How to Form Groups?
Optimizing Group Formation of University Students

Inauguraldissertation

zur Erlangung des akademischen Grades

eines Dr. Phil.,

vorgelegt dem Fachbereich 02

Sozialwissenschaften, Medien und Sport

der Johannes Gutenberg-Universität Mainz

von

Adrienne Mara Aimée Müller,

aus Darmstadt

Mainz, 2023

Referentin: Prof. Dr. Margarete Imhof

Tag des Prüfungskolloquiums: 05.09.2023

Danksagung

Zuerst möchte ich mich bei Margarete Imhof und Henrik Bellhäuser für ihre Unterstützung, die motivierenden Gespräche und die inspirierenden Anregungen bedanken. Darüber hinaus geht ein besonderer Dank an meine Freundinnen und Freunde sowie an meine Eltern, die mich stets unterstützt haben und mich ermutigt und begleitet haben. Ebenfalls bedanke ich mich bei meinen Kolleginnen und Kollegen für unsere wertvolle gemeinsame Zeit in den letzten Jahren. Zu guter Letzt gilt mein Dank noch all denjenigen, die mich auf meinem Weg durch diese spannende Zeit begleitet haben.

Zusammenfassung

Kollaboratives Lernen in Gruppen gilt als äußerst effektive Methode zur Verbesserung der Lernergebnisse und zur Entwicklung wertvoller sozialer Kompetenzen. Gleichwohl birgt die Gruppenbildung einen komplexen Prozess in sich, der sich erheblich auf den Lernerfolg der betreffenden Gruppen auswirken kann. Folglich ist es wichtig zu verstehen, wie Lernende sinnvoll zusammengesetzt werden sollten.

Die Relevanz der vorliegenden Forschungsarbeit ergibt sich aus dem Spannungsverhältnis, dass Gruppenarbeit einerseits als Form und Ziel der universitären Lehre verstanden wird, andererseits aber in der Umsetzung häufig scheitert, ohne auf die dafür relevanten Faktoren Einfluss nehmen zu können. Lernende unterscheiden sich in verschiedenen Aspekten, die die Qualität und Quantität der Interaktionen untereinander beeinflussen. Die vorliegende Dissertation untersucht mit Hilfe algorithmischer Gruppenbildung mögliche relevante Kriterien für eine effektive Gruppenbildung im Rahmen von vier experimentellen Studien in Online-Gruppenarbeit (Studie 1 & 2) und Präsenz-Gruppenarbeit (Studie 3 & 4).

Studie 1 untersuchte experimentell die Ergebnisse von Gruppenbildung anhand der Varianz der Persönlichkeitsmerkmale Extraversion und Gewissenhaftigkeit realisiert in einem vierwöchigen Online-Kurs für angehende Studierende. Die Hypothese war, dass es von Vorteil für die Ergebnisse ist, wenn die Online-Gruppen hinsichtlich der Varianz der Extraversion heterogen und hinsichtlich der Varianz der Gewissenhaftigkeit homogen gebildet werden. Studie 2 ähnelte in der Methodik der Studie 1, hier variierte jedoch ein Gruppenbildungskriterium, sodass basierend auf der Varianz der Extraversion und des Vorwissens Online-Gruppen experimentell gebildet und untersucht wurden.

Zusammenfassend liefern Studie 1 und Studie 2 Hinweise darauf, dass Lernerfolg in Online-Gruppenarbeit eher vom Kursdesign als von individuellen Unterschieden abhängt.

Gleichzeitig ist es für die Studierenden oft schwierig, ihr lernen online in Gruppen zu organisieren, was in solchen Kontexten zu erheblichen Abbruchquoten führen kann.

In den Studien 3 und 4 wurden die Rahmenbedingungen verändert und die Strategien der Gruppen-bildung in Präsenzveranstaltungen untersucht. In Studie 3 wurde untersucht, ob die Heterogenität oder Homogenität der Persönlichkeits-eigenschaft Extraversion, wie sie unter den Gruppenmitgliedern verteilt ist, Auswirkungen auf Ergebnisse wie Zeitaufwand, Zufriedenheit und Leistung hat. Überraschenderweise berichteten Gruppen mit einer homogenen Verteilung der Extraversion über ein höheres Maß an Zufriedenheit verglichen mit heterogenen Gruppen. Studie 4 - methodisch Studie 3 folgend - erweitert die Aussagekraft der Ergebnisse, indem sie unterschiedliche Standorte in verschiedenen Bildungseinrichtungen für die weitere Erforschung nutzt. Die Ergebnisse der Studie zeigen, dass eine homogene Verteilung von Extraversion der Gruppenmitglieder signifikant den Erfolg der Gruppenarbeiten erklärt.

Zusammenfassend tragen die vorgestellten Studien zu einem besseren Verständnis bei, wie algorithmische Gruppenbildung anhand individueller Prädiktoren in verschiedenen Bereichen erfolgreich implementiert und evaluiert werden kann. Implikationen für Forschung und praktische Anwendungen werden aus den Ergebnissen der vier Studien abgeleitet.

Abstract

Collaborative learning through group work is considered a highly effective method for improving learning outcomes and developing valuable social skills. However, group formation is a complex process that can significantly impact the success of collaborative learning. For this reason, it is important to understand, how learners can be effectively grouped together, to ensure successful learning outcomes.

The relevance of the present research arises from the tension that group work is, on the one hand, understood as a form and goal of university teaching, and, on the other hand, often fails in its execution without being able to control the relevant factors. Learners differ in various aspects, that influence the quality and quantity of interactions between them. This dissertation utilizes algorithmic group formation to examine the potentially relevant criteria for effective group formation in four experimental Studies, exploring online group work (Studies 1 & 2) and face-to-face group work (Studies 3 & 4).

Study 1 examines group formation in a four-week online course for prospective students. The group formation was experimentally carried out based on the variance of the personality traits extraversion and conscientiousness. It was hypothesized that it is advantageous regarding the results to have group variances heterogeneous in extraversion and homogeneous in conscientiousness. Study 2 was similar in the methodology of study 1, with variation of one of the grouping criteria. Based on the variance of the personality trait extraversion and prior knowledge, online groups were formed here to experimentally test which form of group formation leads to the best results.

Taken together, study 1 and study 2 provide evidence that successful learning in the online setting is associated with course design and individual differences, among other factors. Simultaneously, students often encounter challenges in organizing their online learning in groups and are prone to experiencing significant dropout rates in such settings.

Therefore, study 3 and study 4 explored the effectiveness of group formation strategies in face-to-face groupwork settings to enhance students' groupwork experiences. Study 3 examines whether the heterogeneity or homogeneity of the personality-trait extraversion, as distributed among group members, affects outcomes such as time spent, satisfaction, and performance. Surprisingly, groups with a homogeneous distribution of extraversion reported higher levels of satisfaction than groups with a heterogeneous distribution. The results of study 4 replicated and extended those of study 3 by investigating the effects of grouping from different perspectives in several student institutions. The results indicated that a homogeneous distribution of extraversion among group members significantly contributed to the success of group work.

The presented studies contribute to a better understanding of how algorithmic group formation and individual predictors can be utilized to promote successful group learning in several domains. Implications for research and practical applications can be derived from the research findings of the four studies.

Table of Contents

List of original Manuscripts

The present dissertation is based on four studies. Please note that the numbering of sections, tables, and figures corresponds to the numbering in the respective original studies to keep references within manuscripts consistent.

Manuscript 1:

Müller, A., Bellhäuser, H., Konert, J., & Röpke, R. (2021). Effects of Group Formation on Student Satisfaction and Performance: A Field Experiment. *Small Group Research, 53*(2), 244–273. https://doi.org/10.1177/1046496420988592

Manuscript 2:

Müller, A., Bellhäuser, H., Konert, J., & Röpke, R.(accepted). Group Formation by the Means of Extraversion and Prior-knowledge as Important Predictor in Higher Education Online. *Journal of Computing in Higher Education*.

Manuscript 3:

Müller, A., Röpke, R., Konert, J., & Bellhäuser, H. (2023). Investigating group formation: An experiment on the distribution of extraversion in educational settings. *Acta psychologica, 242*, 104111. https://doi.org/10.1016/j.actpsy.2023.104111

Manuscript 4:

Müller, A., Goeddeke, A., Kneip, P., Konert, J., Röpke, R., & Bellhäuser, H. (under review). Experiment on Extraversion Distribution in Groups Through a Group-Formation-Algorithm. *Computers and Education Open.*

List of Tables

List of Figures

Part I: Theoretical Background

1. Introduction

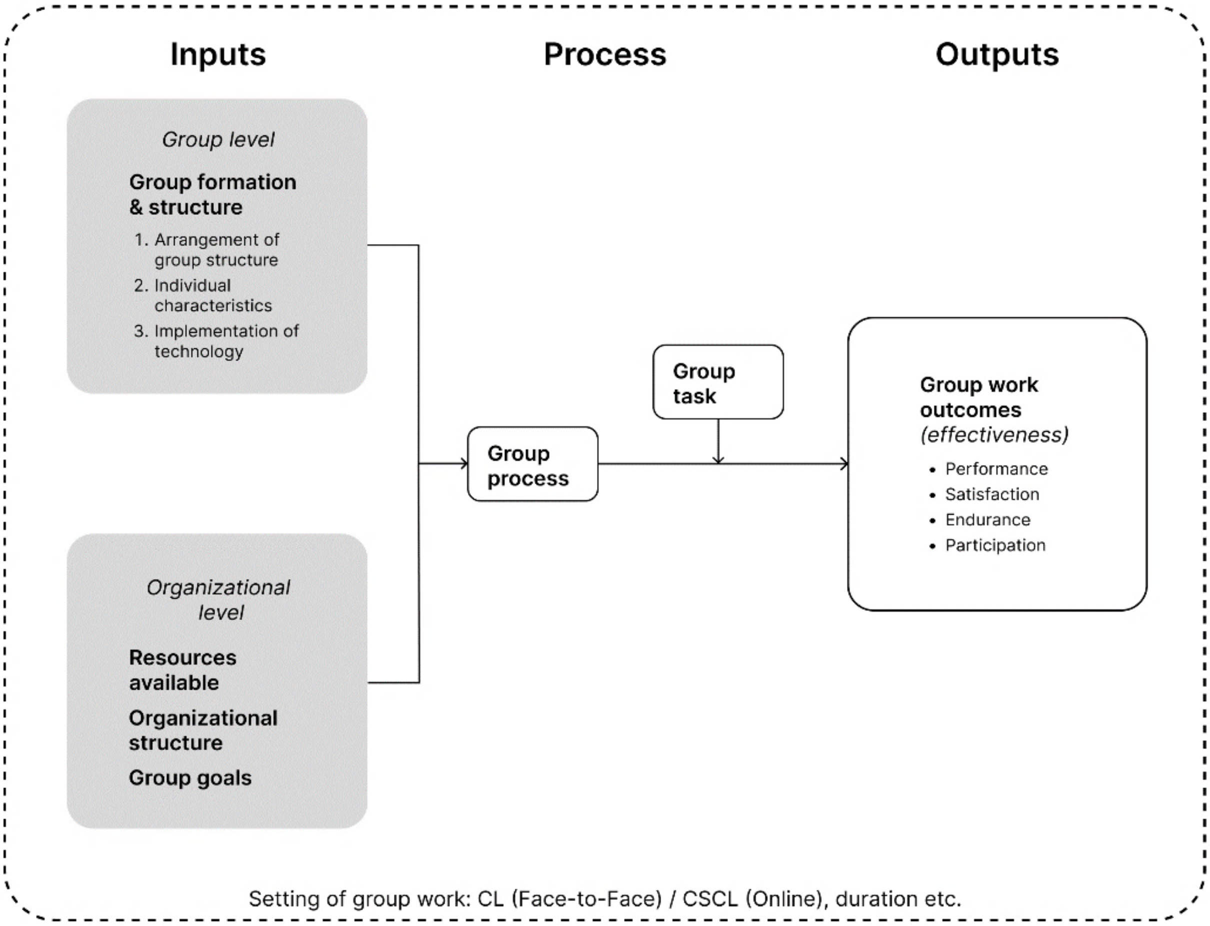
Groupwork represents a compelling field of study that reveals the complexities of human collaboration and interaction. The proposed dissertation aims to delve into this area by investigating the crucial role of group formation in the development of a group throughout its lifecycle. By experimentally exploring the impact of group member characteristics on outcomes in both face-to-face and online interactions, the aim is to further advance our knowledge of the most effective approaches to group formation in educational settings and its relevance in the contemporary world.

To contribute to the development of a theory or model for group formation that can be used to create more successful and satisfied groups systematically and predictably, a modified version of Gladstein's (1984) model of group work provides a good entry point for understanding the influencing factors (see Figure 1). The model provides a logical structure of the group-process, divided into the following phases: Input, process, and output. Tuckman's (1965) model of group development contextualizes the lifecycle of the group, the initial stage of group formation and the progression from its formation to task completion. In combining both models, I want to provide a framework for understanding the different stages and factors involved in the group formation process within a group's lifecycle. By zooming into the input phase, I try to highlight the challenges involved in the group-formation process.

The input phase of the model refers to the actual formation of the group (see section 2.1), including the selection of criteria (see 2.2), and technical supply for group formation (see section 2.3). The process phase focuses on the factors that influence group formation, such as the setting (see section 3) and desired outcome (see section 4). Finally, the output phase looks at the result of the group formation process, including the effectiveness and efficiency of the group (see Part II). As part of the input phase, the dissertation will focus on the role of personality traits as group formation criteria. Personality traits, such as those described by the Big Five model, can influence group work and attitudes towards it (Rammstedt & Danner, 2017).

In summary, this dissertation explores the concept of group work formation in educational settings by reviewing the current literature and identifying key challenges faced by group formation efforts. The analysis will provide a foundation for the experimental Studies and discussions that will be presented later in the dissertation, delving deeper into the terms, concepts, and relevance of group work and collaborative learning in educational contexts.

Figure 1. *Modified Version of the Group-Work-Model by Gladstein (1984)*



Note. Group work porcess from initial formation of a group to task completion

1.1 Defining Collaborative Learning: Key Terminologies

To fully grasp the concept of group formation in educational settings, it is crucial to first clarify the terms and concepts commonly used in research on group work and collaborative learning. Learning is a social activity that involves the cooperation of two or more people to achieve shared objectives and solve problems (Bruffee, 1999; Dillenbourg, 1999; Vygotsky & Cole, 1978). This can occur in face-to-face or online settings, where collaborative learning is enabled with the help of computer tools, called Computer-Supported Collaborative Learning (Maqtary et al., 2019). In this dissertation, the terms groupwork and collaborative learning (CL) will be used interchangeably, with the computer-supported online setting referred to as CSCL.

CL and CSCL are effective when group members interact with one another in a way that enables the group to achieve its goals (Dillenbourg, 1999). However, it is important to note that the term *effective* lacks a clear definition and is subjective, with interpretations changing depending on the research consulted. A dedicated section is included that differentiates and examines various common outcome measures referred to effectiveness (see section 4).

To ensure potentially effective group work, what is certain and essential, however, is to carefully consider the individual attributes, skills, and characteristics of its members (Smith et al., 2005). Hence, group work has to be understood as a complex adaptive system (Ramos-Villagrasa et al., 2018). This dissertation focuses on the initial phase of group formation through experimental manipulation of trait distributions to establish groups and examines its subsequent impact on group work. In the literature, the act of assembling learners into learning groups is described by various terms, including group formation or group composition. These terms essentially share the same meaning, aiming to identify concepts that provide a foundation for the systematic and successful formation of groups. Throughout this dissertation will the term group formation be consistently employed to encompass both terms.

1.2 Advantages and Pitfalls of Groupwork for Collaborative Learning

As outlined in the previous section, group work is a fundamental element of learning (Prince, 2004). Key components of group work that foster a continuous learning process include a sense of belonging, the development of group cohesion and membership, shared norms and values for communication and interaction, and interconnected social roles that lead to emotional and behavioral engagement (Gillen-O'Neel, 2021). Therefore, the promotion of group work is a vital component of lifelong learning (Druskat & Pescosolido, 2002; Noël et al., 2013).

In the educational context, numerous studies have demonstrated that groupwork is effective for student development in various ways. Unlike individual learning, group work not only leads to greater academic success, but also promotes better psychological and social development in learners (Mujkanovic & Bollin, 2019). As a result, members are motivated and energized to actively participate in the learning process (Johnson & Johnson, 2005; Shapiro et al., 2017). Group work provides a rich learning environment that enhances both formal and informal learning (Shibley & Zimmaro, 2002; Smith et al., 2005; Yazici, 2005) and supports social interactions (Curşeu et al., 2020; Johnson et al., 1991; Magnisalis et al., 2011). Through necessary interactions between group members, the social situation naturally promotes communication and helps each member cultivate a thorough comprehension of the learning topic (Johnson et al., 2000; Okdie et al., 2011). Therefore, well-structured group work has a positive impact on understanding, application, analysis, synthesis, and evaluation of learning contents (Williams et al., 2006).

However, group work in practice is not always popular and tends to be associated with negative experiences (Chang & Brickman, 2018). Imbalances in participation, resistance to teamwork, disparities in work speed among group members, and disadvantages for certain individuals are common pitfalls that can arise within a group setting (Crozier & Perkins, 2002; Walker, 2007). The processes that can contribute to the failure of group work have been the subject of previous research, including the loss of motivation and social challenges such as social loafing (Karau & Williams, 1993) as well as the free rider effect (Kerr & Bruun, 1983; Palloff & Pratt, 2005). The question remains as to why these problems occur in certain groups, while in other groups the benefits of group work are more prevalent.

To address social challenges and enhance the outcomes of group-work experiences, a viable solution lies in the application of systematic group formation methods (Feichtner & Davis, 1984; Graf & Bekele, 2006). While the number of studies examining group formation methods may be limited, it is crucial to underscore the significance of these methods for a group's overall success (Anewalt et al., 2003; Christodoulopoulos & Papanikolaou, 2007; Curşeu et al., 2015; Steiner, 1972). Establishing a method to form successful groups necessitates consideration of the individual characteristics of each member and ensuring that meaningful connections are forged between them (LePine et al., 2011). This dissertation aims to bridge the research gap by investigating group formation as a method to ensure that a group fits well and collaborates effectively.

