findings evoke similar results on the importance of “brokerage” and “bonding” ties for publication performance by Badar, Hite, & Ashraf (2015), who suggest that especially younger scholars benefit more from loose ties than established professors. That is to say, that peripheral actors benefit more from more varied and non-redundant ties than more central actors, whose research performance is more consolidated in tenure and expertise.

Other studies have equally emphasized the importance of combining both “bridging” and “bonding” social capital for research and innovation performance (Newell, Tansley, & Huang, 2004; Rost, 2011; Scheidegger, 2010; Stark, 2009).

> *"[…] we suggest that in understanding the relationship between social capital and knowledge integration within a project team, it is necessary to distinguish between two forms of social capital – external bridging social capital and internal bonding social capital." (Stvilia et al., 2011)*

Rost argues that weak network architectures without strong ties have little value. Her results suggest that "[…] actors invested in strong ties and embedded in weak network architectures come up with the most innovative solutions" because they are able to recognize and realize the value of the accessed knowledge (ibid., p.601). Her research combines previous accounts favoring on the one hand either the closure of networks which makes actors willing to share tacit knowledge (Rost, 2011) or concentrating on the availability of assets in sparse networks with many structural holes (Coleman, 1988).

Similar, de Montjoye and colleagues (2014) echo the findings of Rost (ibid.) when they sustain that the problem solving abilities of student teams depend primarily on the strongest ties for both team internal ties as well as external ones. Within-team networks allow for collaboration, mutual engagement and knowing each other well. External ties are relevant for accessing key information and receive valuable feedback. The ability of a given team to perform well needs to be conceived in relation to their team-internal *and* team external ties. However, in both cases, only the strongest ties count: "The strongest ties in both the expressive and instrumental ties networks explain more of the variance than any of the team competencies considered: the mean or maximum of either the self-evaluated competencies, measured competencies, or personality" (ibid., p.3). Instead of weak ties, teams need strong ties within their team and beyond their unit.

As Singh et al. (2015) furthermore observes, the type of knowledge needs to be considered for understanding better the role of "bridging" and "bonding" ties in networks. Collaboration can enhance innovative performance by providing two kinds of knowledge: new knowledge on the one hand, and combinatory knowledge on the other. Whereas "new knowledge" is codified and explicit and as such easily shareable in the form of books, articles, references, or manuals, "combinatory knowledge" is tacit knowledge, i.e. expertise held implicitly by actors that facilitates the recombination and reconfiguration of the "facts" available. Now, as Singh et al. argue in their study of inventors, direct contacts are more important for "bonding" and hence mutual learning of hard to share "combinatory knowledge" whereas heterogeneity of contacts has a much higher impact on innovation through indirect "bridging" ties (ibid.).

On the firm-level, the work of David Stark argues for the combination of "bridging" and "bonding" perspectives into the new concept of "structural folding" (Stark, 2009; Vaan, Vedres, & Stark, 2015; Vedres & Stark, 2010). Structural folds break with the predominant dialectic of brokerage vs. closure in networks by emphasizing the intersection of cohesive groups and closeness of multiple insiders. Actors at the structural fold are insiders to more than one community. They might span wide networks but as such they are at the same time trusted insiders to these distant sub-groups. It is the quality of the social relation that makes all the difference. Structural folds furthermore break with the transmission model of networks in that they emphasize that innovation is not simply the result of accessing and recombining distributed network resources but actually require active "generation". Most innovations and breakthroughs do not originate from a sudden "eureka" moment of the

individual genius; mostly they are the result of a conscious, purposeful, and preservative search for opportunities of improvement (Stark, 2009; Vaan, Vedres, & Stark, 2015; Vedres & Stark, 2010) Structural folds contend that innovation is not simply sitting out there waiting to be discovered but needs active exploration, shaping, constructing. For this active process of "generating" innovation, both trusted relations are as necessary as having non-redundant ideas. As Stark has shown, analyzing actors that form structural folds in Hungarian companies or video games production teams allows to discern high performing and most critically acclaimed products.

Overall, as the meta-analysis of Baer, Evans, Oldham, & Boasso (2015) shows, network size and network strength exhibit a positive, significant relation with innovation (on the individual level) as do brokerage across structural holes and network diversity. Strong ties seem to outweigh their potential negative effects in terms of network closure – reduced access to heterogeneous information and knowledge – especially when the implementation of innovative ideas is required: "[…] two basic features of social networks – the number and average strength of ties – operate through the positional and structural features of brokerage and closure as well as through diversity to impact innovation" (ibid., p.26). This echoes findings by Rost (ibid.) and de Montjoye (ibid.) who emphasize the importance of strong ties.

1.3.3. Boundary Spanning

Work by Ancona & Bresman (2007) and more recently Marrone and collaborators (Marrone, Tesluk, & Carson, 2007; Marrone, 2010) have focused on "boundary spanning" of teams. Rather than using the language of network metrics and concepts, Ancona & Caldwell (1992) distinguished different types of interaction with entities outside the focal team and how these affected team's performance. This literature proposes a certain typology of actions distinguishing between (1) representational functions, (2) coordination of task performance, and (3) general information search. Apart from the already discussed access to external information as an important source of innovation, what this literature advocates is a relational view of team performance in the context of its environment. Boundary spanning activities are as important for team performance as its internal activities because team outcomes need to be aligned to team external necessities. Seldom do teams work in isolation; rather they form part of a development and production chain where their outputs become the input for other teams. Hence, the "performance" of the team depends upon the adequate coordination with external units, including feedback and representation of necessities. As Ancona & Caldwell (ibid). demonstrate, well performing teams are capable of engaging in the right mix of information seeking, coordination and representational activities.

With regard to the representational- and coordination activities, team boundary spanning of research teams emphasis the importance of "science with and for society", i.e. the issue how performance is assessed (see chapter 2). Excellence of scientific output in this sense should not just reflect citation impact of publications for example but also take into account other, societal impact indicators such as "end-user esteem" which implies boundary spanning activities not just as collaboration with other researchers but alignment of research activities with the wider society and its needs.

1.4. The Social Psychology of Work Group Diversity

The literature on work group diversity and small group research in general opens a second important analytic perspective on team performance. Instead of deducing from the distribution of certain "diversity" attributes the performance of the individual node, group or network, this second strand focuses more specifically on the psychological, cognitive and social processes involved. How are group processes rooted in social interaction between a variety of individuals? Which personality traits and which social phenomena seem to foster team processes and help a given group perform better? Team processes and ultimately team outcomes are sustained by an interactive system linking social, psychological and cognitive processes (Salazar et al., 2012). Whereas early research has focused on the cognitive aspects of team work such as constructing shared mental models or elaborating task relevant information, more recent research has also highlighted the role of affective states as a basis for higher social constructs.

1.4.1. Information elaboration – perspective taking – dialogue

Congruent with the preceding paragraphs on structural folding and the importance of "idea" generation, studies on work groups have suggested that the availability of diverse perspectives alone is not sufficient to improve performance. Heterogeneous groups may outperform more homogeneous ones due to a larger pool of task relevant information available to them. However, beyond the mere existence of a wider social network that provides access to more varied information, the important point is to consider how available resources are unlocked and deployed within the group. The Categorization-Elaboration Model (CEM) proposed by van Knippenberg and collaborators (van Knippenberg, De Dreu, & Homan, 2004) emphasizes that inconsistent findings of past research can be explained in part by the insufficient attention to group information processing as the important underlying process of team collaboration. According to the model, teams do not simply benefit from the mere existence of certain diversity attributes but have to actively engage in elaborating task-relevant information. Available information needs to be exchanged, discussed and integrated in order to make a difference for the task at hand:

> *"Arguably, these processes are the key to realizing diversity's potential, because it is not the availability of information per se but the use of this information in group task performance that lies at the basis of diverse groups' potentially superior performance." (van Knippenberg et al 2004, p. 1011)*

The insight into the importance of information elaboration has helped to clarify for example the role of conflict in group decision-making. The literature distinguishes between two types of conflict: whereas relationship conflict appears to be detrimental to group processes in that it absorbs time and energy better deployed on the tasks at hand, task conflict has produced paradoxical findings (van Knippenberg et al., 2004). Task conflict can facilitate better group decision-making by stimulating critical thinking and avoiding premature consensus (de Wit, Greer, & Jehn, 2012). However, this is not always the case. Conflicts between team members regarding a certain task are not automatically beneficial but have to be resolved and handled adequately. Only to the degree that diverging perspectives can be discussed constructively and integrated does task conflict enhance group decision-making. Hence, task conflict needs

to be resolved by making the divergent views explicit and discuss their alternatives, a process which van Knippenberg and collaborators precisely conceived as information elaboration.

A series of authors (Amason, 1996; De Dreu, 2006; Jehn, 1995; Pelled, Eisenhardt, & Xin, 1999; Yong, Sauer, & Mannix, 2014) have further untangled the information elaboration process by analyzing the process of "perspective taking" involved. Perspective taking is defined as a cognitive process through which "an observer tries to understand, in a nonjudgmental way, the thoughts, motives, and (or feelings) of a target, as well as why they think and/or feel the way they do" (S. K. Parker, Atkins, & Axtell, 2008). It is an intentional effort that requires effort to distance oneself from one's own perspective and imagine the view point of the other. Since all meaningful communication requires to some degree an "other" to be addressed and if only silently, "there is almost no aspect of organizational functioning that is not potentially improved by better perspective taking" (ibid., p.8). As Hoever and co-authors now remark, perspective taking is the underlying process which enables first of all the elaboration of task relevant information.

> *"Due to its other-focused nature, the cognitive process of perspective taking has considerable potential to increase the creativity of diverse teams, as it may not only facilitate information exchange (cf. Krauss & Fussell, 1991) but also engender a more comprehensive evaluation of the suggested ideas and an integration of different perspectives." (S. K. Parker et al., 2008)*

Perspective taking is hence the underlying process by which diverse teams can capitalize on their diverse assets because it fosters the emergence of a dialog between the existing viewpoints that ideally converges to a new, integrated solution[3].

Going even further into communication theoretical accounts, perspective taking and maintaining an engaging dialogue are intimately tied together, constituting a mutual foundation for the emergence of new knowledge in organizations (Hoever, van Knippenberg, van Ginkel, & Barkema, 2012, p. 982). By adopting the perspective of the other, i.e. by attempting to understand the other in dialogue, we necessarily alter our own understanding. This has been forcefully shown in relation to stereotypes and bias where adopting the perspective of stigmatized group renders those "Others" more self-similar and hence favors the development of more positive and benign attitudes (Tsoukas, 2009). A self-distancing movement is inherent in all dialogue conceived as a self-referential process whereby past utterances are continually reframed in relation to given responses that indicate how it has been understood. Hence, "[…] self-distanciation occurs through each interlocutor reflexively understanding her own utterances, prompted by the utterances of the other" (Galinsky & Moskowitz, 2000). As such, dialogue with its inherent turn-taking and reframing helps individuals generating alternative hypothesis, explanations and theories. It complexifies individual thinking by provoking conceptual reframings, recombinations and extensions (ibid.).

A similar point has been made in relation to the role of intrinsic motivation for creativity. Intrinsic motivation has been conceived as a key element for being creative since the desire to learn, explore one's interests and follow one's curiosity potentially leads to the discovery

3 The literature on the cognitive aspects of team work is itself huge and will be tackled in a separate section (see Team Cognition page 28)

of novel ideas. Since creativity has been defined as producing not only "novel" but also ideas that are "useful" to someone, Tsoukas (2009, p. 944) argues that creativity is enhanced by "other-focused psychological processes". If team members are motivated to take the perspective of others into account, their ideas are not only novel but also useful, increasing thus their "creativity".

Furthermore, several authors have developed the importance of the related concept of "team reflexivity" for team performance, team innovation and team effectiveness (Konradt, Otte, Schippers, & Steenfatt, 2016; Schippers, Edmondson, & West, 2014; Schippers, West, & Dawson, 2012; West, 2000). Team reflexivity combines "team reflection" and "team adaption", suggesting that new insights generated by the team only become productive once these insights have been adopted. Team reflection in turn is composed of "information seeking" and "information evaluating" activities insisting on the active component of seeking out new information and processing by analyzing and assessing it. Team reflexivity captures the benefits of more self-conscious teams that are able to reflect about their goals, plan strategically for them and evaluate and adjust their internal workings accordingly. As Schippers and co-authors furthermore suggest, the emergence of shared mental models is crucial in this respect (see section Team Cognition page 28).

As the literature suggests, more diverse groups are likely to outperform more homogeneous groups under conditions where team members are able to take the perspective of the other, that is, engage in productive dialogues that express divergent viewpoints and work towards their integration. The research literature has furthermore scrutinized under which social and psychological conditions this exchange of perspective and opinions is most likely to flourish. As the following section will argue, "trust" between team members constitutes a fundamental element.

1.4.2. Trust & Psychological Safety

Many accounts of well performing research groups contain references to the importance of "trust" among its members as an essential ingredient. As Parker & Hackett show in their work on the Resilience Alliance, a high-profile research network in eco-systems ecology, a solid level of "interpersonal trust" and "emotional energy" were key during the formulation of the breakthrough ideas within the group (J. N. Parker & Hackett, 2012). Reiterating the previously mentioned network analytic approaches, "trust" underscores the importance of strong/bonding ties for performance. The central role of trust is furthermore not surprising since research and innovation activities that strive to produce novel ideas are risky business. They have to leave well trodden paths and break with established views that render those who argue for it vulnerable. The core ideas towards which most definitions of "trust" converge to is precisely a "willingness to be vulnerable" and having "confident expectations".

> *"Trust is a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (J. N. Parker & Hackett, 2012).*

To the degree that the production of novel ideas involves taking risks by sharing half-baked ideas and even seemingly ridiculous hypothesis, trust is essential in that it precisely affords these vulnerabilities between individuals and within groups. A parallel argument has been

made from a conversation analytic perspective, where “relational engagement” is at the heart of productive dialogue. In contrast to “calculative engagement” which is sterile and reduced to minimal cooperative behavior, during “relational engagement” participants are able to open up to each other by suspending their irreducible social vulnerability and uncertainty (Rousseau, Sitkin, Burt, & Camerer, 1998, p. 395). They are able to trust each other, harboring favorable expectations regarding the benevolent intentions of the other. Under these conditions, individuals are more willing to freely share information among each other. As participants in a case study of the notoriously innovative 3M corporation sustain, a culture of trust and “psychological safety” was key for allowing employees to take risks and make mistakes without which “innovation” work would not be possible (Möllering, 2006; Tsoukas, 2009).

The role of trust for well functioning teams often evokes the closely related concept of “psychological safety” which has been defined as “individual perceptions about the consequences of interpersonal risks in their work environment. It consists of taken-for-granted beliefs about how others will respond when one puts oneself on the line, such as by asking a question, seeking feedback, reporting a mistake, or proposing a new idea” (Garud, Gehman, & Kumaraswamy, 2011). In contrast to “trust”, psychological safety is an *explicit team level construc*t, that is, members of the same team usually converge towards a shared perception of safety climate within their group. As such it concerns decisions and tactics related to the immediate work environment and time horizon as when employees decide to speak up or remain silent regarding certain issues or tasks. Trust on the other hand is a more dyadic and long-term concept.

Since both concepts describe psychological states that involve perceptions of risk or vulnerability, the effects on team performance have been described in similar terms. As (Edmondson, 2004) argues in her extensive review, research has consistently documented that psychological safety matters for workplace effectiveness. It has consistently been shown to enable performance particularly when work involves uncertain outcomes and creativity such as in R&D teams. Psychological safety also enables organizational learning, especially in hierarchical organizations where people are more likely to offer ideas, ask for help, provide feedback if they feel safe. For example, Edmondson & Lei (2014) showed that trust in management facilitates knowledge sharing in the form of documentation since employees did not fear to become vulnerable to loose their unique value. As the authors further argue, psychological safety also encourages people to speak up (ibid.). Psychological safety is thereby a precondition to make diverse viewpoints emerge and enter the debate.

The positive role of trust for team processes has also been remarked in relation to conflict management. Conflict between group members regarding the content of a given task can be productive to the degree that it prevents premature consensus and “group think” among long standing members of a team. Task conflict resides at the core of the diversity-performance model in the sense that divergent views and conflicting arguments are necessary for the generation of novel ideas. Despite its necessity, it is also a dangerous ingredient from the moment task conflict threatens to degenerate into relationship conflict that does not enhance cognitive diversity but rather undercuts the social fabric of collaboration. As several studies have shown, “trust” is an essential element that prevents task conflict from spiraling out of control and lapsing into relationship conflict (Peterson &

Behfar, 2003; Simons & Peterson, 2000). Trust together with conflict mediates the impact of team diversity on effectiveness (Peterson & Behfar, 2003; Simons & Peterson, 2000). Similar findings are available on the effects of psychological safety, which promotes exploratory and exploitative learning, helps to exploit task conflict and increases team performance (Curşeu & Schruijer, 2010).

1.4.3. Emotion and Group Processes

In recent years, the cognitive bias within the information sharing perspective on diversity has shifted in favor of the role emotions play as foundation for effective group processes (Bradley, Postlethwaite, Klotz, Hamdani, & Brown, 2012; Kostopoulos & Bozionelos, 2011). Although it has been shown that team performance is related to effective information sharing involving "perspective taking", "trust / psychological safety" as described, these constructs depend in turn on the display and perception of emotions in groups. Nonverbal communication through the perception of emotions provides valuable information regarding intentions, preferences, and future behavior and thus ease coordination (Bonaccio, OReilly, OSullivan, & Chiocchio, 2016). Emotional skills at the team level smooth interpersonal communication and cooperation based upon which higher-level constructs such as group cohesion and trust (psychological safety) can emerge.

> "By perceiving, empathizing, and then responding appropriately, people experience greater satisfaction, more positive emotions, and lower stress." (Boroş & Curşeu, 2013; Curşeu, Pluut, Boroş, & Meslec, 2014; Roberge & van Dick, 2010; Stegmann, Roberge, & Dick, 2012).

Positive emotions have also been shown to promote flexibility in thinking thus facilitating problem solving and innovation as well as efficiency and thoroughness in decision making (Bender, Walia, Kambhampaty, Nygard, & Nygard, 2012, p. 40). Emotions are also fundamental to creative processes in that they enable state of heightened consciousness, sharpened attention, and total immersion in the task at hand: "Emotions spark creativity, tighten social bonds, and lower barriers to collaboration" (Isen, 2008). Returning to the previously mentioned role of trust for team work, affective states play an important role in creating, sustaining and ongoing experience of "trust". The experience of trust is an inherently affective phenomena since trust is built in part on expectations that are in part emotional (J. N. Parker & Hackett, 2012, p. 24).

1.4.4. Empathy & Social Sensitivity

A slightly different strand of the literature emphasizes the importance of "empathy" for explaining the potentially positive impact of diversity on group performance (Feyerherm & Rice, 2002; Troth, Jordan, Lawrence, & Tse, 2012). "Empathy" is a promising concept for capturing the positive relation of "diversity" with "performance" because it integrates an affective and a cognitive dimension. On a basic level, empathy is a psychological process through which people come to understand and identify with another persons' situation, perspective or feelings. This "putting oneself into the shoes of others" involves a cognitive dimension as when an individual tries to adopt the perspective of the other person. But it also involves an emotional side: in its basic form, affective empathy is a fast, stimulus-driven

response that aligns motor behavior of observer and observed. Through this activation of shared motor representations, individuals can understand each other's emotions. As previously argued, it is above all this non-conscious, affective empathy where women outperform men.

As a result of the multi-layered nature of “empathy”, it combines two important dimensions for making diversity beneficial to group processes. First, cognitive empathy clearly reaffirms the importance of “perspective taking” for information elaboration. By adopting the perspective of the other, group members set the basis for integrating the available diversity of views and come up with potentially innovative solutions to a given task at hand. Second, affective empathy helps people to relate to each other, reducing stereotypes and potentially discriminatory behaviors. Affective empathy is about “feeling for” someone which has direct effects on pro-social behavior (Roberge & van Dick, 2010; Roberge, 2013; Stegmann et al., 2012). To the degree that one's emotional state is a fundamental dimension of one's psyche, understanding how others feel in a group is a cornerstone of its shared social reality. Being emotionally understood provides the basis upon which higher order constructs such as trust can be built upon.

Research indeed has scrutinized some of the positive impact of “empathy” on group processes. Importantly, the ability to take the perspective of the other and to “feel for” the other reduces stereotypes. Empathy overcomes self-other distinctions and helps to improve the attitude towards the stigmatize group (Roberge, 2013). In that, empathy can directly improve group collaboration by balancing the negative inside/outsider dynamics of social categorization processes (Tajfel, 1982). As Stegmann et al., (2012) summarize, empathy not only can reduce stereotypes and prejudice but also help work groups to prevent and handle conflict. It also increases the likelihood of self-disclosure and hence the likelihood to share new information within the group (ibid.). As the authors furthermore argue, empathy towards specific social groups “can positively influence their attitudes on organizational policies and programs aimed at reducing discrimination” (ibid., p.27).

The research on collective intelligence provides evidence on the importance of empathy, here discussed under the label of “social sensitivity”. As Anita Woolley and collaborates describe, the collective intelligence factor “c” of a group correlates with (a) the individual intelligence factor “g” of each member of the group (but only moderately), (b) the average Theory of Mind (ToM)[4] score of group members, (c) negatively with the average speaking turns by group members, and (c) the proportion of females in groups. What predicts the collective intelligence of the group is not so much the sum of individual intelligence factors but the average social perceptiveness of its group members (Woolley, Aggarwal, & Malone, 2015; Woolley, Chabris, Pentland, Hashmi, & Malone, 2010). Social sensitivity or “social perceptiveness” is measured by how well individuals score on the “Reading the Mind in the Eyes” test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). The test measures individuals' ability to attribute mental states such as goals and beliefs, intentions, desires to others as different from one's own mental states. Originally developed in the context of research on autism, this Theory of Mind (often abbreviated as ToM or mentalizing) describes

4 Theory of Mind is viewed by some as a subset of broader array of skills and abilities associated with the more general concept of emotional intelligence. Theory of Mind is sometimes used as shorthand for the ability to attribute mental states to oneself or another person.

one's ability to reason about the mental states of others. It therefore is closely related to the cognitive dimension of empathy. Although ToM relies on different neural circuity than affective empathy, "social sensitivity", "ToM" and "empathy" define varying angles on a similar phenomena, namely the importance of the ability of group members to recognize other's nonverbal emotional expressions and mental states (Woolley, Aggarwal, & Malone, 2015). What this research on social sensitivity (Woolley et al., 2015) as well as emotional intelligence (Bender et al., 2012) shows, is the fact that this other directed skills are at the foundation of group processes and their effectiveness. In a more recent publication, Engel & Woolley et al. indicate that ToM and collective intelligence is not necessarily restricted to face-to-face situations: "…teams comprised of members with a broader range of ability for perceiving subtle interpersonal cues will be better-equipped to develop higher levels of collective intelligence, especially in less rich, online chat-based environments" (Engel, Woolley, Jing, Chabris, & Malone, 2014, p. 12/16).

How important social sensitivity is for group processes has been shown by Woolley et al. (2010) but also by Hong & Page (2004) who argue that groups of diverse problem solvers outperform groups of highly intelligent ones.

1.4.5. Power, status, hierarchy

An important stream of literature has explored how formal and informal hierarchies in work teams affects performance. Hierarchical relations in the form of formal positions within a company or informal status relations are pervasive (Mannix & Sauer, 2006). As laboratory experiments have shown, hierarchy emerges very quickly and is relatively stable even among groups of strangers (Berger, Cohen, & Zelditch, 1972). Hierarchical relations can have both positive effects on team performance in that they help coordinate tasks and responsibilities and negative ones, in that tree-like command structures are relatively inflexible and can impede the sharing of information (Bunderson, Van der Vegt, Cantimur, & Rink, 2015). There is evidence that high-power teams perform worse than low-power teams due to competition between members around power (Elfenbein, Polzer, & Ambady, 2007). Congruent with this finding, others have argued that the stability of the group hierarchy effects positively conflict management. Relatively stable hierarchies allow individuals to focus on the task at hand instead of spending energy on detrimental relationship conflict (Greer, Caruso, & Jehn, 2011). A relatively recent study Corinne Bendersky & Hays (2012) adds to this literature by showing that status hierarchies are not fixed but dynamically evolving within groups. Importantly, group members conflict over status within the group produces more competitive conflict than any of the other types of process-, task-, or relationship conflict, with detrimental consequences for team performance. The negative impact of status conflict results from the competitive nature of this type of conflict which is especially detrimental to group information sharing (ibid.).

The literature provides evidence how power relations affect groups in three different ways: through the anchorage of shared goals, through risk taking and experimentation, and by conditioning the knowledge transfer between members (Bunderson & Reagans, 2011). As the authors summarize, power and status differences make it difficult for lower status members to see themselves as part of the group picture in terms of shared goals and a common agenda. In hierarchical groups, members have limited opportunities to influence

group decisions and actions and hence feel less accountable and personally involved. This in turn influences the range of informational cues that low status members consider within the group and how comprehensively they consider others.

Second, and more importantly, power and status differences affect the psychological safety of team members. Lower status members perceive their environment to be less safe for learning and risk taking. As argued in the preceding sections, psychological safety however is key for high performing teams especially in an environment where collective learning involves trial and error and risk taking. Since lower status members are more dependent upon higher status members for resources, authorizations, respect and approval, they feel more vulnerable and might engage less in taking initiatives or introducing new ideas with negative consequences for overall team performance.

Third, and most importantly, differences in power and status condition the *content* of shared information. Power relations can limit knowledge transfer between team members in important ways: Individuals at the top of social hierarchies are given disproportionate weight in group decision making while contributions of lower rank members are often overlooked and silenced (Tost, Gino, & Larrick, 2012). In addition to simply silencing contributions from lower-status members, power relations also condition the type of information that gets shared. Because contributing new (unshared) information is risky (needs to be verified), low-status members often repeat accepted/shared information which increases their status and influence in the group. Low-status members also experience a barrier to bring their potentially critical, unshared information into the group – unshared information communicated by low-status members will be less remembered and repeated as when high-status members say something "new" (Cruz, Boster, & Rodríguez, 1997; Larson, Foster-Fishman, & Keys, 1994; Lu, Yuan, & McLeod, 2012; Wittenbaum & Bowman, 2005; Wittenbaum, 2000). Hence, the very power and status differences in teams, including those between low status women and high status (white) men constrain the effective sharing of information between all members. This influences the quality of decision making since important information might be silenced or disregarded. However, as Bunderson & Reagans (2011) argue, power relations can also influence group performance positively when used in "socialized" vs "personalized" ways (see next section on Leadership). See also (Galinsky, Magee, Inesi, & Gruenfeld, 2006)

Power relations have also been discussed from a network analytic perspective. Abbasi et al. (2011) studied the relationship between structural properties of workgroups and their performance in a large, global Telecommunications firm. Hierarchical teams with a tree-like network structure scored worse in terms of manager- and member-rated performance. Groups that work on non-routine tasks and which are organized in a "flat" hierarchy are more flexible and perform better. Similar, greater core-periphery structure of the group network was partially related to worse performance due to marginalization of information. Bunderson (2003) observed that diffuse status cues such as gender were more salient in groups with a more centralized power structure.

> *"In centralized groups, in which members lack a clear motivation to be careful, comprehensive, and accurate in evaluating members' expertise, individuals*

> *should therefore be more likely to rely simply on diffuse status cues in drawing inferences of expertise" (Bunderson, 2003, p. 564)*

1.4.6. Leadership

Closely related to questions of power in groups is the issue of "leadership". Leadership positions are often bound to formal hierarchical positions and are thus often tied to questions of power and status. However, beyond the direct question of power, the literature on leadership has also scrutinized different approaches to team leadership and its effects on performance. Many of the preceding concepts are to be found in the literature on leadership, now with the specific focus how managers can enhance and foster for example perspective taking, psychological safety climate and its potential positive consequences.

Leaders can influence and shape explicitly variables that improve work processes such as group cohesion, a shared goal or team climate (Marshall & Rollinson, 2004; Yanow, 2004). Although the effectiveness of leadership styles has to be seen in relation to the desired outcomes (Nembhard & Edmondson, 2006; Xue, Bradley, & Liang, 2011), empowering leadership has been described as harboring long-term performance increases due to team learning and coordination (Lorinkova, Pearsall, & Sims, 2012; Rosing, Frese, & Bausch, 2011; Somech, 2006; D. Wang, Waldman, & Zhang, 2014). Bunderson & Reagans (2011) argues that the use of power towards "personal" ends negatively affects group learning whereas power used towards "social" ends actually fosters team learning and performance. High status group members can enhance group processes by creating an atmosphere where people not only felt safe but are also explicitly encouraged risk taking. Similarly, the degree to which higher-ranking members overlook contributions of lower status members and mis-recognize their expertise is contingent upon whether these higher-ranking team members have collective or individual goals.

As the meta-analysis by Eagly, Johannesen-Schmidt, & van Engen (2003) suggests (see also (Burke & Collins, 2001; Eagly & Johnson, 1990; Patel, 2013), women leaders tend to exhibit more transformational (participative, empowering, democratic) leadership style compared to men who engaged in a more transactional (directive) style.

> *"This pattern is that women scored higher than men on transformational leadership and contingent reward, whereas men scored higher than women on active and passive management by exception and laissez-faire leadership." (Eagly et al., 2003, p. 579)*

However, other research finds no clear-cut differences between women and men regarding their leadership style. Martin for example neither detects leadership differences based on gender nor on institutional type but rather in terms of years of experience: "the more years of administrative experience respondents had, the more likely they were to use transactional leadership" (Martin, 2015, p. 340). Similar, Murray (2014) argues that leadership style has to be explored in relation to generational aspects.

Transformational leadership is based on inspiring followers commitment and creativity to shape the future whereas transactional or managerial styles use supervision, reward and

punishment to maintain the status quo. Transformational leadership is more congruent with the socially-supportive use of power and the outward-oriented gender roles associated with femininity: it is a participative, horizontal leadership that is built on trust and confidence, which aims to inspire, mentor and empower followers to develop their full potential (Eagly & Johnson, 1990; McKinsey & Company, 2009). In its emphasis on trust and confidence, empowering leadership is a strong facilitator of good team climate, crucial for conflict management and information sharing.

1.5. Team Cognition

Shared cognition among team members is an important ingredient for boosting team performance (Salas, Cooke, & Rosen, 2008; Salas & Fiore, 2004). It constitutes the third element of our perspective on social capital, together with the structural perspective on diversity and the social psychological one as just reviewed.

Forming a closely knit conceptual cluster, team cognition has been discussed in relation to "transactive memory systems", "team mental model", "shared/team cognition", or "group learning". They all examine essentially the ways knowledge is distributed, structured and manipulated among team members. Although many studies have demonstrated the importance of team cognition for sports teams, the military or management teams, especially science and research teams required a shared cognition in order to integrate the distributed resources into new knowledge. In their detailed case study of the logic of innovation at 3M corporation Garud et al. (2011) sustain for example, that a shared system for encoding, storing and retrieving information – Transactive Memory System (TMS) – plays a fundamental role for "holding together" complex innovation processes. In this particular example of teams at 3M, telling stories about past inventions fulfilled the role of TMS in making available knowledge about past projects including involved people and their expertise.