1.3 The Potential of Systematic Group Formation in Higher Education

The most common model in contrast to systematic group formation is unstructured group formation (e.g., random selection of learners). In common practice, unstructured group formation means either student self-selection or arbitrary formation by instructors, neither of which has been proved useful for the group's development process (Chen et al., 2018): It usually leads to uneven participation, off-task behavior, resistance to group work, and different paces of work among learners in the same group (Dillenbourg, 2002). Here, some learners are at a disadvantage, because individual learning preferences, preconditions, or other characteristics are not considered during group formation (Walker, 2007). For instance, shy group members may experience anxiety and worry as a result of being forced into social interaction and feeling pressure to actively communicate (Crozier & Perkins, 2002). However, not only shy group members but also other students reported negative experiences with group work (Forrest & Miller, 2003) and have expressed that a systematic group formation process could be helpful in addressing these issues (Koh & Hill, 2009).

Unstructured group formation may result in groups with substantial differences in desirable or undesirable characteristics among members. Because of that, the desired effect of improved learning for all students can rarely be achieved (Chalmers & Nason, 2005; Johnson & Johnson, 1999). Typically, unstructured group formation by the members themselves occurs based on similarities, friendships, physical proximity, or perceived physical attractiveness (Wax et al., 2017). Even though group homogeneity is often preferred due to reasons such as intergroup relations and social identity, it can be misleading, as the specific similarities may have little to no impact on the success of the group process (Jackson et al., 2019; Tajfel & Turner, 1986). In self-selected groups, members are more prone to manipulate results or be unwilling to report the non-participation of the members belonging to their group (Parmelee & Hudes, 2012; Sibley & Parmelee, 2008). Exemplary, high-performing students tend to join groups with other high-performing students, leaving lower-performing students unsupported and marginalized (Cera Guy et al., 2019). This leads to a well-known phenomenon: Only some groups achieve high performance, whereas the others are far from reaching the expected goals. To avoid such imbalance and ensure that all group members have the support and resources they need to succeed, is it important to transition from unstructured group formation to a more systematic approach. This involves considering individual characteristics and arranging the group structure in a way that creates a balanced and diverse group dynamic.

Determining which individual characteristics to consider and implementing technology to facilitate the group-formation-process are key challenges in this transition. By adopting a systematic approach, it is possible to mitigate the negative effects of unstructured group formation and ensure that all group members have the opportunity to succeed. In the following, I will outline three challenges of systematic group formation: Arranging the group structure, deciding which individual characteristics to consider, and implementing technology to facilitate this process.

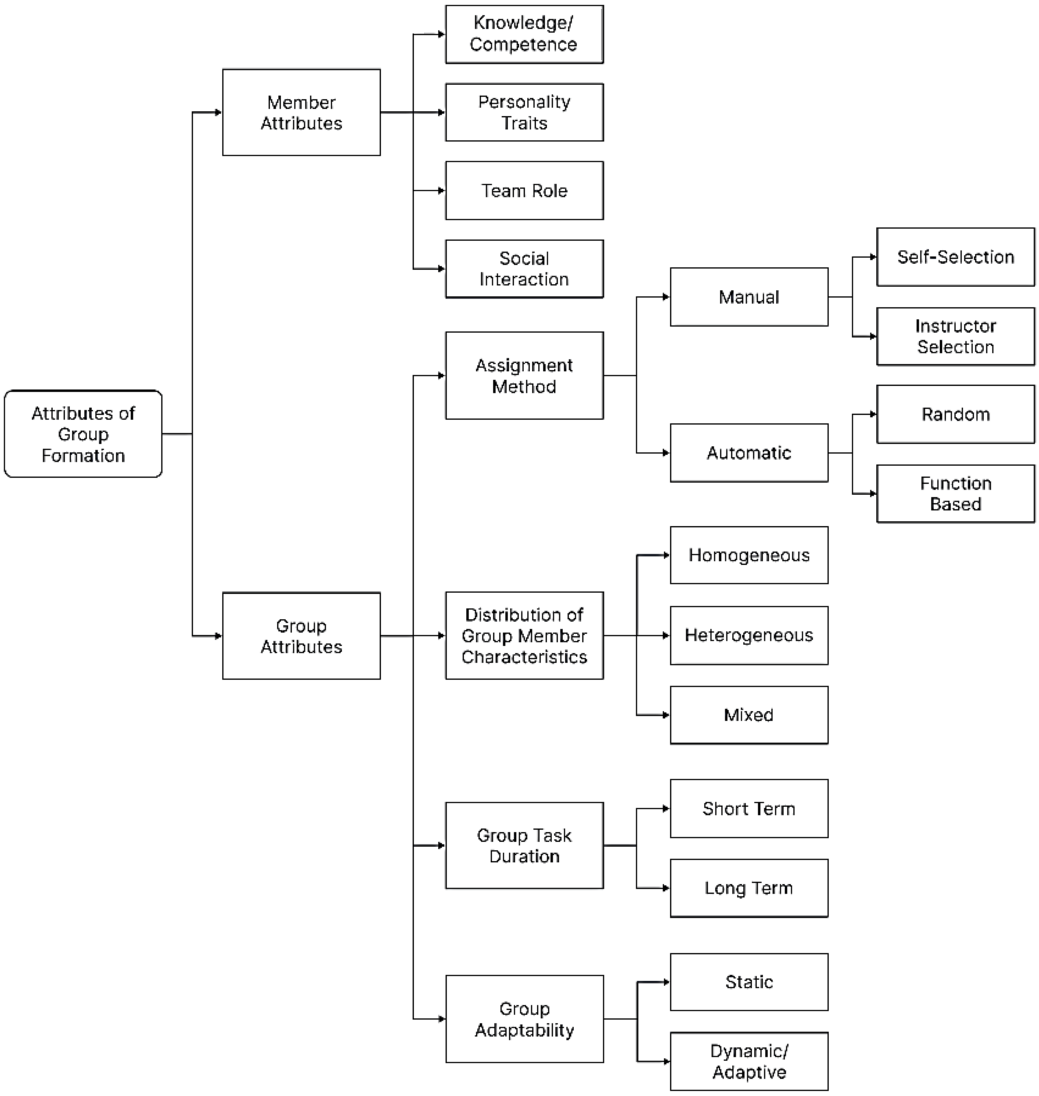
2. The Challenges to Consider when Forming Groups in a Didactically Meaningful Way

The task of forming groups may seem simple at first sight, but it is complex for several reasons. The following questions arise: How should student characteristics be structured within a group? Which characteristics should be considered, and for which specific learning goal? The effectiveness of group formation relies on understanding how multiple factors influencing group interactions can be utilized to prescribe appropriate learning groups and settings that facilitate effective interactions among learners (Dillenbourg, 2002; Strijbos & Fischer, 2007). Important to consider when forming groups are first, the type of arrangement for group structure, secondly, the criteria used to form groups, and third, the technical implementation to perform such a group formation.

In this chapter, I will outline three challenges in group formation. The first challenge is identifying the distribution of potential group characteristics, whether they are homogeneously or heterogeneously distributed within and between groups. The second challenge is determining the nature of these characteristics, whether they are surface-level or deep-level variables such as personality. The third challenge considers the algorithms used to perform group formation. After this, I will outline the setting in which the groupwork will take place (face-to-face or online). To conclude the section, the term *effectiveness* as the intended goal of group work (measured by various outcomes, e.g., performance, satisfaction) will be further elaborated and differentiated.

In the taxonomy of group formation attributes by Maqtary et al. (2019), all relevant decisions to consider during the research preparation stage can be found. Following its use can potentially aid in designing and methodologically describing the group-formation-process, while taking into account the underlying setting and chosen outcomes. The taxonomy, displayed in Figure 2, is adapted to align with this dissertation's research focus, serving as a visual representation guiding through the topics covered in this dissertation. Here, member attributes refer to characteristics of the members of a group, such as surface-level variables (e.g., age, gender) or deep-level variables (e.g., personality traits). Group attributes refer to the characteristics of the entire group, such as its size, the diversity of its members (homogeneous or heterogeneous distribution), and the goals it is working towards (outcome). Additionally, the setting is the first decision to be considered and preconditions groupwork thereafter (e.g., face-to-face vs. online).

Figure 2. *Taxonomy of group formation attributes by Maqtary et al., 2019*



Note. Taxonomy of group formation process decisions, accounting for chapters in respect to column drop

2.1 Challenge 1: Arrangement of Group Structure for Group Formation

In group research, group formation happens randomly, or systematically, based on a study variable, a chosen criteria, e.g., student characteristics, namely traits. When based on such a study variable, there are three options for group formation that consider either the average/mean, variance (power, range, proportion), or extremes (minimum or maximum). Thus, one can either consider the overall mean of the group member's characteristics or the within-group variance of a characteristic, or refer only to the extreme values, with only one group member considered significant (Neuman et al., 1999). Average and extreme distribution only consider the central tendency and the extremes of the data, respectively, and do not take into account the distribution of the data within those boundaries. This can lead to a lack of understanding of the diversity and complexity within the group.

By considering variance in group formation, we can better understand group diversity and how it may impact group dynamics and performance. Therefore, the fit approach focuses solely on trait variance when classifying group members (Seong et al., 2015). Heterogeneous group formation reflects a high level of variance within the group, while homogeneous group formation reflects a low level of variance. In this context, the focus of this dissertation is on groups formed either heterogeneously (complementary fit) or homogeneously (supplementary fit) according to specific characteristics. Here, each characteristic can contribute differently to the group-work-results, depending on its distribution and structure in the groups.

Complementary Fit. The complementary, heterogeneous group formation is characterized by group members, who differ in certain characteristics. The diversity of group members, such as heterogeneity, can be a crucial criterion to promote intensive intragroup interaction and successful learning outcomes. Therefore, it is often mentioned and investigated in research (Graf & Bekele, 2006; Resta & Laferrière, 2007). The diversity of the group members' can then be used to achieve the group's goal (Muchinsky & Monahan, 1987), as each individual benefits the collective (Bekele & Menzel, 2005; Cable & Edwards, 2004). Here, it is assumed that individual characteristics complement each other in such a way that everyone can contribute a certain part to the work (Moore, 2011). The group benefits from this variety of resources (Chiu, 2000), as learners can benefit from different skills, learning styles, knowledge, perspectives, and personality traits (Bradley & Hebert, 1997; Slavin, 1987).

Supplementary Fit. The supplementary, homogeneous group formation is characterized by group members' similarity in certain characteristics. Here, members are likely to share the same values, goals, etc., and thus develop positive attitudes and less conflict within the group, which in turn increases group performance (Cable & Edwards, 2004; Grønkjær et al., 2011). A homogeneous distribution resembles self-selection by members, where similarity is also a driving force: According to the concept of homophily and social group theory, individuals are more likely to seek out and associate with others who share the same attributes for self-affirmation and a sense of belonging (Curry & Dunbar, 2013; McPherson et al., 2001; Tajfel & Turner, 1986). The homogeneous fit increases attraction to and trust in other group members (Ilgen et al., 2005), promotes beneficial attitudes as well as performance (Cable & Edwards, 2004); those members who share goals and values tend to experience fewer task- and relationship conflicts and perform at higher levels (De Dreu & Weingart, 2003). Still, in research, there is criticism, as the homogeneous distribution is often considered the baseline to explore heterogeneous distributions, and therefore independent effects of homogeneity cannot be derived (Apfelbaum et al., 2014).

2.1.1 Exploring the Gaps of the First Challenge

The extent to which heterogeneity or homogeneity in group formation is beneficial for group outcomes is a topic of debate in recent research. The conflicting findings about the respective distribution of different variables in the group-formation-process require explanation: Based on the applied characteristic, either a heterogeneous or homogeneous formation of the respective characteristic is preferable (Parmelee & Hudes, 2012; Sibley & Parmelee, 2008; Sidorenkov et al., 2018). Thereby, surface-level characteristics (e.g., age, gender) and deep-level characteristics (e.g., values, beliefs) can have different degrees of impact on groups.

One potential reason for the conflicting findings in the literature are methodological differences in the studies. For example, some researchers have used non-experimental study designs, which do not permit causal conclusions and final statements to be derived from the respective results compared to experimental designs (Rispens et al., 2021; Van Emmerik & Brenninkmeijer, 2009). In addition, some studies used small or non-representative samples, or have failed to control important variables that could have influenced group performance (Al-Dujaily & Ryu, 2007; Kirschner, 2017). Overall, the relationship between group formation and outcomes is complex and multifaceted, and more research is needed to fully understand the conditions under which the distribution and constellation of certain characteristics may be beneficial or detrimental to group functioning. The author suggests that the effect of the respective distribution of different variables in the group-formation-process may depend on the specific context in which the group is operating, and the specific outcomes being measured Given the complexity of the relationship between group formation and group outcomes, it is important to carefully consider the specific characteristics being examined and the study design when investigating this topic.

2.2 Challenge 2: Individual Characteristics to Consider as Group Formation Criteria

For individual characteristics, that are related to a person and do not change with the learning situation or the tasks to be performed, a distinction can be made between surface-level and deep-level characteristics. Surface-level characteristics are demographic variables that can be realistically determined after a brief observation/comparison of a person's sex, age, and ethnicity. Deep-level characteristics provide information about a person's psychological characteristics, including personality traits, cultural values, and self-efficacy (Bell, 2007). Indeed, the distribution of these variables has various implications for the outcomes of group work and, because of that, can be useful in the formation of groups (Van Emmerik & Brenninkmeijer, 2009). The characteristics applied to form groups are crucial, however, not all characteristics are equally relevant in group formation. Since personality traits have long been the subject of research and can be adequately recorded economically and according to validity requirements, they are ideally suited as group formation criteria. Moreover, studies argue that groups formed considering personality traits show higher shared and individual performance (Bell, 2007).

The studies underlying this dissertation selected personality traits based on the Big Five, extraversion and conscientiousness, and prior knowledge as criteria for group formation. The decision to use these group formation criteria was driven by ethical considerations and a previous literature review, with the aim of ensuring a suitable research project. After reviewing the available literature, extraversion, conscientiousness, and prior knowledge were selected. Extraversion was chosen due to its strong association with social behavior, while conscientiousness and prior knowledge were selected for their strong association with academic performance. The focus on these traits was deemed relevant given the current debate surrounding group hierarchies. Ethical considerations in higher education settings led to the exclusion of neuroticism and agreeableness as group formation criteria. Detailed information on the chosen group formation criteria is provided in the studies consulted for this dissertation and below.

Extraversion. Extraversion is a personality trait that has been linked to various aspects of social behavior and group dynamics. Individuals who score high on extraversion are often characterized as being sociable, exhibiting leadership behaviors, and having a positive attitude towards social interactions such as group work (Hough, 1992; Judge et al., 2002; Neuman et al., 1999). They tend to facilitate group processes in discussions (Mohammed & Angell, 2003), clarify tasks (Forrester & Tashchian, 2010), and exhibit supportive behaviors (Porter et al., 2003), which can enhance the overall performance of the group. On the other hand, individuals who score low on extraversion are often perceived as being reserved, shy, or quiet (Costa & McCrae, 1992). Previous negative experiences with group work may lead them to prefer working alone (Chamorro-Premuzic et al., 2007). A heterogeneous distribution, where group members have varying levels of extraversion, has been demonstrated to lead to improved group dynamics and more positive outcomes. In such groups, introverted members are more likely to participate and contribute, resulting in a more diverse group with a range of perspectives and ideas being shared (French & Kottke, 2013; Humphrey et al., 2007). Furthermore, a heterogeneous distribution can result in effective leadership within the group, as different members bring their own unique leadership styles and strengths to the table (Bass & Riggio, 2005; Kramer et al., 2014; Zaccaro et al., 2018). However, there are also contradictory results: The study by Wilson et al. (2016) found that groups composed of members with similar levels of extraversion showed superior interactions and collaboration. Similarly, the review by Shemla et al. (2016) concludes that group heterogeneity can have both beneficial and harmful effects on group outcomes. While the positive impact of a heterogeneous distribution of extraversion on group outcomes is widely recognized (Bell, 2007; Roney et al., 2012; Thanh & Gillies, 2010), the exact effect of its distribution remains an area for further research.

Conscientiousness. Individuals with low levels of conscientiousness are often perceived as lazy, fickle, and careless (McCrae & John, 1992; Schmutte & Ryff, 1997). On the other hand, individuals with high levels of conscientiousness are characterized as organized and reliable (Barrick & Mount, 1991; McCrae & Costa, 1987). Research has shown that individuals high in conscientiousness tend to apply better learning strategies and distribute their studying, resulting in better academic performance in both face-to-face (O’Connor & Paunonen, 2007; Schneider &

Preckel, 2017) and online learning (Arispe & Blake, 2012; Loya et al., 2015; Varela et al., 2012) settings. Conscientiousness has been found to have the strongest effect on performance among all personality traits (Zell & Lesick, 2022). In group work, individuals with high levels of conscientiousness tend to focus on the learning process and tasks, leading to positive effects on group performance and dynamics (DeChurch & Mesmer-Magnus, 2010). They are also more likely to be elected as group leaders (Feist & Barron, 2003). Still, the exact mode of action through the distribution of conscientiousness is still in question. Despite this, it is important to consider conscientiousness when forming groups, as it can lead to benefits in terms of group performance, positive group dynamics, and academic performance.

Prior Knowledge. Prior knowledge is a critical factor to consider when researching group formation and group work. Studies have shown that cognitive ability and prior knowledge are reliable predictors of academic achievement (Richardson et al., 2012; Schneider & Preckel, 2017). The results of research exploring prior knowledge in CL have been mixed, finding heterogeneous groups to be better, but showing low-ability students to be more engaged in homogeneous groups (Murphy et al., 2017). However, most research suggests that the distribution of prior knowledge and abilities within groups has a positive impact on members' knowledge gains through heterogeneous distribution (Mehta & Kulshrestha, 2014). Diverse prior knowledge can also provide opportunities for learning from each other, but additional support may be needed to ensure equal participation. When forming groups, prior knowledge can be used to match individuals to tasks and identify potential challenges during group work. While prior knowledge is important in group formation, further research is needed to fully understand the ideal distribution of prior knowledge in group formation.

2.2.1 Exploring the Gaps of the Second Challenge

As indicated in the previous section, personality traits are important factors for functioning in groups. However, contradictory research results can be found regarding their distribution. Because of this, the question remains whether an individual trait or multiple of the five traits should be considered, and how those traits should be distributed within a group.