Team cognition tackles shared, cognitive representations among team members. "Transactive memory systems" (Wegner, 1987) focus especially on the distribution of knowledge in teams. The TMS captures "who knows what" and thus captures first and foremost the dispersion and specialization of expertise in groups. "Team mental models" in contrast have been described as shared mental representations of knowledge about key elements of a team's environment (Klimoski & Mohammed, 1994; Mohammed, Ferzandi, & Hamilton, 2010). Whereas TMS capture the importance for being able to locate expertise within teams, shared mental models foreground the developing of a shared understanding of a task or problem. Both cognitive constructs help to understand how knowledge is mentally organized, represented and distributed within the team. In their review on TMS, Ren & Argote (2011) suggest consistent findings on the positive outcomes of TMS on various group outcomes such as team creativity, effectiveness or member satisfaction.

> *"Accurate expertise recognition improves team performance because it facilitates the division of cognitive labor among members, the search and location of required knowledge, the match of problems with the person with the*

> *requisite expertise to solve the problems, the coordination of group activities, and better decisions through the evaluation and integration of knowledge contribute by group members." (Ren & Argote, 2011, pp. 205–206)*

Team cognition enables work group members to develop more accurate predictions about tasks and develop more efficient responses by taking advantage of the available expertise, thereby increase the overall effectiveness and performance. In a meta-analysis of team cognition across 65 independent studies DeChurch & Mesmer-Magnus (2010) found a positive relation between team cognition and behavioral processes, team motivational states and team performance. Evidence from student teams tackling software developments tasks suggests that frequency of communication and team diversity affect team performance (He, Butler, & King, 2014). Apart from "member familiarity", gender diversity exhibited a strong and consistent effect on team performance: mixed gender teams tended to develop higher levels of team cognition (ibid.).

Team cognition provides another powerful control variable: as the literature suggests, teams which would score high on team cognition have good chances to perform better than teams which score low. As such it will provide an important control variable.

1.6. Other aspects: Personality Traits

Among additional important aspects that have not been treated explicitly in the previous sections – although they run through many of the respective contributions – is the impact of personality on team performance. Personality traits[5] are conceived to be especially important for group processes since characteristic behaviors associated with personality traits indicate how people related to each other and interact. Personality traits influence contributions to the team and status. However, the relationship is not as clear-cut as one might expect with laboratory studies generally producing lower effects than field studies LePine et al. (2011). Conscientiousness seems to be highly relevant for individual performance, while "agreeableness" appears to gain in importance due to the collaborative nature of much work. Gloor et al., (2012) found that "openness" correlates with degree centrality and higher levels of trust which in turn influences creativity in teams. Specifically with respect to "agreeableness", the authors note a certain ambiguity: the more agreeable a person is, the more people trust her/him. However, the more agreeable a person is, the more influential people distrust her/him. Agreeable persons seem to be trusted more, but not by influential people.

As already argued in the section on Status Expectation Theory, personality traits is related to the status of individuals in groups. Especially extroversion predicts status in face to face

5 The Big Five Personality Traits are: **Extraversion**: energetic approach to the social and material world and includes traits such as sociability, activity, assertiveness, and positive emotionality. **Agreeableness**: altruism, trust, modesty, tender minded concern for others, i.e. interpersonal characteristics. **Neuroticism**: negative emotionality, i.e. vulnerability to stress, anxiety, depression, negative self-conscious emotions (guilt, shame, embarrassment). **Conscientiousness**: socially prescribed impulse control that facilitates task- and goal-directed behavior: dutiful, hard-working, organized. **Openness to Experience**: breath, depth, originality, and complexity of an individual's mental and experiential life. See (Goldberg, 1993)

groups. Recent research, however questions this, arguing that status of extroverts declines over time in work groups as their real task contributions becomes visible (C. Bendersky & Shah, 2012).

Combining insights from the network literature on team performance and personality traits, a recent meta-analysis argues that persons with “certain personality traits tend to attain structurally advantageous positions and that the occupation of these advantageous positions, in turn, influences people’s performance and career success.” (Fang et al., 2015). Thus, personality traits such as conscientiousness influence team performance by executing tasks better, but also because it helps those people achieve more advantageous network positions.

1.7. Bibliography

Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, *5*(4), 594–607. http://doi.org/10.1016/j.joi.2011.05.007

Abbasi, A., Chung, K. S. K., & Hossain, L. (2012). Egocentric analysis of co-authorship network structure, position and performance. *Information Processing & Management*, *48*(4), 671–679. http://doi.org/10.1016/j.ipm.2011.09.001

Adler, P. S., & Kwon, S.-W. (2002). Social Capital: Prospects for a New Concept. *Academy of Management Review*, *27*(1), 17–40.

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*, *31*(1), 123–148.

Ancona, D., & Bresman, H. (2007). *X-teams : how to build teams that lead, innovate, and succeed*. Boston Mass.: Harvard Business School Press.

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the Boundary: External Activity and Performance in Organizational Teams. *Administrative Science Quarterly*, *37*(4), 634–665.

Badar, K., Hite, J. M., & Ashraf, N. (2015). Knowledge network centrality, formal rank and research performance: evidence for curvilinear and interaction effects. *Scientometrics*, *105*(3), 1553–1576. http://doi.org/10.1007/s11192-015-1652-0

Baer, M., Evans, K., Oldham, G. R., & Boasso, A. (2015). The social network side of individual innovation: A meta-analysis and path-analytic integration. *Organizational Psychology Review*, *5*(3), 191–223. http://doi.org/10.1177/2041386614564105

Ball, S. J. (2012). Performativity, Commodification and Commitment: An I-Spy Guide to the Neoliberal University. *British Journal of Educational Studies*, *60*(1), 17–28. http://doi.org/10.1080/00071005.2011.650940

Bender, L., Walia, G., Kambhampaty, K., Nygard, K. E., & Nygard, T. E. (2012). Social sensitivity correlations with the effectiveness of team process performance. In *Proceedings of the ninth annual international conference on International computing education research - ICER '12* (p. 39). New York, New York, USA: ACM Press. http://doi.org/10.1145/2361276.2361285

Bendersky, C., & Hays, N. A. (2012). Status Conflict in Groups. *Organization Science*, *23*(2),

323–340. http://doi.org/10.1287/orsc.1110.0734

Bendersky, C., & Shah, N. P. (2012). The Downfall of Extraverts and Rise of Neurotics: The Dynamic Process of Status Allocation in Task Groups. *Academy of Management Journal*, *56*(2), 387–406. http://doi.org/10.5465/amj.2011.0316

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*(3), 241–255.

Bonaccio, S., OReilly, J., OSullivan, S. L., & Chiocchio, F. (2016). Nonverbal Behavior and Communication in the Workplace: A Review and an Agenda for Research. *Journal of Management*. http://doi.org/10.1177/0149206315621146

Boroş, S., & Curşeu, P. L. (2013). Would you like to talk about that?' How and when group emotional awareness enhances effectiveness of gender diverse teams. *Psihologia Resurselor Umane*, *11*(2), 45–56.

Bowers, C. A., Pharmer, J. A., & Salas, E. (2000). When Member Homogeneity is Needed in Work Teams: A Meta-Analysis. *Small Group Research*, *31*(3), 305–327. http://doi.org/10.1177/104649640003100303

Bozeman, B., & Boardman, C. (2014). *Research collaboration and team science : a state-of-the-art review and agenda*. Cham (Switzerland): Springer.

Bradley, B. H., Postlethwaite, B. E., Klotz, A. C., Hamdani, M. R., & Brown, K. G. (2012). Reaping the benefits of task conflict in teams: the critical role of team psychological safety climate. *The Journal of Applied Psychology*, *97*(1), 151–8. http://doi.org/10.1037/a0024200

Brass, D. J. (1985). Men's and Women's Networks: A Study of Interaction Patterns and Influence in an Organization. *Academy of Management Journal*, *28*(2), 327–343. http://doi.org/10.2307/256204

Bunderson, J. S. (2003). Recognizing and utilizing expertise in work groups: A status characteristics perspective. *Administrative Science Quarterly*, *48*(4), 557–591.

Bunderson, J. S., & Reagans, R. E. (2011). Power, Status, and Learning in Organizations. *Organization Science*, *22*(5), 1182–1194. http://doi.org/10.1287/orsc.1100.0590

Bunderson, J. S., Van der Vegt, G., Cantimur, Y., & Rink, F. (2015). Different Views of Hierarchy And Why They Matter: Hierarchy as Inequality or as Cascading Influence. *Academy of Management Journal*. http://doi.org/10.5465/amj.2014.0601

Burke, S., & Collins, K. M. (2001). Gender differences in leadership styles and management skills. *Women in Management Review*, *16*(5), 244–257.

http://doi.org/10.1108/09649420110395728

Burt, R. S. (1992). *Structural holes : the social structure of competition*. Cambridge Mass.: Harvard University Press.

Burt, R. S. (1998). The Gender of Social Capital. *Rationality and Society*, *10*(1), 5–46.

Burt, R. S. (2005). *Brokerage and closure : an introduction to social capital*. Oxford ;;New York: Oxford University Press.

Campbell, L. G., Mehtani, S., Dozier, M. E., & Rinehart, J. (2013). Gender-heterogeneous working groups produce higher quality science. *PloS One*, *8*(10), e79147. http://doi.org/10.1371/journal.pone.0079147

Catalyst. (2010). *Why Diversity Matters. Research Studies 2005-2010. Research Studies*. New York, San Jose, Toronto.

Certo, S. T., Lester, R. H., Dalton, C. M., & Dalton, D. R. (2006). Top Management Teams , Strategy and Financial Performance : A Meta-Analytic Examination. *Journal of Management Studies*, *43*(4), 813–839.

Cheruvelil, K. S., Soranno, P. A., Weathers, K. C., Hanson, P. C., Goring, S. J., Filstrup, C. T., & Read, E. K. (2014). Creating and maintaining high-performing collaborative research teams: the importance of diversity and interpersonal skills. *Frontiers in Ecology and the Environment*, *12*(1), 31–38. http://doi.org/10.1890/130001

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, *94*, 95–120.

Conti, A., & Liu, C. C. (2015). Bringing the lab back in: Personnel composition and scientific output at the MIT Department of Biology. *Research Policy*, *44*(9), 1633–1644. http://doi.org/10.1016/j.respol.2015.01.001

Cruz, M. G., Boster, F. J., & Rodríguez, J. I. (1997). The Impact of Group Size and Proportion of Shared Information on the Exchange and Integration of Information in Groups. *Communication Research*, *24*(3), 291–313. http://doi.org/10.1177/009365097024003004

Cummings, J. N. (2004). Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization. *Management Science*, *50*(3), 352–364. http://doi.org/10.1287/mnsc.1030.0134

Cummings, J. N., & Cross, R. (2003). Structural properties of work groups and their consequences for performance. *Social Networks*, *25*(3), 197–210. http://doi.org/10.1016/S0378-8733(02)00049-7

Cummings, J. N., Kiesler, S., Bosagh Zadeh, R., & Balakrishnan, A. D. (2013). Group heterogeneity increases the risks of large group size: a longitudinal study of productivity in research groups. *Psychological Science*, *24*(6), 880–90. http://doi.org/10.1177/0956797612463082

Curşeu, P. L., Pluut, H., Boroş, S., & Meslec, N. (2014). The magic of collective emotional intelligence in learning groups: No guys needed for the spell! *British Journal of Psychology (London, England : 1953)*. http://doi.org/10.1111/bjop.12075

Curşeu, P. L., & Schruijer, S. G. L. (2010). Does conflict shatter trust or does trust obliterate conflict? Revisiting the relationships between team diversity, conflict, and trust. *Group Dynamics: Theory, Research, and Practice*, *14*(1), 66–79.

De Dreu, C. K. W. (2006). When Too Little or Too Much Hurts: Evidence for a Curvilinear Relationship Between Task Conflict and Innovation in Teams. *Journal of Management*, *32*(1), 83–107. http://doi.org/10.1177/0149206305277795

de Montjoye, Y.-A., Stopczynski, A., Shmueli, E., Pentland, A., & Lehmann, S. (2014). The Strength of the Strongest Ties in Collaborative Problem Solving. *Scientific Reports*, *4*, 1–6.

de Wit, F. R. C., Greer, L. L., & Jehn, K. A. (2012). The Paradox of Intragroup Conflict: A Meta-Analysis. *Journal of Applied Psychology*, *97*(2), 360–390.

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: a meta-analysis. *The Journal of Applied Psychology*, *95*(1), 32–53. http://doi.org/10.1037/a0017328

Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, *129*(4), 569–591.

Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin*, *108*(2), 233–256.

Edmondson, A. C. (2004). Psychological Safety, Trust and Learning in Organizations: A Group-Level Lens. In R. Kramer & K. Cook (Eds.), *Trust and Distrust in Organizations: Dilemmas and Approaches* (pp. 239–272). New York: Russell Sage Foundation.

Edmondson, A. C., & Lei, Z. (2014). Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 23–43. http://doi.org/10.1146/annurev-orgpsych-031413-091305

Elfenbein, H. A., Polzer, J. T., & Ambady, N. (2007). Chapter 4 Team Emotion Recognition

Accuracy and Team Performance : Functionality, Intentionality and Morality. In C. E. J. Härtel, N. M. Ashkanasy, & W. J. Zerbe (Eds.), *Functionality, Intentionality and Morality (Research on Emotion in Organizations, Volume 3)* (pp. 87–119). Bingley: Emerald Group Publishing, Limited.

Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the Mind in the Eyes or reading between the lines? Theory of Mind predicts collective intelligence equally well online and face-to-face. *PloS One*, *9*(12), e115212. http://doi.org/10.1371/journal.pone.0115212

Erhardt, N. L., Werbel, J. D., & Shrader, C. B. (2003). Board of Director Diversity and Firm Financial Performance. *Corporate Governance*, *11*(2), 102–111. http://doi.org/10.1111/1467-8683.00011

Fang, R., Landis, B., Zhang, Z., Anderson, M. H., Shaw, J. D., & Kilduff, M. (2015). Integrating Personality and Social Networks: A Meta-Analysis of Personality, Network Position, and Work Outcomes in Organizations. *Organization Science*, *26*(4), 1243–1260. http://doi.org/10.1287/orsc.2015.0972

Feyerherm, A. E., & Rice, C. L. (2002). Emotional Intelligence and Team Performance: The Good, the Bad and the Ugle. *The International Journal of Organizational Analysis*, *10*(4), 343–362. http://doi.org/10.1108/eb028957

Figg, W. D., Dunn, L., Liewehr, D. J., Steinberg, S. M., Thurman, P. W., Barrett, J. C., & Birkinshaw, J. (2006). Scientific collaboration results in higher citation rates of published articles. *Pharmacotherapy*, *26*(6), 759–67. http://doi.org/10.1592/phco.26.6.759

Foster, M. K., Abbey, A., Callow, M. A., Zu, X., & Wilbon, A. D. (2015). Rethinking Virtuality and Its Impact on Teams. *Small Group Research*, *46*(3), 267–299. http://doi.org/10.1177/1046496415573795

Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, *17*(12), 1068–74. http://doi.org/10.1111/j.1467-9280.2006.01824.x

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–24.

Garud, R., Gehman, J., & Kumaraswamy, A. (2011). Complexity Arrangements for Sustained Innovation: Lessons from 3M Corporation. *Organization Studies*, *32*(6), 737–767. http://doi.org/10.1177/0170840611410810

Gloor, P. A., Fischbach, K., Fuehres, H., Lassenius, C., Niinimäki, T., Olguin, D. O., … Putzke, J.

(2011). Towards “Honest Signals” of Creativity – Identifying Personality Characteristics Through Microscopic Social Network Analysis. *Procedia - Social and Behavioral Sciences*, *26*, 166–179. http://doi.org/10.1016/j.sbspro.2011.10.573

Gloor, P. A., Grippa, F., Putzke, J., Lassenius, C., Fuehres, H., Fischbach, K., & Schoder, D. (2012). Measuring Social Capital in Creative Teams Through Sociometric Sensors. *International Journal of Organisational Design and Engineering*, *2*(4), 380–401.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*.

Greer, L. L., Caruso, H. M., & Jehn, K. A. (2011). The bigger they are, the harder they fall: Linking team power, team conflict, and performance. *Organizational Behavior and Human Decision Processes*, *116*(1), 116–128. http://doi.org/10.1016/j.obhdp.2011.03.005

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: recent research on performance and effectiveness. *Annual Review of Psychology*, *47*, 307–38. http://doi.org/10.1146/annurev.psych.47.1.307

Haas, H. (2010). How can we explain mixed effects of diversity on team performance? A review with emphasis on context. *Equality, Diversity and Inclusion: An International Journal*, *29*(5), 458–490. http://doi.org/10.1108/02610151011052771

He, J., Butler, B. S., & King, W. R. (2014, December 8). Team Cognition: Development and Evolution in Software Project Teams. *Journal of Management Information Systems*. Routledge.

Hoever, I. J., van Knippenberg, D., van Ginkel, W. P., & Barkema, H. G. (2012). Fostering team creativity: Perspective taking as key to unlocking diversity’s potential. *Journal of Applied Psychology*, *97*(5), 982–996.

Hollenbeck, J. R., Beersma, B., & Schouten, M. E. (2012). Beyond Team Types and Taxonomies: A Dimensional Scaling Conceptualization for Team Description. *Academy of Management Review*, *37*(1), 82–106. http://doi.org/10.5465/amr.2010.0181

Homberg, F. (2010). *Top Management Team Diversity : A systematic Review*. Zürick.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(46), 16385–9. http://doi.org/10.1073/pnas.0403723101

Hülsheger, U. R., Anderson, N., & Salgado, J. F. (2009). Team-Level Predictors of Innovation at

Work: A Comprehensive Meta-Analysis Spanning Three Decades of Research. *Journal of Applied Psychology*, *94*(5), 1128–1145.

Isen, A. M. (2008). Some Ways in Which Positive Affect Influences Decision Making and Problem Solving. In M. Lewis, J. M. Haviland-Jones, & L. Feldmann Barret (Eds.), *Handbook of Emotions* (3rd ed., p. 548573). New York: The Guilford Press.

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., Northcraft, G. B., & Neale, M. A. (1999). Why Differences Make a Difference : A Field Study of Diversity , Conflict , and Performance in Workgroups. *Aministrative Science Quaterly*, *44*(4), 741–763.

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: shifting impact, geography, and stratification in science. *Science (New York, N.Y.)*, *322*(5905), 1259–62. http://doi.org/10.1126/science.1158357

Joshi, A., & Roh, H. (2007). Context Matters: A Multilevel Framework for Work and Team Diversity Research. *Research in Personnel and Human Resource Management*, *26*, 1–48. http://doi.org/10.1016/S0742-7301(07)26001-3

Joshi, A., & Roh, H. (2009). The Role of Context in Work Team Diversity Research: A Meta-analytical Review. *Academy of Management Journal*, *52*(3), 599–627.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*(1), 1–18. http://doi.org/10.1016/S0048-7333(96)00917-1

Klimoski, R., & Mohammed, S. (1994). Team mental model: construct or metaphor? *Journal of Management*, *20*(2), 403–437.

Konradt, U., Otte, K.-P., Schippers, M. C., & Steenfatt, C. (2016). Reflexivity in Teams: A Review and New Perspectives. *The Journal of Psychology*, *150*(2), 151–72. http://doi.org/10.1080/00223980.2015.1050977

Kostopoulos, K. C., & Bozionelos, N. (2011). Team Exploratory and Exploitative Learning: Psychological Safety, Task Conflict, and Team Performance. *Group & Organization Management*, *36*(3), 385–415. http://doi.org/10.1177/1059601111405985

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the Effectiveness of Work Groups and Teams. *Psychological Science in the Public Interest*, *7*(3), 77–124. http://doi.org/10.1111/j.1529-1006.2006.00030.x

Larson, J. R., Foster-Fishman, P. G., & Keys, C. B. (1994). Discussion of shared and unshared information in decision-making groups. *Journal of Personality and Social Psychology*,

67(3), 446–461. http://doi.org/10.1037/0022-3514.67.3.446

Leahey, E., & Cain, C. L. (2013). Straight from the source: Accounting for scientific success. *Social Studies of Science*, *43*(6), 927–951. http://doi.org/10.1177/0306312713484820

Lee, S., & Bozeman, B. (2005). The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*, *35*(5), 673–702.

Lee, S. M., Koopman, J., Hollenbeck, J. R., Wang, L. C., & Lanaj, K. (2015). The Team Descriptive Index (TDI): A Multidimensional Scaling Approach for Team Description. *Academy of Management Discoveries*, *1*(1), 91–116. http://doi.org/10.5465/amd.2013.0001

LePine, J. A., Buckman, B. R., Crawford, E. R., & Methot, J. R. (2011). A review of research on personality in teams: Accounting for pathways spanning levels of theory and analysis. *Human Resource Management Review*, *21*(4), 311–330. http://doi.org/10.1016/j.hrmr.2010.10.004

Liao, C. H. (2010). How to improve research quality? Examining the impacts of collaboration intensity and member diversity in collaboration networks. *Scientometrics*, *86*(3), 747–761. http://doi.org/10.1007/s11192-010-0309-2

Lin, N., Vaughn, J. C., & Ensel, W. M. (1981). Social Resources and Occupational Status Attainment. *Social Forces*, *59*(4), 1163–1181. http://doi.org/10.1093/sf/59.4.1163

Lorinkova, N. M., Pearsall, M. J., & Sims, H. P. (2012). Examining the Differential Longitudinal Performance of Directive versus Empowering Leadership in Teams. *Academy of Management Journal*, *56*(2), 573–596. http://doi.org/10.5465/amj.2011.0132

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 : An Official Journal of the Society for Personality and Social Psychology, Inc*, *16*(1), 54–75. http://doi.org/10.1177/1088868311417243

Mannix, E. A., & Sauer, S. J. (2006). Status and Power in Organizational Group Research: Acknowledging the Pervasiveness of Hierarchy. *Advances in Group Processes*, *23*, 149–182. http://doi.org/10.1016/S0882-6145(06)23006-6

Marrone, J. A. (2010). Team Boundary Spanning: A Multilevel Review of Past Research and Proposals for the Future. *Journal of Management*, *36*(4), 911–940. http://doi.org/10.1177/0149206309353945

Marrone, J. A., Tesluk, P. E., & Carson, J. B. (2007). A Multilevel Investigation of Antecedents and Consequences of Team Member Boundary-Spanning Behavior. *Academy of Management Journal*, *50*(6), 1423–1439. http://doi.org/10.5465/AMJ.2007.28225967

Marshall, N., & Rollinson, J. (2004). Maybe Bacon Had a Point: The Politics of Interpretation in Collective Sensemaking. *British Journal of Management*, *15*(S1), 71–86. http://doi.org/10.1111/j.1467-8551.2004.00407.x

Martin, J. (2015). Transformational and Transactional Leadership: An Exploration of Gender, Experience, and Institution Type. *Portal: Libraries and the Academy*, *15*(2), 331–351. http://doi.org/10.1353/pla.2015.0015

McFadyen, M. A., & Cannella, A. A. (2004). Social Capital and Knowledge Creation: Diminishing Returns of the Number and Strength of Exchange. *The Academy of Management Journal*, *47*(5), 735–746.

McKinsey & Company. (2007). *Women Matter. Gender Diversity, a Corporate Performance Driver.*

Meeussen, L., & van Dijk, H. (2015). The perceived value of team players: a longitudinal study of how group identification affects status in work groups. *European Journal of Work and Organizational Psychology*, *25*(2), 200–211. http://doi.org/10.1080/1359432X.2015.1028378

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor No More: A 15-Year Review of the Team Mental Model Construct. *Journal of Management*, *36*(4), 876–910. http://doi.org/10.1177/0149206309356804

Möllering, G. (2006). *Trust : reason, routine, reflexivity*. Amsterdam ;;Boston ;;London: Elsevier.

Mortensen, M. (2012). *Boundary Multiplicity: Rethinking Teams and Boundedness in the Light of Today's Collaborative Environment*. *INSEAD Working Paper 2015/31/=HB*.

Mueller, J. S. (2012). Why individuals in larger teams perform worse. *Organizational Behavior and Human Decision Processes*, *117*(1), 111–124. http://doi.org/10.1016/j.obhdp.2011.08.004

Murray, D. (2014). Differences in Leadership Styles and Motives in men and Women: How Generational Theory Informs Gender Role Congruity. In *Proceedings of the European Conference on Management, Leadership* (p. 192).

Nahapiet, J., & Ghoshal, S. (1998). Social Capital, Intellectual Capital, and the Organizational Advantage. *The Academy of Management Review*, *23*(2), 242–266.

Nembhard, I. M., & Edmondson, A. C. (2006). Making it safe: the effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *Journal of Organizational Behavior*, *27*(7), 941–966. http://doi.org/10.1002/job.413

Newell, S., Tansley, C., & Huang, J. (2004). Social Capital and Knowledge Integration in an ERP Project Team: The Importance of Bridging AND Bonding. *British Journal of Management*, *15*(S1), 43–57. http://doi.org/10.1111/j.1467-8551.2004.00399.x

Østergaard, C. R., Timmermans, B., & Kristinsson, K. (2011). Does a different view create something new? The effect of employee diversity on innovation. *Research Policy*, *40*(3), 500–509. http://doi.org/10.1016/j.respol.2010.11.004

Parker, J. N., & Hackett, E. J. (2012). Hot Spots and Hot Moments in Scientific Collaborations and Social Movements. *American Sociological Review*, *77*(1), 21–44. http://doi.org/10.1177/0003122411433763

Parker, S. K., Atkins, P. W. B., & Axtell, C. M. (2008). Building Better Workplaces through Individual Perspective Taking: A Fresh Look at a Fundamental Human Process. In G. P. Hodgkinson & J. K. Ford (Eds.), *International Review of Industrial and Organizational Psychology* (Vol. 23, pp. 149–169). Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9780470773277

Patel, G. (2013). *Gender Differences in Leadership Styles and the Impact within Corporate Boards*.

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the Black Box : An Analysis of Work Group Diversity, Conflict, and Performance. *Administrative Science Quarterly*, *44*(1), 1–28.

Peltokorpi, V., & Hasu, M. (2013). How Participative Safety Matters More in Team Innovation as Team Size Increases. *Journal of Business and Psychology*, *29*(1), 37–45. http://doi.org/10.1007/s10869-013-9301-1

Perry-Smith, J. E., & Shalley, C. E. (2003). The Social Side of Creativity: A Static and Dynamic Social Network Perspective. *Academy of Management Review*, *28*(1), 89–106. http://doi.org/10.5465/AMR.2003.8925236

Perry-Smith, J. E., & Shalley, C. E. (2014). A Social Composition View of Team Creativity: The Role of Member Nationality-Heterogeneous Ties Outside of the Team. *Organization Science*, *25*(5), 1434–1452. http://doi.org/10.1287/orsc.2014.0912

Peterson, R. S., & Behfar, K. J. (2003). The dynamic relationship between performance feedback, trust, and conflict in groups: A longitudinal study. *Organizational Behavior and Human Decision Processes*, *92*(1-2), 102–112. http://doi.org/10.1016/S0749-5978(03)00090-6

Pletzer, J. L., Nikolova, R., Kedzior, K. K., & Voelpel, S. C. (2015). Does Gender Matter? Female Representation on Corporate Boards and Firm Financial Performance - A Meta-Analysis.

PloS One, *10*(6), e0130005. http://doi.org/10.1371/journal.pone.0130005

Podolny, J. M., & Baron, J. N. (1997). Resources and Relationships: Social Networks and Mobility in the Workplace. *American Sociological Review*, *62*(5), 673.

Reagans, R., & Zuckerman, E. W. (2001). Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams. *Organization Science*, *12*(4), 502–517. http://doi.org/10.1287/orsc.12.4.502.10637

Ren, Y., & Argote, L. (2011). Transactive Memory Systems 1985–2010: An Integrative Framework of Key Dimensions, Antecedents, and Consequences. *The Academy of Management Annals*, *5*(1), 189–229. http://doi.org/10.1080/19416520.2011.590300

Renzel, B. (2008). Trust in management and knowledge sharing: The mediating effects of fear and knowledge documentation. *Omega*, *36*(2), 206–220. http://doi.org/10.1016/j.omega.2006.06.005

Richard, O. C., & Miller, C. D. (2013). Considering Diversity as a Source of Competitive Advantage in Organizations. In Q. M. Roberson (Ed.), *The Oxford Handbook of Diversity and Work*. Oxford: Oxford University Press.

Rigby, J., & Edler, J. (2005). Peering inside research networks: Some observations on the effect of the intensity of collaboration on the variability of research quality. *Research Policy*, *34*(6), 784–794. http://doi.org/10.1016/j.respol.2005.02.004

Roberge, M.-É. (2013). A Multi-Level Conceptualization of Empathy to Explain How Diversity Increases Group Performance. *International Journal of Business and Management*, *8*(3), p122. http://doi.org/10.5539/ijbm.v8n3p122

Roberge, M.-É., & van Dick, R. (2010). Recognizing the benefits of diversity: When and how does diversity increase group performance? *Human Resource Management Review*, *20*(4), 295–308. http://doi.org/10.1016/j.hrmr.2009.09.002

Rosing, K., Frese, M., & Bausch, A. (2011). Explaining the heterogeneity of the leadership-innovation relationship: Ambidextrous leadership.

Rost, K. (2011). The strength of strong ties in the creation of innovation. *Research Policy*, *40*(4), 588–604. http://doi.org/10.1016/j.respol.2010.12.001

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393–404. http://doi.org/10.5465/AMR.1998.926617

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors: The Journal of the Human Factors and*

Ergonomics Society, *50*(3), 540–547. http://doi.org/10.1518/001872008X288457

Salas, E., & Fiore, S. M. (2004). Team cognition understanding the factors that drive process and performance. Washington, DC : American Psychological Association,.

Salazar, M. R., Lant, T. K., Fiore, S. M., & Salas, E. (2012). Facilitating Innovation in Diverse Science Teams Through Integrative Capacity. *Small Group Research*, *43*(5), 527–558. http://doi.org/10.1177/1046496412453622

Savino, T., Messeni Petruzzelli, A., & Albino, V. (2015). Search and Recombination Process to Innovate: A Review of the Empirical Evidence and a Research Agenda. *International Journal of Management Reviews*, n/a–n/a. http://doi.org/10.1111/ijmr.12081

Scheidegger, N. (2010). *Der Einfluss von Netzwerkstrukturen auf den Karriereerfolg im Management: Bridging oder Bonding Ties?: Netzwerkanalyse in einem Schweizer Dienstleistungsunternehmen*. Rainer Hampp Verlag.

Schippers, M. C., Edmondson, A. C., & West, M. A. (2014). Team Reflexivity as an Antidote to Team Information-Processing Failures. *Small Group Research*, *45*(6), 731–769. http://doi.org/10.1177/1046496414553473

Schippers, M. C., West, M. A., & Dawson, J. F. (2012). Team Reflexivity and Innovation: The Moderating Role of Team Context. *Journal of Management*, *41*(3), 769–788. http://doi.org/10.1177/0149206312441210

Shore, L. M., Chung-Herrera, B. G., Dean, M. a., Ehrhart, K. H., Jung, D. I., Randel, A. E., & Singh, G. (2009). Diversity in organizations: Where are we now and where are we going? *Human Resource Management Review*, *19*(2), 117–133. http://doi.org/10.1016/j.hrmr.2008.10.004

Shrader, C. B., Blackburn, V. B., & Iles, P. (1997). Women in management and firm financial performance: An exploratory study. *Journal of Managerial Issues*, *9*(3), 355–372.