A positive relationship between extraversion and group effectiveness was demonstrated in a study by Balthazard et al. (2002). Subsequently, another study on the impact of all Big Five personality traits in online group work on outcomes, including behavioral, cognitive, and emotional engagement, revealed that group failure may result from a heterogeneous distribution of extraversion within groups. Consequently, the study recommended homogeneously distributing extraversion, while advocating for heterogeneous distributions of conscientiousness and agreeableness. Furthermore, ineffective group functioning correlated with high levels of neuroticism, whereas effective functioning was linked to high levels of openness (Zhang et al., 2020). Revelo-Sanchez et al. (2021) proposed the best performance on average to be achieved through the heterogeneous distribution of openness to experience, conscientiousness, and neuroticism. Challenging the findings, Kucukozer-Cavdar and Taşkaya-Temizel (2016) study identified no difference in group success, measured by grades, based on either groups homogeneous nor heterogeneous distribution of any personality trait from the Ten-Item Personality Inventory. In a meta-analytical summary of 259 studies, group formation proved most effective only if trait distributions were simultaneously recognized, with the complementary fit of extraversion and the supplementary fit of consciousness together yielding superior results, emphasizing the potential role of trait variance interactions in group formation research (Humphrey et al., 2007).

While personality emerges as a significant factor influencing group functioning and performance (Peeters et al., 2006), uncertainty remains about how the configuration of group members' attributes affects their outcomes in group work. To address these gaps, this dissertation aims to examine the effects of various configurations of group members' attributes. Additionally, it will assess whether heterogeneous or homogeneous group formations should be preferred. Other gaps in group formation research that need to be addressed have emerged from previous research designs. For instance, Balthazard et al.'s (2002) study utilized an experimental design, but it did not target the distribution of personality traits. In contrast, Zhang et al. (2020) opted for a correlational design, assessing both outcomes and personality, after a randomly formed group collaborated online, based on engagement measured through WeChat-activity. Kucukozer-Cavdar and Taşkaya-Temizel (2016) employed a causal-comparative design to examine online working groups. However, these groups were not manipulated, rendering the design non-experimental. The focus was on the heterogeneity versus homogeneity of the overall group and its impact on group success, measured by the final group grade. Yet, to achieve a causal understanding of the influence of member personality on various outcomes, experimental research is necessary, manipulating the distribution of member characteristics. In summary, due to suboptimal design, inconclusive research results suggest that no precise statement can be made about the distribution of various attributes within a group-formation-process.

2.3 Challenge 3: Implementing Technology to Support Systematic Group Formation

In the following section, the question of how to automatically create successful groups with the help of technology will be elaborated. The group-formation approach using technology has attracted attention and the number of technologies that enable CL has increased, especially since the beginning of the COVID-19 pandemic (Houlden & Veletsianos, 2022).

Allocating students to groups is a non-trivial task that requires considerable effort. The problem of possible combinations in terms of the number of students and groups required poses a significant challenge. This challenge is even greater in an educational context and when taking ethics into account, because all the groups formed must be equally capable of overcoming the challenges posed to them. Without technical support, group formation is time- and labor-intensive, and it is therefore almost impossible for an instructor to assemble learning groups that meet several criteria (Hwang et al., 2008). A systematic algorithm-based approach provides a central solution to help apply optimized group formation (Konert et al., 2014; Pai et al., 2014).

Several methods can potentially solve a group-building problem in education, and a variety of group-building techniques have been studied by researchers (Daradoumis et al., 2002; Haller et al., 2000; Martín & Paredes, 2004; Wilkinson & Fung, 2002). One of these methods is to use an algorithm to easily automate and improve group formation for research and practice (Konert et al., 2014; 2016). Such algorithms may become useful for group formation because they can quickly and efficiently process large amounts of data, identify patterns and trends, and make decisions based on that information. In this sense, from a general algorithmic perspective, the group formation challenge can be viewed as a multidimensional optimization problem. In contrast, manual group formation often relies on subjective and potentially biased decision-making, which can lead to suboptimal group formation and dynamics. Various approaches to algorithmic implementation and ways of grouping do exist, for instance, semantic algorithms (Manske & Hoppe, 2017), ant-colony-optimization-algorithms (Graf & Bekele, 2006), or particle-swarm-optimization-algorithms (Yin et al., 2006). The studies underlying this dissertation made use of the Genetic Algorithm and implemented GroupAL as a possible solution to the systematic group-formation problem (Konert et al., 2014; Roepke et al., 2016). GroupAL is provided as a plugin for the learning-management-system-platform Moodle (LMS).

Genetic Algorithm and its Variant GroupAL. The Genetic Algorithm (GA) is a search heuristic used to find near-optimal solutions to search problems and optimization. It is a method for solving optimization and search problems by mimicking the process of natural evolution. It is a population-based algorithm that iteratively improves a set of solutions by applying genetic operators, such as crossover and mutation. The algorithm uses a validity function to evaluate the quality of each solution and a selection operator to select the best solutions for reproduction. The goal of a GA is to find a near-optimal solution to the problem at hand. The genetic algorithm approach to group formation problems consists of a set of students and a set of groups.

GroupAL is a variant of the GA that is designed for group formation and uses the principles of natural selection and genetic variation to evolve a population of candidates as potential solutions, where each solution represents a possible group formation. The algorithm evaluates the fitness of each solution based on a set of predefined criteria, such as group diversity or task performance, and then uses genetic operators, such as crossover and mutation, to generate new solutions. Through this process, the algorithm iteratively improves the quality of the solutions, eventually finding a near-optimal group formation. The goal is to allocate all learners to a group considering the chosen optimal solution (Chen, 2000; Wang & Elhag, 2006; Wang et al., 2007). GroupAL is designed to solve a specific type of group-formation-problem, which is known as the k-partition problem. The k-partition problem is defined as the task of dividing a set of **n** elements into **k** disjoint subsets, such that the elements within each subset have similar characteristics or perform similar tasks. It uses a binary encoding scheme to represent the solutions, where each element is assigned to one of the **k** subsets. The genetic operators, such as crossover and mutation, are then applied to the binary strings to generate new solutions. The algorithm runs until it meets a stopping criterion, such as reaching a maximum number of generations or a satisfactory level of fitness. The final solution represents the best group formation the algorithm could find.

In summary, the GroupAL is a prominent approach that enables mixing different target criteria (such as gender-homogeneous and knowledge-heterogeneous groups) to achieve an overall optimal grouping in a given set of students. In all Studies of this dissertation, the algorithm allows setting up group formation based on predefined criteria, e.g., student characteristics (Roepke et al., 2016). More information can be found in the Studies of this dissertation (see Part II).

2.3.1 Exploring the Gaps of the Third Challenge

The use of algorithms in group formation is still an area of active research and development. Hence, there are ethical and privacy considerations to be considered when using algorithms to make decisions about group membership. As such, it is important for practitioners to carefully consider limitations when using algorithms in group formation. There are various methods in use for automatically forming groups, and many types of data mining techniques that are used to explore and analyze large data sets to discover meaningful patterns (Odo et al., 2019; Srba & Bieliková, 2015). Several other approaches use information from the learner's profile to form groups, e.g., content knowledge, personality, traits, and programming styles (Graf & Bekele, 2006). Still, the formation of groups based solely on learner profiles, technologies, and tasks cannot serve as an indication of well-designed CL-sessions (Resta & Laferrière, 2007; Strijbos et al., 2004). Reviewed articles show the variety of algorithms, but studies on the evaluation of the effectiveness of such algorithms are still missing (Odo et al., 2019). Additional criticism of previous Studies in this area is based on both arbitrarily picked criteria, whose selection is not well justified or whose validity is not guaranteed, as well as a methodologically non-experimental implementation in the study's design, so that it is impossible to conclude causal statements.

3. Considering the Setting: From Conventional to Digital Learning Environments

The literature discussed in the previous sections indicates that group-formation-outcomes are dependent on their specific context. A distinction that is particularly important in this regard is that between online and face-to-face settings. Computer-Supported Collaborative Learning (CSCL) is a form of education that uses technology to facilitate interaction and communication among learners. Group formation plays a vital role in CSCL as it affects communication dynamics and the overall learning experience. In this section, I will compare the differences between online and face-to-face-groupwork, discuss the benefits and challenges and consider the factors crucial to group formation in CSCL-settings, that have an impact on the learning experience.

Online group work has the advantage of forming geographically dispersed groups, making it a viable alternative to traditional face-to-face education (Adedoyin & Soykan, 2020; Contractor, 2013; Laal & Ghodsi, 2012). However, face-to-face group work fosters stronger social bonds, while online group work requires planned communication and may be impersonal (Palloff & Pratt, 2005). Despite challenges such as lower satisfaction and higher dropout rates, CSCL has numerous benefits, including increased motivation, critical thinking skills, social presence, and collaboration (Altınay & Paraskevas, 2007; Joiner, 2004; Rourke & Anderson, 2002; Salovaara, 2005; Voorn & Kommers, 2013). Online groupwork may also be more comfortable for shy learners, as it may require fewer social skills in certain situations (Voorn & Kommers, 2013). However, asynchronous online group work can also lead to feelings of isolation for extroverted learners, who still prefer virtual interaction over no interaction (Carter, 2013; Ortiz de Guinea, 2012). External circumstances such as working from home can also present challenges, as individuals with less favorable home environments may experience difficulties in terms of privacy, noise, or distractions (Wildman et al., 2021).

Crucial to the success of CL and CSCL is, above all, communication (Chou, 2001; Graham & Misanchuk, 2004; Michalsky et al., 2007; Wang et al., 2001). Group formation can impact communication by creating social bonds, influencing the distribution of roles and responsibilities within the group, and establishing norms and expectations (Fiorentino et al., 2021). CL occurs in different settings, each with unique features and challenges. Examining both online and face-to-face settings allows for a more comprehensive understanding of the process and how it differs across different communication platforms (OECD, 2016; Okdie et al., 2011). It also sheds light on how individuals adapt to and navigate these differences (Yulianto et al., 2018). This examination informs the development of strategies and interventions for promoting effective group formation and helps researchers identify the unique challenges and opportunities that each setting presents (Tsovaltzi et al., 2019). Overall, considering both settings offers a more nuanced understanding of group formation and the factors influencing communication and learning.

4. Considering the Outcome: Different Measures of Group-Work-Effectiveness

The term *effective* is used differently among researchers in the field, with the primary focus on group performance (AbuSeileek, 2012; Balthazard et al., 2002; Bayeck et al., 2018; Fenwick & Neal, 2002; Kucukozer-Cavdar & Taşkaya-Temizel, 2016; Mehar & Kaur, 2020; Saqr et al., 2019; Staples & Zhao, 2006; Van Emmerik & Brenninkmeijer, 2009). In addition, group satisfaction counts as an additional definition of effectiveness (Balthazard et al., 2002; Shaw, 2013; Staples & Zhao, 2006; Van Emmerik & Brenninkmeijer, 2009). The problem here is that the term *effective* refers to and manifests itself on various result variables, making it difficult to compare group formation research and implement findings for review (Odo et al., 2019). The development of a shared definition of research group formation effectiveness, including subjective evaluation, i.e., group satisfaction, as well as objective evaluation, i.e., group performance, would benefit future research. In this dissertation, I will use the term *effective* in the context of group work by a range of variables, including group performance, group satisfaction, and other measures such as time investment, participation, and maintenance. This comprehensive and differentiated definition of effectiveness in group work includes both subjective and objective evaluation measures and allows for a more nuanced understanding of group work outcomes. This approach will benefit future research on group formation effectiveness by providing a definition to compare different studies.

5. Summary of Research Aims

Group activities have become increasingly popular as a didactic method to support learning, particularly in higher education. However, researchers have identified a lack of both sufficient technical tools and an underlying psychological theory on the characteristics to use in group formation. This has led to a lack of research on the impact of different student criteria on learning outcomes while collaborating. Additionally, outcome measures differ from study to study, making it difficult to compare and integrate results. To address this issue and further support pedagogically meaningful group formation, a more systematic approach is needed, as well as appropriate design of collaborative learning scenarios, and intelligent support for students to collaborate effectively.

To fully understand the complex dynamics of group work, it is necessary to consider a range of variables and their interrelationships. By using a systematic approach to the group-formation-approach and taking a broader view of the factors that can impact groupwork, we can gain a deeper understanding of how to effectively form and support student groups for collaborative learning. With the increasing number of students and given the large number of potential characteristics, technical advice is needed, as algorithmic assistance appears to be an easy-to-use solution to assist instructors and researchers with the issues involved in group formation. The algorithm employed in the research of this dissertation is intended to experimentally assist in the identification and validation of key criteria for effective group formation, with the research aim of determining key student criteria that can inform the development of personalized group formation recommendations, to improve effectiveness of group work and student performance through CL.

To address the aforementioned research gabs, this dissertation proposes an experimental research design to explore a group formation strategy using specific dimensional trait variables as criteria and examine their effects on various outcomes such as performance, satisfaction, or time investment (see studies 1-4). The selected group formation criteria include conscientiousness (see study 1), prior knowledge (see study 2), and extraversion (see studies 1-4). The study encompasses higher educational settings for online (see studies 1 & 2) and face-to-face group work (see studies 3 & 4). An examination of both settings provides a more nuanced understanding of the outcomes and implications arising from the group formation strategies implemented. In the following chapter, the four studies are presented, that form the basis of this dissertation.

Part II: Original Manuscripts

6. Manuscripts for Dissertation

6.1 Manuscript 1: Müller, A., Bellhäuser, H., Konert, J., & Röpke, R. (2021). Effects of Group Formation on Student Satisfaction and Performance: A Field Experiment. *Small Group Research, 53*(2), 244–273. https://doi.org/10.1177/1046496420988592

Effects of Group Formation on Student Satisfaction and Performance: A Field Experiment

Abstract

This study analyzes the relation of group formation to outcomes of a 4-week-online-course for prospective students. Group formation was experimentally manipulated based on predefined criteria, personality traits, conscientiousness, and extraversion. As research questions, groups were considered advantageous, if they were formed (a) heterogeneously in extraversion, and (b) homogeneously in conscientiousness: As a result, no uniform outcome was identified. Most variance could be explained on group-level, but no significant main effect for experimental grouping was found. Significant interaction between both main effects hint that the results do not provide final answers, but guidance for further research.

Introduction

The term groupwork is referred to in a variety of ways by different practitioners. In the field of educational psychology, the terms groupwork, collaborative learning, and cooperative are most common. Given that there are far more commonalities than differences among these terms, we consider them to be equivalent and will for the rest of this contribution use groupwork and cooperative (Kreijns et al., 2003; Van Merriënboer & Kirschner, 2001). Addressing the difficulty of developing a comprehensive taxonomy of groupwork, Kagan (1994) provides an outline of different forms of groupwork that instructors at various educational levels have implemented. Well-known forms of groupwork include jigsaw, student -team-learning, group-investigation, and learning together; and have been studied and compared in meta-analyses (Johnson et al., 2000; Lou et al., 2001). Meta-analyses show greater achievement for groupwork compared to other learning practices (e.g., individual working processes or competitive ones), and greater academic achievement (Crouch & Mazur, 2001; Hsiung, 2012; Johnson et al., 1991; Lou et al., 2001). Still, the question of why and to what extent achievement increases in collaborative formats remains.

Recent research has focused on how to actively design and manage groups to be more effective (LePine et al., 2011). However, a great deal of confusion prevails about which and why cooperative learning methods affect achievement, and under what conditions cooperative learning has these effects (Lambić et al., 2018).

Group Outcomes

Topics and phenomena linked to the evaluation of groupwork, and thus the selection of outcome measures, include social influence and communication, cooperation and competition, role development and performance, group processes (process gains and losses) and performance behavior. A problem surfaces at the point of trying to classify outcomes as either subjective or objective, and individually or collectively relevant. Reasonable criticism is due here, for the fact that the judging of group success generally depends on one single outcome, namely a measure of performance (Wittenbaum et al., 2004). Due to the lack of a uniform definition of success for groupwork, the present study does not claim to make an evaluation of groupwork that can be considered generally valid; to achieve this, the other indicators of success, such as group satisfaction, need to be examined.

In the groupwork literature, numerous complex models identify different determinants of group effectiveness. Well-known models are, for example, the models devised by Steiner (1972), Hackman (1983), and Gladstein (1984). All three models are dated, but they are still in use today and preferable to newer group models, which are often limited to singular variables without considering their interaction (Ilgen et al., 2005).

One solution to overcome the above-mentioned gap in knowledge could be the Model of Task-Group-Effectiveness (Gladstein, 1984), which will be presented here as an exemplary. This model draws on a classification system that differentiates between subjective and objective performance criteria. Objective performance criteria are independent of participants and refer to criteria related to the objective degree of achievement (performance) in a task. Subjective performance criteria are participant-dependent, because the subjective perception, opinion, and satisfaction within the group and regarding its processes and prerequisites is evaluated here. The model is divided in three categories: Input (structures and prerequisites), process (processes and interaction) and output (results and group effectiveness). The input category differentiates the subordinate group factors on one hand at the group-level, and on the other hand at the organizational level. In the former, it is subdivided into group formation (i.e., appropriate skills, heterogeneity, organizational affiliation, experience) and group structure (i.e., clarity of roles and goals, labor standards, task control, group size, formal leadership). At the organizational level, it distinguishes between available resources (i.e., training, technical advice, goal) and their structuring (i.e., rewards for group performance, monitoring).

Group Formation

Group formation refers to the process of arranging group constellations and is a frequent research topic in computer science. Group composition can be described as a result of group formation (i.e., the configuration of member attributes in a group), and is a topic more frequently investigated in psychology (Tuckman, 1965).

Group formation is thought to have a powerful influence on group processes and outcomes (Hollenbeck et al., 2004; Kozlowski & Bell, 2003; Levine & Moreland, 1990). Further, group composition describes a phenomenon that takes place in parallel through the process of group development, and which could be caused by the previous group formation. Tuckman’s well-known and often quoted psychological theory of the forming-storming-norming-performing model describes a group process which also contributes to group-composition and formation (Tuckman, 1965). In the current study, a first attempt is made to operationalize these constructs. Still, the groups had little time to develop together, because groupwork only lasted 4 weeks and took place exclusively online.

For university teaching, equally good group formation for the entire cohort of participants is a desirable goal. This represents a difference from group research, in which group formation is random, voluntary, or based on a study variable. Natural group formation in most cases occurs according to commonalities, friendship, or seating proximity of learners (Wax et al., 2017). Three ways of grouping exist, considering either the mean value, the variance (performance, range, proportion), or the extremes (minimum or maximum) of an attribute. Thus, one can take into account either the team's overall mean level of the attributes or the within-team variance of one attribute, or just focus on the extremes whereby just one member of the group seems to be important (Neuman et al., 1999). The fit approach to placing members within teams focuses explicitly on the variance while ignoring the others. Considering this, we focus on grouping to be either heterogeneous or homogeneous regarding group member attributes, and therefore representation of criteria in use.

Heterogeneous fit suggests that a group is well-matched, as each individual can fulfill a need by bringing something unique to the collective (Bekele & Menzel, 2005; Moore, 2011). Although individual differences are often used when creating groups (Moynihan & Peterson, 2001), it is not always clear, how to conceptualize characteristics at the team level. Homogeneous fit suggests that people are more comfortable and productive within a group when they are similar to each other (Muchinsky & Monahan, 1987), such that people prefer being around others who share the same goals and values (Cable & Edwards, 2004).

For reasons such as intergroup relations and social identity, most often people are either unconsciously or consciously biased toward group homogeneity, even if the respective commonalities are meaningless for the group process (Jackson et al., 2019). Similarity can increase attraction and trust in other group members (Ilgen et al., 2005); members who share goals and values experience lower levels of task- and relationship conflict along with higher levels of performance (De Dreu & Weingart, 2003). Thus, homogeneous fit promotes both positive attitudes and positive performance (Cable & Edwards, 2004). So, one important issue to consider when generating groups is the ideal mix of individual characteristics to put in the group, structured by demographics, personality, knowledge, skills, and abilities (Halfhill et al., 2005).

There are several things to consider when deciding for or against, including criteria for group formation. In this process, one may divide criteria regarding group formation into two subcategories: surface-level and deep-level criteria. Surface-level criteria refer to overt demographic characteristics that can be reasonably assessed after brief exposure, such as age, race, education level, and organizational tenure. Deep-level criteria refer to underlying psychological characteristics, such as personality factors, values, and attitudes (Bell, 2007). Researchers have suggested that although demographic differences may be important, deep-level criteria can have a stronger influence on group performance (LePine et al., 2011). Previous empirical studies and meta-analyses have contributed to the understanding of the relationships between enduring deep-level formation variables and team performance. Nevertheless, this part of the group formation literature remains fragmented and inconsistent, and conclusions regarding optimal formation for superior group performance are difficult to make (Bell, 2007).

This reveals gaps in knowledge in the current state of research and makes comparison of literature challenging. For example, group is still a very randomly used, undefined term; we have seen it being interpreted in different research studies with differing constellations of people for diverse interactions. As stated above, the respective outcome variables, decisive for the statements of the studies, also differ. In addition, the selection of possible criteria for group formation is almost indescribably large. To cope with the complexity created, more use is made of diverse tools to help form groups (Lambić et al., 2018).

The fact that the findings discussed so far have been derived purely from correlative Studies denotes a serious limitation of the empirical evidence (Klein et al., 2009). Typically, groups were examined, that were either formed by chance or by the students themselves. Considering the theoretical body of group-formation-criteria, it becomes clear that no final recommendation of criteria for the forming of groups exists thus far (Magpili & Pazos, 2018). Hence, our purpose is to enrich this part of hitherto lacking research by an experimental study.

Algorithmically Optimized Group Formation. Here, the benefits of algorithmic group formation are highlighted. Particularly in higher education, the number of participants increases, as well as the demand to perform in groups. To form good groups, algorithmic group formation tools offer an economically beneficial and most promising technical solution to these challenges. Because of its complexity, group formation can only be implemented economically with algorithmic support, since the number of learners is too large to adequately meet their individual requirements and competences. Viable solutions for the selection and weighting of relevant criteria and procedures for grouping through easy-to-use online applications are needed (Pai et al., 2014).

There are several approaches to the composition of learning groups with the help of developed algorithms (Maqtary et al., 2019). For support of computer-supported cooperative learning (CSCL) settings, most algorithms create a maximum of diversity within a group on an algorithmic base concerning multiple criteria.

Group formation based on the combination of personality traits is statistically superior regarding satisfaction and group performance in comparison to a purely random group formation scheme. However, it was subsequently not possible to assess the relative share of the individual criteria in the overall positive effect. For instance, while algorithmic group-formation-tools can be used, no general or evidence-based recommendation or set of consistent criteria for group-formation has been established to feed the algorithms (Magpili & Pazos, 2018). For most of those criteria, it is not possible to create an equally successful distribution simultaneously for all groups. Suggestions regarding optimal distribution, which came forth during the execution of this study—can be found in more detail in the following section.

Group Formation by Personality Traits

The following section will illustrate the potential of personality traits as group formation criteria and discuss the options to form groups, due to specific criteria. Subsequently two personality traits their role in groups are briefly introduced. Personality factor indicators (Costa & McCrae, 1992), characteristic patterns of thought, feeling, and action, could contribute to explaining team performance (Bell, 2007). The personality of the team members is considered an important factor for the functioning and performance of the team (Peeters et al., 2006). Due to good psychometric properties and easily accessible measurement devices, Studies on the influence of extraversion and conscientiousness on group composition are more frequently available than Studies on other factors. Previous research has underestimated the effect of extraversion and conscientiousness variance on performance due to suboptimal design (Humphrey et al., 2007).

Extraversion. Individuals with a high degree of extraversion are described as both dominant (Costa & McCrae, 1992) and sociable (Hough, 1992), which is associated with leadership behavior (Judge et al., 2002). In theory a hierarchical structuring of groups corresponds to the social time picture (i.e., for leadership to be successful, leaders need followers; Meindl, 1995) and this role is filled by team members who are less extrovert as they are more passive (Costa & McCrae, 1992). Therefore, it can be assumed that a heterogeneous distribution within a group is desirable, since this trait is accompanied by leadership behavior, which is not equally pronounced in all group members (Kramer et al., 2014). Additionally, extraversion has been shown to be related to team-processes such as stimulating discussion (Mohammed & Angell, 2003) or supportive behavior (Porter et al., 2003). In turn, there are findings indicating the advantage of homogeneous grouping in extraversion (den Hartog et al., 2019).

Conscientiousness. Conscientious individuals are described as goal-oriented, structured, organized, and self-disciplined (Costa & McCrae, 1992). Characteristics that are associated with performance and willingness to perform as well as with one's own working style (Hurtz & Donovan, 2000). Of all personality traits, conscientiousness has shown the strongest correlation with academic success (Di Fabio & Busoni, 2007). A high level of conscientiousness is essential, not only regarding university studies, but also regarding the working world (Protsch & Dieckhoff, 2011). Since it can be assumed that people with the same objectives and working methods have less potential for conflicts, it is assumed that a homogeneous distribution is desirable (Prewett et al., 2009). Behaviors associated with conscientious team members should also be beneficial for team performance (i.e., fulfillment of task roles; Stewart et al., 2005). Conscientiousness has also been related to backing up behaviors and should be related to processes supportive of task completion and goal attainment (Porter et al., 2003).

In sum, the rationale for the development of this hypothesis was based on the results of several studies, which confirmed that conscientiousness most highly correlated to performance success amongst the studied personality variables (Busato et al., 2000; Di Fabio & Busoni, 2007; Furnham et al., 2003; Lounsbury et al., 2003). Therefore, we conclude and assert as a hypothesis that it is most appropriate to use the homogeneous-fit-perspective with conscientiousness. Importantly, trait-like individual differences that we controlled for in the current study by keeping conditions the same across all groups are team orientation (Bell, 2007), motivation (Nederveen Pieterse et al., 2011) and prior knowledge (Horwitz, 2005).

Groupwork in Online Settings

It is convenient to use virtual learning or distance learning, as this form of education is available to users at their time, place, location, and speed. Thus, for many people, virtual learning offers an alternative to traditional education settings and the possibility to use education in various forms. Suddenly, the potential of online group work is revealed, due to the circumstances of social distancing because of the COVID-19 pandemic.

Moreover, online media provide a tool that facilitates new forms of interaction and knowledge sharing and management and offers new didactic tools to promote interaction and social processes (Kirchner et al., 2009; Liao et al., 2014). Students who work in CSCL-environments report higher levels of learning (Hertz-Lazarowitz & Bar-Natan, 2002), make higher quality decisions, deliver more complete reports (Janssen et al., 2007), engage in more complex and challenging discussions (Benbunan-Fich et al., 2001), and report higher levels of satisfaction compared to students who work in face-to-face-groups (Lou et al., 2001). Alternately, the inability to transmit vocal and nonverbal cues makes online groupwork a colder, less personable experience and is often seen critically by students (Carter, 2013). To sum it up, research reviewed on learning with CSCL indicates that it cannot be compared to groups working face-to-face, especially when focusing on group formation and composition (Hollingshead et al., 1993).

Research Questions

Due to the high relevance of groupwork, the desideratum of research is to investigate group formation as an important factor to better understand the underlying mechanisms of action. Therefore, this study explores potential criteria for successful group formation for students, for them to be as productive and satisfied as possible. Previous research on the effects of homogeneous or heterogeneous group formation has observed greater influence of personality traits such as extraversion and conscientiousness than of demographics such as gender or educational level. And yet studies on group formation considering personality traits are, to the best of our knowledge, constructed as correlative designs (Clark et al., 2019).

In this study we use an experimental study design in respect to group formation based on extraversion and conscientiousness, both considered valid and behaviorally predictive. These two personality traits are neither researched enough nor show definite, unambiguous results when it comes to their structure within groups. Thus, it is not possible to clearly determine, which formation would lead to disadvantages for students in one of the group-formation-conditions. Therefore, the experimental manipulation of traits as group formation criteria in a university context is possible without violating any ethical concerns. More recent research projects question the knowledge gain of result models with linearly verified relationships. They advocate a situation where potential dependencies and relationships cannot be discovered in this conventional way. Instead, curvilinear correlations should be tested (Curşeu et al., 2019). The objectives of this study are to experimentally manipulate the two personality traits (extraversion and conscientiousness) using an established algorithm, to gain more insight into the question of which parameters do matter in the case of group formation. Thus, two hypotheses are presented:

H1: Individuals in groups in which a heterogeneous distribution of extraversion is established algorithmically will report greater satisfaction with group composition and groupwork, invest more time on groupwork, and achieve better results than individuals in groups in which a homogeneous distribution of extraversion is established algorithmically.

H2: Individuals in groups in which a homogeneous distribution of conscientiousness is established algorithmically will report greater satisfaction with group composition and group work, invest more time on groupwork, and achieve better results than individuals in groups in which a heterogeneous distribution of conscientiousness is established algorithmically.

In addition to the main hypotheses, we formulate an open-ended, explorative hypothesis. It states that an interaction effect can be expected to occur between the two experimental variables without predicting its direction.

Method

Sample

The study presents a systematic, fully crossed experimental design with two characteristics (extraversion, conscientiousness) that are manipulated in two levels (homogeneous/ heterogeneous) overall groups. A total of 751 participants (female = 172) were recruited in a voluntary university-preparation-course at a technical university in Germany. Informed consent was obtained in writing from all subjects. No exclusion criteria were determined prior to participation. Those students, who made the decision to work in groups, were included in the analyses. After acquisition, 372 participants were matched to groups of three, leading to 124 groups.

Study Environment

The online pre-course offered, in addition to the 72-module-contents, a diagnostic input-and-output-test, which each of the participants had to take. The entrance test provided individualized computer-supported feedback through an adaptive test scenario and could be differentiated through specific module recommendations to meet the needs of the students in follow-up and further learning. The preliminary course concluded with a final test, which was intended to enable students to visualize their individual learning progress within the framework of self-monitoring.

In this study, a module consists of an appropriate pre-test, the module content (consisting of introduction, explanation, possible applications, as well as error potentials and tasks) and a retest. The module-structure allows for several approaches to processing. Students who want to deal with a topic have the option of starting with the introduction and explanation, whereas students who prefer to merely look up the summary or short information are free to do so. In this way, the different modules ensure a high degree of internal differentiation. Furthermore, the learning groups were given assignments, that had been specially designed for the groups. In addition to the learning modules, the preliminary courses included weekly group tasks, assignments, and so-called warm-ups, where the content differed according to the target group of the preliminary course. Within their learning group, students completed the group assignments and submitted them. Tutors evaluated these tasks. Weekly and thematically different assignments and warm-up tasks were intended to ensure that mathematical content and skills were regularly kept alive and repeated. To be able to cover subject-specific content, the online preliminary course in its realization differentiated between students of mathematics and computer science and students of engineering. The differentiation was based on different tasks and focuses within the preliminary course and a varying recommendation regarding the sequence of the 72 different module contents.

Instruments

We chose Moodle-Software to provide the online learning environment. Moodle is open-source based, therefore freely available, and a global software-development-project. Moodle is both a course management system and a learning platform that can be installed on any computer operating system. The software offers the possibilities to foster the use of cooperative teaching and other learning methods. Students can use these virtual meeting rooms to access working materials and learning activities. Additionally, the software can be extended by implementing and installing plugins. For practical purposes the plugin MoodlePeers, which implements the algorithm named GroupAL, was developed.

MoodlePeers is an activity-based plugin that takes over the algorithmic formation for learning groups. The plugin is published as Open-Source-Project under GNU General Public License and is available in several versions at Moodle.org. In MoodlePeers instructors can choose between three types of learning-group-scenarios, as well as the contents of the questionnaire and the criteria to be considered for grouping. Additionally, instructors can determine the maximum number of groups created or the maximum group size. The complete questionnaires can be edited only with developer support. Instructors can however add a question regarding prior knowledge. After the initial configuration of the group formation activity, the students can take the selected questionnaire, after they have agreed to the conditions of data collection, to participate in the group formation process. As demonstrated as part of this study, MoodlePeers is an activity managing group formation by gathering and evaluating the required data using questionnaires.

Students participating were asked to fill out a demographic and psychological online questionnaire via Moodle at the beginning of the course, which included questionnaires regarding their personality (Big Five, BFI-K (Rammstedt & John, 2005)), prior knowledge (self-estimation on every subdimension of mathematic content from school), their motivation for the course (Motivation, EVC (Kosovich et al., 2015)), and team orientation. The personal questionnaire was shown to have robust reliabilities in this setting (extraversion: eight items, $\alpha = 0.89$; conscientiousness: nine items, $\alpha = 0.83$; openness: five items, $\alpha = 0.74$; neuroticism: four items, $\alpha = 0.77$; agreeableness: four items, $\alpha = 0.61$). Motivation was measured within four subscales: expectations (four items, e.g., "I know that I can learn the contents of the preliminary course," $\alpha = 0.80$), use (five items, e.g., "I understand how important the preliminary course is for my future," $\alpha = 0.80$), cost (six items, e.g., "The time required for the preliminary course seems great to me," $\alpha = 0.83$), and interest (seven items, e.g., "I'm looking forward to the preliminary course," $\alpha =$

0.77). Team orientation was measured using three items (e.g., “If I have a choice, I’d rather work in a team than alone,” $\alpha = 0.89$). All questions were rated online, using a scale from 1 (not true) to 6 (true).

After all students completed the questionnaires, the results were used to complete the group formation based on previously defined criteria with the help of GroupAL. The answers to their questions are shown to the students as individual feedback (i.e., score of a personality trait). When data collection was completed, the results of the questionnaire can be used to calculate the various criteria used in the group formation process and display the resulting groups on Moodle.

GroupAL depicts an algorithm to optimize matching individuals based on different criteria in maximally efficient learning groups. It is a nonlinear optimization algorithm optimizing two performance indices. The underlying structure allows comparison of different generated group formations and is robust against variations on the number of criteria or changes in the underlying cohort of participants. It makes use of a numerical technique based on N-dimensional vectors per criterion. The weighted criteria can be optimized accordingly: homogeneously (1) or heterogeneously (0). For this purpose, a distance metric is used, which calculates in pairs the fit (homogeneous) or complement (heterogeneous) of the group members in the criterion dimensions.

First, the GroupAL-Algorithm divided the students randomly into four conditions. Afterwards, groups were heterogeneous or homogeneous in extraversion, and simultaneously heterogeneous or homogeneous in conscientiousness. Groups were matched homogeneously in regard to their motivation and prior knowledge, to start with the same conditions regarding those characteristics. The GroupAL algorithm was intended to divide consistently, coming out evenly in groups of three people in the ideal case, whereby the four different group conditions were established (see Table 1).

Table 1. *Grouping Scheme, the Algorithm Applies to Match in Groups*

	Conscientiousness		
Extraversion	**Homogeneous**	**Heterogeneous**	**Total**
Homogeneous	$N = 108$ ($K = 36$)	$N = 108$ ($K = 36$)	$N = 87$ ($K = 29$)
Heterogeneous	$N = 90$ ($K = 30$)	$N = 87$ ($K = 29$)	$N = 177$ ($K = 59$)
Total	$N = 198$ ($K = 66$)	$N = 174$ ($K = 58$)	$N = 372$ ($K = 124$)

Note. N = students; K = groups.

The final evaluation of the GroupAL-Tool resulted in better outcomes than other algorithms, offered better cohort performance indices, and more tools for group formation with higher quality under the chosen conditions and with the selected data sets (Konert et al., 2014).

Assignments

This section describes the group tasks and shows how the dependent variable performance is derived from the result. The group assignments aimed to ensure group participants practiced mathematical communication, problem solving, modelling, and argumentation. The two online course rooms (mathematics and computer science; engineering) received different group assignments, adjusted to the students' requirements. Whereas for students of engineering, mathematical modeling and basics were selected as the focus, for students of mathematics and computer science, mathematical argumentation and the introduction of logical inference constituted the focus.

There were three different group-assignments in each of the two course rooms, each of which was made available to students in 1-week intervals in Moodle. Engineering students, for example, were required to develop a proposal on how to construct a volume-maximized open cube from an A4-sheet of paper. The variety of solutions to the task was intended to stimulate group processes and communication. In the same week, students of mathematics and computer science were given the group task of examining and discussing the Dirichlet function on different definition sets regarding consistency and differentiability. The tasks were intended to deal with basic technical content on a higher, formal level. The motive for these choices was to create occasions for discussion that demand and promote the skills of communication and argumentation, but also to enable an approach to university-ways of thinking and presentation. Each group exercise was corrected by a tutor based on an evaluation key provided in advance. Each exercise yields a maximum of 10 points, which were awarded for various necessary argumentations and/or calculations. In the case of open tasks with no clear-cut solution, the type and quality of the students' argumentation and calculations was assessed. The evaluation keys were prepared in advance by the pre-course management.

Procedure

The course lasted for 4 weeks between September 10th and October 5th, 2018, before the participants' first semester started as a chance to improve their mathematics skills. The students had access to the course structure via Moodle, where they found all the instructions and tests. Participants were encouraged to use the group tools in a welcoming video, and open and complex weekly assignments should force groupwork being necessary. Students were able to choose, if they wanted to work online in a group, to work online alone, or to work in groups in a face-to-face course. The last option though failed. After completion of the first questionnaire, participants were informed about the formation of their learning group, and each member of the group received access to an online bulletin board for group communication via chat or forum.