Simons, T. L., & Peterson, R. S. (2000). Task conflict and relationship conflict in top management teams: the pivotal role of intragroup trust. *The Journal of Applied Psychology*, *85*(1), 102–11.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews*, *30*(6), 855–63. http://doi.org/10.1016/j.neubiorev.2006.06.011

Singh, H., Kryscynski, D., Li, X., & Gopal, R. (2015). Pipes, pools, and filters: How collaboration networks affect innovative performance. *Strategic Management Journal*, n/a–n/a. http://doi.org/10.1002/smj.2419

Slaughter, S., & Rhoades, G. (2000). The Neo-Liberal University. *New Labor Forum*, *6*, 73–79.

Somech, A. (2006). The Effects of Leadership Style and Team Process on Performance and Innovation in Functionally Heterogeneous Teams. *Journal of Management*, *32*(1), 132–157. http://doi.org/10.1177/0149206305277799

Stark, D. (2009). *The Sense of Dissonance. Accounts of Worth in Economic Life.* Princeton, NJ: Princeton University Press.

Stegmann, S., Roberge, M.-É., & Dick, R. (2012). Getting tuned in to those who are different: The role of empathy as mediator between diversity and performance. *Zeitschrift Für Betriebswirtschaft*, *82*(S2), 19–44. http://doi.org/10.1007/s11573-011-0543-y

Stewart, G. L. (2006). A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance. *Journal of Management*, *32*(1), 29–55. http://doi.org/10.1177/0149206305277792

Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008). The science of team science: overview of the field and introduction to the supplement. *American Journal of Preventive Medicine*, *35*(2 Suppl), S77–89. http://doi.org/10.1016/j.amepre.2008.05.002

Stvilia, B., Hinnant, C. C., Schindler, K., Worrall, A., Burnett, G., Burnett, K., … Marty, P. F. (2011). Composition of scientific teams and publication productivity at a national science lab. *Journal of the American Society for Information Science and Technology*, *62*(2), 270–283. http://doi.org/10.1002/asi.21464

Sundstrom, E., de Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist*, *45*(2), 120–133.

Tajfel, H. (1978). Social categorization, social identity and social comparison. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 61 – 76). London: Academic Press.

Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, *33*, 1–39. http://doi.org/10.1146/annurev.ps.33.020182.000245

The Lehman Brothers Centre for Women in Business. (2007). *Innovative Potential : Men and Women in Teams*. *Women in Business*.

Tost, L. P., Gino, F., & Larrick, R. P. (2012). When Power Makes Others Speechless: The Negative Impact of Leader Power on Team Performance. *Academy of Management Journal*, *56*(5), 1465–1486. http://doi.org/10.5465/amj.2011.0180

Troth, A. C., Jordan, P. J., Lawrence, S. A., & Tse, H. H. M. (2012). A multilevel model of

emotional skills, communication performance, and task performance in teams. *Journal of Organizational Behavior*, *33*(5), 700–722. http://doi.org/10.1002/job.785

Tsoukas, H. (2009). A Dialogical Approach to the Creation of New Knowledge in Organizations. *Organization Science*, *20*(6), 941–957.

Vaan, M. de, Vedres, B., & Stark, D. (2015). Game Changer: The Topology of Creativity. *American Journal of Sociology*, *120*(4), 1144–1194.

van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *The Journal of Applied Psychology*, *89*(6), 1008–22. http://doi.org/10.1037/0021-9010.89.6.1008

van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, *58*, 515–41. http://doi.org/10.1146/annurev.psych.58.110405.085546

Vedres, B., & Stark, D. (2010). Structural Folds: Generative Disruption in Overlapping Groups. *American Journal of Sociology*, *115*(4), 1150–90.

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, *34*(10), 1608–1618. http://doi.org/10.1016/j.respol.2005.08.002

Wang, D., Waldman, D. A., & Zhang, Z. (2014). A meta-analysis of shared leadership and team effectiveness. *Journal of Applied Psychology*, *99*(2), 181–198.

Wang, J., & Hicks, D. (2015). Scientific teams: Self-assembly, fluidness, and interdependence. *Journal of Informetrics*, *9*(1), 197–207. http://doi.org/10.1016/j.joi.2014.12.006

Webber, S. S., & Donahue, L. M. (2001). Impact of highly and less job-related diversity on work group cohesion and performance: a meta-analysis. *Journal of Management*, *27*(2), 141–162. http://doi.org/10.1177/014920630102700202

Wegner, D. M. (1987). Transactive Memory: A Contemporary Analysis of the Group Mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of Group Behavior* (pp. 185–208). New York, NY: Springer New York. http://doi.org/10.1007/978-1-4612-4634-3

West, M. A. (2000). Reflexivity, revolution and innovation in work teams. In M. M. Beyerlein, D. A. Johnson, & S. T. Beyerlein (Eds.), *Product development teams* (Vol. 5, pp. 1–29). JAI Press.

Whitfield, J. (2008). Collaboration: Group theory. *Nature*, *455*(7214), 720–3. http://doi.org/10.1038/455720a

Williams, K. Y., & O'Reilly, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. In B. M. Staw & L. L. Cummings (Eds.), *Research In*

Organizational Behavior (Vol. 20, pp. 77–140). JAI Press.

Wittenbaum, G. M. (2000). The Bias Toward Discussing Shared Information: Why Are High-Status Group Members Immune. *Communication Research*, *27*(3), 379–401. http://doi.org/10.1177/009365000027003005

Wittenbaum, G. M., & Bowman, J. M. (2005). Member Status and Information Exchange in Decision-Making Groups. In M. C. Thomas-Hunt (Ed.), *Status and Groups (Research on Managing Groups and Teams, Volume 7)* (pp. 143 – 168). Emerald Group Publishing Limited.

Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective Intelligence and Group Performance. *Current Directions in Psychological Science*, *24*(6), 420–424. http://doi.org/10.1177/0963721415599543

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, *330*(October), 686–688.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science (New York, N.Y.)*, *316*(5827), 1036–9. http://doi.org/10.1126/science.1136099

Xue, Y., Bradley, J., & Liang, H. (2011). Team climate, empowering leadership, and knowledge sharing. *Journal of Knowledge Management*, *15*(2), 299–312. http://doi.org/10.1108/13673271111119709

Yanow, D. (2004). Translating Local Knowledge at Organizational Peripheries*. *British Journal of Management*, *15*(S1), 9–25. http://doi.org/10.1111/j.1467-8551.2004.t01-1-00403.x

Yong, K., Sauer, S. J., & Mannix, E. A. (2014). Conflict and Creativity in Interdisciplinary Teams. *Small Group Research*, *45*(3), 266–289. http://doi.org/10.1177/1046496414530789

Zheng, W. (2008). A Social Capital Perspective of Innovation from Individuals to Nations: Where is Empirical Literature Directing Us? *International Journal of Management Reviews*, *12*(2), 151–183. http://doi.org/10.1111/j.1468-2370.2008.00247.x

GEDII GENDER DIVERSITY IMPACT – Improving research and innovation through gender diversity

WWW.GEDII.EU

D1.1. Conceptual Framework

Chapter 2 – Research Assessment

Sandra Klatt, Ulf Sandström

31st of May 2016 – Version 1.0

Table of Contents

2. Research Assessment

The assessment of research performance or research output in general is connected with a variety of problems and foremost with the need to define concepts of assessment, specify approach and methodology, and define research performance and output. At the end, assessment results have to be interpreted in light of the applied assessment exercise.

Science assessment in terms of quality of research, assessment of impact, and what methodologies are to be applied is under permanent discussion and evaluation. Research as ´production of knowledge´ used to be located at the scientific institutions foremost for the academic interests, characterized as hierarchical, disciplinary, homogeneous and academically (described in literature as Mode 1); a change in research landscape took place which is referred to as Mode 2 (Gibbons et al., 1994). This „new production of knowledge“ (Gibbons et al., 1994) is characterized as collaborative, trans-disciplinary and an emphasize of more applicable science (in contrast to basic research). Mode 2 also implies the inclusion of societal aspects, users and stakeholders. “Based on these new understandings of knowledge production, research is viewed as a continuum, involving the whole process of discovery and spanning the spectrum from curiosity-driven to user-led, from blue-sky to practice-based.“ (European Commission, 2010, p. 24)

Another interpretation and understanding of the different activities usually named “research” and often held under one heading would be that they cover very different things, on the one hand scientific activities and on the other hand technological activities (de Solla Price, 1965). Each of these has a specific way of creating legitimacy and should therefore be kept apart. While the organization and the rationality behind science are to search for explanation –“why is that…?” – the organization and rationality for technological activities are to create functions for users – “does this work for our needs…?”. The group that gives legitimacy is in the first case the scientific community and in the second case the market or the users (Bunge, 1979). According to traditional sociology of science (Merton, 1973), the only group that can participate in the scientific process at the research front would be the scientists themselves. It would be a contradiction in terms to involve end-users at that stage, as no one knows in what direction valuable results are going to turn up. Exactly this is the argument for scientific self-organization – in the end an open process is the best way to find the unknown good. A huge body of research “sociology of science” is dealing with the role of the scientific community in society, its boundaries as well as the place of the scientific world within society. In this context, for instance it is explored how and why science is established in society although society mostly does not comprehend the scientists work or cannot profit directly from their achievements in science (Ben-David, 1971; Merton, 1973).

As defined by Gibbons (1999), “science has until recently flourished partly because of a stable, underlying agreement between its practitioners and the rest of society. In other words, there has been a social contract between science and society“ (Gibbons, 1999, p. C81). The scientific community was asked to give reliable knowledge that was merely communicated to society. However, the introduction of a new social contract ist to be noted, ensuring „that scientific knowledge is 'socially robust', and that its production is seen by society to be both transparent and participative.“ (Gibbons, 1999, p. C81)

With upcoming of the New Public Management[1], evaluation of research activities became a major task. Government and policy strives to measure research performance especially as part of management of research funding, that is funding was related to performance and quality of research efforts. For example, in the United Kingdom since the 80s a Research Assessment Exercise (RAE) takes place on a regularly basis, Australia has a national evaluation system (Research Quality Framework RQF, now Excellence in Research for Australia ERA), the United States use the Program Assessment Rating Tool (PART) (Grant, Brutscher, Kirk, Butler, & Wooding, 2010). In addition to that, several university rankings can be found, such as the Shanghai Ranking, Times QS World University Ranking, Webometrics, U-Multirank or Leiden Ranking, which have been established as a benchmark to research excellence at universities (European Commission, 2010). However, the rankings result from different assessment approaches, indicators and weighting of these (Aguillo, Bar-Ilan, Levene, & Ortega, 2010; Phillips & Maes, 2012). The rankings stand on their own and are widely known and used.

An expert group on the Assessment of University-Based Research (AUBR) installed by the EU presented a "Multidimensional Research Assessment Matrix" (European Commission, 2010, p. 53) that suggests specific indicators and methods respectively for certain purposes. The indicators are categorized as follows:

- Research Productivity (for example research publications and outputs, research outputs per 'Research Academic' staff)
- Quality and Scholarly Impact (for example number and percentage of publications in top-ranked, high impact journals, citations, number keynote addresses at national and/or international conferences)
- Innovation and Social Benefits (for example external research income, number and percentage competitive grants won, end-user esteem)
- Sustainability and Scale (for example postgraduate research student load, number of collaborations and partnerships)
- Research Infrastructure (for example research active academics, research infrastructure and facilities)

The paper of the expert group state: "The optimum method of research assessment is to combine qualitative and quantitative indicators, e.g. indicator-based data with peer or end-user review." (European Commission, 2010, p. 52)

The research environment is changing. For example, the numbers of academics and researchers is growing, the funding of research has increased and also the scientific output has grown (Vincent-Lancrin, 2006). In addition to that, internationalization and the mobility of researcher have grown (Vincent-Lancrin, 2006). This has an impact on the conditions and possibilities for research assessment. The ICT revolution with introduction of new

1 Guiding Principle for the reform and modernization of state and government since the early 1990s: Market and competitive orientation , target and result oriented control (outcome and output orientation), applying of business models in governance

technologies such as Internet and big data processing provides new possibilities to analyze data by computer and databases (Web of Science, Scopus, Google Scholar) and enables new bibliometric analyzes; ICT and the growth of Social Media (Twitter, Facebook, blogs, etc.) enabled new types of communication; more interdisciplinary and international collaboration among scientists and researchers (virtual teams) can be found; researchers have access to more publications and other scientific output than ever before (Ding, Levin, Stephan, & Winkler, 2010); much more articles/papers are being published (for example via online publishing) (Bornmann & Mutz, 2014) and also, new metrics are developing (altmetrics) (J. Priem, Taraborelli, Groth, & Neylon, 2010).

2.1. Methods of Research Assessment

Research assessment exercises are dominated in general by three different methods. Bibliometrics, peer review (as traditional metrics) and a mixture of these two aforementioned methods, which is known for example as informed peer review (relatively fresh to the peer review scenery). In addition to that, new indicators (altmetrics) are prevailing the assessment scenery. Although citation based assessment is initially a quantitative exercise, the derived impact indicators are often seen as equals for quality (van Raan, 2005). Peer Review is acknowledged as a ´gold standard´ for assessment of research quality and can be found established for article selection of journals, as exercise for decisions of funding or career recruitment/promotions, and as research assessment on national level. However, research assessment can also be categorized in relation to the unit that is to be assessed. The following levels are being used: nation / country, institution, departments, research group and/or individual researcher (Moed & Halevi, 2015). Depending on the assessment level and purpose of the assessment, research assessment approaches and methods must be appropriate (European Commission, 2010; Moed & Halevi, 2015).

In the next sections the types of assessment are briefly presented and discussed, with main focus on gender awareness and bias respectively.

2.1.1. Bibliometrics or Quantitative Citation-based Assessment

Well-acknowledged outputs of research are papers, books, and articles in journals or contributions to conferences. Next to mere counting of output, that is counting publications and/or other outputs per person or per group, Scientometrics has become important as a “study of the quantitative aspects of science and technology (...). Some of the main themes include ways of measuring research quality and impact, understanding the processes of citations, mapping scientific fields and the use of indicators in research policy and management” (Mingers & Leydesdorff, 2015, p. 1). One of the core topics is the measurement of citations. Back in the 1950s the first concepts of the Science Citation Index emerged, originally as a way to systematize search options in literature (Garfield & Sher, 1963). Soon further possibilities how to analyze the indexed citations on a statistical basis have been detected. At the same time, the social importance of measuring science and research output was pushed by governments and policy makers, who were keen to measure science in order get decision guidance for research grants and funding (van Raan, 2005).

Mingers & Leydesdorff (2015) as well as Durieux & Gevenois (2010) give a profound overview

on bibliometric indicators that can be measured. At the same time advantages and flaws are being discussed in detail (see Durieux & Gevenois, 2010; Mingers & Leydesdorff, 2015). The most important bibliometric indicators that measure impact are shortly presented in the following section.[2]

First of all, the data source of the analyzes should be briefly reflected, since for applying the indicators, several sources are available, however these sources vary in content, extent and data quality. The oldest data source Web of Science (WoS) or ISI Web of Knowledge is a database with citation indices in different disciplines; it covers journals, conference proceedings and books. Scopus (launched 2004) is similar to WoS, however the coverage differs. The third great database is Google Scholar (launched 2004) and covers available scholary literature on the web; Google Scholar also includes theses, technical reports and such into the data base. In contrast to WoS and Scopus, it is free accessible via Google on the internet.

The Journal Impact Factor (JIF) goes back to Garfield & Sher (Garfield & Sher, 1963) and describes how often on average an article is published in this journal is cited from other scientific articles in two years and five years respectively. WoS has also introduced a metric called cited half-life. Although a very dominant measure, it was acknowledged that many defencies came along with using the JIF like depending heavily on research field (publishing and citation habits), depending on the data resources (WoS or Scopus), manipulating the JIF by the journals themselves or no possibility to assess an individual researcher or paper (Mingers & Leydesdorff, 2015). As response to these defencies, more advanced metrics have been developed, for example including a prestige factor (called Eigenfactor at WoS, article influence AI at Scopus). In addition to that, there are further advanced, complex developments of indicators that include normalization factors or relationships of journals (SJR, SJR2 at Scopus). However, these complex indicators are hardly transparent any more.

On the level of the researcher or research group, one can find also metrics for research performance based on citations.

Citations are counted usually for a certain collection of publications in a certain period of time. The easiest way is to count the total numbers of citations, whereby the total number is of course strongly affected by the number of publications, i.e. productivity. The traditional indicator, which cuts out the total number of papers, is the cites per paper (cpp) (van Raan, 2005). This indicator was used as a basis for the “crown indicator”, which measured – normalized – the impact of researcher or research unit by taking into account factors like for instance research domain, type of paper, language or region (Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011). The crown indicator was further developed as mean normalized citation score (Waltman et al., 2011) and is basis of the Leiden Ranking methodology.[3]

Another highly influential indicator is the so-called h-index, introduced by Hirsch in 2005 (Hirsch, 2005). Main advantages are a simple way to combine impact (i.e. citations) and

2 In this context, we refer to scientometrics, a study field in ist own right, discussing and developing methodologies, indicators in scientometrics, see also Journal of Scientometrics, established in 1978.

3 Leiden Ranking measures the performance of universities based on a set of indicators, see www.leidenranking.com

productivity (i.e. number of papers); the h-index can by applied not only to an individual researcher, but also to units or journals themselves. It is robust to poor data and is seen as a suitable metric for comparing researchers. However, as the h-index works as a simple enough metric, there are flaws attached, for instance the h-index gives no information about the total number of citations, so that two researchers could differ extremely in the number of citations, but would have the same h-index. Because of its disadvantages, the h-index is subject to numerous suggestions to modify the calculation model, over time the g-index, m-index or r-index have been proposed (Mingers & Leydesdorff, 2015).

Nonetheless, probably because of its simplicity, the h-index is well established. For example, all three above mentioned databases are providing the h-index as a benchmark among other indices.

An important discussion in literature is about normalization methods for indicators in order to be able to compare and classify impact across disciplines, but also within disciplines when citation rates differ significantly. Also there is evidence that citations of a paper increase as times go by, thus only normalization would enable comparable results (Waltman & van Eck, 2013). Normalization methodologies refer to field classifications like the mean normalized citation score, to source normalization for instance like audience factor or fractional counting of citations.

In addition to the before mentioned methods, percentile-based approaches are being used. In order to avoid mean-based measures, non-parametric measures based on percentiles have been suggested for research groups, researchers and also journals. Publications that are being evaluated are assigned to percentile rank classes. These rank classes have been built on reference papers, which are split percentile wise like top 1 percent, top 10 percent etc., and for each percentile rank class a minimum of citations necessary to get into the rank class is specified. This method is used for example by the US National Science board (Bornmann, Leydesdorff, & Mutz, 2013).

Few papers receive many citations and many papers few citations. Depending on the discipline, citation and publication practices differ immensely (even within disciplines). Therefore, mean value based indicators cannot reflect the importance of publications. In addition to that, the number of citations increases over time and the distribution of citations per year is usually not a linear, but skewed. A good, but fairly simple method would be median measure or a percentile measure, such as top percentage number x% as an indicator of citation distribution. If there is a huge proportion of publications in the top x%, a huge impact of the scientific output can be assumed (Swedisch Environmental Protection Agency, 2014).

As state of the art indicators, advanced size-independent field-normalized indicators are being used to measure performance. However, bibliometricians use also size-dependent citation indicators, as they may give valuable rsults for discussing hypotheses. (Sandström & Wold, 2015; van den Besselaar & Sandström, 2015)

2.1.2. Limitations to bibliometric measurements

Bibliometric indicators are to be used with caution, and for statistical handling, evaluations appropriate expertise is required. Furthermore, for the interpretation of results the overall

context of analyzes is to be considered. Wouters et al. (2015) summarize these necessary considerations when conducting bibliometric analyzes as follows: “differences between academic subjects/disciplines; coverage of sources within databases; the selection of the appropriate unit of analysis for the indicator in question; the question of credit allocation where outputs may include multiple authors, and accounting for self-citations.” (Wouters et al., 2015, p. 30)

Literature also discusses the “gaming” with bibliometric measures (Haustein & Larivière, 2015). For example because bibliometric indicators are basis for funding decisions, researchers and institutions try to manipulate the relevant indicator to their favor, through salami publishing, self-citation or even citation cartels. The upcoming of a lot of new (often open access) journals - with a sometimes-questionable threshold of quality that publish quite an amount of articles - dilute the quality of the indicator and therefore questions the evidence of the indicator (Beall, 2012; Harzing, 2012).[4] In order to raise awareness to such questionable journals and publishers, Beall listed “potential, possible, or probable predatory scholarly open-access publishers” (Beall, 2016).

Although bibliometricians see citation rates as an appropriate measure of academic impact (Leydesdorff & Milojević, 2015), this may represent only one dimension of academic quality. “Quality needs to be seen as a multidimensional concept that cannot be captured by any one indicator, and which dimension of quality should be prioritized may vary by field and mission.” (Wouters et al., 2015, p. 47)

2.1.3. Peer review assessment

Peer review can be classified according to the objectives of the review:

- assess the quality of research results, outcomes, projects and programmes;
- determine the level of performance, either in absolute terms or comparatively, of (parts of) the scientific and innovation system;
- promote accountability;
- contribute criteria and evidence for resource allocation;
- contribute criteria and evidence for science and technology policy making;
- contribute criteria and evidence for career decisions and human resource policies (after Geiser & Wager 2002, in: Wouters et al., 2015).

Peer review is widely accepted as a methodology to assess the quality of research output such as publications (Abramo, D’Angelo, & Di Costa, 2011; van Raan, 2005), and it has long been exercised for assessing article manuscripts for publication in journals; some countries use it as method for allocating national research funding, for example United Kingdom (Wilsdon et al., 2015).

There are different ways of conducting a peer review process such as double / single blind review or open reviews (Manchikanti, Kaye, Boswell, & Hirsch, 2015); flaws and limitations of

4 Harzing for example identified a “super author”, who published nine of Essential Science Indicators´ (ESI) most highly-cited articles in Economics & Business in two years, however these articles were published by the same publisher as well as 40 articles overall in these two years. With help of Journal self citation rate, publisher self citation rate and author co-citation (Harzing, 2012)

peer review methods are exhaustive discussed in literature, especially peer review bias such as conservatism bias, conflict of interests or gender bias. A comprehensive overview is given by Lee, Sugimoto, Zhang, & Cronin (2013)(2013). Gender bias is discussed in chapter 3 in detail.

However, in order to improve peer review processes, combinations with bibliometric indicators are suggested (Abramo et al., 2011), even the replacement of peer review in context of grant decisions with bibliometric methodologies is proposed, as studies show that bibliometric approaches serve the purpose of finding excellence better (Abramo et al., 2011). In contradiction to this, the Higher Education Funding Council for England (HEFCE) discussed the introduction of citation-count based methodologies instead of peer review in United Kingdom, but at last decided against it, since in the latest report of the HEFCE the methodology of peer reviews was evaluated and confirmed in general. Expert peer review should be applied as the main method since this method “is the least worst form of academic governance we have“ (Wilsdon et al., 2015, p. 136). As in the opinion of the HEFCE, quantitative indicators are being seen as not feasible to replace a peer review process. Main objective is that only with qualitative methods such as peer review and impact case studies the context of research outcome can be taken into account properly (Wilsdon et al., 2015).

From a gender perspective, certain types of peer review seem to be gender biased with a negative outcome for women. Wenneras & Wold (1997) published their findings that peer reviews favor men and nepotism. They showed that if the reviewer in the peer review system consisted mainly of men, the percentage of successful male applicants is significant. The same seems to apply to review exercises in assessing journal articles. Male authors are more likely to be granted a publication in a journal than female authors. The introduction of double blind reviews, where neither reviewer nor author identity is known to each other, at the Behavioral Ecology journal showed that the share of female authors increased; these findings support the suspicion that female authors are disadvantaged (Budden et al., 2008). In contrast, other studies found no significant gender bias (Marsh, Bornmann, Mutz, Daniel, & O’Mara, 2009); Lee et al. (2013) give a short overview on contradictory studies.

The supporters of bibliometric assessments name the bias of peer review as well justified reasons to shift science and research assessment to bibliometrics (Abramo et al., 2011), at least to combine the advantages of bibliometric measurements with peer reviews. In practice, peer review remain one of the most accepted methodologies for assessment for publications, however, in cases of assessment of applicants or institutions/universities bibliometric measures have been introduced as a supportive or supplementary method for the reviewers (see for example United Kingdom, Wilsdon et al., 2015).

There seems to be a dilemma. On the one hand, the percentage of women in science and research needs to be improved, given the rates of female students and graduates, but the phenomenon of the “leaky pipeline” (Shaw & Stanton, 2012) persists. On the other hand, despite numerous efforts to enhance women in science, there are hints that the knowledge of the applicant’s gender may influence the assessment in that way that female authors or applicants have greater hurdles to pass than their male counterparts.

It is well acknowledged that research disciplines are diverse in respect to their measurable outcomes, their (societal) impact and also habits of producing results such as publishing.

Therefore, metrics or indicators for appropriate assessment must consider the diversity of research and the precisely characteristics of the research output in question.

2.1.4. Patents

Another approach of research assessment is the analytics of patents. Patents are used foremost as indicator for R&D Output. OECD and also EU are publishing patent statistics as innovation indicators (OECD, 2008), however primarily for the private sector. For example, in 2003 – 05 more than 80% of patent applications were filed by the private sector, whereas 4% were only filed by universities (OECD, 2008).

Patents enable the measurement of inventive activity across various technology sectors and at many different levels including individuals (inventors), institutions, regions and countries. Thus, patents are a measure of economic impact of research (IP Australia, 2013).

As to measure impact with patent data, patent citation is suggested as metrics in literature. Thereby, citations of patents by other patents (applications) and non-patent literature can be found as impact indicator (Giuri et al., 2007; Glänzel & Meyer, 2003). Bornmann (2013) for example classified patent citation as a good measure for economic and therefore societal impact.

Patents are part of measuring intellectual property in context of research performance exercises in statistics, evaluation or allocation of grants. While in statistics mere counting of patents is being done, in evaluation (for example Spain university-industry collaboration) or allocation (for example Sweden ´The Academic Ranking & Rating Agency´ for allocating basic grants) derived patent indicators are being used (Mortensen, 2011).

Several studies have been conducted to analyze patent data from a gender perspective.

Naldi and Vannini Parenti (2002) presented the first study based on a huge database that took sex-disaggregated data analysis of patents (and publications) into consideration. For this purpose they created a First Name Data Base (FNDB), which is used as basis for analysis exercises to date. They gave a comprehensive overview about the inventors' scenery and showed among others things that women inventors are highly underrepresented.

Ding et al. (2006) showed that in the discipline of Life Science a gender gap in patenting exists. By analyzing 4227 life scientists over a 30-year period they found that women patent 40% of the rate of men. The gender gap improved over time but remained very significant.

Frietsch et al. (2009) analyzed patent applications at the European Patent Organization (EPO) of five priority years and overall more than 2.4 million inventors. The study showed that the contribution of women increased during the years, however, remains on a very low level (highest rate in Spain with 12.3% in 2003-2005 over all disciplines, highest rate of female inventors in the field of pharmaceuticals with 21%). For patent activities, the study pointed to the correlation that "(.) [the] higher the income of (public and) private researchers, the lower the representation of women. Or to put it the other way around: if researchers are paid comparatively well, the representation of men is higher" (Frietsch et al., 2009, p. 597).

Busolt and Kugele (2009) also analyzed patent data of the EPO. In addition to Frietsch et al., they looked into the input-output-relation and found a significant research productivity gap.

That gap „will increase if the percentage of female researchers is increased without improving their output productivity by providing more resources and better working conditions at the same time“ (Busolt & Kugele, 2009, p. 120).

Sugimoto et al. (2015) analyzed 4.6 million patents from 1976 to 2013 issued by the United States Patent and Trade Office (USPTO) and found that “female patenting is proportionally more likely to occur in academic institutions than in corporate or government environments. However, women’s patents have a lower technological impact than that of men, and that gap is wider in the case of academic patents” (Sugimoto et al., 2015, p. 1). The findings apply to all disciplines. They also found evidence that women, especially academic women, are more collaborative and multidisciplinary than their male counterparts and that the patenting rates of women in academia are higher compared to industry, government and individuals. Sugimoto et al. (2015) concluded that women find more supporting conditions for patenting in academia than in the other sectors.

2.1.5. Altmetrics

As problems with judging the value of scientific articles have been observed and the value of metrics such as the impact factor have been questioned (in terms of their limitations), the search for new tools for measuring impact was launched (Neylon & Wu, 2009). New measures for assessment of research include metrics such as data analysis of downloads, view counts, web citation or link analysis. The web was discovered as profound source information related to scientific activities, that may bridge the limitations and flaws of the traditional metrics, for example web-based metrics may predict which publications are later more likely cited or give more attention to publications or scientific activities in disciplines with less citation practice (Zahedi, Costas, & Wouters, 2014).

Since especially the time gap of reliable citation-based measures is unsatisfying, Neylon & Wu (2009) suggested to count downloads in research libraries such as Mendeley, CiteUlike or Zotero with the argumentation that such reference managers document active interest in the publication.

Priem (2010) introduced the term “altmetrics” – another term is for example scientometrics 2.0 (Priem & Hemminger, 2010) - for describing alternative metrics, retrieved from the web. He claims the following advantages:

- Researchers use more and more the web, Zotero and Mendeley store more than 40 Millions publications, so that activities of scholars can be traced in the web respectively in blogs, twitter, and on the online research libraries.
- Altmetrics include a broader view on the impact such as raw science (data sets, codes), semantic publishing (arguments or passages instead of an whole article) or self-publishing formats (blogging, comments on existing work).
- Altmetrics are diverse, so that altmetrics are ideal for reflecting diverse landscape of research disciplines.
- Altmetrics are fast, data could be retrieved within days or weeks.
- Altmetrics are open (data, scripts, algorithms).

Priem postulates the vision that altmetrics could “crowd-source peer review” (J. Priem et al.,

2010). “In the future, greater participation and better systems for identifying expert contributors may allow peer review to be performed entirely from altmetrics. Unlike the JIF, altmetrics reflect the impact of the article itself, not its venue. Unlike citation metrics, altmetrics will track impact outside the academy, impact of influential but uncited work, and impact from sources that aren’t peer-reviewed.” (J. Priem et al., 2010)

The idea of altmetrics as viable alternative or supplement to traditional metrics is discussed in the scientific community. Studies show that there is a wide range of possible web tools and that they measure research impact, however, a systematic way of research assessment by altmetrics seems still to be missing. Even though the possibilities of alternative metrics seem very promising, Wouters & Costas (2012) state a lack of critical consideration of the quality and the reliability of data and new indicators respectively. Another aspect is the individual (author) level that altmetrics usually cover. Other studies have investigated the correlations of altmetrics with traditional metrics. These studies have found moderate to significant correlations when comparing citations and citation and reads on Mendeley (among others, see Thelwall, Haustein, Larivière, & Sugimoto, 2013), but rather weak correlations to bookmarks or users tags for physical journals (Zahedi et al., 2014). Studies found that Mendeley readership (as indicator) seems to be a major player in the altmetrics tool set, since the online reference manager covers high percentiles of publications of diverse disciplines as well as multidisciplinary journals (Zahedi et al., 2014).

Summarized one can say that altmetrics or new metrics promise new possibilities to measure impact, especially in a broader sense (that would also support the measurement of societal impact, for example measurement of interactions between science and stakeholders, see below). However, as altmetrics are still in the early stages of development and acceptance, a general acknowledged conceptual framework is needed for finding a proper place in the assessment scenery. Furthermore, it is still research to be done if and how altmetrics are suitable for gender equitable measurements.