During the 4-week-course, participants completed a test independently at the very beginning of the course as well as at the end of the course in addition to three group-assignments (number of turn in [max. 3]; number of points [max. 10 points per assignment]). An overall performance score was calculated for the combined results of all 3 weeks of assignments, as well as the number of submissions of the three assignments. Students needed to answer a short evaluation with the turn-in of each assignment. There they were asked questions about their satisfaction regarding groupwork (e.g., “I am satisfied with the cooperation in my group”), productivity of the group (e.g., “Our group has worked productively”), time investment (e.g., “How much time [in minutes] did you personally spend on individual preparation?”), frequency of meetings (e.g., “How many personal meetings have you had with your group in the last week?”) and communication (e.g., “How many members of your study group did you communicate with in any way this week?”).

Results

Analyses were conducted to test the hypotheses, as to say that heterogeneous group formation in extraversion and homogeneous group formation in conscientiousness will lead to positive outcomes concerning assignments, turn in, achieved points in assignments, satisfaction, and frequency of communication. Additionally, we provide two different analyses to answer the question, if either heterogeneous or homogeneous group composition of conscientiousness and extraversion, their interaction, or any other explorative predictor leads to any effect on the outcome variables. We added analyses to explore, if variance can be explained by groups or the individual-level, and if the experimental formation of groups would yield successful outcomes. We calculated rmANOVAs using SPSS23 V5 and MLMs using MPlus V5.

Results of the rmANOVAs

First, we focus on the effect of the experimental condition’s extraversion and conscientiousness and their effect on the performance-related variables and outcome measures and note that here the grouping shows no significant major effects. An interaction effect of extraversion × conscientiousness on the homework evaluation is to be mentioned: $F(2,16) = 4.7$, $p = .03$, $\eta 2 = 0.03$. This suggests that groups homogeneous in both traits achieved the best rating (~3 points), followed by heterogeneous trait grouping. The lowest scores were from groups that were homogeneous in extraversion and heterogeneous in conscientiousness (~2.3 points). We cannot report significant results on individual performance (pretest, posttest).

No significant effects of experimental condition and self-rated evaluation of time investment could be found; therefore, the grouping showed no effect on the indication of the frequency of communication and participation of all members. However, the experimental conditions showed significant interaction effects on satisfaction. More specifically, belonging to a group with a different experimental design had an effect on the self-rated satisfaction of the subjects in the evaluation (satisfaction, measurement time point two, $F(2,15) = 9.25$, $p < .01$, $\eta 2 = 0.01$); and satisfaction, measurement time point three, $F(2,99) = 10.28$, $p < .01$, $\eta 2 = 0.01$). At the time of the first survey, the level of satisfaction was not high, but satisfaction increased from the second to the third evaluation. The self-rated satisfaction was highest in groups with heterogeneously distributed conscientiousness and heterogeneously distributed extraversion within the groups; the lowest degree of satisfaction was found in groups with heterogeneously distributed extraversion and homogeneously distributed conscientiousness.

Multilevel Models

When researchers apply standard statistical methods to multilevel data such as the regression model, the assumption of independent errors is violated. For example, if we have points awarded on homework as part of students' respective groups, it would be reasonable to assume that members of the same group will have points, that are more highly correlated between one another in contrast to those in other groups. This within-group correlation would be due, for example, to the shared experience, the teaching curriculum, and a single set of administrative policies. The within-group correlation will in turn result in an inappropriate estimate of the standard errors for the model-parameters, which will lead to errors of statistical inference, such as p-values smaller than they should be, and the resulting rejection of null effects above the stated Type I-error rate for the parameters.

In addition to the underestimation of the standard error, another problem with ignoring the multilevel structure of data is that we miss important relationships involving each level in the data. Therefore, they allow us to answer substantive questions about sources of variation in our data (Hox et al., 2018). In sum, three-level longitudinal models (MLM) offer a choice of which level to use to randomize to experimental conditions. MLM, with time (Level 1) nested for individuals (Level 2), and individuals nested within groups (Level 3), was tested. The term nested is used, as each student only learns in one group, and each group is doing evaluations and group tasks over a period (three events to be nested). Additionally, MLMs were run between the four different group compositions, due to the established variable of interaction (multiplication of experimental conditions; coded 0, 1) to show whether grouping in regard of heterogeneous or homogeneous contributions of extraversion and conscientiousness had any effect on the above-mentioned outcomes. It was additionally controlled for as a selection effect of the sample by adding a random slope for maintenance. Models of the slope are not included in the result section since the result did not differ marginally from model two. Significant predictors remain their significance, so we assume there will not be a selection effect. First, we calculated the errors of the covariant structure. Secondly, we sort the indicator for the group-level (Group ID). Thirdly, group means were centered for variables that had not a natural zero point (e.g., personality, satisfaction, or the estimation of the frequency of time spent together). The decision, how to center data, depends on the researchers' theory-driven considerations, corresponding to their respective postulated, suppositionally focused, intended prediction of results. Analyzing the data, we used grand mean centering to simplify the interpretation of results. The main reason for doing so corresponded to the research question of how individuals in defined group formations evaluate their situation as well as how they perform. As a result, we focused solely on the individual-level. Nonetheless, the use of multilevel structures allowed us to consider group structures in the analyses of nested data. Opposingly, the group-mean-centered model leads to a second level-coefficient, where effects at the individual-level can be mistaken for effects at the group/context-level (Wu & Wooldridge, 2005).

As an initial research step, we chose not to alter the data's information by any centering, but rather to present it in accordance with its structure. Yet, an additional reason behind our decision to take such a course of action arose from disagreements between the different centering methods published in the various literature (Braun et al., 2020; den Hartog et al., 2019; English et al., 2004). The empty-or-null-level-model describes the partition between variance at the student's level and at the group-level.

In our data set, the empty model for satisfaction partitions approximately 30% of the variance into the individual-level and approximately 39% of the variance into the group-level. In the case of the variable frequency of communication the empty model partitions approximately 13% of the variance into the individual-level and approximately 27% of the variance into the group-level. For the turn in of assignments, approximately 4% could be explained at the individual-level, and 40% could be explained at the group-level. For performance on assignments, the empty model partitions approximately 0% of the variance into the individual-level and 41% of the variance into the group-level, approximately, since assignments were part of the groupwork, turn-in and grades did not differ much across group members. The turn-in of the assignment did not explain any more variance and was therefore not included. Given the hypotheses that either heterogeneously structured extraversion or homogeneous structure of conscientiousness will lead to positive outcome measures, we first focused on factors at the group-level, as to say the manipulation of the experimental variables within groups. But only a minor impact was revealed. Their structure and their interaction did not show any significant effects on the outcomes on the group-level.

The results for the models are shown in Tables 2 to 4. Table 2 presents the results for the scores on the homework assignments, Table 3 presents the frequency of communication, and Table 4 shows satisfaction. If a relationship between two variables is nonlinear, the rate of increase or decrease can change simultaneously with changes in the values of a variable, resulting in a curved pattern in the data. Such a curvilinear trend may be better modeled with a nonlinear function, such as a quadratic or cubic function, or the relationship may be made linear by a transformation. In other words, the positive effect of a variable decreases again at high values, resulting in a u-shaped (or inverse u-shaped) relationship. To control for nonlinearity of our data, we additionally included the quadratic predictors in Model 3. All tables are subdivided in individual-level and group-level predictors. On the group-level the experimental conditions are tested. The sizes of the coefficients reflect the relative importance of the explanatory variables in the model.

Table 2. *Individual-Level and Group-Level Predictors of Points on Homework.*

Predictors	Model 1	Model 2	Model 3
Intercept	1.16 (0.07)*	1.19 (0.37)**	1.17 (0.39)**
Level I (Time)			
Level II (Individual)			
Gender		−0.02 (0.07)	−0.03 (0.08)
Age		0.00 (0.01)	−0.00 (0.01)
Prior knowledge		0.00 (0.00)	0.00 (0.00)
Extraversion		−0.03 (0.03)	−0.03 (0.03)
Conscientiousness		−0.02 (0.04)	−0.02 (0.04)
Agreeableness		0.07 (0.03)*	−0.08 (0.04)*
Neuroticism		0.01 (0.03)	0.01 (0.03)
Openness		0.01 (0.03)	0.01 (0.03)
Team orientation		0.01 (0.02)	0.04 (0.03)
Maintenance		0.03 (0.09)	0.03 (0.09)
$Extraversion^2$			0.05 (0.02)*
$Conscientiousness^2$			0.06 (0.04)
$Agreeableness^2$			0.02 (0.02)
$Neuroticism^2$			−0.02 (0.02)
$Openness^2$			−0.01 (0.03)
Team $orientation^2$			0.01 (0.02)
Level III (Group)			
Criterion Extraversion		0.02 (0.20)	0.20 (0.15)
Criterion Conscientiousness		0.10 (0.18)	0.33 (0.16)
Interaction		−0.07 (0.27)	−0.26 (0.23)
Variance			
GL variation of the DV	0.00 (0.02)	0.35 (0.04)**	0.36 (0.06)**
GL variation in IL effect of DV	0.41 (0.06)**	0.00 (0.01)	0.07 (0.04)
GL variation in IL effect of DV2		0.23 (0.05)**	0.11 (0.04)**
−2 log-likelihood	754,040	732,612	729,224

Note. Gender coded: 1 for men and 2 for women. Criterion of experimental manipulation for group formation (Level III); Criterion_Extraversion/Conscientiousness = Intervention for group formation; heterogeneous = 0 homogeneous = 1. Personality type coded 3 for high level, −3 for low level. Squared values marked². Unstandardized coefficients reported. Robust standard errors reported in the parentheses. Missing data handled with case deletion. *p = .05; **p = .01.

Table 3. *Individual-Level and Group-Level Predictors of Self-Rated Frequency of Communication.*

Predictors	Model 1	Model 2	Model 3
Intercept	−0.08 (0.08)	−0.43 (0.48)	1.98 (0.51)**
Level I (Time)			
Level II (Individual)			
Gender		0.12 (0.11)	0.15 (0.10)
Age		−0.01 (0.02)	−0.01 (0.02)
Prior knowledge		0.00 (0.00)	0.00 (0.00)
Extraversion		0.11 (0.05)*	0.10 (0.04)*
Conscientiousness		−0.00 (0.06)	−0.00 (0.06)
Agreeableness		0.06 (0.07)	0.07 (0.07)
Neuroticism		0.12 (0.05)*	0.14 (0.05)**
Openness		−0.03 (0.04)	−0.04 (0.04)
Team orientation		0.07 (0.05)	0.08 (0.05)
Maintenance		0.12 (0.11)	0.12 (0.10)
Extraversion²			0.02 (0.04)
Conscientiousness²			0.12 (0.05)*
Agreeableness²			0.05 (0.05)
Neuroticism²			−0.01 (0.03)
Openness²			−0.07 (0.04)
Team orientation²			−0.01 (0.03)
Level III (Group)			
Criterion Extraversion		0.18 (0.15)	0.20 (0.15)
Criterion Conscientiousness		0.32 (0.16)*	0.33 (0.16)
Interaction		−0.21 (0.23)	−0.26 (0.23)
Variance			
GL variation of the DV	0.13 (0.07)**	0.37 (0.06)**	0.36 (0.06)**
GL variation in IL effect of DV	0.27 (0.08)**	0.08 (0.04)	0.07 (0.04)
GL variation in IL effect of DV2		0.12 (0.04)**	0.11 (0.04)**
−2 log-likelihood	891,948	852,936	845,654

Note. Gender coded: 1 for men, 2 for women. Criterion of experimental manipulation for group formation (Level III); Criterion_Extraversion/Conscientiousness = Intervention for group formation; heterogenous = 0, homogeneous = 1. Personality type coded 3 for high level, −3 for low level. Squared values marked². Unstandardized coefficients reported. Robust standard errors reported in the parentheses. Missing data handled with case deletion. **p* = .05; ***p* = .01.

Table 4. *Individual-Level and Group-Level Predictors of Self-Rated Satisfaction.*

Predictors	Model 1	Model 2	Model 3
Intercept	−0.08 (0.08)	−0.02 (0.79)	3.83 (0.79)**
Level I (Time)			
Level II (Individual)			
Gender		−0.07 (0.13)	0.02 (0.14)
Age		−0.02 (0.03)	−0.01 (0.03)
Prior Knowledge		0.03(0.00)	0.00 (0.00)
Extraversion		−0.13 (0.08)	−0.14 (0.07)*
Conscientiousness		−0.03 (0.08)	0.02 (0.07)
Agreeableness		0.05 (0.07)	0.06 (0.08)
Neuroticism		−0.10 (0.06)	−0.09 (0.06)
Openness		0.07 (0.06)	0.07 (0.06)
Team orientation		0.19 (0.04)**	0.20 (0.05)**
Maintenance		0.25 (0.17)	0.25 (0.16)
Extraversion²			−0.05 (0.05)
Conscientiousness²			0.19 (0.06)**
Agreeableness²			0.00 (0.06)
Neuroticism²			0.01 (0.04)
Openness²			0.05 (0.05)
Team orientation²			0.02 (0.03)
Level III (Group)			
Criterion Extraversion		−0.15 (0.21)	−0.12 (0.21)
Criterion Conscientiousness		−0.35 (0.21)	−0.34 (0.21)
Interaction		0.41 (0.30)	0.41 (0.29)
Variance			
GL variation of the DV	0.30 (0.00)	0.24 (0.04)**	0.25 (0.04)**
GL variation in IL effect of DV	0.39 (0.00)**	0.23 (0.06)**	0.20 (0.06)**
GL variation in IL effect of DV2		0.23 (0.09)**	0.24 (0.09)**
−2 log-likelihood	894,798	849,384	840,318

Note. Gender coded: 1 for men and 2 for women. Criterion of experimental manipulation for group formation (Level III); Criterion_Extraversion/Conscientiousness = Intervention for group formation; heterogeneous = 0, homogeneous = 1. Personality type coded 3 for high level, −3 for low level. Squared values marked². Unstandardized coefficients reported. Robust standard errors reported in the parentheses. Missing data handled with case deletion. *p = .05; **p = .01

Discussion

The present study aimed to investigate students' online groupwork mechanisms. We tested, what kind of group formation yielded the best outcomes, by implementing an algorithm that allows optimal member matching criteria based on personality traits, extraversion and conscientiousness concerning their standard deviation. Hypotheses stated that better outcomes, and, therefore, better group formation, will be established through (1) heterogeneously matched extraversion and (2) homogeneously matched conscientiousness within the groups. Contrary to expectation, the hypotheses of the study were not supported. Regarding the results of the multilevel-model-analyses, it appears that the high variance explained on group-level is due to other, probably uncollected, and/or uncontrolled variables at group-level, which differ between groups, and not due to the structuring of experimental variables and group formation being used. However, the low level of impact on learning of the group-level factors, compared to student-level factors, is an important finding.

Interestingly, some of the individual predictors explain parts of the variance and are significant. For example, for the outcome measure of frequency of communication, high extraversion and high neuroticism tend to lead to high frequency of communication within the groups. It could be supposed, that in a computerized environment students may feel freer to indicate their personal attributes with time, more confidently and without interference by others. Accordingly, it was found that CSCL may reduce students' anxiety that stems from face-to-face cooperative debate and motivates the shy students to be self-conscious and more eager to work in this unthreatening environment (AbuSeileek, 2007, 2012). Again, results for the outcome-measure-frequency of the manipulation of conscientiousness were found to be of marginal significance. Team orientation is a significant trait on the individual-level of the evaluation over time for self-rated satisfaction with the groupwork. Overall, it can be said that the matching does not show a uniform result on the outcome variables. In addition, Springer et al. (1999), indicate that conflicting results relating to the significance of factors affecting student performance in mathematics can be found.

The additional third model of multilevel analyses, in which potential curvilinear relationships were considered, had a better model fit for all three outcome variables. Nevertheless, the significant predictors to be reported differed only slightly from those in Model 2. The deviations that occurred when observing the curvilinearity of the model are reported in the following. For the scoring of homework, a significant effect is shown for the squared variable of extraversion; for the frequency of communication the squared predictor variable conscientiousness is significant. In the model for self-reported satisfaction, extraversion and the squaring of conscientiousness are significant predictor variables, due to the consideration of curve-linearity. Even if these differences are not substantial, this indeed suggests that the relationship between personality traits and the outcome of group work cannot be assumed to be linear per se. Results support previous studies, which have argued that the investigation of the relationship between personality and work behavior should go beyond linearity-assumptions (Curşeu et al., 2019).

It appears to be beneficial for performance to consider the level of performance and personality traits of group members when forming groups, as opposed to self-selection or random grouping (Bekele & Menzel, 2005). Greater diversity in personalities within the group has proved to have a positive influence on overall performance (Roberge & van Dick, 2010). Contrary to this, results of these analyses indicate that successful and unsuccessful teams were homogeneous in terms of different characteristics (Wax et al., 2017). Even though individual Studies often show marked differences between homogeneous and heterogeneous groups, the integration of results by a meta-analysis reveals the combined effect sizes of these Studies not to be significant. It is possible that significant effects that have been found appear to be explained by the type and difficulty of the task used in the Studies (Bowers et al., 2000).

However, it is suggested that ideal grouping could not be carried out following a conceptual model, stating that the nature of specific traits needs to be considered for optimal results. It follows that, in order to create well performing groups, many variables of the individual group members may be considered to optimize the group match. Consequently, researchers and practitioners have a multitude of opportunities to create different grouping models considering individual trait-constellations, while setting the focus on different characteristics than the two specified in the full theoretical model of this study (Humphrey et al., 2007). Considering that groups are often formed with the intention to perform tasks in the same constellation throughout a longer period. Variables for team composition may transform over time, expression of character traits intensify and gain relevance, and thereby alter team performance. Team conscientiousness in particular seems to affect team performance over time, seeing as high levels of conscientiousness within the group has individuals in dedicated task-role-behavior, promoting completion of the task at hand (Peeters et al., 2006). Bearing in mind that our study setting had a time frame of 4 weeks, we cannot account for longer-term development of the group dynamics, as characteristics may not have had the time to unfold and establish themselves. A further limitation to this may be manifested in the importance of established group roles and their influence on performance about the various elements of the task, which denote to be higher than formerly presumed (Lisak & Erez, 2009).

A prominent focus should also be set on the processes, as opposed to merely outcomes, of learning and interaction, especially regarding CSCL. Quality of interaction and learning outcomes show substantial variations according to a wide range of research results (Strijbos et al., 2004), caused by and large by differences in group sizes, the utilized technology, extent of the study, units of analysis and the research methodology (Lipponen, 2002).