2.2. Societal Impact of Research

As the measurement of scientific output and performance has been developed and further evolved as scientific impact measurement after World War II, the demand for evidence and evaluation of research for society arose in the 1990s. The measurement of impact on scientific knowledge alone was not enough anymore, but it was asked for evidence for the value for society (Bornmann, 2013). This development in science policy is well described by the new social contract between science and society (see introduction to the chapter, p. 1).

But what is exactly societal impact of research? And how can it be measured? In the following section main findings in literature are briefly summarized.[5] The definition of societal impact varies in the literature; it is named as third stream activities, social benefits or social quality, public values or social relevance (Bornmann, 2013). Summarized it is described as economic, social, environmental and cultural benefits (European Commission, 2010):

„Economic Benefits, e.g. improved productivity; adding to economic growth and wealth creation; enhancing the skills base; increased employment; reduced costs; increased

5 Please be referred also to Bornmann (2013) who provides a comprehensive literature review on this topic.

innovation capability and global competitiveness; improvements in service delivery; as well as unquantifiable economic returns resulting from social and public policy adjustments.

Social Benefits, e.g. improving people's health and quality of life; stimulating new approaches to social issues; changes in community attitudes; influence upon developments or questions in society at large; informed public debate and improved policy-making; enhancing the knowledge and understanding of the nation; improved equity; improvements in health, safety and security; improved social attachment; and improvements in the level and security of political rights.

Environmental Benefits, e.g. improvements in environment and lifestyle; reduced waste and pollution; improved management of natural resources; reduced consumption of fossil fuels; uptake of recycling techniques; reduced environmental risk; preservation initiatives; conservation of biodiversity; enhancement of ecosystem services; improved plant and animal varieties; and adaptation to climate change.

Cultural Benefits, e.g. supporting greater understanding of where we have come from, and who and what we are as a nation and society; understanding how we relate to other societies and cultures; stimulating creativity within the community; contributing to cultural preservation and enrichment; and bringing new ideas and new modes of experience to the nation." (European Commission, 2010, pp. 41–42)

The aforementioned listing shows that societal impact consists of multiple dimensions which cannot bee seen separately, but often as a mixture or at least with fuzzy boundaries (Spaapen, Shinn, Msh-paris, & Marcovich, 2011). In addition to that, when discussing societal impact one must mention the different geographical dimensions (for example impact is not limited to a certain region or country when assessing climate change), and the broad groups of stakeholders, who are affected by societal impact; they are identified as:

„1. policy makers at the intermediary or government level, whose goal is either to use research for their own policies, or to facilitate the transfer of knowledge from science to society;

2. Professional users (profit and nonprofit); that is, industry and societal organizations that want knowledge to develop products and services (this may refer to researchers who profit from developments in other disciplines);

3. end users; that is, the public at large or individual target groups (for example farmers, aids victims)." (Spaapen, Dijstelbloem, & Wamelink, 2007, p. 79)

As a consequence, several problems are linked with the assessment of societal impact (after Bornmann, 2013, p. 219):

(a) Causality problem: It is not clear what impact can be attributed to what cause.

(b) Attribution problem: it is not clear what share of impact can be allocated to a specific research.

(c) Internationality problem: Research is international, so that a clear allocation of input and impact is impossible.

(d) Evaluation time problem: Premature impact measurement may result in policy

actions that contribute to only short-term benefits.

(e) Assessment through peer experts instead of indicators is difficult to conduct because of lack of volunteers. “Scientists generally dislike impacts considerations” and it “takes scientists beyond the bounds of their disciplinary expertise” (Holbrook & Frodeman, 2011, p. 244).

(f) Different sciences such as medicine or social science have different research results and may affect many different aspects of society, so that it will not be possible to have an assessment concept that fits all (Martin, 2011; Molas-Gallart, Salter, Patel, Scott, & Duran, 2002).

(g) Societal impact assessment has to be set in cultural and national context (Molas-Gallart et al., 2002).

(h) Societal impact can be negative or positive (Martin, 2011).

Evidently the assessment of societal impact is a very complex business and there are neither accepted frameworks nor universally valid indicators for societal impact assessment. However, one can find societal impact assessment exercises on national level (Netherlands, UK), with grant peer review processes (US National Science Foundation), in studies (with primary focus on economic dimension), and not least projects with the purpose to develop indicators for this very societal impact assessment.

Summarized three types of measurements can be differentiated: Measurement with data (applicable especially to economic dimension), method of surveys and a case study approach.

In this context, the scope of the European Commission is to “measure what counts” (European Commission, 2010). In the report on university ranking, societal impact is also discussed and possible indicators are suggested to assess the societal impact, that is external research income, number and percentage competitive grants won, research income per academic staff or FTE, employability of PhD graduates, commercialization of research generated intellectual property (IP), end-user esteem, number and percentage funding from end-users (e.g. industry, professions, government, community). Most of the indicators describe an interaction with extern stakeholders of university. This concept of measurement can be found also in other studies, for example in the SIAMPI-project as key approach (Spaapen et al., 2011). Spaapen and van Drooge specified the relevant interactions as “productive interactions (…) in which knowledge is produced and valued that is both scientifically robust and socially relevant.“ (Spaapen & van Drooge, 2011, p. 212) By assessing these interactions in terms of direct, indirect and financial interactions, the (potential of) societal impact could be identified. “By discussing social impact in terms of the three kinds of ‘productive interactions’, it became clear that data could be gathered that can give useful information about steps that are necessary to achieve social impact and about the relationship with various stakeholders. Also, the researchers became more aware of the value of their contributions to social impact. (…) [The] enlightenment function also works for stakeholders.“ (Spaapen & van Drooge, 2011, p. 217) By this approach, researchers as well as stakeholders and society are becoming more aware of the contribution of research for society.

A similar approach is undertaken by the project SISOP, which tried to conceptualize a model for the system Science-Society and Science Production System in a broad context. Interactions are the key concept of the project. In the science society systems, scientists and scientific institutions interact with other actors (e.g. governments, decision-makers) to produce, distribute and use knowledge and artifacts (i.e. products like papers). “The performance of science-society and science production systems depends significantly on relationships and information flows between actors within the system and belonging to different systems. Policy actions may influence and modify these relationships. The performance of these systems depends on a broad range of social, cultural, economic and legal factors. These constitute the context in which the system operates.“ (Walker et al., 2013, p. 23) Measuring impact was done by knowledge sharing, trajectories and also automated text analyses as well as scientometric tools. However, the project investigated on case study level and therefore used one of the three main measurement types for societal impact (discussed above).

Another approach in literature is the attempt to measure the societal impact by indicators, analogously to bibliometrics, which have developed well-sophisticated indicators (as discussed earlier). It is not surprising that for that purpose indicators that are already used are being evaluated and tested. Bornmann (2014) used altmetrics (twitter counts by altmetrics.com) in combination with the F1000 peer review system. His analyzes show that papers with the tag “good for teaching” (interpreted as tag for impact beyond science, in a broader sense) show a higher count of altmetrics, whereas papers with the tag of “new finding” (seen as scientific tag) are much more cited according to traditional indicators. He concludes that papers aiming for readers outside the research community unfold a certain social impact, measurable by altmetrics. “However, it is not clear which kind of impact is measured: Does it measure social, cultural, environmental and/ or economic impact? With evaluating citations in university text books (impact on education), patents (impact on industry) and clinical guidelines (impact on clinical praxis), there are already some approved instruments available for the reliable societal impact measurements which could be complemented by altmetrics.” (Bornmann, 2014, p. 31)

A further suggestion is to write assessment reports especially addressed to society (Bornmann & Marx, 2014). These reports could be a service from science for society, by making reliable knowledge available. Bornmann & Marx present „(…) an alternative for the measurement of societal impact which, unlike other proposals (see an overview in Bornmann 2012, 2013), has a number of advantages, particularly compared to carrying out case studies.“ (Bornmann & Marx, 2014, p. 217) The approach is seen as relatively simple, at the same time based directly on the formal scientific communication process. It could be applied in every research field and would be therefore comparable. However, the claim that the preparing of such assessment reports may be no additional effort has to be questioned and also, the societal impact of such assessment reports must still be measured, for example by altmetrics. Even if the proposed assessment reports are activities of societal nature, the question of measurement of social impact is not really answered.

2.3. Responsible Research and Innovation

In recent years, the framework for science is discussed as responsible research and innovation. From a more detached science the way leads to a science that is more connected with society. The shifting strategy of the European Union can be seen in shifting goals of "Science and Society" (6th Framework Programme) to "Science in Society" (7th framework programme) and lately to "Science with and for Society" in Horizon 2020. The interaction and influence of science seems to change and society has a growing influence on science by "re-contextualizing, by ethicisation or simply by advocating democratization" (European Commission, 2009, p. 19). It is asked for the value of science for society and therefore what society values (Owen, Macnaghten, & Stilgoe, 2012; see also Wilsdon, Wynne, & Stilgoe, 2005).

2.3.1. Responsible Metrics

As consequence to the demand of responsible research, responsible metrics are proposed. Responsible metrics can be understood in terms of a number of dimensions (see Wilsdon et al., 2015; Wouters et al., 2015):

- Robustness: basing metrics on the best possible data in terms of accuracy and scope;
- Humility: recognizing that quantitative evaluation should support – but not supplant – qualitative, expert assessment;
- Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results;
- Diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system;
- Reflexivity: recognizing and anticipating the systemic and potential effects of indicators, and updating them in response.

As counterpart to responsible metrics, a deterrent example has been introduced: Bad metrics shall be made public in order to raise awareness for an informed and appropriate handling of metrics (www.responsiblemetrics.org).

2.3.2. Critical voices regarding research assessment: DORA and Leiden Manifesto

The San Francisco Declaration on Research Assessment (DORA) was initiated by American Society for Cell Biology (ASCB) and a group of editors and publishers of scientific journals (American Society for Cell Biology, 2013). The declaration criticized the frequent practice of research assessment, that is the use of impact factors as proxies for scientific merit. DORA promotes the following core themes:

- "the need to eliminate the use of journal-based metrics, such as Journal Impact Factors, in funding, appointment, and promotion considerations;
- the need to assess research on its own merits rather than on the basis of the journal in which the research is published; and
- the need to capitalize on the opportunities provided by online publication (such as relaxing unnecessary limits on the number of words, figures, and references in articles, and exploring new indicators of significance and impact)."

The declaration has got to the heart of uneasiness and criticism concerning assessment exercises; a wider discussion of research quality and merits has been kicked off (for example Misteli, 2013; Pulverer, 2015; Wilsdon et al., 2015).

One crucial problem of the use of metrics becomes clear: If the indicators are not properly used (as intended) and interpreted, so that a “misuse” is to be stated, then the results of application of these metrics are meaningless or misleading. These findings lead to another declaration or manifesto that appeals to the right use of metrics in order to get high quality results: the so-called Leiden Manifesto.

The scientific community of scientometrics is well aware of problems attached to quantitative measurements and use of indicators. Therefore, a manifesto was issued. The manifesto summarizes the known crucial aspects for dealing with metrics in science to ascertain a proper use and interpretation as follows (Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015):

1) Quantitative evaluation should support qualitative, expert assessment.

2) Measure performance against the research missions of the institution, group or researcher.

3) Protect excellence in locally relevant research.

4) Keep data collection and analytical processes open, transparent and simple.

5) Allow those evaluated to verify data and analysis.

6) Account for variation by field in publication and citation practices.

7) Base assessment of individual researchers on a qualitative judgment of their portfolio.

8) Avoid misplaced concreteness and false precision.

9) Recognize the systemic effects of assessment and indicators.

10) Scrutinize indicators regularly and update them.

The manifesto claims “research evaluation can play an important part in the development of science and its interactions with society. Research metrics can provide crucial information that would be difficult to gather or understand by means of individual expertise” (Hicks et al., 2015, p. 431). However, the authors also note that beside quantitative metrics (“robust statistics” Hicks et al. 2015, p.431) also qualitative assessment is asked for.

Bibliography

Abramo, G., Cicero, T., & D'Angelo, C. A. (2015). Should the research performance of scientists be distinguished by gender? *Journal of Informetrics*, *9*(1), 25–38. http://doi.org/10.1016/j.joi.2014.11.002

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). National research assessment exercises: a comparison of peer review and bibliometrics rankings. *Scientometrics*, *89*(3), 929–941. http://doi.org/10.1007/s11192-011-0459-x

Aguillo, I., Bar-Ilan, J., Levene, M., & Ortega, J. (2010). Comparing university rankings. *Scientometrics*, *85*(1), 243–256.

American Society for Cell Biology. (2013). The San Francisco Declaration on Research Assessment (DORA). Retrieved from http://www.ascb.org/dora/

Beall, J. (2012). Predatory publishers are corrupting open access. *Nature*, *489*(7415), 179.

Beall, J. (2016). Scholary Open Access. Retrieved February 15, 2016, from https://scholarlyoa.com/

Ben-David, J. (1971). *The scientist's role in society: a comparative study*. Prentice-Hall.

Bornmann, L. (2013). What is societal impact of research and how can it be assessed? a literature survey. *Journal of the American Society for Information Science and Technology*, *64*(2), 217–233. http://doi.org/10.1002/asi.22803

Bornmann, L. (2014). Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime. *Journal of Informetrics*, *8*(4), 935–950.

Bornmann, L., & Marx, W. (2014). How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons. *Scientometrics*, *98*(1), 211–219.

Bornmann, L., & Mutz, R. (2014). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. Digital Libraries; Physics and Society; Applications. Retrieved from http://arxiv.org/abs/1402.4578

Bornmann, L., Leydesdorff, L., & Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits. *Journal of Informetrics*, *7*(1), 158–165. http://doi.org/10.1016/j.joi.2012.10.001

Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favours increased representation of female authors. *Trends in Ecology & Evolution*, *23*(1), 4–6. http://doi.org/10.1016/j.tree.2007.07.008

Bunge, M. (1979). *Treatise on Basic Philosophy: Ontology II: A World of Systems*. Springer Netherlands. Retrieved from https://books.google.de/books?id=4hpNzUzH1E4C

Busolt, U., & Kugele, K. (2009). The gender innovation and research productivity gap in Europe. *International Journal of Innovation and Sustainable Development*, *4*(2-3), 109–122.

Cole, J. R., & Zuckerman, H. (1984). The productivity puzzle: Persistence and change in patterns of publication of men and women scientists. *Advances in Motivation and Achievement*, *2*(2), 1.

de Solla Price, D. J. (1965). Is technology historically independent of science? A study in statistical historiography. *Technology and Culture*, 553–568.

Ding, W. W., Levin, S. G., Stephan, P. E., & Winkler, A. E. (2010). The Impact of Information Technology on Academic Scientists' Productivity and Collaboration Patterns. *Management Science*. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1100.1195#.Vu_SisaOxVo.mendeley

Ding, W. W., Murray, F., & Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *Science*, *313*(5787), 665–667.

Durieux, V., & Gevenois, P. A. (2010). Bibliometric Indicators: Quality Measurements of Scientific Publication. *Radiology*, *255*(2), 342–351. http://doi.org/10.1148/radiol.09090626

European Commission. (2009). *Challenging Futures of Science in Society: Emerging trends and cutting-edge issues*. *Report of the MASIS Expert Group setup by the European Commission. EUR 24039*.

European Commission. (2010). *Assessing Europe's University-Based Research*. *Research Policy*. http://doi.org/10.2777/80193

Frietsch, R., Haller, I., Funken-Vrohlings, M., & Grupp, H. (2009). Gender-specific patterns in patenting and publishing. *Research Policy*, *38*(4), 590–599. http://doi.org/10.1016/j.respol.2009.01.019

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, *14*(3), 195–201. http://doi.org/10.1002/asi.5090140304

Gibbons, M. (1999). Science's new social contract with society. *Nature*, *402*, C81–C84.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. Sage.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., … Harhoff, D. (2007). Inventors and invention processes in Europe: Results from the PatVal-EU survey. *Research Policy*, *36*(8), 1107–1127.

Glänzel, W., & Meyer, M. (2003). Patents cited in the scientific literature: An exploratory study of'reverse'citation relations. *Scientometrics*, *58*(2), 415–428.

Grant, J., Brutscher, P.-B., Kirk, S. E., Butler, L., & Wooding, S. (2010). Capturing Research Impacts: A Review of International Practice. Documented Briefing. *Rand Corporation*.

Harzing, A.-W. (2012). How to become an author of ESI Highly Cited Papers. *University of Melbourne*.

Haustein, S., & Larivière, V. (2015). The use of bibliometrics for assessing research: possibilities, limitations and adverse effects. In *Incentives and Performance* (pp. 121–139). Springer.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden manifesto for research metrics. *Nature*, *520*, 429–431.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* , *102* (46), 16569–16572. http://doi.org/10.1073/pnas.0507655102

Holbrook, J. B., & Frodeman, R. (2011). Peer review and the ex ante assessment of societal impacts. *Research Evaluation*, *20*(3), 239–246. http://doi.org/10.3152/095820211X12941371876788

IP Australia. (2013). *Research Performance of University Patenting in Australia: A Pilot Assessment*. Retrieved from http://www.ipaustralia.gov.au/uploaded-files/publications/Research_Performance_of_University_Patenting_in_Australia.pdf

Larivière, V., Ni, C. Q., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Global gender disparities in science. *Nature*, *504*(7479), 211–213.

Larivière, V., Vignola-Gagné, E., Villeneuve, C., Gélinas, P., Gingras, Y., Larivière, V., … Gélinas, P. (2011). Sex differences in research funding, productivity and impact: an analysis of Québec university professors. *Scientometrics*, *87*, 483–498. http://doi.org/10.1007/s11192-011-0369-y

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, *64*(1), 2–17.

Leydesdorff, L., & Milojević, S. (2015). Scientometrics. In Micheal Lynch (Ed.), *International Encyclopedia of Social and Behavioral Sciences*. Elsevier.

Maliniak, D., Powers, R., & Walter, B. F. (2013). The gender citation gap in international relations. *International Organization*, *67*(04), 889–922.

Manchikanti, L., Kaye, A. D., Boswell, M. V, & Hirsch, J. A. (2015). Medical journal peer review: process and bias. *Pain Physician*, *18*(1), E1–E14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25675064

Marsh, H. W., Bornmann, L., Mutz, R., Daniel, H.-D., & O'Mara, A. (2009). Gender effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research*, *79*(3), 1290–1326.

Martin, B. R. (2011). The Research Excellence Framework and the "impact agenda": are we creating a Frankenstein monster? *Research Evaluation* , *20* (3), 247–254. http://doi.org/10.3152/095820211X13118583635693

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, *246*(1), 1–19.

http://doi.org/10.1016/j.ejor.2015.04.002

Misteli, T. (2013). Eliminating the impact of the Impact Factor. *The Journal of Cell Biology*, *201*(5), 651–2. http://doi.org/10.1083/jcb.201304162

Moed, H. F., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology*, *66*(10), 1988–2002.

Molas-Gallart, J., Salter, A., Patel, P., Scott, A., & Duran, X. (2002). Measuring Third Stream Activities: Final Report to the Russel Group of Universities. *SPRU-Science and Technoly Policy Research*, (April), 85.

Mortensen, P. S. (2011). *Patentometrics as performance indicators for allocating research funding to universities*. Citeseer.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479.

Naldi, F., & Vannini Parenti, I. (2002). *Scientific and Technological Performance by Gender A feasibility study on Patent and Bibliometric Indicators Vol . II : Methodological Report*, *EUR 20309* (Vol. II).

Neylon, C., & Wu, S. (2009). Article-level metrics and the evolution of scientific impact. *PLoS Biology*, *7*(11), e1000242. http://doi.org/10.1371/journal.pbio.1000242

OECD. (2008). *Compendium of patent statistics 2008*. *Organization for Economic Cooperation and Development* (Vol. 6). Retrieved from message:<A0AEA587-BDC0-4BF9-8DED-262E5485D5BD@gmail.com>\npapers2://publication/uuid/69B3F9ED-4260-4294-BD7B-F44452E05719

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* , *39* (6), 751–760. http://doi.org/10.1093/scipol/scs093

Phillips, M., & Maes, K. (2012). *Research universities and research assessment*. Retrieved from http://www.leru.org

Priem, J., & Hemminger, B. H. (2010, July 2). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*. Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/2874/2570

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto. Retrieved December 12, 2015, from http://altmetrics.org/manifesto

Pulverer, B. (2015). Dora the Brave. *The EMBO Journal*, *34*(12), 1601–1602. http://doi.org/10.15252/embj.201570010

Sandström, U., & Wold, A. (2015). Centres of Excellence: reward for gender or top-level research. *Thinking Ahead: Research*, *Funding and the Future. In RJ Yearbook*, *16*.

Shaw, A. K., & Stanton, D. E. (2012). Leaks in the pipeline: separating demographic inertia from ongoing gender differences in academia. *Proceedings. Biological Sciences / The Royal Society*, *279*(1743), 3736–41. http://doi.org/10.1098/rspb.2012.0822

Spaapen, J., & van Drooge, L. (2011). Introducing “productive interactions” in social impact assessment. *Research Evaluation*, *20*(3), 211–218. http://doi.org/10.3152/095820211X12941371876742

Spaapen, J., Dijstelbloem, H., & Wamelink, F. (2007). *Evaluating Research in Context: A Method for Comprehensive Assessment*. The Hague : Consultative Committee of Sector Councils for Research and Development (COS).

Spaapen, J., Shinn, T., Msh-paris, C., & Marcovich, A. (2011). *SIAMPI final report Executive summary*.

Sugimoto, C. R., Ni, C., West, J. D., & Larivière, V. (2015). The academic advantage: gender disparities in patenting. *PloS One*, *10*(5), e0128000. http://doi.org/10.1371/journal.pone.0128000

Swedisch Environmental Protection Agency. (2014). Bibliometric evaluation of SEPA-funded large research programs 2003–2013.

Symonds, M. R. E., Gemmell, N. J., Braisher, T. L., Gorringe, K. L., & Elgar, M. A. (2006). Gender differences in publication output: towards an unbiased metric of research performance. *PloS One*, *1*(1), e127. http://doi.org/10.1371/journal.pone.0000127

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS One*, *8*(5), e64841. http://doi.org/10.1371/journal.pone.0064841

van den Besselaar, P., & Sandström, U. (2015). Early career grants, performance, and careers: A study on predictive validity of grant decisions. *Journal of Informetrics*, *9*(4), 826–838.

van Raan, A. F. J. (2005). Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), (pp. 19–50). Dordrecht: Springer Netherlands. http://doi.org/10.1007/1-4020-2755-9_2

Vincent-Lancrin, S. (2006). What is changing in academic research? Trends and futures scenarios. *European Journal of Education*, *41*(2), 169–202.

Walker, R., Alexander, J., Barros, B., Cochrane, M., Guzman, E., Geuna, A., … Göhnert, T. (2013). Requirements document and Specifi cation of conceptual model, 1–47. Retrieved from http://sisob.lcc.uma.es/repositorio/deliverables/SISOB-D22.pdf

Waltman, L., & van Eck, N. J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics*, *7*(4), 833–849. http://doi.org/10.1016/j.joi.2013.08.002

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*(1), 37–47. http://doi.org/10.1016/j.joi.2010.08.001

Wenneras, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, *387*(6631), 341–343. Retrieved from http://dx.doi.org/10.1038/387341a0

West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The role of gender in scholarly authorship. *PloS One*, *8*(7), e66212.

http://doi.org/10.1371/journal.pone.0066212

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., … Johnson, B. (2015). *The metric tide : Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from www.nationalarchives.gov.uk/doc/open-government-licence/version/2

Wilsdon, J., Wynne, B., & Stilgoe, J. (2005). *The Value of Science*. London: Demos.

Wouters, P., & Costas, R. (2012). *Users, narcissism and control: tracking the impact of scholarly publications in the 21st century*.

Wouters, P., Thelwall, M., Kousha, K., Waltman, L., de Rijcke, S., Rushforth, A., … Wouters, P. (2015). *The Metric Tide: Literature Review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management)*.

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of “alternative metrics” in scientific publications. *Scientometrics*, *101*(2), 1491–1513.

GEDII GENDER DIVERSITY IMPACT – Improving research and innovation through gender diversity

WWW.GEDII.EU

D1.1. Conceptual Framework

Chapter 3 – Gender Research

Anne-Charlott Callerstig, Jörg Müller

31st of May 2016

Table of Contents

3. Gendering Diversity, Science, Organizations and Teams

The aim of this particular review is to synthesize and discuss insights from the field of gender research regarding gender diversity in research teams and the effects of gender diversity on research outputs. In a way it revisits insights from the preceding chapters now from a gender perspective. To which degree do we have to conceive of science, its disciplines and organizations themselves as gendered? In which way is teamwork affected by gender diversity and gender bias and how does this impact the assessment of research performance? And to which degree is the very concept of “diversity” or the emphasis on “performance” in need of a critical revision when a gender perspective is taken into account?

3.1. Diversity and Diversity Research

A basic assumption underlying the conceptual framework and the project itself stipulates that the world is diverse. Assuming that reality is fundamentally heterogeneous is an ontological baseline without which the quest for “diversity” within science would hardly make sense. Assuming that reality is diverse suggest striving for diversity as an adequate principle for science. This is not necessarily self-evident nor has it historically always been the case. Quite on the contrary, science was dominated by the “grand narrative” of its progress, continuously unraveling the mysteries of nature. Mathematics, as the queen discipline of science, believed until the beginning of the 20th century in the possibility of a solid foundation (being consistent, complete and decidable) before this supposed bliss-point of mathematics was shattered through the work of Kurt Gödel or Alan Turing. Equally, the dream of physics to derive the known universe from a set of simple universal laws seems continuously receding as with each new discovery, “dark” new questions arise. Science, rather than producing a more and more integrated and complete picture of a homogeneous reality must be said to complexify our knowledge with each new discovery.

Feminist philosophers such as Donna Haraway (Haraway, 1991) and Karen Barad (2007) have contributed to this post-modern understanding of knowledge, insisting on its always partial, framed and fragmented nature. Feminist perspectives on knowledge are congruent with other post-structuralist critiques of our logocentric tradition, undermining the belief in an essential reality, that is, an origin that is simple, homogeneous and present to itself. However, if our contemporary assumptions make room for an ontology of “complexity”, “mediation”, “difference” or “impurity” among other concepts, then “diversity” promises to become a more adequate concept for contemporary science – and its activities. This is the underlying ontological assumption upon which a quest for diversity in research teams can build.

In contemporary Gender Studies, the notion of gender as a binary construction, such as woman-man or female-male, has also been problematized, and thus, from this viewpoint, “gender diversity”, along with terms such as "genderqueer", "non-binary gender", "gender-expansiveness", "gender ambiguity", "gender pluralism" or "overlapping gender", describe more fluid understanding of gender. In such perspectives “gender” and “sex” are not seen as interchangeable or even directly connected, but rather intersecting in various and different ways, to take account of such variations as cis-gender, transgender, intersex, agendered or

even postgender people and positions. In line with this thinking gender is to a large extent seen as contextual and enacted or performed (Butler, 1990; Hearn, 2015; Hill & Mays, 2011; Monro, 2005).

Interestingly, in a move to free the literature on diversity from its logocentric inheritance, Phillips & Apfelbaum (2012) propose to put the standard assumption on the "benefits" of homogeneity on its head: where much diversity researcher implicitly compares the effects of diversity to the *de facto* standard of homogeneity, the authors suggest to assume homogeneity as the aberration and diversity as the new standard! This implies that homogeneous teams, even when they work well, do not work "better." Homogeneous groups without sufficient level of conflict and an elevated sense of cohesion might enjoy the advantage of smooth teamwork but this comes at the price of being "systematically inattentive to critical information, perspectives and viewpoints" that do indeed exist but are suppressed by group bias and conformity. This perspective, based on the ontological assumption of the primacy of diversity, suggest that we should not judge well functioning teams according to the "wrong" standard of their homogeneity, but rather on the grounds of their approximation to disruptive engagements with reality and with each other.

Examining more closely the literature on diversity itself, it appears that diversity research is a growing field of interest (Ahonen, Tienari, Merilainen, & Pullen, 2013; Knights & Omanović, 2015) including both more critical and more mainstream approaches. There has been a call for a "revitalization" of diversity research, particular in relation to what is being understood as the marginalization of the "social justice" perspective as a driver in the field (Ahonen et al., 2013; Calás, Holgersson, & Smircich, 2009; B. D. Metcalfe & Woodhams, 2008, 2012; T. Miller & Triana, 2009; Zanoni, Janssens, Benschop, & Nkomo, 2010). One central concern targets the uncritical usage of diversity categories or attributes that are seen as something existing *a priori*, and that can be studied and/or managed in order to produce desirable outcomes for organizations.

> *"Identities are conceptualized as ready-made, fixed, clear-cut, easily measurable categories [...] ready to be operationalized as the independent variable to explain the specific phenomenon under study" (Zanoni et al., 2010, p. 13)*

The essentializing of differences goes hand in hand with approaching them instrumentally, i.e. as a potential source of value that needs to be activated by virtue of the employment relation. What thereby often get neglected however, are the questions of "value": who decides what are valid outcomes of diversity? Towards which ends is diversity beneficial? These questions can only be addressed as long as context specific factors and power relations are taken into account. Specifically for GEDII one has to remember that questions of scientific excellence, i.e. productivity or quality of research results, are highly political concepts that change across disciplines, countries or organizations.

Mainstream diversity approaches have been criticized for reproducing often cultural and organizational contexts where white heterosexual middle-class men are the norm and provide the baseline against which other minorities are judged and managed (Ostendorp & Steyaert, 2009; Sinclair, 2000). In the context of gendered organizations, this echoes the findings of Joan Acker (1992) regarding the multiple ways in which organizations and work

arrangements themselves have to be understood as gendered and following the masculine norm. A more critical approach to conceptualizing diversity therefore needs to consider "the ways power and context have (not) been conceptualized and operationalized into research practice in the field" (Acker 1992, p.2).

A starting point for such an approach is available consulting Harrison & Klein (2007) that articulate diversity a three-folded concept that includes *disparity*, *separation*, and *variety*. Diversity as *disparity* concerns the (un)equal distribution of power. In terms of gender and science, *disparity* captures differences in the possession of assets and resources: differences in pay, funding, prestige, or (decision making power). Diversity as *separation* reflects stand point or position in terms of values, beliefs, attitudes, or orientation. It concerns differences in value judgement that can include deep cultural values and stereotypical assumptions regarding the role and competencies of women and men in science. And finally, diversity also alludes to *variety* which refers to the distribution of views, perspectives, ideas and information in general. Specifically for research teams, variety in terms of access to diverse and non-redundant information is key for producing innovative results. Often, these diverging connotations of diversity are mixed up in research questions on teams and team processes. As Harrison and Klein (2007, p.1209) note, different dimensions of diversity are addressed by examining for example power relations among teams (disparity), or how different knowledge caches between women and men might spark creativity and innovation (variety), or when the effect of cohesion and identification upon team processes are addressed (separation).

Congruent with these critical accounts of the diversity literature, where not only the findings of mainstream diversity research has been questioned but also underlying ontological and epistemological assumptions underpinning them, we will take into account the various ways in which gendered power relations gendered value judgments and gendered contextual factors (organization and scientific discipline) shape teamwork.

3.2. Locating Gender Diversity

The question of "gender diversity" in research teams implies to address the current state of the art regarding gender differences and similarities. A more detailed discussion of where differences are located is indeed necessary since empirical research all too often reduces diversity to counting simply the proportion of women and men in work groups. This practice implies that differences easily get essentialized and conceived of as fixed attributes that all women and men possess per se based on their sex. On the other hand, a more constructionist position argues that diversity is not a fixed attribute of the individual but an emergent result of social interaction and communication (for an overview, see for example Hearn & Louvrier, 2015). The effects of diversity (for teams) are thus not inherent to the individual group members but rather need to be located on the level of perceptions and beliefs about the "other".

Diversity based on sex differences are often seen as static biological differences between women and men. Diversity from this perspective would be conceived as permanently enhancing or inhibiting team performance based on those fixed biological attributes of women and men. Gender differences on the other hand are understood as shaped by

societal factors and transmitted through socialization and educational processes, themselves rooted in distinct (stereotypical) beliefs regarding women and men. However, research has increasingly shown that gender and sex differences and similarities have to be viewed as mutually interacting as when social and cultural beliefs shape biological differences and vice versa (Wood & Eagly, 2012). What more, the distinction or difference between gender vs sex itself needs to be considered as gendered and reproduced in part in the literature on how gender effects teamwork.

Here, Ridgeways definition of gender is helpful, as she defines gender as an "institutionalized system of social practices within society that constitute people as two significantly different categories, men and women, and organize relations of inequality on the basis of this difference" (Ridgeway, 2007, p. 312). Since gender is an "institutionalized system of social practices" it implies that science and research itself have to be conceived as forming part of these social practices. Thus, scientific knowledge production is gendered including the very production of sex-differences in much biological or medical research. One example is how biological differences are overlooked in favor of conforming these rather fluid differences to a neat two-sex male/female model (Fausto-Sterling, 2000). Wood and Eagly (2012) show in their extensive review of *Biosocial Construction of Sex Differences and Similarities in Behavior* that sex differences vary across historical time and cultures. Rather than documenting a fixed set of large or small sex differences, the authors argue for the "variability of sex differences and similarities across contexts" which emerges "as men and women regulate their behavior through proximal social psychological and biological processes" (ibid., p.93). Given the evidence on the relation between sex- and gender differences, the available literature has been scanned to isolate those differences that potentially affect team performance.

Reported sex differences between women and men exist for example as differences in health outcomes. Insights are available from gender-inspired critical accounts in medicine or transport research. The members of the EUGenMed[1] described in a detailed manner the differences regarding manifestation and outcomes of cardiovascular disease between statistical groups of men in comparison to women on group level (Regitz-Zagrosek et al., 2015). Ignoring such differences has lead in the past to wrong and late diagnosis of heart attacks especially for women. Similar, there is converging evidence on sex differences on group level in pain: gonadal steroid hormones such as estradiol and testosterone modulate sensitivity to pain and analgesia (Craft, Mogil, & Aloisi, 2004; Greenspan et al., 2007).

Considering research on sex differences related to team collaboration and communication, research from neuroscience regarding emotion processing and empathy have been flagged up as relevant (see section Gender and Sex Differences in Social Sensitivity/ Empathy on page 89). Sex differences in empathy provide the foundation for higher level social psychological constructs to emerge which in turn affect the quality of the overall group communication. As research by Woolley and colleagues has shown (Woolley, Chabris, Pentland, Hashmi, & Malone, 2010), the proportion of women in groups improves information sharing and turn-taking in conversations due to the higher social sensitivity of women.

Another important strand of the literature to be consulted within the context of the present project is research on sex differences in cognitive abilities. Current research has shown that

1 See http://www.eugenmed.eu

sex differences regarding cognitive abilities are basically non-existent, suggesting the hypothesis of gender similarities rather than differences (Hyde, 2005, 2014). That is to say, that sex differences in cognitive abilities precisely fall short of explaining the huge differences we observe in terms of study and career choices or vertical segregation in the current science system. It seems much more promising to consult at this point the role of gendered social beliefs and attitudes and its effects on research teams.

In fact, the effects of potential sex differences need to be set side by side with the strong social effects regarding gender differences in work groups. Leslie et al., for example states that it is the *expectations* of brilliance in disciplines that underlies the distribution of men and women in these disciplines (Leslie, Cimpian, Meyer, & Freeland, 2015). It's not sex differences in terms of ability that explains why more men are in STEM fields but rather social expectations regarding the necessity of innate talent for success. Since women are stereotyped to lack such talent, they do not enter these fields. Similar, Nosek et al., (2009) show, that gender-science stereotypes and not actual abilities predict national differences in science and math achievement. Similar, for a long time research has scrutinized the impact of surface-level diversity aspects such as age, gender, race on work group performance – with mixed results (Qin, Muenjohn, & Chhetri, 2014). However, recent findings suggest that what's more important than visible diversity in groups is the "perceived diversity" by team members (Harrison, Price, Gavin, & Florey, 2002; Hobman, Bordia, & Gallois, 2004; Homan, Hollenbeck, Humphrey, Stephen E. Knippenberg, & Ilgen, 2008). That is to say, teamwork is more affected by beliefs about differences by the group members than actual surface differences. Team collaboration is affected by the ways team members perceive and communicate about differences. The social dimensions of gender diversity perception and communication then also puts the spotlight on the potential impact of organizational measures to palliate stereotype threat and create favorable "diversity climate" among employees.

Considering gender diversity as a social phenomenon implies to consider the value judgments attached to gender roles. Here it is important to note that we thereby switch registers in terms of how diversity affects teamwork; whereas diversity as information/variety is conceptualized as an asset for producing innovative ideas, gender diversity as separation instead explores how gendered stereotypes introduce bias into research and science. Rather than boosting performance, gender bias affects research teams negatively by undermining optimal information sharing. This is important to bear in mind; both negative as well as positive effects in relation to gender diversity will be examined in their impact on research performance through the GEDII project.

3.3. Gendering Science, Technology and Innovation

As gender scholars have shown, there are many ways that Science, Technology and Innovation (STI) needs to be addressed as gendered (Hearn & Husu, 2011; Walby, 2007). Although this is not the place to even try to rehearse the vast knowledge about gender and science, some key ideas nevertheless need to be briefly introduced in order to identify those aspects that affect teamwork.

First, science is gendered in terms of its vertical and horizontal segregation. Vertical

segregation means that relatively few women reach the top of the career ladder: 80% of all grade A (single highest) positions for researchers are occupied by men and 90% of rector positions (European Commission, 2013). First and foremost, gender inequality implies here the unequally distribution of resources and access to power between men and women within STI. Besides access to decision making positions, this also includes for example the wage gap which, according to She Figures 2012, for women in science was higher than the overall pay gap in EU-27, namely 25% on average (European Commission, 2009, 2011). The vertical segregation and lower access to decision making positions constitutes a key factor to bear in mind when considering diversity effects in teams because power differentials skew information sharing and competency recognition as the chapter on “Team Science” has argued.

Secondly, a tight association exists between certain aspect of science and more specifically technology and masculine culture (Kiesler, Sproull, & Eccles, 1985; Wajcman, 2007). A reliance on objectivity, meritocratic and gender-neutral approach to science is quite common while research itself is often assumed to be conducted in "a culture without culture" (Egeland, 2001). Deep cultural values and stereotypical assumptions regarding women run across the whole STI spectrum, in the form of (unconscious) bias, that undervalue, neglect and disadvantage women. As Husu puts it:

> *"What is happening can really be that "nothing happens” or that something that should happen in your career does not happen: you are not made visible, heard, read, cited or quoted, invited, encouraged, get no support and will not be validated." (Husu, 2005, p. 23)*

Subtle stereotypes operate to the disadvantage of women regarding excellence, career progression, decision making, work-life balance and knowledge making (European Commission, 2003; Fausto-Sterling, 1981; Xie & Shauman, 2003). These unexamined stereotypical assumptions, give rise to procedures and practices that tacitly sanction or perpetuate gender discrimination. As the section on Gendered Competency Expectations from page 87 onwards will argue, those stereotypes constitute a key factor undermining the effective information sharing in groups.

Analyzing the forms and consequences of gender bias in research content (Schiebinger et al., 2016) is another prominent expression of the relative invisibility of women in R&D. Consistently, professions, knowledge fields and activities coded as feminine enjoy less visibility, prestige and status than those coded as masculine.

Thirdly, innovation and more precisely innovation policy and funding are gendered: whereas regional economic development through knowledge transfer activities, innovation or research grants tend to focus on masculine, high-tech innovation that values growth-oriented profits, women and feminized types of knowledge targeting equality or community well-being in such areas as education and health are under-valued (Blake & Hanson, 2005; Gray & James, 2007; Parken, 2010; Pecis, 2016; Pettersson, 2007). There exists a gender gap in funding (see section on Gender Gaps in Research Productivity and Citations on page 83). These structural disadvantages (wage gap, publication gap, funding gap) are important when examining the research performance of mixed-teams, since the systematic performance “penalty” that is attached to being a woman in science needs to be considered and

controlled when comparing the research performance between mixed-gender groups.

3.3.1. Critical Reflections on "Performance"

The gendered subtext of gender equality policy itself needs to be considered. Gender equality policy on the European level that has arguably been shifted away from a rights-based approach to a more market-oriented one that succumbs to an economic imperative promoting a narrow view of well-being and equality (Elomaki, 2015). Garforth & Kerr (2009) write about an "uncomfortable juxtaposition of economistic and democratic goals" (ibid, p. 384) within gender and science politics. Women are often either seen as an underused human resource in economical terms or an underrepresented group with references to equal opportunities, which is in line with a liberal democratic equality discourse. The consequences of this argument is to address social justice demands basically through market mechanisms, that is to say that "… diversity and equality of opportunity are to be brought about through participation in the labor market, not by correcting its structural inequalities." (ibid, p.386).

Feminist scholars have criticized the acceleration of academic life that goes hand in hand with the marketization of universities and research performing organizations (Mountz et al., 2015). Joining calls to resist the "neoliberal and elitist pressures with the academy" with its emphasis on efficiency, productivity and excellence (Ball, 2012; Slaughter & Rhoades, 2000), the collective of authors call for a "slow scholarship" that cultivates a caring academic culture and processes. Against the production of marketable outputs of quantifiable nature, other ways of counting and collective processes are invented to opt-out of the narrowly defined metric frameworks of publication counts and citation impacts. In relation to the GEDII project, this literature speaks directly to the concerns of finding alternative "performance" indicators that are sensitive to this feminist critique of the neoliberal university. What this literature (see also Menzies & Newson, 2007) also highlights is the fact creative processes can never by totally controlled. A certain "slowness" is necessary to wait for events driven by serendipity and opportune moments, key elements of creative processes (Garud, Gehman, & Kumaraswamy, 2011).

> *"Under certain circumstances failing, losing, forgetting, unmaking, undoing, unbecoming, not knowing my in fact offer more creative, more cooperative, more surprising ways of being in the world" (Halberstam, 2011, p. 3)*

Similar, Rönnblom (2009) reminds us that the focus on economic growth regarding gender equality depoliticizes understandings of gender and their concomitant power relations. The potential depoliticization of gender relations is clearly visible for example through the emphasis on performance. Measuring performance implies a comparative perspective that systematically associates input with diverging outputs. This type of benchmarking however depoliticizes gender equality agendas by "translating political problems of collective action into statistical issues of quantification", embedded in a matrix of competitiveness and efficiency that lose track of the complex and contested nature of social injustice (Bruno, 2009).

The fact that performance is an equally central as well as contested concept is also apparent in relation to teamwork itself. Mainstream research on teamwork has been argued to be

largely underwritten not only by a gender neutral, but a masculinist discourse of performance, management and organization ignoring that managerial practices and organizing arrangements have been created for men, by men (B. Metcalfe & Linstead, 2003). A large body of research on teams links both the definition and effectiveness of teams to performance. Metcalfe and Linstead (2003) show how team output is associated with masculinity and performance, conjuring imagery of planning, physical action and achieving (Cheng, 1996). Constructions of masculinity are connected to gender performances specifically within the work-place (Collinson & Hearn, 1994, 1996; Connell, 1987; Siedler, 1990). Indeed, masculine identity work rests on playing down emotional sensitivities and relations, since stereotypical beliefs regarding masculinity require disengagement from feelings and demonstration of control. The notion of control furthermore implies that the notion of lack of control can be negatively associated with the feminine. In a team context this might mean that exposing personal sensitivities and expressing 'intimate' feelings may be perceived as threatening and that in order to maximize team performance team members must at all times play down emotions and sensibilities because this may raise concerns about team competence and the team's ability to be in control and to perform (ibid.).

Gendered discourses in team environments can furthermore be used as a resource during concertive control (Hawkins, 2013) i.e. in interactions between team members in which they exert pressure on their peers to act in line with a negotiated consensus about team values (Barker, 1993, p. 411). These interactions serve to reify the team values as specifically masculine. Team members negotiate core team values, translate these into specific actions and regulate these actions through concertive control interactions. Hawkins found three ways in which gender acts as a resource for these concertive control processes. These are: team members' assumptions about men's and women's relative skills and capacities, the 'tough' masculinity of the haulage industry (in which one of the studied teams operates) and the regulation of performances of heterosexuality during customer interactions (ibid.). She found that gender were not only embedded in the values and managerial style associated with teamwork but also integrated into the collaborative process of teamworking itself. Hawkins emphasizes that social categories like gender become resources in the regulation of conduct at work and can reify hierarchies even in so-called participative practices like self-managed teamwork.

The insights into the gendered nature of science and gender equality policy, its economic underpinnings and tight association with performance suggests a careful revision of the gendered nature of research assessment in particular (see the section on Gendering Research Assessment from page 81 onwards).

3.4. Gendering the Research Context

Teams constitute the fundamental unit of analysis for GEDII. However, as already developed, teams are embedded into a wider organizational and disciplinary context that has important consequences for their internal processes. A rich tradition exists within gender research that has analyzed the gendered subtext of organizations within and outside of the science system. Similarly, many insights exist on how disciplines are more or less gendered. In the following section, key ideas are briefly reviewed in order to prepare the ground for conceptualizing

how disciplinary and organizational context needs to be incorporated into our research.

3.4.1. The Organizational Context of Research Teams

The research on gender and organizations has since the 1970´s very convincingly shown that organizations, including academic organizations, are gendered in many respects. The field of gender and organization is today multi-faceted and includes so many theoretical perspectives that it is not possible to speak about it as one theory but rather as several different theories and also different research approaches (Calás & Smircich, 2006; Hearn & Parkin, 2003). Different ways to analyze how gender operates in organizations have been suggested, often along the lines of investigating a gender perspective in relation to organizational representations, relations, symbols, culture and power hierarchies as well as the "gendered self" as described in the early and influential work of Joan Acker (1992). Conceiving organizations as gendered involves examining: the systematic disparities between participants in power and control over goals, resources, and outcomes; workplace decisions such as how to organize work; opportunities for promotion and interesting work; security in employment and benefits; pay and other monetary rewards; respect; and pleasures in work and work relations (Acker, 2000, 2006). One example, of how inequalities operates in organizations, is through the distributive powers of middle-managers, that is, how decisions made about the distribution of tasks and duties to employees affect wages, career opportunities, and in more general terms the possibility to influence their working situation, the possibility of a career in an organization (Andersson, Amundsdotter, Svensson, & Däldehög, 2009). In terms of teamwork, this places interest on those who have the power to distribute tasks and responsibilities among team-members and their implicit norms that may inform their decisions.

An exemplary article that sets out how gendered organizational structures are woven into team processes has been published by Yvonne Benschop and Hans Dooreward (Benschop & Doorewaard, 1998). Looking at the Dutch banking sector, the authors compare hierarchical (Tayloristic) organizations with flatter, team-based work settings that supposedly offer better opportunities for a high quality working life for both women and men. As their results show, however, gender inequality is (re)produced in both settings. Benschop and Dooreward attribute this to the "gender subtext" of organizations; a set of power-based processes (re)producing gender distinction in social practices through organizational and individual arrangements. Central for Benschops and Doorewards analysis of teamwork is the notion of the abstract and disembodied ideal worker (Acker 1992). In both the Tayloristic and the team based approaches the findings of the studies show how the gender subtext of a 'disembodied worker' leads people to assume that men are better suited for certain tasks and women for others. Implicit connotations, which ascribe care responsibilities to women and qualification profiles to men produce and reinforce unequal opportunities for women and men to climb the career ladder. Here, a supposedly more gender egalitarian teamwork turns out to repeat deeply entrenched ideas regarding care responsibilities (women) and qualifications (men).

Furthermore, although organization studies with an explicit gender perspective are generally becoming more common they often lack an intersectional perspective (Holvino, 2010). One problem with analyzes of gender and gender (in)equality in organizations that do not have an

intersectional perspective is that they risk creating a false and very normative picture for all individuals of a certain category. In her later work, Acker has pointed to the importance of an intersectional perspective in order to change organizational inequalities such as gender inequality, arguing that entrenched economic (class) interests, the legitimacy of class interests, and allegiances to gendered and racialized identities and advantages makes change towards gender equality more difficult (Acker, 2006, p. 460). This is also emphasized by Hancock (2007).who calls for policy solutions that are attuned to the multiple marginalization of race, class, gender, or sexual orientation at the individual and institutional levels that create social and political stratification.

How gender intersects with other social categories in academia has been criticized for being under-researched and invisible in existing research even though some studies are making visible and analyze intersections (Lamont, Mallard, & Guetzkow, 2006; van den Brink & Benschop, 2012). An interesting exceptions to this observation is the work of DiTomaso and colleagues (2007) who argue that among research teams in the US, being white and male is associated with higher status than being female or belonging to an ethic minority. Similar, (Joshi & Knight, 2015) argue for the complex interplay between surface level demographic attributes (gender, ethnicity) with other status cues such as educational level and tenure. Social affinity with partners who shared similar educational level, gender and ethnicity contributed to deference beyond actual task contributions, concentrating mainly status among a close network of white males.

The often unreflexive “doing” of gender (C. West & Zimmerman, 1987) puts the spotlight furthermore on policy measures and organizational initiatives that raise awareness of the importance of gender among employees. Much, perhaps most, of the practicing of gender in an organization is unintentional and unreflexive (Martin, 2006). Hence, organizational initiatives such as gender equality plans can make a huge difference for women (and men) in raising awareness about the existence and effects of those unconscious biases. Nishii (2012) for example exposes how organizational policies for building a “climate for inclusion” can reduce interpersonal bias and reduce conflict associated with gender diversity. Similar, Ely & Thomas (2001) show, how different ways of framing racial diversity on the organizational level affects group functioning and conflict management. Many of those organizational policies address the reduction of bias and stereotype threat. As the review by Walton and co-authors argues, “[H]igh levels of stereotype and social identity threat can undermine people's performance and thus an organization's success” (Walton, Murphy, & Ryan, 2015). Organizations can either change cues present in the environment that favor stereotypes (e.g. increase representation of people belonging to a minority) or they can change how people interpret these cues (e.g. valuing diversity, constructing a diversity climate). These initiatives within organizations can change diversity beliefs, i.e. how demographic diversity is perceived as either threatening or a beneficial asset (Hentschel, Shemla, Wegge, & Kearney, 2013; van Dick, van Knippenberg, Hagele, Guillaume, & Brodbeck, 2008).

3.4.2. The Disciplinary Context of Research Teams

Apart from the organizational context, research has also highlighted the importance of the scientific discipline to which research groups pertain and within which they develop their

work. Research teams in the field of humanities differ from research teams in natural sciences in many ways including implicit values that affect the gendered identity of the "ideal scientist" (van den Brink & Benschop, 2012). The disciplinary context affects also the recognition of expertise (of minorities within this discipline). As Joshi shows, in contexts where women are the minority, the majority (of men) fail to recognize women's expertise:

> *"Conversely if the discipline in which the team is embedded is male-dominated, team members may not have had exposure to visible symbols of female success. Team members may assume that female team members are generally less qualified than men, and gender may therefore significantly predict expertise recognition and utilization (e.g., Ibarra 1992; Ely, 1994)" (Joshi, 2014, p. 5)*

The expertise of the minority is then misjudged with overall detrimental consequences for the team:

> "Teams with greater proportion of highly educated women were significantly more productive in gender-balanced disciplines than in male-dominated disciplines. These findings support the argument that the level of gender integration in any given discipline can shape the salience of gender as basis for the status differences or role expectations among men and women in science and engineering" (Joshi, 2014, p. 27).

As Joshi observed, the proportion of women in a group (or discipline) influences how well their expertise is recognized. The recognition of expertise is also dependent on the strength of the gender identification among men. "Male actors who strongly identify with their gender are more likely to favor men irrespective of their educational status [...]" penalizing women with high educational status (Joshi, 2014, p. 19).

Interesting insights are also available from the Elsevier report (Elsevier, 2015) regarding disciplinary differences in terms of publication count and choice of topics in Germany. The productivity gap (see next section) disappears for women researchers in clearly male dominated disciplines such as Engineering or Computer Science where they publish on average more articles than their male colleagues. Skewed ratios between men and women regarding the scientific discipline also affected the choice of research topics: where women are under-represented, they tend to choose similar topics as men; in disciplines with more balanced gender rations, women tend to focus on different topics.

3.5. Gendering Research Assessment

GEDII aims to assess the effects of gender diversity on team performance. For science and research teams that means first and foremost bibliometric analysis regarding publication/patent (productivity) and quality (citations). However, assessing team performance from a gender perspective implies first and foremost to account for any systemic inequalities inherent in the science system itself. If publications are *the* means for assessing research performance, we need to account for any systematic disadvantages that prevent women from being as productive as men. As the existing literature demonstrates, there are indeed structural gender gaps when it comes to social capital, access to funding or

publications that translate into the lower performance of women and difficulties for career advancement.

3.5.1. Social Capital

A vast literature documents disadvantages of women in terms of climbing the career ladder that is explained in part due to differences in terms of social capital. Social capital lies here at the root of other equally systematic disadvantages such as access to funding or collaboration. The structure of women's social networks has been identified as an important factor for their persistent under-representation in top positions (Etzkowitz, Kemelgor, & Uzzi, 2000). Two concepts are key: first, “homophily” explains differences in social capital between men and women. Homophily refers to the propensity of actors to form network ties with “similar” persons in terms of “race”, “ethnicity”, “sex”, “education”, “religion”, or any other dimension of interest (McPherson, Smith-Lovin, & Cook, 2001). Contact between similar persons occurs at a higher rate than dissimilar people, forming closed social relationships to the detriment of the excluded minority. Especially among senior managers and entrepreneurs, men tend to have more sex homophilious networks than women do, forming “old boys networks” with detrimental consequences for women career advancements (Kanter, 1977). Subtle ways of “doing gender” persist in how networks are build and maintained in academia (L. Berger, Benschop, & van den Brink, 2015; Sadl, 2009).

Second, even for similar positions within networks, women and other low status groups face disadvantages in converting their social capital into real advantages such as career promotion. That is to say, that women have less advantageous network ties, and even when they reach equal or more advantageous position than men, they face difficulties in cashing them in (Ibarra, 1992, 1993). Thus, Burt (1998) showed that similar structural positions of men and women in terms of their networks within the company had differential outcomes, disadvantaging women in their career advancement. Women managers basically lacked “legitimacy” in order to cash in their structural position as their male colleagues did: only when being connected to high status male managers and “borrowing” their social capital did they fare similar in terms of promotion. Interestingly, Burt (ibid.) also highlights the intersectional dimension of his finding, arguing that entry-rank men had a similar problem as women do in terms of being deemed “suspect.”

3.5.2. Gender and Research Collaboration

The importance of social capital can be spelled out when looking at the gender gaps in collaboration among scientists and its subsequent impact in terms of research productivity - which ultimately translates into assets for career advancement.

As Larivière et al. (2013) argues, women's publication portfolios are more domestic than their male colleagues with important consequences in terms of citation impact. As they argue, “in the most productive countries, all articles with women in dominant author positions receive fewer citations than those with men in the same positions. And this citation disadvantage is accentuated by the fact th*at women's publication portfolios are more domestic than their male colleagues – they* profit less from the extra citations that international collaborations accrue. Given that citations now play a central part in the evaluation of researchers, this situation can only worsen gender disparities.” (ibid., p. 212

emphasis added).

Looking at the available evidence regarding gendered collaboration differences, however, the results are not very consistent. Cole & Zuckerman (1984) or Lee & Bozeman (2005) suggest that men outperform in women in terms of research collaboration whereas (Bozeman & Gaughan, 2011) or (Abramo, D'Angelo, & Murgia, 2013) suggest that women scientists have a greater capacity to collaborate than men. Gaps persist apparently on the international level where women engage less in international collaborations than men (Abramo et al., 2013; Uhly, Visser, & Zippel, 2015). When comparing these insights with the study published by Elsevier (2005), a different picture emerges: for Germany, female-only publications are the most internationally collaborative while mixed-gender publications are more interdisciplinary and less internationally collaborative.

Considering domestic collaboration, Kegen (2013) argues in her work on the German "Excellenz Initiative" that gender homophily did not significantly correlated with research collaboration patterns. Unlike the common assumption, "sex does not play a considerable role in joint activities of scientists in cutting-edge research institutions" (ibid., p.73).

3.5.3. Gender Gaps in Research Productivity and Citations

A persistent gender gap exists with regard to research productivity, i.e. in differences between women and men regarding publication output (Abramo, D'Angelo, & Caprasecca, 2009; Symonds, Gemmell, Braisher, Gorringe, & Elgar, 2006; Xie & Shauman, 1998) and patents (Busolt & Kugele, 2009). Recent data from the Netherlands suggest that these gaps are closing for younger generations of scientists (van Arensbergen, van der Weijden, & van den Besselaar, 2012).

A study of Larivière et al. (2011) found that women at universities in Quebec receive less research funding than men and that women "are generally less productive in terms of publications, have a more restricted and local network of collaborators, and are at a slight disadvantage in terms of the scientific impact of their publications as measured by citations." (ibid., p. 494) The study observed the tendency of women to collaborate more with local partners than those from abroad. This finding supports the idea that family responsibilities "could reduce the mobility of researchers and their levels of international collaboration." (ibid., p. 494)

A study by West et al. (2013) analyzed the phenomenon of self-citation, looking into 1.6 million papers with 40 million citations, of which one million were from scholars referring to their own work. The findings showed that men cited 56% more likely their own work than women. In some male-dominated fields such as mathematics, the gender gap was even more evident, with men who cited themselves about 84% more likely than women. As the authors argue, gender inequity is far from disappearing in science looking at publications:

> *"For instance, even where raw publication counts seem to be equal between genders, close inspection reveals that, in certain fields, men predominate in the prestigious first and last author positions. Moreover, women are significantly underrepresented as authors of single-authored papers. Academics should be aware of the subtle ways that gender disparities can occur in scholarly authorship." (ibid., p. 1)*

Maliniak and co-authors (2013) conducted a study of citation and publication patterns focusing on the gender in the field of political science on 12 peer-reviewed publications between 1980 and 2006, that is 3,087 articles. They show a very clear citation gap between men and women where men are much more cited than women. The study shows „that women will have lower citation counts all else equal. Moreover, the bias stems not from a difference in quality or choice of research strategy, but from underlying behaviors (fewer self-citations by women, and more within-gender citations)“ (ibid., p. 27). Women are promoting themselves much less than men, as they are doing for example much less self-citations; the citation patterns show also that citations are rather along gender lines, that means that men favor citations of men and women favor citations of women (ibid.).

Interesting findings are also available from the Elsevier report “Mapping Gender in the German Research Area” (Elsevier, 2015). Consistent with other studies, the report highlights that women researchers in German tend to be less productive than their male colleagues and have a lower citation impact. However, women publish more than men in Energy, Engineering, Computer Science, Material Science, the Earth and Planetary Science, Physics and Astronomy – all of which are male dominated disciplines. Regarding citation impact, gender gaps occur primarily at the junior and middle-senior level but are more balanced at the senior level.

Several authors have also addressed the reasons for this “productivity puzzle” (Cole & Zuckerman, 1984), that refers to gender disparities in science and technology publication rates in terms of “whether this is caused by gender bias, childbearing and caring responsibilities, other variables – or a combination of these.” (Wilsdon et al., 2015, p. 92). In general, several explanations have been proposed to explain this gender gap, including scientific ability, self-selection, social selection, and accumulated disadvantage.

Sex differences regarding women's and mens differing cognitive and other abilities usually do not satisfactorily explain the gender gap in productivity. This finding is supported by the fact that girls tend to outperform boys at all educational levels, which also contradicts this assumption (Buchmann, DiPrete, & McDaniel, 2008; Pekkarinen, 2008). A combination of other factors which are partly choice (self-selection) and partly based on social selection further contribute to explanations regarding the gender gap in scientific productivity. Child rearing and other care responsibilities that affect women more than men translate into less time for research and therefore a lower scientific performance in the early career (Long, 1992; Symonds et al., 2006). While being able to catch up more or less with male researchers later on in their career, the initial lower productivity has a negative effect on women's careers (Fuchs, von Stebut, & Allmendinger, 2001; Hunter & Leahey, 2010; Karamessini, 2003; Prozesky, 2008). However, other studies indicate that the academic career of female scientists does not suffer from parenthood (Cole & Zuckerman, 1991; Fox & Xiao, 2013).

Women researchers have a lower degree of specialization, tend to work in other disciplines, tend to focus more on teaching, tend to work at universities and departments with lower reputation, and have a less developed international collaboration and co-authoring network (Allison & Long, 1990; Bland, Center, Finstad, Risbey, & Staples, 2006; Carayol & Matt, 2006; Dundar & Lewis, 1998; Leahey, 2006; Lee & Bozeman, 2005; Prpić, 2002; Puuska, 2010), which all affects performance and career in a negative way.

Leahey suggests that specialization is conditioned by differences in professional ties: men's wider and more diverse professional networks may allow them to find collaborators whose interest overlap more easily and thus specialize whereas women's smaller and more homogeneous networks require them to branch out to other substantive areas inhibiting disciplinary specialization affecting their overall research productivity (Leahey, 2006, p. 774).

Furthermore, women receive less academic support and mentoring than men (Fuchs et al., 2001; Landino & Owen, 1988). This may be a disadvantage for women too, as academic careers depend on support by academic mentors (Van Balen, Van Arensbergen, Van der Weijden, & Van den Besselaar, 2012). A slower career progress recursively may lead to less research and more teaching (Snell, Sorensen, Rodriguez, & Kuanliang, 2009), and consequently to less scholarly productivity. The accumulation of all these self and social selections over time leads to cumulative disadvantage.

However, over time gender roles and responsibilities in family life are changing (Prozesky, 2008; Xie & Shauman, 1998) , which may also influence work and career orientation of women. Recent data suggest that the performance gap has been closed for PhD students (Ceci, Ginther, Kahn, & Williams, 2014; D. I. Miller & Wai, 2015). If this is indeed a slow and generational process, one would wonder whether this gender performance turn has also reached the research system, first of all the generation of early career researchers.

3.5.4. Gender (Un-)Biased Metrics?

That bias in research assessment plays a role was very clearly put on the agenda when Wennerås and Wold (1997) published their results of a study on grant decision-making. They showed that not meritocracy was the standard, but cronyism: having friends in the relevant committees proved to help considerably in getting a grant. On top of that, they also showed the role of gender bias: women needed a substantially (160 per cent) higher performance than male researchers to be successful in biomedical grant applications. Replicating that study some ten years later, Sandström and Hällsten Sandström & Hällsten (2007) found again nepotism, but no sexism anymore: female researchers even had a slightly better chance than male researchers. Obviously, the council studied in both papers changed its gender policy in the meantime.1 After the Wennerås and Wold study, quite a few other studies found that that gender matters, but some evidence suggests differently in different disciplines. For example, a Dutch study showed that in science fields female researchers received positive evaluations and high success rates, even higher than could be expected from past performance. In contrast, the life sciences were characterized by negative gender bias, as no substantial differences between track records of unsuccessful women and successful men were found (Brouns, 2000).