Limitations

In this study, groupwork took place exclusively in the virtual form, with relatively low intensity. Homogeneity and heterogeneity of the group may play a less important role with this set-up because members will not reach a level of familiarity with each other that allows them to compare the similarities and differences among them. A designated leader, characterized by a combination of high extraversion and high conscientiousness, to motivate and initiate the group tasks, could prove essential in this setting, where groups, assigned heterogeneously in extraversion and conscientiousness, have a higher probability to contain a suitable person.

A lack of consistency in previous Studies may have various reasons, and may relate to the limitations of the study design. Formation effects in groupwork for instance have been studied mostly in correlation-designs, and external validity could have been compromised by the reduced sample size after dropout. In addition to the limitations outlined, most research findings derived exclusively from questionnaire formats based on student self-reports. Presumably, our future research will be to include the learning environment presented, the nature of the task, and the establishment of an in-group outcome goal. In conclusion, the limitations reveal opportunities for further research, which we propose to integrate in future study designs.

Ultimately, we have reason to assume that the way in which personality traits unfold in groups as roles could not develop completely within the short study frame and distance between learners; therefore, measuring it appears redundant. One could argue that different personality traits than those that were found to impact face-to-face groupwork might be even more important in an online setting, as was suggested by the explorative findings of effects from neuroticism.

Conclusion

The strength of the present study is certainly the experimental matching of groups, which allows to draw causal conclusions from the experimental matching of groups before the actual groupwork. The results of the study indicate that additional variables beyond the selection that we had considered in our measures could have a greater influence on the effect. The advantages of multi-level modelling need to be mentioned. Thus, in contrast to regular regression, the average variation between levels is not ignored. Individual regression can be associated with sampling problems and lack of generalization. Further research is needed on the algorithmic formation of individuals in groupwork. More attention needs to be paid to the creation and allocation of learning environments, the identification of dysfunctional groups, and the implementation of guidance and the correct assessment of successful learning outcomes. An inclusive overall solution with guidelines for effective groupwork and its formation is needed.

6.1.1 Summary of Study 1 and Motivation for Study 2

Study 1 investigated the impact of conscientiousness and extraversion as criteria for group formation in a four-week online group-work-course. However, the results did not align with the hypotheses, revealing no significant main effect of experimental grouping, but suggesting interaction effects between extraversion and conscientiousness. However, methodological limitations, including a high dropout rate and potential statistical power issues, influenced the study's outcomes. Notably, caution in interpretation was necessary due to acknowledged limitations linked to the learning environment, task complexity, and the absence of a defined group goal.

Building upon the identified limitations of study 1, the design of study 2 aimed for a methodically stronger exploration of online groupwork. The research was conducted with a new cohort of students in the subsequent academic year following study 1. Given the suspected selection effects associated with conscientiousness (Verbeek & Nijman, 1996) in study 1, study 2 opted for prior knowledge, a well-established and widely-known predictor of success in group work. While study 2 shared similarities in research design with the preceding study 1, the research design here was simplified to a 2 × 2 format, excluding random conditions. The aim was not to compare the results of random group formation, but to examine the effects of homogeneous and heterogeneous distributions of prior knowledge and extraversion on group work outcomes, thereby breaking down the research design to its essential components. Additionally, in response to the high dropout rate observed in study 1, study 2 addressed this issue by attempting to create a larger number of experimental groups to mitigate its impact.

6.2 Manuscript 2: Müller, A., Bellhäuser, H., Konert, J., Röpke, R., & Genc, Ö. (accepted). Group Formation by the Means of Extraversion and Prior-knowledge as Important Predictor in Higher Education. *Journal of Computing in Higher Education*.

Group Formation by the Means of Extraversion and Prior Knowledge as Important Predictor in Higher Education Online

Abstract

The study investigates how the 2 × 2 configuration of homogeneous and heterogeneous distributions of extraversion and prior knowledge influences group outcomes, including satisfaction, performance, and stability Based on the standard deviation of extraversion and prior knowledge, groups were established to test experimentally, what form of grouping leads to best outcomes. The randomized controlled trial took place in the context of an online course with 355 prospective students, working in 82 groups. The two characteristics extraversion and prior knowledge were distributed algorithmically, either homogeneously or heterogeneously. Results showed no superiority of heterogeneous formation, yet there were systematic interaction effects by the experimental group formation on satisfaction and performance. Due to the increasing relevance of online groupwork, explorative results are reported and integrated. Ideas for future research on group formation as an important influencing factor are discussed. Findings supports knowledge about cooperative online learning by optimizing the selection of group members using a therefore implemented algorithm.

1. Introduction

The objective of this study presented is to investigate the effects of algorithm-based group formation by homogeneous and heterogeneous distributions of prior knowledge and extraversion on online group-work-outcomes, including satisfaction, performance, and group stability. For this purpose, we utilize a 2 x 2 research design to assess which type of group formation leads to the best results by the preset criteria for matching group members.

Group work has long been an important didactic tool for promoting learning at various levels and has already proven its worth as such (Lin et al., 2016; Mujkanovic & Bollin, 2019). In the wake of the Covid-pandemic, the need for pedagogical methods for Computer-Supported-Collaborative-Learning (CSCL) has increased (Hodges et al., 2020). For online learning to be successful, it is important to include social elements (Gillen-O'Neel, 2021; Wildman et al., 2021). A key solution to creating good starting conditions for online learning for each student is to create groups (Gillies, 2004). Potentially disadvantaged students could also be identified and targeted through research on forming groups based on various criteria (Chahal et al., 2022; Hachey et al., 2022). However, we already know from previous research that groupwork does not always benefit every learner (Chang & Brickman, 2018).

A crucial aspect seems to be the way in which groups are formed, which has a major impact on their success or failure (Bellhäuser et al., 2018; Müller et al., 2022). Therefore, there is increasing interest in research on what criteria could be used for group formation. This research trend on group formation has already been shaped by the increasing prevalence of online-based learning with large numbers of users (e.g., Massive Open Online Courses [MOOCs]), which require algorithmic support for group formation.

Research on group formation can be found predominantly in the field of computer science (Borges et al., 2018; Liang et al., 2021; Maina et al., 2017; Odo et al., 2019). Still, a critical look at this research literature from a psychological point of view reveals that the measuring instruments often do not meet common psychometric requirements (Kirschner, 2017). Additionally, in many cases, the correlative research designs used do not allow causal conclusions. Experimental research on the outcomes of group formation of students at the university is still missing (Bell, 2007; Nijstad & de Dreu, 2002; Thanh & Gillies, 2010). As a desideratum of research, we deduce that interdisciplinary approaches are necessary to meet an optimal group-formation-challenge (Bellhäuser et al., 2018; Houlden & Veletsianos, 2022; Müller et al., 2022). The aim of this research is therefore to systematically evaluate, how homogeneous and heterogeneous distributions of extraversion and prior knowledge, configured in a 2 × 2 research design, affect satisfaction, performance, and group stability.

1.1. Virtual Learning in Groups

The COVID-19 pandemic abruptly compelled many students to transition to remote learning, making online groupwork a pertinent and promising tool (Houlden & Veletsianos, 2022). Groups, viewed as complex adaptive systems (Ramos-Villagrasa et al., 2018), exhibit internal cohesion termed a „we-feeling" (Stürmer et al., 2013). A virtual group, defined as individuals geographically, organizationally, and/or temporally dispersed, collaborating on organizational tasks (Powell et al., 2004), surpasses constraints of time zones, distances, and organizational boundaries (Lipnack & Stamps, 1999). Previous research strongly advises against comparing online and face-to-face learning groups, particularly regarding group formation and outcomes (Atchley et al., 2013). The pandemic underscored the significance of online group work, necessitating further research (Williams & Castro, 2010) to delve into the social interactions influencing virtual groups (Hwang et al., 2013; Montoya-Weiss et al., 2001).

1.2. Challenges and Opportunities in Online-Group-Work-Research

Despite extensive research on group work and attempts at improvement through formation (Borges et al., 2018), a standard model for group formation lacks consensus in the literature, hindering consistent and beneficial comparisons (Clark et al., 2019; Magpili & Pazos, 2018). Computer-Supported Collaborative Learning (CSCL) is a widely used strategy in online-supported university teaching, yielding performance advantages (Johnson et al., 1991) and enhancing emotional motivation (Cleveland-Innes & Campbell, 2012). It also has the potential to mitigate social isolation prevalent in digital learning contexts, positively influencing learner satisfaction (Liu et al., 2020; Mehall, 2021). The online environment's anonymity can aid in overcoming social anxiety, encouraging participation by silent or shy members (Kerr & Tindale, 2004).

However, socio-technical challenges may arise in online working groups (Montoya-Weiss et al., 2001). Virtual learning environments exhibit higher dropout rates compared to traditional settings (Diaz, 2002; Yang et al., 2013). Students in virtual spaces may feel lonely and isolated, demotivating them and increasing the likelihood of course abandonment (LaRose & Whitten, 2000). Issues such as low engagement from others when questions are posed may impede a sense of belonging, potentially leading to failed group work (Conole et al., 2008; Erez et al., 2013). Prolonged videoconferencing can result in „zoom fatigue" (Nesher Shoshan & Wehrt, 2021). Online environments often neglect to support social processes (Krejins et al., 2002), crucial for collaborative task-solving (Lou et al., 2001). Incorporating cooperative learning elements increases interactivity, reduces feelings of isolation, and can counteract low participation and high dropout rates in online courses (Khalil & Ebner, 2014; Liu et al., 2020).

1.3. Group Formation Based on Algorithmic Assistance

Group formation significantly influences learning-group success (Bell, 2007; Halfhill, et al., 2005). It involves assembling groups through criteria-based member selection, while group composition refers to processes post-formation (Tuckman, 1965). Relevant criteria for group composition include demographic aspects, personality traits, attitudes, and cognitive preconditions, with either homogeneous or heterogeneous distributions considered advantageous depending on the criterion (Bowers et al., 2000). Productive interaction among learners often does not occur spontaneously, necessitating criteria-based group formation. However, such research is limited due to its association with challenging selection procedures, prompting a need for interdisciplinary research on group formation, including development, criteria, and algorithm evaluation (Dincă et al., 2021).

CSCL is prevalent across disciplines, and recent research has focused on group formation in computer science and interdisciplinary projects. The demand for online tools, including MOOCs, requires algorithmic support, but experimental studies are scarce, and common recommendations for group formation are lacking (Magpili & Pazos, 2018). Standardized reporting methods for studying online collaboration are also absent in the literature (Hachey et al., 2022; Pai et al., 2014).

1.4. Characteristics used for Group Formation

Group formation within collaborative learning environments is influenced by two subcategories of variables: surface and deep-level criteria. Surface-level variables, including demographics such as gender and age, are deemed less critical for group success compared to deep-level variables, which encompass personality factors, values, and attitudes (Bell, 2007; Harrison et al., 1998, 2002; LePine et al., 2011). While demographics provide insights into group composition, it is the deeper traits that significantly impact group performance.

The selection of single-group-member-attributes as criteria for group formation can result in either homogeneous or heterogeneous constellations. Homogeneous groups, characterized by similarity among members, often foster comfort, productivity, and friendly behavior, leading to a preference for homogeneous fit (Ilgen et al., 2005; den Hartog et al., 2019; Muchinsky & Monahan, 1987). Conversely, heterogeneous groups, with diverse member attributes, contribute unique skills to the collective, enhancing overall performance (Bekele & Menzel, 2005; Cable & Edwards, 2004; Moore, 2011). While research on group formation varies, studies emphasize the significance of personality traits over demographics (Martin & Paredes, 2004).

This study examines group formation based on the combination of personality trait extraversion and prior knowledge skill level. Beyond optimal group distribution, within-group determinants of extraversion and prior knowledge are explored under homogeneous and heterogeneous fits. The study evaluates these determinants' importance through objective outcome measures like academic performance and subjective outcomes such as satisfaction, forming a theoretical framework for the criteria used to form groups. In exploring group dynamics within collaborative learning environments, the integration of diverse demographic and personal information variables, including age, gender, average math grade and final school grade, is crucial for understanding group dynamics. Demographic variables offer insights into group interactions and outcomes, with age diversity linked to increased creativity and gender diversity associated with improved decision-making (Woolley et al., 2010). Additionally, academic indicators like math grade and final school grade influence individual contributions and interactions within groups (Cohen & Lotan, 1995). Personality traits, particularly the Big Five traits—extraversion, conscientiousness, openness, neuroticism, and agreeableness—shape group formation and dynamics. Extraversion impacts communication and social interactions, while conscientiousness influences task-oriented behaviors and productivity (Barrick & Mount, 1991). Motivation, assessed through factors like expectations and interest, drives engagement in collaborative learning experiences (Deci & Ryan, 1985; Kosovich et al., 2015). Additionally, team orientation preferences influence group dynamics, communication, and effectiveness (Harvey et al., 2019).

In conclusion, the incorporation of demographic information, personality traits, motivation factors, and team orientation variables provides a comprehensive understanding of the factors influencing group dynamics and collaboration within collaborative learning environments. By considering these variables, researchers can optimize group interactions and enhance learning outcomes in collaborative settings.

1.4.1. Extraversion

Extraversion is considered relevant to the formation of a group (Humphrey et al., 2007). It symbolizes a very interesting personality trait and is thus associated strongly with effectiveness (Hogan et al., 1994) and leadership behavior (Driskell et al., 2006; Judge et al., 2002; Nonaka et al., 2016). People with a low level of extraversion seem to be reserved and less involved in social situations (Power & Pluess, 2015). Since extroverts are more likely to assert themselves in groups, it follows that these individuals often take on leadership roles when working with other people (McCabe & Fleeson, 2012; Taggar et al., 2006). Concerning personality traits as group formation criteria, the literature assumes a heterogeneous distribution of extraversion within a group. Extraversion goes together with leadership-behavior and is differently pronounced among members (Kramer et al., 2014).

1.4.2. Prior Knowledge

Due to an ambiguous definition of prior knowledge, we should emphasize first of all prior knowledge's multidimensionality (Williams et al., 2008). Prior knowledge has had a major influence on the outcome of groups (Horwitz, 2005). The extraordinary role of prior knowledge, especially the activation of prior knowledge for learning, can be verified for the success of learning processes of young children (Saalbach & Schalk, 2011). Additionally, prior knowledge seems to be a decisive predictor of academic success (Riazy et al., 2021). Group members can share and increase prior knowledge within a group through contribution. Superiority of it in single members can cause a great added value for the whole group (Weinberger et al., 2007) and those synergy effects between the participants can significantly influence student achievement (Hailikari et al., 2008). We can assume approximately the same results for online groups (Engel et al., 2014, 2015).

Drawing a valid picture of group formation using the criterion of prior knowledge, we must consider the individual's preconditions. Most likely, low-ability students are more motivated to learn in heterogeneous groups, average-ability students perform better in homogeneous groups, and high-ability students show equal results when examining the effects of heterogeneous or homogeneous groups on the performance of pupils with high, average, or lower abilities (Saleh et al., 2005). Consequently, high, mid, and low-competence students would differ between heterogeneous and homogeneous groups in their learning outcomes (Donovan et al., 2018). Researchers found improved cooperative skills and performance in heterogeneously formed student groups based on intelligence and gender (Mehta & Kulshrestha, 2014). Although researchers assumed positive effects of determinants in children and young adults in the school-leaving age, the question arises, whether their observation allows for the prognosis for students.

1.5. Aim of Study

Interdisciplinary approaches are essential to address the challenge of optimal group formation. Since previous studies on online group formation techniques are scarce or limited to correlative settings, findings on the significance of extraversion and prior knowledge in online groups are lacking (Odo et al., 2019). Comparing results to understand the nature of groupwork collectively is complex (Magpili & Pazos, 2018), necessitating more attention and research in this area, as advocated by some authors, who called for further investigation in the link between personality in groups and group outcomes (LePine et al., 2011).

The assessment of group-work-outcomes should encompass various levels. It is not only performance that matters, but also the satisfaction of group members with their group dynamics and processes, as well as the group's duration over the course of the project. Course completion often serves as a measure of effectiveness (Hachey et al., 2013, 2022), highlighting the need for institutions to predict online students' persistence to address dropout rates. However, many studies focus on individual outcome measures exclusively, leading to a lack of standardization in the literature. Reporting diverse outcome measures is crucial for comparing study results. Thus, we examine the effect of extraversion and prior knowledge distribution on outcome variables such as satisfaction with group work, assignment performance, and group retention (referred to as "group stability"). This study investigates extraversion and prior knowledge as criteria for online group work in a university context. Building on previous findings, we hypothesize that heterogeneous group formation will be advantageous in a similar setting. Experimental studies exploring these criteria are currently unavailable. We have formulated our hypotheses accordingly:

H1: Individuals in groups, assigned heterogeneously in extraversion by algorithm, will be more satisfied with the group composition, produce better results in the assignment and spend more time on group work than those in homogeneously extraverted groups.

H2: Individuals in groups assigned heterogeneously in prior knowledge by algorithm will achieve better outcome measures (see above) compared to homogeneous grouping.

Additionally, we formulated an open, explorative research question assuming an interaction effect of both above-described measures.

RQ1: There will be an interaction effect between extraversion and prior knowledge for the outcome measures (see above).

2. Materials and Methods

2.1. Sample

We recruited participants from a four-week-online-preparation-math-course at the university in September 2019. This online math course is mainly for beginners of science, technology, engineering, and mathematics (STEM) subjects. It focuses on repeating the mathematical basics from school to improve the scholastic aptitude and reduce the heterogeneity of knowledge among the students. Students can do all the topics and tasks in the online math course voluntarily and in any preferred order; participating in the online math course does not result in a grade. We recruited participants (female = 172) and obtained their informed consent in writing. We did not determine exclusion criteria before participation. We included those students who made the decision to work in groups in our computation. After the acquisition, we matched 375 participants to groups of three, leading to 125 groups. To maximize the number of formed groups, we chose a group size of three members. We asked participants to work on weekly assignments and fill in evaluations of the quality of their groupwork. We also conducted a test at the beginning and the end of the course and a final evaluation after the course.

2.2. Study Design

The study presents a systematic, fully crossed experimental design with two factors (extraversion, prior knowledge) manipulated in two stages (homogeneous, heterogeneous) in all groups. We, therefore, have a between-subject-design with two factors (personality trait extraversion and prior knowledge) with two levels, respectively.