The main problem with most studies arguing that no gender bias exists is that they do not take performance into account at all. This is also the case with recent meta-reviews (Marsh, Bornmann, Mutz, Daniel, & O 'mara, 2009; Mutz, Bornmann, & Daniel, 2012), claiming that gender bias no longer exists in peer review. These studies do not refer to performance, but look only at the success rates.

Partly based upon these experiences, there are also attempts to create gender unbiased metrics. Given the differences of men and women in the quantity of research output, that is

“on average, males publish more papers than their female counterparts, a trend that is consistent across scientific disciplines and exists even when obvious mitigating factors are taken into consideration” (Symonds et al., 2006, p. 1), the authors suggest “that female researchers produce fewer but higher quality publications“ (ibid.) and therefore, quantity and quality of the articles should be taken into consideration when assessing performance. For that purpose they proposed a derived h-index, namely a residual h-index that considers the relation of quality and quantity in order to create an unbiased metric.

Abramo, Cicero, & D’Angelo (2015) investigated if it would be helpful to separate the measurements of research performance of scientists by gender. They looked into the performance of Italian professors and compared the ranking lists with and without separating gender. As a result, female professors moved up in the ranking lists in 70% of the covered disciplines: the performance of these women would be rated more positive if compared only with peers of the same gender. However, they did not recommend this approach, but left it open to decision-makers and policy if a segregated assessment is desirable or not (ibid.).

3.5.5. Gendering Social Impact

As the review on research assessment argues, determining the social impact of research is complicated. There is no simple method or process that would allow to determine the value of research beyond the standard bibliometric indicators. Partly, the problem has to do with the “worth” of research which always depends on the needs it actually aims to address. And since end-user needs change in time as well across different social groups, the social impact can seldom be assessed in a straightforward manner.

However, when looking at social impact from a gender perspective, the integration of a sex and gender dimension into research needs to be considered. One of the central insights put forward by Schiebinger et al., (2016) is the fact that the integration of a gender perspective affects directly the quality, relevance and applicability of research and development. By integrating more diverse viewpoints during development processes and especially taking into account gender aspects, the end-product (or research outcome) addresses also a more varied user needs which improves its relevance and applicability (Pollitzer & Schraudner, 2015). From this perspective, the integration of a gender perspective is considered by some funding agencies as a quality criteria to be taken into account during funding decisions. The gender perspective also forms part of the wider framework of Responsible Research and Innovation, that precisely aims to bind research closer back to societal challenges and needs[2].

Another attempt to measure social impact is based on Almetrics, i.e. the impact certain publications have throughout the social networks. Despite the already mentioned difficulties of Altmetrics, except an article published by Paul-Hus et al., (2015) little evidence is available regarding the a potential gender gap in social media metrics.

2 See for example RRI Tools http://www.rri-tools.eu

3.6. Gendering teamwork

The following section focuses specifically on the social psychological aspects of teamwork and how gender needs to be considered from this analytical perspective.

3.6.1. Gendered Competency Expectations

As described in the chapter on Team Science, formal and informal hierarchies in work teams can affect group performance in different ways. One approach that details with great care how gender affects and is affected by hierarchical relations in groups is expectation-states theory as developed by Cecilia Ridgeway (Ridgeway & Smith-Lovin, 1999; Ridgeway, 2007, 2011). Status, as an informal hierarchical relation among group members, conditions the competency expectations among members which in turn conditions which information is elaborated and how it is shared. Status has to be seen as related but distinct from formal hierarchies in organizations where authority and power are institutionally fixed. Whereas formal status differences are tied to institutional positions that individual may or may not occupy, informal status is often linked to diffuse status characteristics of actors – such as gender, age, race, ethnicity – around which evaluations and beliefs are organized (Berger, Rosenholtz, & Zelditch, 1980). The effects of gendered status expectations resemble the effects of power relations within groups at large: they limit information sharing in important ways, silencing often non-redundant and most valuable information form low-status/low-power members.

According to expectation-states theory, inequalities in status evolve rapidly in groups of strangers and once established prove to be highly persistent (Berger, Cohen, & Zelditch, 1972). In groups where women and men interact, the status of group members initially is derived from pre-existing cultural stereotypes. Gender stereotypes have diffuse general status beliefs attached to them – men are seen as more agentic and competent than women. Women appear as less competent, but better in care giving and communal tasks, although these tasks are less valued (Fiske, Cuddy, Glick, & Xu, 2002; Glick et al., 2004). Thus, being white and a man is associated with a higher status than being a woman or belonging to a minority (DiTomaso et al., 2007; Haines, Deaux, & Lofaro, 2016).

Different evaluations associated with gender thus generate different gender biased competence expectations in groups that structures opportunities to participate in discussions, claim credit, influence group decisions or evaluate group results. Since status differentials distribute power and resources in groups, it has implications for group dynamics and communication, even though status differentials might be totally unrelated to the task at hand. “The theory implies that compared with those lower in status, high-status actors are granted more (and take more) action opportunities, make more task contributions (performance outputs), have their contributions evaluated more positively, and exercise more influence.” (Simpson, Willer, & Ridgeway, 2012, p. 154). Diffuse status beliefs such as those based on gender operate independently of the actual competencies available and required: often, competencies are biased because they are under- (in the case of women) or over-estimated (in the case of men) in relation to the actual task-relevant expertise (Berger et al., 1980). The competency expectations of individual members in groups is the product of these diffuse/background status beliefs and explicit/foreground status beliefs that reflect actual task related experience. Importantly, the relative salience of diffuse competency

expectations is dependent upon the context in question.

The strength of these biasing effects is proportional to the relevance of gender to the goals of the setting (Berger, Fiske, Norman, & Zelditch, 1977; Wagner & Berger, 1997). The effects of status beliefs are already detectable in mixed sex groups. However, cultural beliefs about gender bias competence and status expectations operate most powerfully in contexts that are strongly linked to one sex or the other, such as childcare, the military, nursing, or engineering. Research shows that gendered competency expectations vary across scientific disciplines, depending precisely upon the relative minority status of women. When certain team members are a minority, such as for example women in many engineering disciplines, gender stereotypes regarding competency expectations remain relatively strong – as previously mentioned.

However, besides the gender composition of the environment (discipline, organization, team) there are further factors that influence the status of individual group members. Status attainment is linked to trait dominance of persons as well. Members with a more assertive character emerge easier as group leader, even though they do not necessarily possess more task competence. Among the Big Five personality dimensions "…Extraversion was the most important individual difference in predicting status in face-to-face groups" (Anderson, John, Keltner, & Kring, 2001). Neuroticism was dependent upon gendered role expectations: neurotic men were perceived as less capable leaders. The strength of the personality trait clearly emerges when considering that trait dominant people are perceived as competent even when they are not – to the potential detriment of more competent but less assertive individuals of the group (Anderson & Kilduff, 2009). The importance of trait dominance can also be seen by the fact that it predicts more consistently than any other individual-difference dimension who emerges as leader in the group (Lord, de Vader, & Alliger, 1986). The importance of personal traits for status attainment implies that in order to identify the role of gender for biased competency expectations, it needs to be set side by side with the influence of personality traits.

In a recent research article, Joshi & Knight (2015) show that education and tenure, two explicit status cues, predict asymmetric deference at the individual level in science teams much more reliably. Highly tenured and educated team members were more likely to received consultations from others and see their recommendations valued than lower tenured and educated members. Other, non task relevant demographic attributes including gender and ethnicity did not function as powerful status cues and did not directly predict individual deference. Although this contrasts to a certain degree with the predictions made by expectation states theory, it actually confirms it, since gender and ethnicity are diffuse status cues whose salience depends to a large degree on context: team members that have been working together and know each other are likely to ground status beliefs upon more task relevant criteria such as education and tenure than non-related demographic cues such as gender. However, gender and ethnicity nevertheless influenced who defers to whom through social affinity ties. Team members tend to develop social affinity ties (friendship, stronger interpersonal relations) to others with whom they share the same demographic attributes which biases in turn deference, "above and beyond perceived task contributions" (ibid., p.78). As a result, when deference is based on social affinity instead of perceptions of task contributions, team performance is diminished instead of enhanced. That is to say,

those teams that manage to rely on task-based deference tend to perform better since the bias inherent in homophily based networks is avoided.

Overall, research has shown that gendered competency expectations tied to status differentials can affect interaction of group members in several ways (Ridgeway & Bourg, 2004). Status differentials can become apparent by observing the (a) frequency of participation, (b) the frequency of interruptions, (c) gaze and gesture dominance, or (d) the usage of tentative speech. For example, research with sociometric badges has evidenced differences in gendered participation in group tasks as a function of tie strength and group size. Women tend to talk much more than men in gender-homogeneous small groups but less then men gender mixed large groups (Onnela, Waber, Pentland, Schnorf, & Lazer, 2014).

The differences in status beliefs can be observed through differences in communication and interaction styles. Women show greater communality and other-directedness and men greater agency and self-assertion (Carli, 1989, 2006, 2010). Men display more status asserting, dominant and negative communication while women are socialized to display warmth, support, empathetic and related positive social behaviors, a fact that also shows in gendered language uses (Mulac, Bradac, & Gibbons, 2006; Newman, Groom, Handelman, & Pennebaker, 2008). Women use more tentative, deferential speech and this speech forms makes the speaker appear less convincing. As Linneman (2012) argues, women use “uptalk” more frequently, for signaling uncertainty or as a compensatory strategy to assuage conflicting gender roles as the authors argue.

In sum, gendered competency expectations provide an important element for conceptualizing the influence of gender on teamwork. Power relations in the form of gendered status beliefs or formal rank condition competency expectations that can become manifest through interaction styles and communication patters. Gender bias in working teams hinders the optimal elaboration of information as when valuable ideas from low status members get silenced.

3.6.2. Gender and Sex Differences in Social Sensitivity/ Empathy

The second strand of research for conceptualizing the impact of gender on research teams deals with sex differences regarding social sensitivity and empathy. Some research evidence is available on sex differences regarding empathy that points to an advantage of women in terms of recognizing and interpreting emotions of others. These processes seem to be linked to better performance of “perspective taking”, a fundamental ingredient for integrating diverse perspectives.

One research strand whose argumentation runs along these lines comes out of work carried out by Anita Woolley and colleagues who have observed that the “collective intelligence” factor of human groups is predicted by the proportion of females in the group (Woolley et al., 2010). As mentioned, what predicts the collective intelligence of the group is not so much the sum of individual intelligence factors but the average social perceptiveness of group members (Woolley, Aggarwal, & Malone, 2015). The fact that the proportion of women was also a strong predictor for the collective intelligence was explained by the fact that women usually score better in the “Reading the Mind in the Eyes” test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Women appear to have a higher ability to recognize other's

nonverbal emotional expressions and mental states.

Publications on ex differences in empathy provide an entry point to explain the higher scores on social perceptiveness of women. Empathy can be understood as a three-fold construct that involves (1) emotion recognition, (2) perspective taking and (3) affective responsiveness (Derntl et al., 2010). Men and women seem to activate different neural network regions when solving emotional tasks: “while females seem to recruit more emotion and self-related regions, males activate more cortical, rather cognitive-related areas” (ibid.). An encompassing review on “Empathy: Gender effects in brain and behavior” (Christov-Moore et al., 2014) argues for the phylogentic and ontogentic roots of sex differences in empathy. Although motivational differences have been recognized to affect study outcomes (“females are motivated to try harder to understand other people's feeling during the task if they think that what is measured is relevant to a stereotypical female role”), they cannot explain female advantages in automatic/unconscious nonverbal perception, or female advantages in populations that do not exhibit social desirability biases, such as nonhuman animals and infants. Women seem to have a greater ability to perceive and respond to positive facial emotion at an automatic processing level compared to men (Donges, Kersting, & Suslow, 2012). Study on conscious face recognition found similar results: Women more readily recognize images as faces but also more often; showing superior skills in facial affect recognition (Pavlova, Scheffler, & Sokolov, 2015). However, research in this area is contested also when using recent neuro-imagining technologies for detecting differences in neuronal activity patterns (Kaiser, Haller, Schmitz, & Nitsch, 2009). Both, the combination of different socialization processes for women to more communal behavior as well as certain sex differences regarding empathy and emotion processing could explain the observed differences in social sensitivity in Woolley's et al. (ibid.) experiment.

There are also some studies that describe the role of women for developing a collective emotional intelligence (Curşeu, Pluut, Boroş, & Meslec, 2014). Collective emotional intelligence is defined as the ability of a group to develop a set of norms that promote awareness and regulation of member and group emotions. Due to their higher social sensitivity and relational orientation, women promote the emergence of collective competencies by facilitating the coordination of individual competencies during social interactions. “We find that the percentage of women in the group positively influences collective emotional intelligence” (ibid.).

In summary, in the same way that gendered status beliefs inhibit information sharing in research teams, now it is argued that sex and gender differences in social sensitivity can be supportive of group processes. Gender balanced groups appear to have better chances to create an integrative communicative environment.

3.6.3. Sexual Harassment

From the arguments of Metcalfe and Linstead (2003) about teamwork and the strong association to a masculinist discourse emphasizing effectiveness and performance above as well as in (Hawkins, 2013) accounts on gender discourse as a mean for concertive control mechanisms in teams, it follows that team members and indeed mainstream teamwork research seldom addresses the link between individuals gendered experience as team members with that of other team members or team processes and how this may be linked to

productivity or performance. One example of how individual experiences of gender oppression in organizations operate on a team level is sexual or gender harassment processes. Earlier research has shown that harassment is an “organisational stressor” that has significant, negative outcomes for targets including low job satisfaction, psychological distress, anxiety, and depression, job loss, career interruption, increased turn over, and absenteeism (Fitzgerald, Drasgow, Hulin, Gelfand, & Magley, 1997; Raver & Gelfand, 2005). According to Fitzgerald, Gelfand, & Drasgow (1995) sexual harassment is a behavioral construct consisting of three dimensions:

> (1) gender harassment: insulting verbal and nonverbal behaviors conveying insulting, hostile, or degrading attitudes toward women;
>
> (2) unwanted sexual attention: verbal and nonverbal behaviors that are offensive, unwanted and unreciprocated (e.g., unwanted touching or grabbing; repeated and nonreciprocal requests for dates), and
>
> (3) sexual coercion: behaviors using bribes or threats, and/or making job-related benefits contingent upon sexual cooperation.

Later, Magley, Waldo, Drasgow, & Fitzgerald (1999) have defined gender harassment as consisting of two distinct dimensions: sexist hostility, which comprises insulting verbal and nonverbal behaviors based on gender (e.g. making offensive sexist remarks), and sexual hostility, which comprises insulting, explicitly sexual verbal and nonverbal behaviors (e.g., repeatedly telling sexual stories).

Raver and Gelfand (2005) proposed that high levels of “ambient sexual harassment” (see Glomb et al., 1997) would be associated with high levels of team conflict, low levels of team cohesion, and low levels of team citizenship behaviors. It was further proposed that these mediating processes would subsequently be associated with teams' financial performance. Discretionary stimuli are “characteristics of a work environment that are transmitted to individuals differentially; ambient stimuli pervade a group setting and are potentially available to all group members” (ibid., p.388). Whereas earlier research has suggested that ambient sexual harassment explained variance in individuals' outcomes that went beyond the variance accounted or by being a direct target of sexual harassment. Raver and Gelfand study provides the basis for conceptualizing sexual harassment as having team level influences and direct effects on teamwork outputs.

Bibliography

Abramo, G., Cicero, T., & D'Angelo, C. A. (2015). Should the research performance of scientists be distinguished by gender? *Journal of Informetrics*, *9*(1), 25–38. http://doi.org/10.1016/j.joi.2014.11.002

Abramo, G., D'Angelo, C. A., & Murgia, G. (2013). Gender differences in research collaboration. *Journal of Informetrics*, *7*(4), 811–822.

Abramo, G., D'Angelo, C., & Caprasecca, A. (2009). Gender differences in research productivity: A bibliometric analysis of the Italian academic system. *Scientometrics*, *79*(3), 517–539.

Acker, J. (1992). Gendering Organizational Theory. In A. J. Mills & P. Tancred (Eds.), *Gendering Organizational Analysis* (pp. 248–260). London: SAGE Publications.

Acker, J. (2000). Gendered Contradictions in Organizational Equity Projects. *Organization*, *7*(4), 625–632.

Acker, J. (2006). Inequality Regimes: Gender, class, and race in organizations. *Gender & Society*, *20*(4), 441–464.

Ahonen, P., Tienari, J., Merilainen, S., & Pullen, A. (2013). Hidden contexts and invisible power relations: A Foucauldian reading of diversity research. *Human Relations*, *67*(3), 263–286. http://doi.org/10.1177/0018726713491772

Allison, P. D., & Long, S. J. (1990). Departmental Effects on Scientific Productivity. *American Sociological Review*, *55*(4), 469–478.

Anderson, C., John, O. P., Keltner, D., & Kring, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology*, *81*(1), 116–132. http://doi.org/10.1037/0022-3514.81.1.116

Anderson, C., & Kilduff, G. J. (2009). Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology*, *96*(2), 491–503. http://doi.org/10.1037/a0014201

Andersson, S., Amundsdotter, E., Svensson, M., & Däldehög, A.-S. (2009). *Middle Managers as Change Agents*. Stockholm.

Ball, S. J. (2012). Performativity, Commodification and Commitment: An I-Spy Guide to the Neoliberal University. *British Journal of Educational Studies*, *60*(1), 17–28. http://doi.org/10.1080/00071005.2011.650940

Barad, K. (2007). *Meeting the universe halfway : quantum physics and the entanglement of*

matter and meaning.

Barker, J. A. (1993). Tightening the Iron Cage - Concertive Control in Self-Managing Teams. *Administrative Science Quarterly*, *38*(3), 408–437.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251. http://doi.org/10.1111/1469-7610.00715

Benschop, Y., & Doorewaard, H. (1998). Six of One and Half a Dozen of the Other: The Gender Subtext of Taylorism and Team-based Work. *Gender, Work & Organization*, *5*(1), 5–18.

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*(3), 241–255.

Berger, J., Fiske, H. M., Norman, R. Z., & Zelditch, M. (1977). *Status characteristics and social interaction : an expectation-states approach*. New York: Elsevier Scientific Pub. Co.

Berger, J., Rosenholtz, S. J., & Zelditch, M. (1980). Status Organizing Processes. *Annual Review of Sociology*, *6*(1), 479–508. http://doi.org/10.1146/annurev.so.06.080180.002403

Berger, L., Benschop, Y., & van den Brink, M. (2015). Practising Gender When Networking: The Case of University-Industry Innovation Projects. *Gender, Work & Organization*, *22*(6), 556–578. http://doi.org/10.1111/gwao.12104

Blake, M. K., & Hanson, S. (2005). Rethinking innovation: context and gender. *Environment and Planning A*, *37*(4), 681–701. http://doi.org/10.1068/a3710

Bland, C. J., Center, B. A. (Bruce A., Finstad, D. A., Risbey, K. R., & Staples, J. (2006). The Impact of Appointment Type on the Productivity and Commitment of Full-Time Faculty in Research and Doctoral Institutions. *The Journal of Higher Education*, *77*(1), 89–123. http://doi.org/10.1353/jhe.2006.0002

Bozeman, B., & Gaughan, M. (2011). How do men and women differ in research collaborations? An analysis of the collaborative motives and strategies of academic researchers. *Research Policy*, *40*(10), 1393–1402. http://doi.org/10.1016/j.respol.2011.07.002

Brouns, M. (2000). The Gendered Nature of Assessment Procedures in Scientific Research Funding: The Dutch Case. *Higher Education in Europe*, *25*(2), 193–199. http://doi.org/10.1080/713669261

Bruno, I. (2009). The “Indefinite Discipline” of Competitiveness Benchmarking as a Neoliberal Technology of Government. *Minerva*, *47*(3), 261–280. http://doi.org/10.1007/s11024-009-9128-0

Buchmann, C., DiPrete, T. a., & McDaniel, A. (2008). Gender Inequalities in Education. *Annual Review of Sociology*, *34*(1), 319–337.

Burt, R. S. (1998). The Gender of Social Capital. *Rationality and Society*, *10*(1), 5–46.

Busolt, U., & Kugele, K. (2009). The gender innovation and research productivity gap in Europe. *International Journal of Innovation and Sustainable Development*, *4*(2/3), 109. http://doi.org/10.1504/IJISD.2009.028066

Butler, J. (1990). *Gender trouble : feminism and the subversion of identity*. New York: Routledge.

Calás, M. B., Holgersson, C., & Smircich, L. (2009). “Diversity Management”? Translation? Travel? *Scandinavian Journal of Management*, *25*(4), 349–351. http://doi.org/10.1016/j.scaman.2009.09.006

Calás, M. B., & Smircich, L. (2006). From the “ Woman”s Point of View’ Ten Years Later : Towards a Feminist Organization Studies. In S. R. Clegg, C. Hardy, T. B. Lawrence, & W. R. Nord (Eds.), *The Sage Handbook of Organization Studies* (pp. 284–346). London: Sage Publications.

Carayol, N., & Matt, M. (2006). Individual and collective determinants of academic scientists’ productivity. *Information Economics and Policy*, *18*(1), 55–72.

Carli, L. L. (1989). Gender differences in interaction style and influence. *Journal of Personality and Social Psychology*, *56*(4), 565–576. http://doi.org/10.1037/0022-3514.56.4.565

Carli, L. L. (2006). Gender Issues in Workplace Groups: Effects of Gender and Communication Style on Social Influence. In M. Barrett & M. J. Davidson (Eds.), *Gender and Communication at Work* (pp. 69–83). Aldershot: Ashgate Pub Ltd.

Carli, L. L. (2010). Gender and Group Behavior. In *Handbook of Gender Research in Psychology* (p. 132). Springer.

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest*, *15*(3), 75–141. http://doi.org/10.1177/1529100614541236

Cheng, C. (1996). *Masculinities in organizations*. Thousand Oaks Calif.: Sage Publications.

Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: gender effects in brain and behavior. *Neuroscience and Biobehavioral*

Reviews, *46 Pt 4*(P4), 604–27. http://doi.org/10.1016/j.neubiorev.2014.09.001

Cole, J. R., & Zuckerman, H. (1984). The Productivity Puzzle: Persistence and change in Patterns of Publication of Men and Women Scientists. In M. W. Steinkempt & M. L. Maehr (Eds.), *Advances in Motivation and Achievement A Research Journal Women in Science* (Vol. 2, pp. 217–258). Greenwich, Connecticut: JAI Press.

Cole, J. R., & Zuckerman, H. (1991). Marriage, Motherhood, and Research Performance in Science. In H. Zuckerman & J. R. Cole (Eds.), *The Outer Circle: Women in the Scientific Community* (pp. 157–170). New York: Norton.

Collinson, D. L., & Hearn, J. (1994). Naming men as men: implications for work, organization and management. *Gender, Work & Organization*, *1*(1), 2–22.

Collinson, D. L., & Hearn, J. (1996). Breaking the silence: on men, masculinities and management. In D. L. Collinson & J. Hearn (Eds.), *Men as Managers, Managers as Men. Critical Perspectives on Men, Masculinities, and Management* (pp. 1–24). London: SAGE Publications.

Connell, R. (1987). *Gender and power: society, the person, and sexual politics*. Stanford Calif.: Stanford University Press.

Craft, R. M., Mogil, J. S., & Aloisi, A. M. (2004). Sex differences in pain and analgesia: the role of gonadal hormones. *European Journal of Pain (London, England)*, *8*(5), 397–411. http://doi.org/10.1016/j.ejpain.2004.01.003

Curşeu, P. L., Pluut, H., Boroş, S., & Meslec, N. (2014). The magic of collective emotional intelligence in learning groups: No guys needed for the spell! *British Journal of Psychology (London, England : 1953)*. http://doi.org/10.1111/bjop.12075

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology*, *35*(1), 67–82. http://doi.org/10.1016/j.psyneuen.2009.10.006

DiTomaso, N., Post, C., Smith, D. R., Farris, G. F., & Cordero, R. (2007). Effects of Structural Position on Allocation and Evaluation Decisions for Scientists and Engineers in Industrial R&D. *Administrative Science Quarterly*, *52*(2), 175–207. http://doi.org/10.2189/asqu.52.2.175

Donges, U.-S., Kersting, A., & Suslow, T. (2012). Women's greater ability to perceive happy facial emotion automatically: gender differences in affective priming. *PloS One*, *7*(7), e41745. http://doi.org/10.1371/journal.pone.0041745

Dundar, H., & Lewis, D. R. (1998). Determinants of research productivity in Higher Education.

Research in Higher Education, *39*(6), 607–631.

Egeland, C. (2001). Bergsklätterskan och de militanta lebiska feministerna: om könskarriärer och jämställdhet inom akademin. *Kvinnovetenskaplig Tidskrift*, *1*, 53–67.

Elomaki, A. (2015). The economic case for gender equality in the European Union: Selling gender equality to decision-makers and neoliberalism to women's organizations. *European Journal of Women's Studies*, *22*(3), 288–302. http://doi.org/10.1177/1350506815571142

Elsevier. (2015). *Mapping Gender in the German Research Arena*. *Analytical Services Report*. Elsevier.

Ely, R. J., & Thomas, D. A. (2001). Cultural Diversity at Work: The Effects of Diversity Perspectives on Work Group Processes and Outcomes. *Administrative Science Quarterly*, *46*(2), 229. http://doi.org/10.2307/2667087

Etzkowitz, H., Kemelgor, C., & Uzzi, B. (2000). *Athena unbound : the advancement of women in science and technology*. Cambridge ;;New York: Cambridge University Press.

European Commission. (2003). *Waste of talents: turning private struggles into a public issue. Women and Science in Enwise countries. Science*.

European Commission. (2009). *She Figures 2009: statistics and indicators on gender equality in science*. *Economy and Society*. Luxembourg: Office for Official Publications of the European Communities.

European Commission. (2011). *The pay gap for women in decision making positions: increasing responsibilities, increasing pay gap*. Luxembourg.

European Commission. (2013). She figures 2012 gender in research and innovation : statistics and indicators. Luxembourg: Publications Office of the European Union.

Fausto-Sterling, A. (1981). Women and science. *Women's Studies International Quarterly*, *4*(1), 41–50. http://doi.org/10.1016/S0148-0685(81)96343-0

Fausto-Sterling, A. (2000). *Sexing the body : gender politics and the construction of sexuality*. New York NY: Basic Books.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902.

Fitzgerald, L. F., Drasgow, F., Hulin, C. L., Gelfand, M. J., & Magley, V. J. (1997). Antecedents and consequences of sexual harassment in organizations: a test of an integrated model. *The Journal of Applied Psychology*, *82*(4), 578–589.

Fitzgerald, L. F., Gelfand, M. J., & Drasgow, F. (1995). Measuring Sexual Harassment: Theoretical and Psychometric Advances. *Basic and Applied Social Psychology*, *17*(4), 425–445.

Fox, M. F., & Xiao, W. (2013). Perceived chances for promotion among women associate professors in computing: individual, departmental, and entrepreneurial factors. *The Journal of Technology Transfer*, *38*(2), 135–152. http://doi.org/10.1007/s10961-012-9250-2

Fuchs, S., von Stebut, J., & Allmendinger, J. (2001). Gender, science, and scientific organizations in Germany. *Minerva*, *39*(2), 175–201.

Garforth, L., & Kerr, A. (2009). Women and Science. What's the Problem? *Social Politics*, 379–403.

Garud, R., Gehman, J., & Kumaraswamy, A. (2011). Complexity Arrangements for Sustained Innovation: Lessons from 3M Corporation. *Organization Studies*, *32*(6), 737–767. http://doi.org/10.1177/0170840611410810

Glick, P., Lameiras, M., Fiske, S. T., Eckes, T., Masser, B., Volpato, C., … Wells, R. (2004). Bad but bold: Ambivalent attitudes toward men predict gender inequality in 16 nations. *Journal of Personality and Social Psychology*, *86*(5), 713–28. http://doi.org/10.1037/0022-3514.86.5.713

Glomb, T. M., Richman, W. L., Hulin, C. L., Drasgow, F., Schneider, K. T., & Fitzgerald, L. F. (1997). Ambient Sexual Harassment: An Integrated Model of Antecedents and Consequences. *Organizational Behavior and Human Decision Processes*, *71*(3), 309–328. http://doi.org/10.1006/obhd.1997.2728

Gray, M., & James, A. (2007). Connecting gender and economic competitiveness: lessons from Cambridge's high-tech regional economy. *Environment and Planning A*, *39*(2), 417–436. http://doi.org/10.1068/a37406

Greenspan, J. D., Craft, R. M., LeResche, L., Arendt-Nielsen, L., Berkley, K. J., Fillingim, R. B., … Traub, R. J. (2007). Studying sex and gender differences in pain and analgesia: a consensus report. *Pain*, *132 Suppl* , S26–45. http://doi.org/10.1016/j.pain.2007.10.014

Haines, E. L., Deaux, K., & Lofaro, N. (2016). The Times They Are a-Changing … or Are They Not? A Comparison of Gender Stereotypes, 1983-2014. *Psychology of Women Quarterly*, 0361684316634081–. http://doi.org/10.1177/0361684316634081

Halberstam, J. (2011). *The queer art of failure*. Durham: Duke University Press.

Hancock, A.-M. (2007). When Multiplication Doesnt Equal Quick Addition: Examining Intersectionality as a Research Paradigm. *Perspectives on Politics*, *5*(1), 63–79.

Haraway, D. (1991). *Simians, cyborgs, and women : the reinvention of nature*. New York: Routledge.

Harrison, D. A., & Klein, K. J. (2007). What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations. *Academy of Management Review*.

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, Teams, and Task Performance: Changing Effects of Surface- and Deep-Level Diversity on Group Functioning. *Academy of Management Journal*, *45*(5), 1029–1045.

Hawkins, B. (2013). Gendering the Eye of the Norm: Exploring Gendered Concertive Control Processes in Two Self-Managing Teams. *Gender, Work & Organization*, *20*(1), 113–126. http://doi.org/10.1111/j.1468-0432.2011.00588.x

Hearn, J. (2015). *Men of the world : genders, globalizations, transnational times*.

Hearn, J., & Husu, L. (2011). Understanding Gender: Some Implications for Science and Technology. *Interdisciplinary Science Reviews*, *36*(2), 103–113. http://doi.org/10.1179/030801811X13013181961301

Hearn, J., & Louvrier, J. (2015). Theories of Difference, Diversity and Intersectionality: What Do They Bring to Diversity Management? In B. Regine, I. Bleijenbergh, E. Henttonen, & A. J. Mills (Eds.), *The Oxford Handbook of Diversity in Organizations* (pp. 62–82). Oxford ; New York: Oxford University Press.

Hearn, J., & Parkin, W. (2003). The gendered organisation. A positive critique. *Comportamento Organizacional E Gestão*, *9*(2), 125–146.

Hentschel, T., Shemla, M., Wegge, J., & Kearney, E. (2013). Perceived Diversity and Team Functioning: The Role of Diversity Beliefs and Affect. *Small Group Research*, *44*(1), 33–61. http://doi.org/10.1177/1046496412470725

Hill, M. R., & Mays, J. (2011). *The Gender Book*.

Hobman, E. V., Bordia, P., & Gallois, C. (2004). Perceived Dissimilarity and Work Group Involvement: The Moderating Effects of Group Openness to Diversity. *Group & Organization Management*, *29*(5), 560–587. http://doi.org/10.1177/1059601103254269

Holvino, E. (2010). Intersections: The simultaneity of race, gender and class in Organization Studies. *Gender, Work and Organization*, *17*(3), 248–277.

Homan, A. C., Hollenbeck, J. R., Humphrey, Stephen E. Knippenberg, D. van, & Ilgen, D. R. (2008). Facing differences with an open mind : openness to experience, salience of intragroup differences, and performance of diverse work groups. *Academy of*

Management Journal, *51*(6), 1204–1222.

Hunter, L. A., & Leahey, E. (2010). Parenting and research productivity: New evidence and methods. *Social Studies of Science*, *40*(3), 433–451. http://doi.org/10.1177/0306312709358472

Husu, L. (2005). *Dold könsdiskriminering på akademiska arenor – osynligt*, *synligt*, *subtilt* (No. 41R). Stockholm.

Hyde, J. S. (2005). The gender similarities hypothesis. *The American Psychologist*, *60*(6), 581–92. http://doi.org/10.1037/0003-066X.60.6.581

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, *65*, 373–98. http://doi.org/10.1146/annurev-psych-010213-115057

Ibarra, H. (1992). Homophily and Differential Returns: Sex Differences in Network Structure and Access in an Advertising Firm. *Administrative Science Quarterly*, *37*(3), 422–447.

Ibarra, H. (1993). Personal Networks of Women and Minorities in Management: a Conceptual Framework. *Academy of Management Review*, *18*(1), 56–87. http://doi.org/10.5465/AMR.1993.3997507

Joshi, A. (2014). By Whom and When Is Women's Expertise Recognized? The Interactive Effects of Gender and Education in Science and Engineering Teams. *Administrative Science Quarterly*, *59*(2), 202–239. http://doi.org/10.1177/0001839214528331

Joshi, A., & Knight, A. P. (2015). Who Defers to Whom and Why? Dual Pathways Linking Demographic Differences and Dyadic Deference to Team Effectiveness. *Academy of Management Journal*, *58*(1), 59–84. http://doi.org/10.5465/amj.2013.0718

Kaiser, A., Haller, S., Schmitz, S., & Nitsch, C. (2009). On sex/gender related similarities and differences in fMRI language research. *Brain Research Reviews*, *61*(2), 49–59. http://doi.org/10.1016/j.brainresrev.2009.03.005

Kanter, R. (1977). *Men and women of the corporation*. New York: Basic Books.

Karamessini, M. (2003). *The Cost of Non Gender Equality Legislation: The Greek Report*.

Kegen, N. V. (2013). Science Networks in Cutting-edge Research Institutions: Gender Homophily and Embeddedness in Formal and Informal Networks. *Procedia - Social and Behavioral Sciences*, *79*, 62–81. http://doi.org/10.1016/j.sbspro.2013.05.057

Kiesler, S., Sproull, L., & Eccles, J. S. (1985). Pool Halls, Chips, and War Games: Women in the Culture of Computing. *Psychology of Women Quarterly*, *9*(4), 451–462. http://doi.org/10.1111/j.1471-6402.1985.tb00895.x

Knights, D., & Omanović, V. (2015). Rethinking Diversity in Organizations and Society. In B. Regine, I. Bleijenbergh, E. Henttonen, & A. J. Mills (Eds.), *The Oxford Handbook of Diversity in Organizations* (pp. 83–108). Oxford: Oxford University Press.

Lamont, M., Mallard, G., & Guetzkow, J. (2006). Beyond Blind Faith: Overcoming the Obstacles to Interdisciplinary Evaluation. *Research Evaluation*, *15*(1), 43–55.

Landino, R. A., & Owen, S. V. (1988). Self-efficacy in university faculty. *Journal of Vocational Behavior*, *33*(1), 1–14.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, *504*(7479), 211–213. http://doi.org/10.1038/504211a

Larivière, V., Vignola-Gagné, E., Villeneuve, C., Gélinas, P., & Gingras, Y. (2011). Sex differences in research funding, productivity and impact: an analysis of Québec university professors. *Scientometrics*, *87*(3), 483–498. http://doi.org/10.1007/s11192-011-0369-y

Leahey, E. (2006). Gender Differences in Productivity: Research Specialization as a Missing Link. *Gender & Society*, *20*(6), 754–780. http://doi.org/10.1177/0891243206293030

Lee, S., & Bozeman, B. (2005). The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*, *35*(5), 673–702.

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265. http://doi.org/10.1126/science.1261375

Linneman, T. J. (2012). Gender in Jeopardy!: Intonation Variation on a Television Game Show. *Gender & Society*, *27*(1), 82–105. http://doi.org/10.1177/0891243212464905

Long, J. S. (1992). Measures of Sex Differences in Scientific Productivity. *Social Forces*, *71*(1), 159–178. http://doi.org/10.1093/sf/71.1.159

Lord, R. G., de Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, *71*(3), 402–410.

Magley, V. J., Waldo, C. R., Drasgow, F., & Fitzgerald, L. F. (1999). The Impact of Sexual Harassment on Military Personnel. *Military Psychology*, *11*(3), 283–302.

Maliniak, D., Powers, R., & Walter, B. F. (2013). *The Gender Citation Gap in International Relations*. *International Organization* (Vol. 67).

Marsh, H. W., Bornmann, L., Mutz, R., Daniel, H.-D., & O 'mara, A. (2009). Gender Effects in

the Peer Reviews of Grant Proposals: A Comprehensive Meta-Analysis Comparing Traditional and Multilevel Approaches. *Source: Review of Educational Research*, *79*(3), 1290–1326.

Martin, P. Y. (2006). Practising gender at work: Further thoughts on reflexivity. *Gender, Work and Organization*.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. http://doi.org/10.1146/annurev.soc.27.1.415

Menzies, H., & Newson, J. (2007). No Time to Think: Academics' life in the globally wired university. *Time & Society*, *16*(1), 83–98. http://doi.org/10.1177/0961463X07074103

Metcalfe, B. D., & Woodhams, C. (2008). Critical perspectives in diversity and equality management. *Gender in Management*, *23*(6), 377–381.

Metcalfe, B. D., & Woodhams, C. (2012). Introduction: New Directions in Gender, Diversity and Organization Theorizing - Reimagining Feminist Post-colonialism, Transnationalism and Geographies of Power. *International Journal of Management Reviews*, *14*(2), 123–140. http://doi.org/10.1111/j.1468-2370.2012.00336.x

Metcalfe, B., & Linstead, A. (2003). Gendering Teamwork: Re-Writing the Feminine. *Gender, Work & Organization*, *10*(1), 94–119. http://doi.org/10.1111/1468-0432.00005

Miller, D. I., & Wai, J. (2015). The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: A 30-year analysis. *Frontiers in Psychology*, *6*(FEB).

Miller, T., & Triana, M. del C. (2009). Demographic Diversity in the Boardroom : Mediators of the Board Diversity – Firm Performance Relationship Toyah Miller and María del Carmen Triana. *Journal of Management Studies*, *46*(5), 755–786.

Monro, S. (2005). *Gender Politics Citizenship, Activism and Sexual Diversity*. London: Pluto Press.

Mountz, A., Bonds, A., Mansfield, B., Loyd, J., Hyndman, J., Walton-Roberts, M., … Curran, W. (2015, August 18). For Slow Scholarship: A Feminist Politics of Resistance through Collective Action in the Neoliberal University. *ACME: An International E-Journal for Critical Geographies*.

Mulac, A., Bradac, J., & Gibbons, P. (2006). Empirical support for the gender-as-culture hypothesis. *Human Communication Research*, *27*(1), 121–152. http://doi.org/10.1111/j.1468-2958.2001.tb00778.x

Mutz, R., Bornmann, L., & Daniel, H.-D. (2012). Does Gender Matter in Grant Peer Review?

Zeitschrift Für Psychologie, *220*(2), 121–129. http://doi.org/10.1027/2151-2604/a000103

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, *45*(3), 211–236. http://doi.org/10.1080/01638530802073712

Nishii, L. H. (2012). The Benefits of Climate for Inclusion for Gender-Diverse Groups. *Academy of Management Journal*, *56*(6), 1754–1774. http://doi.org/10.5465/amj.2009.0823

Nosek, B. a, Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10593–7. http://doi.org/10.1073/pnas.0809921106

Onnela, J.-P., Waber, B. N., Pentland, A., Schnorf, S., & Lazer, D. (2014). Using sociometers to quantify social interaction patterns. *Scientific Reports*, *4*. http://doi.org/10.1038/srep05604

Ostendorp, A., & Steyaert, C. (2009). How different can differences be(come)?: Interpretative repertoires of diversity concepts in Swiss-based organizations. *Scandinavian Journal of Management*, *25*(4), 374–384. http://doi.org/10.1016/j.scaman.2009.09.003

Parken, A. (2010). Gender, Knowledge Dynamics and Regional Policy. In H. Halkier, M. Dahlström, L. James, J. Manniche, & L. S. Olsen (Eds.), *Knowledge Dynamics* , *Regional Development and Public Policy*. Aalborg Ost: Department of History, International and Social Studies, Aalborg University.

Paul-Hus, A., Sugimoto, C. R., Haustein, S., & Larivière, V. (2015). Is There a Gender Gap in Social Media Metrics?

Pavlova, M. A., Scheffler, K., & Sokolov, A. N. (2015). Face-n-Food: Gender Differences in Tuning to Faces. *PLOS ONE*, *10*(7), e0130363. http://doi.org/10.1371/journal.pone.0130363

Pecis, L. (2016). Doing and undoing gender in innovation: Femininities and masculinities in innovation processes. *Human Relations*, 0018726716634445–. http://doi.org/10.1177/0018726716634445

Pekkarinen, T. (2008). Gender differences in educational attainment: Evidence on the role of tracking from a finnish quasi-experiment. *Scandinavian Journal of Economics*, *110*(4), 807–825.

Pettersson, K. (2007). Men and Male as the Norm? A Gender Perspective on Innovation Policies in Denmark, Finland and Sweden. *Innovation*. Stockholm: Nordregio.

Phillips, K. W., & Apfelbaum, E. P. (2012). Delusions of Homogeneity? Reinterpreting the Effects of Group Diversity. In E. A. Mannix & M. A. Neale (Eds.), *Looking Back, Moving Forward: A Review of Group and Team-Based Research* (Vol. 15, pp. 185–207). Bingley: Emerald Group Publishing. http://doi.org/10.1108/S1534-0856(2012)0000015011

Pollitzer, E., & Schraudner, M. (2015). Integrating Gender Dynamics into Innovation Ecosystems. *Sociology and Anthropology*, *3*(11), 617–626.

Prozesky, H. (2008). A career-history analysis of gender differences in publication productivity among south african academics. *Science Studies*, *21*(2), 47–67.

Prpić, K. (2002). Gender and productivity differentials in science. *Scientometrics*, *55*(1), 27–54.

Puuska, H. M. (2010). Effects of scholar's gender and professional position on publishing productivity in different publication types. Analysis of a Finnish university. *Scientometrics*, *82*(2), 419–437.

Qin, J., Muenjohn, N., & Chhetri, P. (2014). A Review of Diversity Conceptualizations: Variety, Trends, and a Framework. *Human Resource Development Review*, *13*(2), 133–157. http://doi.org/10.1177/1534484313492329

Raver, J. L., & Gelfand, M. J. (2005). Beyond the Individual Victim: Linking Sexual Harassment, Team Processes, and Team Performance. *The Academy of Management Journal*, *48*(3), 387–400.

Regitz-Zagrosek, V., Oertelt-Prigione, S., Prescott, E., Franconi, F., Gerdts, E., Foryst-Ludwig, A., … Stangl, V. (2015). Gender in cardiovascular diseases: impact on clinical manifestations, management, and outcomes. *European Heart Journal*. http://doi.org/10.1093/eurheartj/ehv598

Ridgeway, C. L. (2007). Gender as a Group Process: Implications for the Persistence of Inequality. In S. J. Correll (Ed.), *Social Psychology of Gender* (Vol. 24, pp. 311–333). Bingley: Emerald Group Publishing Limited.

Ridgeway, C. L. (2011). *Framed by Gender. How Gender Inequality Persists in the Modern World*. New York: Oxford University Press.

Ridgeway, C. L., & Bourg, C. (2004). Gender as Status: An Expectation States Theory Approach. In A. H. Eagly, A. E. Beall, & R. J. Sternberg (Eds.), *The Psychology of Gender* (2nd ed., pp. 217–241). New York NY: Guilford Press.

Ridgeway, C. L., & Smith-Lovin, L. (1999). The Gender System and Interaction. *Annual Review of Sociology*, *25*(1), 191–216. http://doi.org/10.1146/annurev.soc.25.1.191

Rönnblom, M. (2009). Bending towards growth: Discursive constructions of gender equality in an era of governance and neoliberalism. In E. Lombardo, P. Meier, & M. Verloo (Eds.), *The Discursive Politics of Gender Equality Stretching, Bending and Policy-Making* (pp. 105–120). London: Routledge.

Sadl, Z. (2009). “We Women Are No Good at It”: Networking in Academia. *Sociol. Cas.*, *45*, 1239–1263.

Sandström, U., & Hällsten, M. (2007). Persistent nepotism in peer-review. *Scientometrics*, *74*(2), 175–189. http://doi.org/10.1007/s11192-008-0211-3

Schiebinger, L., Klinge, I., Paik, H. Y., Sánchez de Madariaga, I., Schraudner, M., & Stefanick, M. (2016). Gendered Innovations in Science, Health & Medicine, Engineering, and Environment. Retrieved from http://genderedinnovations.stanford.edu

Siedler, V. (1990). Men, feminism and power. In J. Hearn & D. Morgan (Eds.), *Men, Masculinities and Social Theory* (p. ??). London: Unwin Hyman.

Simpson, B., Willer, R., & Ridgeway, C. L. (2012). Status Hierarchies and the Organization of Collective Action. *Sociological Theory*, *30*(3), 149–166. http://doi.org/10.1177/0735275112457912

Sinclair, A. (2000). Women within diversity: risks and possibilities. *Women in Management Review*, *15*(5/6), 237–246. http://doi.org/10.1108/09649420010372850

Slaughter, S., & Rhoades, G. (2000). The Neo-Liberal University. *New Labor Forum*, *6*, 73–79.

Snell, C., Sorensen, J., Rodriguez, J. J., & Kuanliang, A. (2009). Gender differences in research productivity among criminal justice and criminology scholars. *Journal of Criminal Justice*, *37*(3), 288–295.

Symonds, M. R. E., Gemmell, N. J., Braisher, T. L., Gorringe, K. L., & Elgar, M. A. (2006). Gender differences in publication output: towards an unbiased metric of research performance. *PloS One*, *1*(1), e127. http://doi.org/10.1371/journal.pone.0000127

Uhly, K. M., Visser, L. M., & Zippel, K. S. (2015). Gendered patterns in international research collaborations in academia. *Studies in Higher Education*, 1–23. http://doi.org/10.1080/03075079.2015.1072151

van Arensbergen, P., van der Weijden, I., & van den Besselaar, P. (2012). Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics*, *93*(3), 857–868. http://doi.org/10.1007/s11192-012-0712-y

Van Balen, B., Van Arensbergen, P., Van der Weijden, I., & Van den Besselaar, P. (2012). Determinants of success in academic careers. *Higher Education Policy*, *25*, 313–334.

van den Brink, M., & Benschop, Y. (2012). Slaying the seven-headed dragon: The quest for gender change in academia. *Gender, Work and Organization*, *19*(1), 71–92.

van Dick, R., van Knippenberg, D., Hagele, S., Guillaume, Y. R. F., & Brodbeck, F. C. (2008). Group diversity and group identification: The moderating role of diversity beliefs. *Human Relations*, *61*(10), 1463–1492. http://doi.org/10.1177/0018726708095711

Wagner, D. G., & Berger, J. (1997). Gender and Interpersonal Task Behaviors: Status Expectation Accounts. *Sociological Perspectives*, *40*(1), 1–32. http://doi.org/10.2307/1389491

Wajcman, J. (2007, June 1). From women and technology to gendered technoscience. *Information, Communication and Society*. Routledge.

Walby, S. (2007). *Gendering the knowledge economy : comparative perspectives*. Basingstoke [England] ;;New York: Palgrave Macmillan.

Walton, G. M., Murphy, M. C., & Ryan, A. M. (2015). Stereotype Threat in Organizations: Implications for Equity and Performance. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 523–550. http://doi.org/10.1146/annurev-orgpsych-032414-111322

Wenneras, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, *387*(6631), 341–3. http://doi.org/10.1038/387341a0

West, C., & Zimmerman, D. H. (1987). Doing Gender. *Gender & Society*, *1*(2), 125–151. http://doi.org/10.1177/0891243287001002002

West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The Role of Gender in Scholarly Authorship. *PLoS ONE*, *8*(7).

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., … Johnson, B. (2015). *TheMetric Tide. Report of the Independent Review of the Role of Metrics in Research Assessment and Management*.

Wood, W., & Eagly, A. H. (2012). Biosocial Construction of Sex Differences and Similarities in Behavior. In *Advances in Experimental Social Psychology* (Vol. 46, pp. 55–123). Elsevier. http://doi.org/10.1016/B978-0-12-394281-4.00002-7

Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective Intelligence and Group Performance. *Current Directions in Psychological Science*, *24*(6), 420–424. http://doi.org/10.1177/0963721415599543

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, *330*(October), 686–688.

Xie, Y., & Shauman, K. A. (1998). Sex Differences in Research Productivity: New Evidence about an Old Puzzle. *American Sociological Review*, *63*(6), 847–870.

Xie, Y., & Shauman, K. A. (2003). *Women in science : career processes and outcomes*. Cambridge Mass.: Harvard University Press.

Zanoni, P., Janssens, M., Benschop, Y., & Nkomo, S. (2010). Guest Editorial: Unpacking Diversity, Grasping Inequality: Rethinking Difference Through Critical Perspectives. *Organization*, *17*(1), 9–29. http://doi.org/10.1177/1350508409350344

GEDII GENDER DIVERSITY IMPACT – Improving research and innovation through gender diversity

WWW.GEDII.EU

D1.1. Conceptual Framework

Chapter 4 – Integration and Operationalization

Jörg Müller, Sandra Klatt, Anne-Charlott Callerstig, Ulf Sandström

31st of May 2016

Table of Contents

4. Integrating Insights from Team Science, Research Assessment, and Gender Studies

The literature review so far has targeted three knowledge areas separately: we have gathered evidence on the factors that influence team work, how research assessment should be carried out and how Science, Technology and Innovation can be understood as gendered. In a subsequent step, these loosely connected areas need to be combined into a coherent conceptual framework specifically targeting how gender (diversity) affects team work and their performance. It is the intersection of these three knowledge fields that interest us now and that not only needs to be spelled out but also operationzalied for our empirical work. The formulation of the conceptual framework at this stage also needs to specify existing gaps within the literature and open questions in order to anticipate problematic issues for our own project but also identify possible contributions for downstream research.

The conceptual framework will have to guide the case studies, the development of the questionnaire and the Gender-Diversity Index (GDI). The operationalization of the identified concepts therefore can vary quite substantially. During the case studies, the key dimensions of team performance will be mainly implemented through concrete questions in the semi-structured interview guides. In addition, quite limited survey measures will be used to contextualize sociometric badges data. For the cross-national survey on the other hand, extensive scales will be deployed to collect data on team demographics, team climate, network structure, and other than bibliometric performance measures.

The next section will give a quick summary of the key dimensions for assessing the impact of gender diversity on research performance. This includes on the one hand variables that directly affect team processes and on the other hand control variables in order to distinguish gender effects from all the other factors that affect team performance.

4.1. Delineating “Teams” and “Research Collaboration”

A crucial question to be addressed by the conceptual framework concerns the criteria that define team membership. Since research teams are the basic reference unit for measuring research “performance”, membership criteria need to be defined with great care. Where we draw the line between “core” team members vs. external collaborators for example will affect substantially the “performance” score of the group. Both, collaborations among core members but also with members of other teams or other individuals, influences idea generation and work processes within the team. Advice networks often span well beyond the actual research team and constitute important sources of diverse knowledge and information assets that has an impact on the overall performance. Thus, the logic upon which we decide that a certain person pertains to the “core” team or rather is an external collaborator needs to be clearly specified. This is all the more important, since a clear-cut and widely accepted definition of what constitutes a “team” or a research collaboration does not exist.

As outlined, for the purpose of the current project, team membership will be based on a formal – organizational criteria, preferably based on labor contract, stipends, fellowships. This includes team members across the whole available staff categories from lab assistants to team leaders. For research groups operating in private institutions, membership will be defined through some form of paid working relationship (which also includes for example more or less permanent freelance collaboration with team members). Making team membership dependent upon an organizational criteria also emphasizes and acknowledges the role of the organizational context for team work. Endorsing furthermore the importance of additional team-external ties, a "collaboration" will be defined as any type of exchange of ideas, the provision of feedback or information that happens with non-team members. This corresponds with Bozeman & Boardman (2014, p. 2) suggestion that the provision of resources alone does *not* qualify someone as collaborator; however the simple discussion of an idea – however fleetingly – does qualify as a collaboration and should be considered. Research projects such as European ones, are conceived as research collaborations but not as teams.

4.2. Team Performance and Gender Diversity

A major challenge for the current research project consists of isolating the influence of "gender diversity" among all the other factors that affect team performance. Hence, the proposed integrated framework delineates the positive and negative consequences of gender diversity on teams but also specifies a host of other variables affecting research performance that are only indirectly or not at all related to gender.

Put bluntly, team performance is dependent upon available resources and a team's capacity to use and transform these input resources into research output. A gender perspective becomes relevant where issues such as the unequal distribution and access to resources (e.g. funding) are considered. It also extends to a wider analysis of the various ways science, technology and innovation are gendered. More specifically, however, our primary concern focuses on how gender bias and stereotypes affect research groups in their ability to recognize and utilize the expertise available among their members. In what follows, the main issues relevant from a gender perspective are briefly summarized.

4.2.1. Gendered Status Beliefs

First, expectation states theory shows that gender stereotypes can affect how competent people appear to others. The underlying reason for this misjudgment is that gender stereotypes imply status hierarchies: being white and a man is associated with a higher status than being a woman or belonging to a minority (DiTomaso, Post, Smith, Farris, & Cordero, 2007; Haines, Deaux, & Lofaro, 2016). Especially in contexts where gender is salient – such as meetings between strangers or for strongly gender-typed tasks – gendered status beliefs project competency expectations upon team members that do not necessarily match their actual task expertise. Research has also shown that these competency expectations are contingent upon a series of team- and organizational context variables. Diffuse status cues such as gender or ethnicity are more salient in more centralized and shorter-tenured groups

whereas specific status cues such as task relevant expertise are more prevalent in decentralized and longer-tenured teams (Bunderson, 2003). In addition to tenure and centralization of power, the organizational context has been identified as affecting information sharing within groups (see below).

4.2.2. Homophily & Social Affinity

Second, gender can also affect team performance through *social affinity* ties (Joshi & Knight, 2015). Social identity theory highlights how insider/outsider categorization affects work groups, for example through "group think", "homophily" or network closure in general. Their research demonstrates on the one hand that attributes that are directly relevant to the task such as education and tenure are more relevant for deference than surface level attributes such as gender or ethnicity. However, social gender identity and the concomitant categorization of group "insider" vs. "outsider" affects team performance due to social affinity ties. Homophily based on educational level, gender and ethnicity affect who defers to whom with detrimental consequences for performance: "Teams that rely on task-based deference perform better than those that rely on social affinity-based deference" (Joshi & Knight, 2015, p. 79) In contrast, Kegen (2015) observed no homophily relations among women scientists of the German Excellence Initiative. What this suggests is that research teams that manage to share their expertise less on secondary, diffuse status cues such as gender perceptions or other social affinity attributes and focus instead on task-relevant cues such as task-contribution can pool their knowledge and skills more adequately and perform better.

4.2.3. Formal Power Relations

Third, *power* relations affect teams in various ways. Since women are under-represented at the top of the career ladder, vertical segregation potentially skews which information is voiced and shared within teams and the organization. Since women have usually less influence in group interaction than men (Mannix & Sauer, 2006), gendered hierarchies restrict women's contributions in important ways. Discussion in groups tends to focus on shared information that all members know rather than unshared and novel information (Carli, 2010). High status members participate more and are more dominant in groups, silencing lower status members. This observation is different from the gendered status expectations – which also imply hierarchical relations. Independent of specific gendered status expectations that condition the recognition of expertise, vertical segregation shapes the sharing of information in teams no matter who is on top and at the bottom of the hierarchy. Since women in this case are over-represented at the lower ranks, we expect that the unequal distribution of power will affect team performance.

4.2.4. Social sensitivity

Fourth, research on social sensitivity and empathy in groups has shown that the proportion of women positively affects turn-taking in groups and by that performance. The proportion

of women in teams facilitates information exchange among its members and contributes to a favorable team climate. Overall, the issue of social sensitivity has to be seen in relation to emotional "management" within teams. Groups that are able to provide a "healthy" socio-emotional environment are better placed to develop cognitive capacities for complex research tasks on top of it. Transformational leadership, which could be considered a more feminine style in that it is more supportive, less vertical and more empowering has been shown to increase team learning and coordination. That is to say, whereas gender stereotypes subvert optimal expertise recognition and team learning when left unquestioned, certain affective and communication oriented sex/gender differences catalyze group processes.

4.2.5. Access to Resources & Opportunities

Fifth, gender affects the availability of resources to teams, less in terms of expertise but in terms of funding available or supporting research collaborations. There is abundant evidence on the funding gap for women. There is also evidence on a gap between women and men in terms of *international* research collaboration. Women and men show similar collaboration patterns on the national levels, however not on the international. These factors condition the overall team performance, since both funding as well as collaboration has been linked to increase research performance. The productive gap of women in scientific publications needs to be considered in using gender sensitive assessment indicators.

4.2.6. Organizational and Disciplinary Context

Sixth, disciplines and organizational contexts are gendered – and affect working conditions of individuals to the better or the worse which directly strengthen or weaken individual capacity, motivation and overall team performance. The influence of the disciplinary context has also been shown in relation to gendered status beliefs, diversity and gender equality policies can make a difference by fostering "diversity beliefs" which influence the recognition and value of diverse team members within organizations (Hentschel, Shemla, Wegge, & Kearney, 2013). Where women are a minority – in their organization or discipline – they find it harder to get their expertise and contributions recognized and valued.

4.2.7. Additional factors affecting team performance

In addition to these gender factors we know of several other factors that are related to team performance but where it is difficult to establish any relation to either sex or gender differences. Thus, "*transactive memory*" has been linked to affecting group performance; however, apart from status expectations regarding expertise, no further gender related factors have been discerned that affect it.

An area that has received relatively little attention from a gender perspective are *temporal processes*. Although research increasingly explores the temporal dimension of group processes, it is not clear to which degree gender differences play a role. Thus, one could speculate that the restricted availability of women due to higher care responsibilities affects

group dynamics differently. This could also affect *kairos*, the mutual folding of creative and innovative processes that enables individuals to attend and wait for creative ideas and then implement them.

4.3. Implications and Operationalizations for Case Studies

The following section will examine the implications of the literature review for the case studies. Our case studies with medical engineering on the one hand and transport/mobility research groups on the other will offer the opportunity to gain a better understanding of these two disciplinary fields and its actors. We aim at deeper insights how individuals interact on the team level and how gendered hierarchies and bias conditions team interaction and the sharing of information. The case studies will develop a methodology for studying how gender stereotypes and behavioral cues in interaction condition information sharing; they provide the foundation within GEDII and beyond to relate micro-level *gendered* group dynamics to research performance.

The case studies deploy a mix of research methods, including a) desktop research on the organizational and disciplinary environment, b) semi-structured interviews with representatives of the organization, the team leader and team members, c) sociometric badges, and d) short survey measures. The detailed operationalization for each method is developed in D2.1 Fieldwork Manual.

4.3.1. Desktop Research

Compiles all available documents (online and offline) to construct the profile of the organization including the disciplinary and/or professional context of the team participating in the case study.

4.3.2. Interviews

Semi-structured interviews will be held with two members of the management / HR level, the team leader and two team members. The interview guides are structured according to thematic blocks of questions, soliciting information regarding:

- personal and professional background of the interviewee, current position within the team.
- the team characteristics including its composition, key events in its history, strength and weaknesses, team climate and organization of team work itself.
- gender and gender equality, including observed differences between men and women members, opinions about advantages and disadvantages for women vs. men in science, the organization and the team.
- Scientific excellence including quality indicators and disciplinary differences.

Interviews with the team leader will ask additional questions regarding team leadership and the trajectory of the group. Interviews with management level and Human Resources will tackle issues regarding the overall organization of the research institute and the presence (or not) of gender- and diversity measures. A detailed policy/activities sheet forms part of the interview guidelines for the HR members.

4.3.3. Survey measures for case studies

A short questionnaire will be administered to each team member. This serves two broad goals: first, to compile data about the team composition on the one hand and to interpret the sociometric data in the light of established measurement scales.

First, team composition will be measured according to the following items:

- Gender composition of teams
- Age
- Highest educational qualification (pre-coded)
- Experience / Tenure: overall experience in the field.
- Team tenure: duration in month of team membership
- Current position: (Phd student, lab assistant, group leader, researcher, etc.)
- Disciplinary background (In case of doubt: of PhD if available) based on the Frascatti classification.

Second, and most importantly, the gender identity and bias of team members need to be established, since it provides the primary control variable. As section 1.5 will describe, many measurement scales for sexism exist. Despite the variety of measurement scales, there is not a specific scale that would fit our needs 100%. The “Neosexim” scale (Tougas, Brown, Beaton, & Joly, 1995) is relatively short (11 items) and yields to a large degree the results of the “Modern Sexism Scale” (Swim, Aikin, Hall, & Hunter, 1995) except for its added emphasis on reactions to affirmative action. Importantly, the Neosexism Scale correlates with pro-male bias in evaluation women's and men's competence, an important bias that affects information sharing in groups as previously described. However, these measures are quite old and not specific to gender issues in science.

Third, in order to arrive at a more reliable judgment regarding the relative dominance of certain team members during interactions and group discussions, the relative “weight” of each member during conversations needs to be contrasted with their personality traits. Personality traits have been shown to influence status expectations, the propensity to trust others and be trusted, or to speak out during group discussions. More extrovert people enjoy higher status, are more central in a network and are more trusted (Gloor et al., 2011, 2012). Importantly, as Gloor et al. has shown, personality traits correlate significantly with network centrality measures captured by sociometric badges. The Ten-Item Personality Iventory (TIPI) takes one minute (Gosling, Rentfrow, & Swann, 2003) and has been recommended to assess personality traits of persons in research settings where personality is not the primary target.

Fourth, we need a reliable measure of the perceived status of each member within the group. Joshi & Knight (2015) for example measure the perception of "task contributions" with an item "This person makes valuable work-related contributions to the lab". Since status in the sense of Ridgeway is conferred to individuals based on expectations to contribute to group goals, this measure provides a potential bridge that links gendered status expectations to task contributions. The perceived status could follow a round-robin format (all team members judge all other team members). This requires to reduce each scale to a single item in order to keep the burden of participants manageable.