2.3. Instruments

After having consented to participate in the study, participating students were asked to fill out a demographic and psychological questionnaire at the beginning of the preparation course, which included questionnaires regarding their personality, prior subject knowledge, motivation for the course and team orientation. Participants answered the questionnaires online, using a Likert-scale from 1 (“does not apply”) to 6 (“does completely apply”).

2.3.1. Experimental Variables

Extraversion. The short version of the Big Five Inventory (BFI-K; Rammstedt & John, 2005) was used to assess the extraversion of the participants. The BFI-K was developed as a quick-response questionnaire that, with an average duration of processing of less than 2 minutes, can be considered extremely economical. It measures extraversion with 8 items, answered on a 6-point Likert-Scale, ranging from very inaccurate to very accurate. The validity between the BFI-K and the NEO-PI-R (Costa & McCrae, 1992) was established by Rammstedt and John (2005). Exemplary items for extraversion are "I am very enthusiastic" and "I am outgoing, sociable." Cronbach's alpha was $\alpha = 0.89$.

Prior Knowledge. We measured prior subject knowledge with participants’ self-estimation on every subdimension of mathematical content from the school. We based matching concerning previous knowledge on the result of the entrance tests, i.e., participants completed the entrance test before the end of the group formation. The entrance test focused on mathematical tasks students should solve to succeed in the first mathematical lectures. The entrance test is adaptive, so that each of the participants works on a different set of tasks based on whether they solve tasks correctly or incorrectly (Konert et al., 2016). The participants can score x points out y possible points, where x and y can differ for all participants. We then added the following two questions for the group formation and asked the participant to describe his “achievements score on the test” and the “maximum possible score on the test.” We used the entrance-test-score to calculate the quotient x/y of the number of points achieved and the number of points achievable. We used this score for grouping as the value for previous knowledge of the participants.

2.3.2. Control Variables

Demographics and Personal Information. We asked for participants’ age, gender, average math grade and average final school grade, as well as confirmed consent to participate in the current study.

Personality. The Big Five personality questionnaire (BFI-K; Rammstedt & John, 2005) demonstrated robust reliabilities in this setting (extraversion: 8 items, alpha = .89; conscientiousness: 9 items, alpha = .83; openness: 5 items, alpha = 0.70; neuroticism: 4 items, alpha = 0.79; agreeableness: 4 items, alpha = 0.64).

Motivation. We measured motivation (EVC; Kosovich et al., 2015) within four subscales: expectations (4 items, e.g.,” I know that I can learn the contents of the preliminary course,” alpha = .86), use (5 items, e.g.,” I understand how important the preliminary course is for my future,” alpha = .78), cost (6 items, e.g.,” The time required for the preliminary course seems great to me,” alpha = .83), and interest (7 items, e.g.,” I’m looking forward to the preliminary course,” alpha = .80). Reliabilities of the motivation scales were high.

Team orientation. We measured team orientation using three questions (e.g., “If I have a choice, I’d rather work in a team than alone,” alpha = .86.)

Honesty. We recorded the honest answering of the questionnaires with the question, “I have concentrated the questions and answered them honestly,” with the possible answers: “Yes, completely concentrated and honest,” “Yes, mainly concentrated, and honest,” and “No, not concentrated and honest at all.”. Only the last option led to the exclusion from participants.

2.3.3. Dependent Variables

The evaluation questionnaire contained questions of mainly satisfaction. Additionally asked in the evaluation were question regarding involvement and time spent (e.g.,” How much time (in minutes) did you personally spend on individual preparation?”,” How many personal meetings have you had with your group in the last week?”. A communication question included:” How many members of your study group did you communicate with in any way this week?”) which were not all included in the result section.

Satisfaction. The evaluation of satisfaction was done with an online evaluation questionnaire filled out by participants, as a precondition to group assignments turn-in. Questions included for satisfaction were for example: “I am satisfied with the cooperation in my group”,” Our group has worked productively”. We used the overall score of all answers regarding participant satisfaction as a result measure of satisfaction ranging from: 1 ("low satisfaction") to 6 ("high satisfaction").

Assignment. Homework handed in was graded for quality of the proposed solution by different previously trained student tutors. Homework needed to be turned in three times during the course. Grade point ranged from 0 to 10.

Group Stability. In addition, we used the number of all group homework assignments to be handed in during the preliminary course (absolute value = 3) as a key figure to obtain an objective measure of group stability and the possibility of making the stability of group cooperation measurable and portrayed over time.

2.3.4. Algorithm in Use to Perform the Group Formation

Moodle is an online e-learning platform used at the university where we conducted our study. To facilitate the chosen study design, we developed the plugin MoodlePeers, which implements the algorithm named GroupAL. The plugin is published as an Open-Source-Project and is available in several versions at Moodle.org.[1] For the two-factorial and two-stage experimental design, the algorithm has to meet the following objectives: extendable modelling and exchangeability of criteria, support for the formation of mixed homogeneous and heterogeneous groups across multiple criteria, and normed quality metrics for group formation and differences between the formed learning groups (Konert et al., 2016).

MoodlePeers has shown that non-linear optimization is a preferable method to semantic, ontology-based approaches for achieving these goals. Consequently, the GroupAL is also based on this optimization and uses n-dimensional vectors to represent the criteria. To assign participants to groups, the algorithm relies on three metrics that build on each other: the PairPerformanceIndex (PPI), which shows the suitability of two participants to each other, the GroupPerformanceIndex (GPI), which measures how all participants in a group match each other, and the CohortPerformanceIndex (CPI), which indicates the difference or similarity of all groups in relation to each other. Users can optimize the weighted criteria, based on either homogeneously (1) or heterogeneously (0). For this purpose, the PPI uses a weighted normalized distance function as the basis for matching in terms of fit (homogeneous) or complementarity (heterogeneous) of group members on criterion dimensions. The evaluation of the MoodlePeers tool showed better results than other non-linear optimized algorithms in terms of the quality of group formation both within groups and between groups in the resulting cohort (Konert et al., 2016). Consequently, it was possible to realize the planned experimental design in which the cohort of participants was divided into small groups within four segments of equal size (see Table 1).

Table 1. *Grouping Scheme, Algorithm Applied to Match in Groups*

	Prior knowledge (PK)	
Extraversion (E)	**Heterogeneous group formation**	**Homogeneous group formation**
Heterogeneous group formation	Heterogeneous PK & heterogeneous E 1	Homogeneous PK & heterogeneous E 2
Homogeneous group formation	Heterogeneous PK & homogenous E 3	Homogeneous PK & homogeneous E 4

Note. Experimental algorithmically established groups by heterogeneous/ homogenous extraversion/ prior knowledge

We therefore divided participants into groups, that were similar or dissimilar in the two traits of extraversion and prior knowledge. Individuals with similar quotients were thus matched, to create homogeneity in the groups with respect to their levels of prior knowledge and extraversion. Matching participants with similar quotients in the two experimental variables ensure that there is homogeneity in the groups with respect to prior knowledge. The same is true for extraversion. Groups that are matched very differently in the quotient of these experimental variables (extraversion, prior knowledge) are in turn maximally different in these characteristics, i.e., heterogeneous within their group. Here, the algorithm tries to generate the largest possible distance to the group mean, and thus a high standard deviation within the members of this group across the entire population.

2.4. Data Analysis

We grand mean-centered the personality traits as well as motivation subscales for better interpretation. We used block randomization in randomly assigning each participant to one of the four conditions. As mentioned above, the algorithm randomly divides the whole sample in four equally large parts. It then makes sure that all four parts are comparable in their distribution of the relevant personality trait and attribution of prior knowledge.

2.4.1. Data Exclusion

The algorithm will not match participants, who have not filled in the questionnaires honestly with a group. Instead, it puts them together with people with missing data and forms random groups. We excluded from analyses participants who forgot their participation codename or misspelled it in the posttest, since we could not match data from pre- and posttest. Additionally, we detected questionnaire data for traces of careless responses and eliminated them when there were obvious cases (Meade & Craig, 2012).

2.4.2. Explorative Analysis

As part of our exploratory data analysis, multilevel models were created for each of the three outcome variables (satisfaction, assignment, group stability). In doing so, we included different variables in each model, similar to a regression procedure (Moerbeek et al., 2003), to show their proportional effect on the respective outcome variable. After prior construction of the null model, we stepwise selected gender, age, and conscientiousness, in addition to the experimental conditions of group formation by extraversion and prior knowledge. We decided to include conscientiousness in our models because this variable showed strong correlations with prior knowledge (Meyer et al., 2022). Results from these exploratory analyses can be used for hypothesis-building in future projects.

3. Results

We conducted our analyses in light of our hypotheses; that is, we looked at whether heterogeneous grouping in extraversion and heterogeneous grouping in prior knowledge led to positive outcomes regarding the group members satisfaction, achieved points in assignments, and total number of assignments submitted (group stability). We also examined the interaction effects of the heterogeneous and homogeneous group formation. We also explored whether variances at the group- or individual-level and whether the experimental groups would lead to successful outcomes. We analyzed the data using SPSS 23.2, and R.

3.1. Descriptive Analyses of the Data Structure

We start the presentation of results with a brief presentation of the underlying data structure, computed with SPSS. The description of the sample now includes the dropout analysis of the study. Most students, who filled out the questionnaire at measurement time point 1 before group formation, participated only in the first measurement time of group work. Due to this and the overall high dropout, we used the evaluation of the first measurement time point of groupwork only. We included the satisfaction with group work in the first evaluation (Satisfaction), the performance quality of the first homework (Assignment), and a measure of group stability as an outcome, including the sum of submitted group homework across all time points (Group stability). Table 2 illustrates the data structure at the selected result variables.

Table 2. *Descriptive Measures of Main Dependent Variables*

Dependent variable	***Mean***	***Skewness***	***Kurtosis***
Satisfaction	4.89	-1.12	3.10
Assignment	3.08	0.68	1.56
Group stability	0.92	0.70	1.86

Note. N = 172

3.2. Univariate Analysis of Variance with Two Factors

We are interested in confirming or rejecting the posed question: Will either the heterogeneous or homogeneous grouping in the two manipulated variables of extraversion and prior knowledge affect the three outcome measures: satisfaction, performance, and group stability? We conducted ANOVAs to investigate changes in mean-value-differences and if changes were by chance or systematic and significant. We found no significant main effect on the outcome measure of the dependent variable satisfaction for both factor extraversion F(1,72) = 0.24, *p* = .63, η2 = 0.01, and factor prior knowledge F(1,72) = 0.10, *p* = .75, η2 = 0.01. There was also no significant interaction effect: F(1,72) = 0.26, *p* = .61, η2 = 0.04. Additionally, effect sizes are negligibly small. The main effect of criterion extraversion on the dependent variable first assignment is also not significant: F(1,235) = 1.80, *p* = .18, η2 = 0.08. Like the main effect of prior knowledge on the dependent variable assignment, F(1,235) = 1.61, *p* = .27, η2 = 0.07, the interaction effect is not significant: F(1,235) = 1.23, *p* =.27, η2 = 0.01. The main effect of extraversion showed no significance on the dependent variable group stability F(1,235) = 0.16, *p* = .69, η2 = 0.01 and the main effect of prior knowledge F(1,235) = 0.07, *p* = .79, η2 = 00. Thus, the interaction effect is significant despite the small effect size F(1,235) = 4.15, *p* = .04, η2 = 0.02.

3.3. Data Analyses: Considering the Group Structure

As an explorative part of our analyses, we used R-package "1me4," version 1.1-18.1 to calculate multi-level models (MLM), taking into account the structure of data, where individuals were nested in groups (Bates et al., 2020). Traditional multiple regression techniques treat the units of analysis as independent observations. One consequence of failing to recognize hierarchical structures is that standard errors of regression coefficients will be underestimated, leading to an overstatement of statistical significance. As in our study, mostly the key research question regarding group formation research concerns the extent of grouping in individual outcomes and the identification of 'outlying' groups. In evaluations of group performance, for example, interest centers on obtaining 'value-added' group effects on students' attainment. Such effects correspond to group-level residuals in a multilevel model, which adjusts for prior attainment (Hox et al., 2018; Van Landeghem et al., 2005).

We used a step-up modeling strategy to address the different problems and structures of the outcome variables. The special features of the result variables are now first listed, and then the respective solution for each result variable is shown in a model. We can report group variance using the Interclass Correlation Coefficient, ICCs. This was done by first setting up the empty or null-level-model without any explanatory variables, which describes the partition between variance at the student-level and at the group-level. ICCs that are nontrivial and greater than .05 must be considered (Hox, 2010). It is important to mention that the variances could be misleadingly high, due to the small group size and slight variation of outcomes on the individual-level. Thereby, the group could explain 54% of the variance on variable assignment, 34% of the variance in variable satisfaction, and 60% in group stability. The calculated linear equation models are shown in Tables 3, 4, and 5. We added the predictor's age (age of the students), average grade (as the self-stated average grade in math during school), personality traits, extraversion, agreeableness, neuroticism, and conscientiousness in each model. The proportional values of the explanatory variables in the model are represented by the respective sizes of the coefficients. The best model fit can be identified by AIC or BIC. Asterisks mark the significant predictors.

Satisfaction. For the dependent variable satisfaction with the group (" Satisfaction "), we assumed normal distribution and linear equation models were calculated. Table 3 shows the results. Model fit was best in Model 1 and 4 showing lowest BIC and AIC values. In the models, experimental conditions and personality traits, extraversion and conscientiousness are shown to be more important predictors, then demographics such as gender and age. In the models, experimental conditions and personality traits extraversion and conscientiousness are shown to be more important predictors, then demographics such as gender and age. No significant predictor for satisfaction could be revealed.

Table 3. *Individual-level and Group-level Predictors of Dependent Variable Satisfaction*

Predictors	Model 1	Model 2	Model 3	Model 4
Criterion_Extraversion	0.14			0.13
	(0.26)			(0.29)
Criterion_Prior knowledge	-0.08			0.08
	(0.26)			(0.29)
Gender		0.13		
		(0.27)		
Age		-0.02		
		(0.03)		
Average Grade		0.03		
		(0.02)		
Extraversion			0.23	0.18
			(0.17)	(0.16)
Conscientiousness			-0.09	
			(0.23)	
Agreeableness			0.05	
			(0.17)	
Neuroticism			0.01	
			(0.14)	
Constant	5.02**	4.69**	4.73**	4.84**
	(0.23)	(0.78)	(0.16)	(0.26)
Observations	76	86	75	65
Log Likelihood	-115.82	-138.64	-123.96	-101.41
AIC	241.64	289.28	261.93	214.83
BIC	253.29	304.00	278.15	227.87

Note. Gender coded 1 for men and 2 for women. Criterion of experimental manipulation for group formation; Criterion_Extraversion/_Prior knowledge = Intervention for Group formation; heterogenous = 0, homogeneous = 1. Unstandardized coefficients reported. Robust standard errors reported in the parentheses. Missing data handled with case deletion. *p =.05. ** p =.01.

Assignment. We established and adapted a hierarchical linear model for not normally distributed variables for dependent variable assignment. Model fit does improve from Model 1 to Model 4 with model 4 having the best fit. Again, only conscientiousness is revealed as a significant predictor for assignment. Table 4 reports the results.

Table 4. *Individual-level and Group-level Predictors of Dependent Variable Assignment*

Predictors	Model 1	Model 2	Model 3	Model 4
Criterion_Extraversion	-0.56		-0.39	
	(0.77)		(0.77)	
Criterion_Prior knowledge	0.70		0.67	
	(0.77)		(0.77)	
Gender		-0.26		
		(0.46)		
Age		0.09		
		(0.06)		
Average Grade		-0.04		
		(0.03)		
Extraversion			-0.02	-0.12
			(0.25)	(0.27)
Conscientiousness				0.90**
				(0.34)
Agreeableness				-0.10
				(0.27)
Neuroticism				0.16
				(0.22)
Constant	2.93**	2.08	2.82**	3.02**
	(0.66)	(1.60)	(0.65)	(0.36)
Observations	239	254	216	233
Log Likelihood	-642.91	-688.20	-587.43	-634.36
AIC	1,295.82	1,388.41	1,186.86	1,282.72
BIC	1,313.20	1,409.63	1,207.12	1,306.88

Note. Gender coded 1 for men and 2 for women. Criterion of experimental manipulation for group formation; Criterion_Extraversion/_Prior knowledge = Intervention for Group formation; heterogenous = 0 homogeneous =1. Unstandardized coefficients reported. Robust standard errors reported in the parentheses. Missing data handled with case deletion. *p = .05. **p = .01.

Group stability. Table 5 shows the individual and group-level predictors of the dependent variable of overall submitted assignments ("Group stability") as an indicator of group work endurance. Model fit constantly improved from Model 1 to Model 4. The decision was made to calculate a generalized linear mixed model (GLMM). In contrast to simple regression analysis and multiple regression analysis, the dependent variable here can be binary with only two values: 0 for "not delivered" and 1 for "delivered". This means that it is not the value of the dependent variable that is predicted here, but the probability that the dependent variable takes on the value 1. Furthermore, the conditions are less restrictive than in linear regression analysis. Still, any postulated causal relationship must be theoretically justified (Hox et al., 2017, 2018). Most of the independent variables have no influence on the probability that the dependent variable "group stability" takes the value 1, i.e., that the group stability remains. Only conscientiousness turns out to be a significant predictor of group stability.

Table 5. *Individual-level and Group-level Predictors of the Dependent Variable Group Stability*

Predictors	Model 1	Model 2	Model 3	Model 4
Criterion_Extraversion	-0.05	-0.03		
	(0.15)	(0.15)		
Criterion_Prior knowledge	0.04	0.02		
	(0.15)	(0.15)		
Gender		-0.16		
		(0.17)		
Age		0.04		
		(0.02)		
Average Grade		-0.03**		
		(0.01)		
Extraversion			0.03	0.01
			(0.08)	(0.08)
Agreeableness			0.12	0.06
			(0.08)	(0.09)
Neuroticism			0.00	0.02
			(0.07)	(0.07)
Conscientiousness				0.23*
				(0.11)
Constant	0.92**	0.87	0.87**	0.87**
	(0.13)	(0.58)	(0.07)	(0.07)
Observations	239	236	233	233
Log Likelihood	-375.77	-364.57	-355.73	-353.43
AIC	757.53	741.14	719.45	716.86

Note. Gender coded 1 for men and 2 for women. Criterion of experimental manipulation for group formation; Criterion_Extraversion/_Prior knowledge = Intervention for Group formation; heterogenous = 0, homogeneous = 1. Unstandardized coefficients reported. Robust standard errors reported in parentheses. Missing data handled with case deletion. *p = .05. ** p = .01.