4.3.4. Sociometric badges

A primary objective of the case study is to assess the feasibility of using sociometric data for detecting the influence of gender on team work. To our knowledge, sociometric badges have not been used for conducting research on gender. This implies to explore to which degree patterns detected with the sociometric data correlate with relevant established measurement scales or not (see previous section). The sociometric data need to be interpreted in the light of these standard measures in order to assess the reliability for what they are supposed to measure, namely the influence of gender-related constructs on team interaction and information sharing. Gendered status beliefs conditions interaction in important ways. It affects:

- the likelihood that people speak up in a situation; frequency of participation
- whether others attend them if they speak up; frequency of interruptions
- how their suggestions are evaluated and recognized
- to which extend they can claim credit for their suggestions
- whether they become influential and eventually rise to the top as leaders
- extend to which they and others are willing to infer high ability upon themselves (which in turn influences how they perform certain tasks)
- extend to which people accept men or women as leaders and how they are received when they are in power

Some of these dimensions of interaction can be directly inferred from sociometric data:

Conversational turn-taking provides insights into the relative contributions of members during structured meetings. Its analysis suggests relative dominance of members during discussions. This includes successful- and unsuccessful interruption attempts among members. As the literature review has suggested, gender status beliefs have been shown to influence the relative "on-air" time of men and women in teams. Apart from gender status beliefs, personality traits are an important dimension to bear in mind when considering conversational turn-taking. Turn-taking will give us also insights into "social sensitivity" of team members.

Physical proximity including face-to-face conversations data provides insights into the information flow among team members. This measure potentially indicates social affinity and the strength of network ties among team members – which has been shown to affect

information sharing and group performance. The analysis includes standard network measures such as degree centrality and betweeness centrality. A contribution index of each actor can be calculated based on the propensity of actors to be a looked at versus looking at others (Gloor, Laubacher, Dynes, & Zhao, 2003). In addition, network can be analyzed for "homophily" along various dimensions, including gender, tenure, educational level, status.

Body movement – "honest signals" detection provides insights into the "intensity" with which team members communicate. According to Pentland (2008) different types of "Honest Signals" can be distinguished:

1. "influence", which means the extent to which someone causes the other person's pattern of speaking to match their own pattern
2. "mimicry", the automatic copying of one person by another during a conversation
3. "activity", represented by the energy and the time spent in a conversation
4. "consistency", which is measured through low variability in the speech signal.

Honest signals could indicate how convincing certain actors achieve to be. Theory predicts that high status members are more visible and convincing in group interactions. "Honest signals" might also relate to "social sensitivity" of team members, i.e. how well their body language is synchronized. As described, affective empathy involves the activation of shared motor representations by which individuals understand each other's emotions. However, no prior research indicates how gender processes might influence these "honest signals."

Speech features – such as volume and tone of voice may provide insights into the hierarchy within the group. As recent research has shown, hierarchy "affects speakers' acoustics and the acoustics of perceivers used to perceive hierarchical rank of speakers" (Ko, Sadler, & Galinsky, 2015, p. 11). Linneman (2012) has furthermore argued, that "uptalk" - when the typical intonation of questions is used for expressing normal statements – is gendered. Women seem to use uptalk more frequently, for signaling uncertainty or as a compensatory strategy to assuage conflicting gender roles as the authors argue (ibid.). Speech features therefore might give insights in the gendered use of tentative, deferential speech.

4.4. Implications and Operationalizations for GDI / Survey

The following section gives an overview of potential variables and scales that could be deployed during the cross-national survey. Since the survey will be developed from month 12 onward, the paragraphs are thought to prepare this upcoming work rather than delivering a final questionnaire. It recollects measurement scales without making final decisions which ones to use.

Taking into account the recent meta-analytical findings (Hülsheger, Anderson, & Salgado, 2009), in order to assess the "performance" of teams the following list of key variables should to be considered:

- Team process variables, including external and internal communication, vision, support for innovation, task orientation, and cohesion are especially conducive to innovation.

- Team input variables, although displaying smaller mean effect sizes are somewhat conducive to team innovation: goal interdependence, job-relevant diversity, team size
- Moderated by measurement method and level: self-ratings show higher effects; team (vs. individual level) level show higher effects.

In more detail:

Team composition – within team composition, several sub-dimensions need to be integrated.

- **Team size** has been shown to have a positive relationship with innovation. Team size is relatively easy to assess, either through our own field-preparatory work or during the questionnaire / case studies directly.
- **Gender composition** of team
- **Highest educational level** of team members.
- **Overall Experience** (tenure) in the field
- **Team tenure** – average length of time of members with team. This is also important in relation to the salience of gendered role expectations. It is plausible that gendered role expectations recede into the background in teams that have been working sufficiently long together in order to perceive task-competence.

Network of external collaborators – has been shown to positively influence to a certain degree the innovation capacity. This could be formalized by contrasting team members listing with co-author papers or patents. It could also be based upon a questionnaire item inquiring about the number of external (non-team member) collaborating contacts. The network of external collaborators can then be contrasted with team-internal member diversity and collaboration intensity. From a gender perspective, women have less international collaborations.

Resource allocation – on the individual level this might include different types of resources such as for example the time available for research activities vs. teaching or administrative responsibilities. Importantly, resource allocation is highly gendered, for example in terms of funding but also available international networks. On the team and organizational level this includes previous funding and research experience of the university but also, multi-university collaborations and top-tier universities are important predictors of research productivity and impact (Han, Han, & Brass, 2014).

Team climate – this measure integrates various important dimensions regarding the quality and nature of the social relations – including their psychological consequences – within the team. The composite construct focuses primarily on those aspects of the team environment that foster innovation. The original measure as proposed by Anderson & West (1998) consists of 38 items measuring four basic dimension:

- **shared vision**: common understanding of team objectives, display high commitment to goals. In part this could be related to “team cognition” - not just a shared vision of

the objectives but also knowledge about "who knows what" and shared mental models of problems and issues. Transactive memory scale, see (Lewis, 2003; West, 1990)

- **participation safety** as previously argued which is especially important to voice deviant solutions and unheard of, new information within groups. Psychological safety important in recent study in Google (Duhigg, 2016). Related also to "group cohesion", i.e. the commitment and desire to maintain group membership. Team cohesion (see Ancona 1992, p.655)

- **task orientation**: also called **climate for excellence** and highest standards of performance. Task orientation implies "team reflexivity", i.e. mutual performance monitoring and feedback which is meant to lead to exploration and assessment of opposing opinions and their integration. Task orientation seems also to be equivalent to intrinsic motivation (Salas, Cooke, & Rosen, 2008)

- **support for innovation**: team members are receptive to new ideas and change. In part this is related to "diversity beliefs". In teams that have stronger beliefs in the positive effects of group diversity, members have a greater disposition to take the perspective of others into account and thus integrate diverse viewpoints.

Recently, Lee et al., (2015) have proposed the **Team-Descriptive-Index** scale that situates teams on a rigorous scale along the three fundamental Team Scaling dimensions, namely (1) skill differentiation, (2) authority differentiation, and (3) temporal stability as proposed by Hollenbeck, Beersma, & Schouten (2012). The authors developed and tested a long questionnaire (21 items) and short form (three items). Although the scales produce reliable results, questions seem to be answered indirectly as when one person answer subjectively regarding the overall team. For example, temporal stability is assessed on scale from 1 to 9 in a descriptive way.[1] However, such information could be retrieved much more reliably by asking each team member for their past and future membership-time.

Leadership

There exist many different scales for measuring the leadership style within teams. For a start, the item developed by Lee et al. (2015) just mentioned on "authoritative differentiation" could be used. But there exist many other, more specific scales. For example the *Multifactor Leadership Questionnaire* (MLQ) developed by Bruce Avolio and Bernard Bass examines the full range of leadership by measuring the rates of laissez-faire, transactional, and transformational leadership styles used by a leader (Avolio & Bass, 2004). However, for a research project where leadership is not the main objective, the 45 items of the MLQ are too large.

1 For example: "This team was likely to have a future together (one to two years), and had a history of working together (one to two years). New members were added to the team occasionally, but for the most part, it was the same people all the time."

4.4.1. Gender stereotypes / gender role expectations

As previously argued, we stipulate that gendered status expectations influence information sharing and elaboration in groups with consequences for team performance. Hence, a reliable measure regarding the prevalence of gender stereotypes and bias will be required.

Research on gender stereotypes in social psychology has developed several, widely used measurement scales (for reviews see Beere, 1990; Fiske & North, 2015; McHugh & Frieze, 1997). Among the most widely used are the *Modern Sexism Scale* (Swim et al., 1995), the *Neosexism Scale* (Tougas et al., 1995), the *Ambivalent Sexism Inventory* (Glick & Fiske, 1996). More specifically related to the context of women and science, the *Women in Science Scale* (Erb & Smith, 1984; Owen et al., 2007) exists which measures attitudes of adolescents towards women in science. These self-report instruments on explicit attitudes towards gender are complemented by implicit tests measuring reaction times to presented stimulus such as for example in the *Implicit Association Test* (IAT) (Nosek et al., 2007).

Work by Rashotte & Webster presented a measurement scale explicitly focused on Gender Status Beliefs as developed by Ridgeway (see Rashotte & Webster, 2005). This scale takes into account that although general sex stereotyping has changed little, there is certainly a greater awareness and caution displaying these gender stereotypes. It consists of two parts: the first part assesses status significance of gender while the second part assesses tendencies to respond in socially desirable ways to behavior related to gender. However, the scales and test appear quite cumbersome and artificial in relation to the teams and their work; participants are asked to judge photographs of anonymous women and men and how “capable” they appear. They are also asked to judge how well they think these persons perform different gender typed tasks. For the current research it seems more practical to inquire about gender stereotypes more directly and judge expertise in relation to the actual members of the group. So far, it appears that no other research has used these scales since 2005 which also makes it difficult to assess its reliability in more detail.[2]

Recently, a new publication regarding the “separate spheres” model regarding women and men has been proposed (Miller & Borgida, 2016). It measures the beliefs in stereotypes regarding the “separate sphere” to which men and women respectively pertain. The Separate Spheres Ideology (SSI) predicts attitudes regarding workplace flexibility accommodations, income distribution within families between women and men partners, distribution of labor between work and family spheres, and discriminatory workplace behaviors. The scale targets specifically work-life balance issues in relation to gender and measures reliably and consistently individuals attitudes in this respect. However, it is less related to gendered role and status expectations regarding (research) team work.

A reliable scale targeting more specifically the issue of women in science is the “Women in Science Scale” (Owen et al., 2007). It measures attitudes toward women in science. A re-evaluation study of the original WiSS scale allowed to shorten it from 27 to 14 items with two underlying factors, namely Equality and Sexism. The advantage of this scale consists of an integration of general beliefs about equality with sexist attitudes, specifically in the context of science. However, as Owen et al. state, this 14-item scale would need further testing to

2 As confirmed by email correspondence with the first author of the scale.

consolidate its reliability.

Joshi & Knight (2015) use a measure for Gender Identification, proposed by (Derks, van Laar, & Ellemers, 2009). This is an 8 item scale which measures the importance of an individual's social group membership to his or her identity. The measure makes sense within the design of Joshi's study where social affinity was an important predictor variable for deference. However, within the context of the present study, we will not assess gender identification of individual members in relation to their deference and it is likely that this scale is of limited use.

Overall, the best measurement scale would allow us to detect (gender) bias in competency expectations regarding highly skilled (scientific) work in teams. The “Neosexism” scale appears to be the most reliable candidate while correlating precisely with bias in evaluating women's and men's competence.

4.4.2. Attitudes towards Diversity

Positive attitudes towards diversity in teams act as precondition for capitalizing its benefits. Diversity climate surveys often capture various dimensions: individual, group and organization wide profiles that favor diversity.

“Diversity beliefs” capture the attitudes towards team- and organizational diversity that makes a difference for team performance (Hentschel et al., 2013; Pirola-Merlo, Härtel, Mann, & Hirst, 2002; van Dick, van Knippenberg, Hagele, Guillaume, & Brodbeck, 2008). On an individual level, it captures people's beliefs about the value of working in diverse teams. In teams that believe stronger into the positive effects of group diversity, members have a greater disposition to take the perspective of others and thus integrate diverse viewpoints. However, in much of the research on diversity, gender is just one dimension among other types. Hentschel et al. for example measures “diversity beliefs” with three items without specifying the dimension of diversity at all. Respondents indicate on a 7-point Likerty-type scale their opinion on: “Teams are more effective when they include people who are different from on another”, “I prefer to work with people who are different from me”, and “Teams perform better when they include people who are different from one another”. Which type of diversity people thus express their views about is not clear.

Several scales that measure attitudes towards diversity exist. See for example (Hentschel et al., 2013; Kossek & Zonia, 1993; van Dick et al., 2008; van Knippenberg, van Ginkel, & Homan, 2013).

4.4.3. Gender Diversity Policy / Human Resource Diversity Policies

On the organizational level it has been shown that company policies on diversity in general and gender equality in particular have a positive impact.

The simplest check could be to query the presence of a gender quality plan or discrimination policy. More extension checks could inquire individual items regarding:

- Recruitment and progression practices
- Equality based career structures
- Pay and reward management
- Working time arrangements
- Training and continuing professional development
- Formal equality and diversity measures
- Gender in research content policy

For overviews of diversity climate measures see (Goyal & Shrivastava, 2013) and (Burkard, Boticki, & Madson, 2002).

4.4.4. Personality Traits

An important factor for team interaction and information sharing are personality traits. Survey measures for personality traits are available. The most common is the Big Five personality dimensions (Goldberg, 1990, 1993). Shorter versions which reduce the original 44 item scales to 10 items are available as well (Gosling et al., 2003; Rammstedt & John, 2007). Rammstedt & John's 10 item scale is translated in German and English. For an application of the NEO-FFI in relation to badges see (Gloor et al., 2012).

The recommended 10 item scale for measuring the personality trait asks respondents to rate on a scale ranging from “1” (Disagree strongly) – to “7” (Agree strongly) the following items:

1. Extraverted, enthusiastic.
2. Critical, quarrelsome.
3. Dependable, self-disciplined.
4. Anxious, easily upset.
5. Open to new experiences, complex.
6. Reserved, quiet.
7. Sympathetic, warm.
8. Disorganized, careless.
9. Calm, emotionally stable
10. Conventional, uncreative.

The scale takes less then one minute to complete (see Gosling et al., 2003).

4.5. Overview of Research Assessment Metrics

The project GEDII needs to take into consideration metrics that are valuable in terms of quantity and quality. They build the basis of the index to be proposed. Furthermore, indicators need to be robust, reliable and possible to measure. And, for the purpose of the project scope, metrics need to take gender aspects into account or should be at least gender sensitive if possible. Gender sensitive metrics can appear as gender inclusive and gender specific indicators. Gender inclusive indicators measure comparable information on women´s

and men´s figures, while gender specific indicators measure progress or the impact that are special related to women, such as the She Figures, which are published since 2003.[3]

Indicator		What to measure	Aggregation level	Gender (sensitive) aspects
Research publications and other outputs				
Quantity / Productivity	Books, book chapters, journal articles, conference publications	Total count of numbers	Individual, Team, Organization	Broken down by gender / composition
	Technological output such as patents	Total count of numbers, broken down by gender and team composition respectively	Individual, Team, Organization	Broken down by gender / composition
	Possibility to distinguish certain types of output	International publications, mixed author / inventor team, output directed towards society, Output socio-economic benefit	Individual, Team, Organization	
Quality / scientific, scholarly impact				
	Citations	Total citation counts	Individual, Team	Broken down by gender / composition
		Normalized citation counts	Individual, Team	Broken down by gender / composition
		Journal impact factor	Individual, Team	
		h-factor	Individual, Team	
		(Gender sensitive factor?)	Individual, Team	
Top ranked journal publications	To be checked if feasible in the areas of medical engineering or transport	Total number	Individual, Team	Broken down by gender / composition

3 *She Figures* are updated and released every 3 years and draws a picture of the progress towards gender equality in research & innovation in the EU. The newest edition is *She Figures 2015*.

Indicator		What to measure	Aggregation level	Gender (sensitive) aspects
Altmetrics	Altmetrics indices (has to be specified/ confirmed)	Downloads to reference, Blogs, Twitter	Individual, Team	Broken down by gender / composition
Peer Esteem				
		Key contributions to national and/or international conferences	Individual, Team, Organization	Broken down by gender / composition
		Editorial or refereeing exercises for national / international Journals / Publishers, Curator, Jury memberships Appointment to advisory groups, research evaluation committees	Individual, Team, Organization	Broken down by gender / composition
Academic Teaching and Research				
	PhD completion	Numbers of graduates	Individual, Team, Organization	Broken down by gender / composition
	Research Student load	Numbers per staff	Individual, Team, Organization	m/f students allocation by gender / composition
	Academic Teaching load	Hours per staff	Individual, Team, Organization	Broken down by gender / composition
	Grants won	Number of funded projects, research income per year	Individual, Team, Organization	Funding success rate by gender
	Research Content		Individual, Team, Organization	Gender Dimension

Indicator	What to measure	Aggregation level	Gender (sensitive) aspects
Collaboration activities			
Co-publications	Total numbers	Individual, Team level	Broken down by gender / composition
Research income by external sources	Per year	Individual, Team level, Organization	Broken down by gender / composition
Formal collaboration with other research teams, researchers	Formal, informal, national, international, one-time-collaboration, longtime collaboration	Individual, Team level, Organization	Broken down by gender / composition
Interdisciplinarity		Individual, Team level	Broken down by gender / composition
Organizational			
Women in leadership	Total number / share of women as team leader	Team level, Organization	Gender specific / gender composition
	Total number / share of women on boards	Organization	Gender specific / gender composition
Gender equality plans	Existence of equality plans	Organization	
	Knowledge, adoption of equality plans	Individual, Team level	Broken down by gender / composition
End user esteem – wider society (Societal Impact measure)			
Interaction with stakeholders and societal institutions or functions	Appointments to organization, committees, councils, involvement in policy, expert services	Individual, Team level, Organization	Broken down by gender / composition

Indicator	What to measure	Aggregation level	Gender (sensitive) aspects
	(similar to peer esteem, however with society as target)		

4.5.1. Research publications and other outputs

The total count of numbers of publications and other output as a quantitative measure gives a first overview about research activity, the most important disciplines and societal impact by clustering the types of publications. The share / composition of gender in author / inventor teams can be measured.

4.5.2. Indicators of quality based on citations

Citations are the basis of a range of indicators, which represent acknowledged indicators of quality. The most common indicators are going to be used such as the JIF, the h-index in a normalized form, taking into consideration the limitations of these indicators as described further above. The share / composition of gender in author / inventor teams can be measured.

4.5.3. Top ranked journal publications

This is taken on the basis of the journal impact factor provided by Thomson Reuters. Advantage of this indicator is the widely accepted concept that published articles in the top journals are indeed a reference of quality. The High Impact Journals are available for each research field, however the rankings of journals in social science and humanities are in conflict with ranking of experts (European Commission, 2010, p. 73). As for that, the feasibility for the sectors transport and medical engineering needs to be assessed. The share / composition of gender in author / inventor teams can be measured.

4.5.4. Altmetrics

As a new metrics that is relatively easily available we will consider using altmetrics score. The chosen altmetrics are going to be an indicator of societal impact. Almetrics is based the impact of published articles in social networks. Hence, the share / composition of gender in author / inventor teams can be measured through the article, provided that gender allocation is possible.

4.5.5. Peer esteem

Peer esteem measures how research and researchers are acknowledged firstly in the scientific community, but it also measures end-user esteem in a wider sense since the

indicators also reach out to the wider society. However, the diverse indicators of peer esteem are difficult to compare across disciplines. Also the value of certain indicators is difficult to estimate. For example, for some prizes or awards the value has been consented upon such as Nobel Prizes. The share / composition of gender in author / inventor teams can be measured, providing gender allocation is possible.

4.5.6. Teaching and research Activities

Teaching activities are used as indicator from a gender perspective since these activities describe a quality dimension of research activities more linked to women. The share / composition of gender in author / inventor teams can be measured, providing gender allocation is possible.

4.5.7. Collaboration Activities

Collaborative activities can be measured on different levels, apart from counting total numbers to specific individuals or teams, it can be distinguished for example geographically, i.e. regional, national, international or by discipline, i.e. interdisciplinary collaboration and/or by gender.[4] Ancona & Caldwell (1992) proposed in their work to distinguish between different type of boundary spanning activities, including activities that represent the team to "others" vertically within the organizational hierarchy or horizontally to other teams or end-users. Collaboration activities in this sense are also closely aligned to the issue of social impact (see next section), in that collaboration with NGOs, associations or other external, non-scientific groups promotes a closer fit between research and the wider society.

4.5.8. End User Esteem & Social Impact

In accordance with the policy mission statement of the EU "Science with and for society", the measurement of end-user esteem is an important field of assessment exercises as a reflection of the societal benefit of science. The incorporation of a gender dimension into research content can be assessed easily through a specific questionnaire item. However, this does not indicate anything regarding the "real" impact this might have. Especially from a gender perspective, other forms of "counting" more social impact should be contemplated. As Mountz et al., (2015) suggest, instead of counting "articles published or grants applied for, what if we accounted for thank you notes received, friendships formed, collaborations forged?" What about ways to make visible and value those forms of care within the university and research that might not result in direct publications but provide and sustain the very social fabric in which intellectual and creative work can strive and prosper?

4 In this context be referred to Larivière et al who state that "female researchers could find themselves excluded from an extended network (a basic network being essentially academic in nature) of increasing importance in the current context of scientific production. This implies a need for more scholarly work to understand the participation of women in the 'third' (societal, community-based or entrepreneurial) mission of universities, which tends to be increasingly important in comparison to—if not in direct competition with—the fundamental missions of a university: teaching and research." (Larivière et al., 2011, p. 495)

4.5.9. Organizational Indicators

Women in leading positions, such as team leader on team level or women on decision boards are indicators to be determined. As for gender politics, the existence of gender equality plans, knowledge and enforcement of them are also important to take into consideration.

Bibliography

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the Boundary: External Activity and Performance in Organizational Teams. *Administrative Science Quarterly*, *37*(4), 634–665.

Anderson, N., & West, M. A. (1998, May 1). Measuring climate for work group innovation: development and validation of the team climate inventory. *Journal of Organizational Behavior*.

Avolio, B. J., & Bass, B. (2004). *Multifactor Leadership Questionnaire*. Redwood City Calif.: Mind Garden.

Beere, C. A. (1990). *Gender Roles: A Handbook of Tests and Measures*. New York: Greenwood Press.

Bozeman, B., & Boardman, C. (2014). *Research collaboration and team science : a state-of-the-art review and agenda*. Cham (Switzerland): Springer.

Bunderson, J. S. (2003). Recognizing and utilizing expertise in work groups: A status characteristics perspective. *Administrative Science Quarterly*, *48*(4), 557–591.

Burkard, A. W., Boticki, M. A., & Madson, M. B. (2002). Workplace Discrimination, Prejudice, and Diversity Measurement: A Review of Instrumentation. *Journal of Career Assessment*, *10*(3), 343–361. http://doi.org/10.1177/10672702010003005

Carli, L. L. (2010). Gender and Group Behavior. In *Handbook of Gender Research in Psychology* (p. 132). Springer.

Derks, B., van Laar, C., & Ellemers, N. (2009). Working for the self or working for the group: how self- versus group affirmation affects collective behavior in low-status groups. *Journal of Personality and Social Psychology*, *96*(1), 183–202. http://doi.org/10.1037/a0013068

DiTomaso, N., Post, C., Smith, D. R., Farris, G. F., & Cordero, R. (2007). Effects of Structural Position on Allocation and Evaluation Decisions for Scientists and Engineers in Industrial R&D. *Administrative Science Quarterly*, *52*(2), 175–207. http://doi.org/10.2189/asqu.52.2.175

Duhigg, C. (2016, February 25). What Google Learned From Its Quest to Build the Perfect Team. *The New York Times Magazine*.

Erb, T. O., & Smith, W. S. (1984). Validation of the attitude toward women in science scale for early adolescents. *Journal of Research in Science Teaching*, *21*(4), 391–397. http://doi.org/10.1002/tea.3660210407

Fiske, S. T., & North, M. S. (2015). Measures of Stereotyping and Prejudice: Barometers of Bias. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of Personality and Social Psychological Constructs* (pp. 684–718). Amsterdam et al: Elsevier. http://doi.org/10.1016/B978-0-12-386915-9.00024-3

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512.

Gloor, P. A., Fischbach, K., Fuehres, H., Lassenius, C., Niinimäki, T., Olguin, D. O., … Putzke, J. (2011). Towards “Honest Signals” of Creativity – Identifying Personality Characteristics Through Microscopic Social Network Analysis. *Procedia - Social and Behavioral Sciences*, *26*, 166–179. http://doi.org/10.1016/j.sbspro.2011.10.573

Gloor, P. A., Grippa, F., Putzke, J., Lassenius, C., Fuehres, H., Fischbach, K., & Schoder, D. (2012). Measuring Social Capital in Creative Teams Through Sociometric Sensors. *International Journal of Organisational Design and Engineering*, *2*(4), 380–401.

Gloor, P. A., Laubacher, R., Dynes, S. B. C., & Zhao, Y. (2003). Visualization of Communication Patterns in Collaborative Innovation Networks - Analysis of Some W3C Working Groups. *Proceedings of the Twelfth International Conference on Information and Knowledge Management - CIKM ’03*, 56. http://doi.org/10.1145/956873.956875

Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. http://doi.org/10.1016/S0092-6566(03)00046-1

Goyal, S., & Shrivastava, S. (2013). Organizational Diversity Climate: Review of Models and Measurement. *Journal of Business Management & Social Sciences Research*, *2*(5).

Haines, E. L., Deaux, K., & Lofaro, N. (2016). The Times They Are a-Changing … or Are They Not? A Comparison of Gender Stereotypes, 1983-2014. *Psychology of Women Quarterly*, 0361684316634081–. http://doi.org/10.1177/0361684316634081

Han, J., Han, J., & Brass, D. J. (2014). Human capital diversity in the creation of social capital for team creativity. *Journal of Organizational Behavior*, *35*(1), 54–71. http://doi.org/10.1002/job.1853

Hentschel, T., Shemla, M., Wegge, J., & Kearney, E. (2013). Perceived Diversity and Team Functioning: The Role of Diversity Beliefs and Affect. *Small Group Research*, *44*(1), 33–

61. http://doi.org/10.1177/1046496412470725

Hollenbeck, J. R., Beersma, B., & Schouten, M. E. (2012). Beyond Team Types and Taxonomies: A Dimensional Scaling Conceptualization for Team Description. *Academy of Management Review*, *37*(1), 82–106. http://doi.org/10.5465/amr.2010.0181

Hülsheger, U. R., Anderson, N., & Salgado, J. F. (2009). Team-Level Predictors of Innovation at Work: A Comprehensive Meta-Analysis Spanning Three Decades of Research. *Journal of Applied Psychology*, *94*(5), 1128–1145.

Joshi, A., & Knight, A. P. (2015). Who Defers to Whom and Why? Dual Pathways Linking Demographic Differences and Dyadic Deference to Team Effectiveness. *Academy of Management Journal*, *58*(1), 59–84. http://doi.org/10.5465/amj.2013.0718

Kegen, N. V. (2015). Cohesive subgroups in academic networks: unveiling clique integration of top-level female and male researchers. *Scientometrics*, *103*(3), 897–922. http://doi.org/10.1007/s11192-015-1572-z

Ko, S. J., Sadler, M. S., & Galinsky, A. D. (2015). The sound of power: conveying and detecting hierarchical rank through voice. *Psychological Science*, *26*(1), 3–14. http://doi.org/10.1177/0956797614553009

Kossek, E. E., & Zonia, S. C. (1993). Assessing diversity climate: A field study of reactions to employer efforts to promote diversity. *Journal of Organizational Behavior*, *14*(1), 61–81. http://doi.org/10.1002/job.4030140107

Lee, S. M., Koopman, J., Hollenbeck, J. R., Wang, L. C., & Lanaj, K. (2015). The Team Descriptive Index (TDI): A Multidimensional Scaling Approach for Team Description. *Academy of Management Discoveries*, *1*(1), 91–116. http://doi.org/10.5465/amd.2013.0001

Lewis, K. (2003). Measuring transactive memory systems in the field: scale development and validation. *Journal of Applied Psychology*, *88*(4), 587–604.

Linneman, T. J. (2012). Gender in Jeopardy!: Intonation Variation on a Television Game Show. *Gender & Society*, *27*(1), 82–105. http://doi.org/10.1177/0891243212464905

Mannix, E. A., & Sauer, S. J. (2006). Status and Power in Organizational Group Research: Acknowledging the Pervasiveness of Hierarchy. *Advances in Group Processes*, *23*, 149–182. http://doi.org/10.1016/S0882-6145(06)23006-6

McHugh, M. C., & Frieze, I. H. (1997). The Measurement of Gender-role Attitudes. A Review and Commentary. *Psychology of Women Quarterly*, *21*(1), 1–16. http://doi.org/10.1111/j.1471-6402.1997.tb00097.x

Miller, A. L., & Borgida, E. (2016). The Separate Spheres Model of Gendered Inequality. *PloS One*, *11*(1). http://doi.org/10.1371/journal.pone.0147315

Mountz, A., Bonds, A., Mansfield, B., Loyd, J., Hyndman, J., Walton-Roberts, M., … Curran, W. (2015, August 18). For Slow Scholarship: A Feminist Politics of Resistance through Collective Action in the Neoliberal University. *ACME: An International E-Journal for Critical Geographies*.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., … Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36–88. http://doi.org/10.1080/10463280701489053

Owen, S. V., Toepperwein, M. A., Pruski, L. A., Blalock, C. L., Liu, Y., Marshall, C. E., & Lichtenstein, M. J. (2007). Psychometric reevaluation of the Women in Science Scale (WiSS). *Journal of Research in Science Teaching*, *44*(10), 1461–1478. http://doi.org/10.1002/tea.20187

Pentland, A. (2008). *Honest signals how they shape our world*. Cambridge, Mass. : MIT Press,.

Pirola-Merlo, A., Härtel, C., Mann, L., & Hirst, G. (2002). How leaders influence the impact of affective events on team climate and performance in R&D teams. *The Leadership Quarterly*, *13*(5), 561–581. http://doi.org/10.1016/S1048-9843(02)00144-3

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. http://doi.org/10.1016/j.jrp.2006.02.001

Rashotte, L. S., & Webster, M. (2005). Gender status beliefs. *Social Science Research*, *34*(3), 618–633. http://doi.org/10.1016/j.ssresearch.2004.05.004

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 540–547. http://doi.org/10.1518/001872008X288457

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and Racism: Old-Fashioned and Modern Prejudices. *Journal of Personality and Social Psychology*, *68*(2), 199–214.

Tougas, F., Brown, R., Beaton, A. M., & Joly, S. (1995). Neosexism: Plus Ca Change, Plus C'est Pareil. *Personality and Social Psychology Bulletin*, *21*(8), 842–849. http://doi.org/10.1177/0146167295218007

van Dick, R., van Knippenberg, D., Hagele, S., Guillaume, Y. R. F., & Brodbeck, F. C. (2008). Group diversity and group identification: The moderating role of diversity beliefs. *Human Relations*, *61*(10), 1463–1492. http://doi.org/10.1177/0018726708095711

van Knippenberg, D., van Ginkel, W. P., & Homan, A. C. (2013). Diversity mindsets and the performance of diverse teams. *Organizational Behavior and Human Decision Processes*, *121*(2), 183–193. http://doi.org/10.1016/j.obhdp.2013.03.003

West, M. A. (1990). The social psychology of innovation in groups. In M. A. West & J. L. Farr (Eds.), *Innovation and creativity at work: Psychological and organizational strategies* (pp. 309–333). Chichester, UK: John Wiley & Sons.