4. Discussion

In this study, we implemented group formation based on the personality trait extraversion, aligning with prior knowledge regarding the corresponding standard deviation. The working hypotheses posited that superior results in subjective satisfaction, performance, group stability, and overall better group formation would be achieved through (1) a heterogeneously formatted group in extraversion and (2) a heterogeneously formatted group in prior knowledge. However, our study results did not substantiate these hypotheses. Consequently, the study hypotheses are rejected. However, it is essential to underscore that the rejection of hypotheses still constitutes a significant finding. Our research yielded no significant results, indicating the absence of a main effect of extraversion or prior knowledge on group outcomes. Yet we observed an interaction-effect of extraversion and prior knowledge on group stability: interactions with a heterogeneous distribution of extraversion and a homogeneous distribution of prior knowledge demonstrated the highest retention in groups. Although these differences lack statistical significance, they suggest that assuming a direct relationship between personality traits and group-work-outcomes may be unwarranted. These results align with prior studies advocating for research on the personality-work-behavior relationship to transcend linearity assumptions (Curşeu et al., 2019).

The literature suggests that different ability-types among students yield benefits from working in either heterogeneous or homogeneous groups (Saleh et al., 2005). Given the overall low value of prior knowledge in our sample, one might infer that heterogeneous groups are generally superior to homogeneous ones, due to the possibilities inherent in group formation. The variability in mean values between homogeneously grouped prior knowledge groups could be another contributing factor to the absence of significant results. A crucial finding is that, despite the initial online nature of group work, the group level could statistically account for most variances. This underscores the importance of group formation and development in understanding groups. Exploratory findings reveal significance in the model for all performance-related outcomes in the predictor conscientiousness. Conscientious individuals, characterized as goal-oriented, structured, organized, and self-disciplined (Costa & McCrae, 1992), are associated with better performance (Hurtz & Donovan, 2000). Prior studies have consistently affirmed that conscientiousness exhibits the highest correlation with performance success among other personality traits (Busato et al., 2000; Di Fabio & Busoni, 2007; Furnham et al., 2003; Lounsbury et al., 2003) and displays the strongest correlation with academic success (Di Fabio & Busoni, 2007; Protsch & Dieckhoff, 2011). Consequently, behaviors associated with conscientious team members are likely to be beneficial for group performance, including the fulfillment of task roles, as evidenced in our study.

Considering the achievement level and personality characteristics of group members in group formation, as opposed to self-selection or random group formation as suggested by Bekele and Menzel (2005), where greater diversity of personalities within the group positively affects overall performance (Roberge & van Dick, 2010). In contrast, other studies suggest that successful and unsuccessful teams are homogeneous with respect to various characteristics (Wax et al., 2017). Similar studies revealed significant effects due to group formation based on standard deviation. However, results of a meta-analysis showed that this formation was unrelated to performance (Devine & Philips, 2001). It is possible that the significant effects found in the studies may be explained by the nature and difficulty of the task (Bowers et al., 2000). Additionally, literature on MOOCs has shown that participants are more likely to complete to obtain a certificate (Liu et al., 2020). Both the voluntary participation in the course and the lack of relevant evaluation of the course could be reasons for the students in the study not completing the course.

Students are familiar with online tools and generally show a positive attitude toward learning with them. However, problems can occur when creating their own online learning environment (Lim & Newby, 2020). Here, group-working methods could be a promising tool. In online group work, we assumed that the group participants were unacquainted with each other before the group work and probably did not meet personally during the process. Through enhanced cohesion, group members built a stronger bond within the learning group. The resulting affiliation to the group—

prompting a desire for continued group membership—could promote higher participation, crucial for positive development in virtual teams (Williams et al., 2006). Computer-based asynchronous programs cannot transmit gestures, non-verbal subtleties, or symbolic content (Montoya-Weiss et al., 2001). This limitation can make communication more challenging, as we are accustomed to communicating with these aids from our everyday life and may impact group-problem-solving efficiency within this study. Nevertheless, significant effects found in earlier studies could be due to the type and difficulty of the task used in the analyses (Bowers et al., 2000).

In regard to the multitude of individual and group-level variables affecting CSGBL-processes and the challenges in predefining independent static conditions, we propose a looser observation set-up (Strijbos et al., 2004). Students deem direct communication as the most informative, and less informative, text-based online work negatively affects communication and social interaction (Okdie et al., 2011; Straub, 1997). Students are found to be less likely to engage in collaborative learning, interactions, and discussions in online settings compared to traditional classroom settings (Dumford & Miller, 2018), which might have hindered interaction in our study. This type of online work may have taken place in our study, disproving the assumption that social interaction will inevitably transpire with the provision of adequate technology. Technology encourages communication by offering more appropriate means to complete the task, but it does not guarantee the required social exchange (Kreijns et al., 2003). Technology knowledge positively correlates with technology acceptance. Addressing students' technology proficiency and acceptance is an important step for designing online courses and group work (Nami & Vaezi, 2018). In future studies, prior knowledge should be replaced by technology-knowledge. For online-group-learning in university settings, we need to distinguish the outcome: differentiating the development of knowledge and the learning process of individuals to gain knowledge (van

Merriënboer & Kirschner, 2001). Using group constellations complicates the measurability of both outcomes. Several key questions arise, such as whether predefining independent static learning or instruction conditions are a feasible possibility in a grouping, and whether we can control all relevant conditions that affect group interaction and individual knowledge gain.

It is noteworthy, that while stimulating group collaboration and fostering communal learning, educational techniques may not have the ability to establish it all together. We observed that groups might not have been actively working together during the execution of the study. Creating a sociable CSGBL-atmosphere could be a possible solution to this issue. The solution could include generating an environment that allows for interpersonal, social, non-task-related exchange and provides external bonding opportunities. It also includes intensifying the number of task-related and non-task encounters, resulting in a more constant presence and awareness of the group members (Lin et al., 2010; Strijbos et al., 2004). For future projects using the Moodle-platform, we might use an Online-Course-Design-Checklist (OCDC) and integrate an analytics-framework for detecting students at risk of dropping out (Baldwin & Ching, 2019; Monllaó Olivé et al., 2020).

4.1. Strengths and Limitations

The study's notable strength lies in its experimental design, facilitating the determination of causal relationships. It represents a well-designed field-study conducted in a real-world-environment and an authentic teaching-scenario that assumes the students' natural behavior. Consequently, the results boast higher external validity, enhancing their generalizability. Moreover, we can report a substantial initial sample size. The assessment of homework processing serves as an objective measure, free from dropout bias, as we also considered absentee records during data analysis. Evaluations, being a prerequisite for submitting homework, were processed more frequently by students compared to previous studies. Alternative frameworks, such as teaching students from home or extending enrollment periods, pose potential avenues for future exploration. An intriguing question emerges regarding whether a robust expression of conscientiousness and prior knowledge yields similar effects.

Regarding data analysis, a notable strength is the utilization of multi-level modeling, allowing the investigation of variance at both the individual and group levels. This approach contrasts with regular regression, which may encounter sampling problems and lack generalization. However, a limitation of the study is its reliance on a virtual groupwork setting, which was still uncommon at the time of the research. Given that the study predates the "corona pandemic," this virtual setting was unfamiliar to many prospective students. Additionally, voluntary participation in the course contributed to a high dropout rate, negatively impacting the entire group and disrupting the group-process. Despite the challenges posed by the unfamiliar situation, experimental studies in the online context have become indispensable today due to the pandemic, offering valuable guidance and design recommendations for institutions navigating the shift to online teaching. The study's small group size and the simultaneous existence of a large number of groups may have led to a potentially artificially high Intraclass Correlation Coefficient (ICC) (Hox et al., 2017). However, this circumstance serves the purpose of group comparisons, aligning with the study's objectives and strengthening the power of the results. It also allowed for an increase in the number of groups, enhancing the sample size at the group level. This trade-off is a recurring challenge in university and educational research, where the number of available subjects is not limitless.

As previously mentioned, the overall low level of prior knowledge may have influenced the formation of experimental groups. On average, the algorithm had to form homogeneous groups from individuals with low prior knowledge to correctly generate heterogeneous groups, potentially explaining the absence of significant results and the presence of only interaction effects. Despite this limitation, it is reasonable to assume that prospective students opting for a prerequisite course to enhance their mathematical skills likely had lower prior knowledge, introducing a selection effect. Thus, this restriction is considered acceptable within the context of the study's setting.

In the context of groupwork occurring within a short timeframe and an unfamiliar online setting, it is plausible that the desired group dynamic did not have sufficient time to develop. While this limitation is inherent in studies on online group work, it remains an assumption that cannot be verified, but warrants consideration. Furthermore, the study's group-formation-aspect should be replicated over an extended period to reveal potential effects over time. In addition to the limitations, most research findings are derived exclusively from self-report questionnaire formats. This poses a notable overall limitation to (virtual) group-work-research, emphasizing the need for more objective-dependent variables, such as quality and quantity of group discussions in forums, log-files, and videography. The study attempted to address this limitation by incorporating both objective and subjective outcome measures.

4.2. Implications

Considering the current situation caused by COVID-19, studies exploring didactically online learning settings for students and how they may actively foster participation and continued engagement are mandatory (Wildman et al., 2021). The rapidly advancing digitization of university education demands that students take the initiative and display conscientious self-organization to progress in their attainment of further knowledge. Group work has proven to be very beneficial in this regard. Even though most studies support the importance of personality traits, such as extraversion and conscientiousness, and cognitive aspects, such as prior knowledge, in education, the question of how to make use of them as criteria for group formation arises.

Taken together, the benefits offered by the group formation algorithm are highly relevant for universities, as it allows first-year students to form remote learning groups according to criteria relevant to them. Thus, a group formation tool, has the great potential to create social interaction and thereby a sense of belonging for students despite social distance. Such an algorithm can additionally be useful for various other settings. In addition to the benefits of the group-formation-tool for university, it also has potential benefits for the didactics of schools in the business contexts, as well as for the leisure sector and thus for private-group-design. The question remains open as to how other characteristics, or more precisely other personality traits, play out in this context. Other settings, such as groupwork that does not take place online or hybrid formats, should also be investigated regarding other group formation characteristics or those used here. What has already been clearly found in this study is that the outcomes of group work can be explained at the group level, and that group formation is thus an important and, moreover, economical means of choice to enable the success of group work. However, more research is needed on the characteristics used in group formation, the settings in which they lead to success, and the outcome variables on which they affect. Attention should be paid to personality traits, as group formation can lead to positive and negative outcomes depending on their structure, and we already have evidence that the traits studied here are quite relevant. Moreover, the algorithm used here can be successfully used in other follow-up-projects to study-group formation.

For future work, we plan to repeat the described experiment under different conditions, both in face-to-face courses and in virtual environments that promote CSCL. Another upcoming work is to experimentally manipulate additional student characteristics, e.g., other personality traits such as conscientiousness, as well as a replication of the present work, where we would use previous student grades as a criterion for group formation, rather than prior knowledge queried selectively, to look at the outcome that collaborative learning has for previously low-performing students. Other characteristics that could be influential are factors such as prior technical knowledge among students and their motivation regarding specific learning activities.

4.3. Conclusion

Our study introduced an experimental approach to group formation with promising criteria, thoroughly researched. Not only did the study demonstrate that the proposed experimental research method and the applied algorithm successfully achieved the goal of obtaining homogeneous and heterogeneous groups, but it also revealed that the interaction of characteristics, specifically heterogeneous extraversion, and homogeneous prior knowledge, positively influenced the development of activities within the collaborative learning context and served as an indicator of group stability. Furthermore, we explored the potentially crucial role of conscientiousness for online working groups. In this context, it is crucial to emphasize that the inclusion of specific student characteristics requires careful consideration, preferably guided by methodology-grounded psychological insights. This underscores the necessity of considering numerous variables of individual group members to optimize group fit for well-functioning groups.

Contrary to expectations, the study's hypotheses could not be supported. It appears that the high variance explanation at the group level is attributed to other group-level variables differing between groups, rather than the structuring of the experimental variables and group formation. Nevertheless, the impact of group-level factors compared to student-level factors is a noteworthy finding. This underscores the importance of investigating group formation criteria, because the results of group work can be influenced by the group formation processes. This is a significant outcome, highlighting that the mix of group member characteristics is more pivotal to the results than the characteristics of individual members alone. Researchers and practitioners have diverse approaches to construct different grouping models, considering individual trait constellations and focusing on traits beyond those examined in this study. Our approach presents an opportunity for scientists to conduct future research on group formation, enriching the body of knowledge on online group formation. Such studies are crucial for educational institutions and other professional domains. Given factors such as spatially distributed and interdisciplinary group work, digitalization, increasing demands for mobility in the working world, and current considerations like social distancing and flexibilization, the workload is escalating while available time is diminishing, reflecting the transformative nature of the way we work.

6.2.1 Summary of Study 2 and Motivation for Study 3

Studies 1 and 2 examined the impact of group formation on various outcome measures in an online pre-course for prospective mathematics students. Both Studies analyzed the distribution of variance in the personality trait extraversion, along with either conscientiousness (study 1) or prior knowledge (study 2). The project faced challenges with data acquisition fluctuations and group instability, due to its reliance on a voluntary course structure conducted in an unusual online setting. Data was collected by prospective students, who were not given any incentives or external motivations, resulting in high dropout rates. The fragmentation or disbandment of existing groups resulted in limitations and inconclusive results in experimental studies. To address these limitations, studies 3 and 4 modified the experimental setting by making mandatory face-to-face seminars the new research target. This was done with the goal of generating more continuous and controllable data for the experiment, with only extraversion as the criterion for forming the experimental group. The aim of this strategic adjustment was to minimize dropout rates and ensure continuous data collection.

The motivation for study 3 was to investigate the role of group hierarchies in student groups, providing insights in face-to-face group dynamics and suggesting potential improvements in group formation strategies within such settings. Data was collected through obligatory seminars, with an emphasis on constant groupwork over an extended period, to allow for the natural evolution of group roles. This new implementation aims to overcome previously identified challenges, contributing to a nuanced understanding of face-to-face group work and offering insights for potential enhancements through refined group formation strategies.

6.3 Manuscript 3: Müller, A. M., Röpke, R., Konert, J., & Bellhäuser, H. (2023). Investigating Group Formation: An Experiment on the Distribution of Extraversion in Educational Settings. ACTA Psychologia*, 242*, 104111. https://doi.org/10.1016/j.actpsy.2023.104111

Investigating Group Formation:

An Experiment on the Distribution of Extraversion in Educational Settings

Abstract

Group formation plays a crucial role in enhancing collaborative learning experiences. This study investigates the impact of extraversion as a criterion for group formation on collaborative learning outcomes. A total of 180 students participated in the experiment and were assigned to groups, that were homogeneously or heterogeneously distributed in terms of extraversion. The groups met weekly and worked on group assignments throughout the semesters. The first hypothesis posed the outcomes to be explainable at the group-level. Surprisingly, the results show that groups with a homogeneous distribution of extraversion reported higher levels of group work satisfaction than those with a heterogeneous distribution, in contrast to the second hypothesis and the group hierarchy theory. These findings emphasize the potential of considering personality traits when forming groups and extend the existing literature on group formation. The study takes a critical stance by addressing normative definitions of leadership. Future research is suggested to further enhance collaborative learning experiences using similar interdisciplinary and experimental methods.

1. Introduction

1.1. Group Work in Higher Education Settings

Group work and collaborative learning significantly enhance student learning, motivation, and satisfaction in higher education (Blasco-Arcas et al., 2013; Johnson et al., 1991; Magnisalis et al., 2011; Ryan & Patrick, 2001). However, challenges like uneven workload distribution, poor communication, and low attendance can lead to potential group failure (Chang & Brickman, 2018; Hall & Buzwell, 2013; Rummel & Spada, 2005). Group formation has emerged as a potential solution to address these challenges equally for all students (Amara et al., 2016; Kozlowski & Bell, 2013; Srba & Bielikova, 2015). The process of group formation, involving the selection and organization of individuals into groups, is crucial for fostering effective group experiences (Borges et al., 2018). Ensuring equal and successful group learning among students necessitates the design of effective group formation approaches (Damşa, 2014; Fazal-e-Hasan et al., 2021). Various methodologies have been investigated, ranging from computer-assisted learner-group formation (Bekele, 2006) to sophisticated algorithms like genetic algorithms (Zheng et al., 2018) and multi-objective ant-colony systems (Fahmi & Nurjanah, 2018).

1.2. Unveiling the Impact of Group Formation: Distinguishing Group-Level and Individual-Level Factors

Understanding how group- and individual-level factors interact is crucial for comprehending group behavior and implementing effective interventions (Kozlowski & Bell, 2013). Identifying whether group formation independently shapes outcomes or if individual traits wield greater influence is fundamental in perfecting group formation techniques (Hitt et al., 2007). The intricate relationship between individual traits and their impact within groups significantly shapes group dynamics and eventual outcomes (Blanco-Fernández et al., 2023).

Multilevel data structures pose challenges in capturing contextual features encompassing both individual characteristics and their surrounding contexts. While conventional analyses often focus on individually measured outcomes (Raudenbush & Bryk, 2002), the composition of trait expressions within groups profoundly influences group dynamics and outcomes in higher education (Blanco-Fernández et al., 2023). This emphasizes the significance of recognizing both contextual and individual influences on group outcomes (LePine et al., 2011).

Previous Studies highlight the impact of group-level factors on decision-making, problem-solving, creativity, and communication patterns, all significantly influencing group performance (Gawande et al., 2003; Loignon et al., 2018; Mannix & Neale, 2016; Van Knippenberg & Schippers, 2007; Voltmer et al., 2022; Zennouche et al., 2014). These factors intricately link with individual personality traits, underscoring the need to consider these traits when forming groups (Maqtary et al., 2019; Moynihan & Peterson, 2001; Yannibelli & Amandi, 2011). The complex interrelationship between individual and group-levels of extraversion warrants deeper exploration to understand, how an individual's traits are influenced within the broader group context. Investigating the interaction between levels of extraversion within and between groups holds promise in offering invaluable insights into optimizing group dynamics and enhancing outcomes.

1.3. The Distribution of Personality Traits Applied as Group-Formation Criteria

The widely used "Big Five" framework effectively measured and described personality traits (Hough & Oswald, 2000, 2005), yet inconsistencies persist regarding its role in group formation (Bell, 2007; Driskell et al., 2006; Lykourentzou et al., 2016). The distribution of traits within groups prompts critical questions about optimal trait compositions, rooted in the person-environment (P-E) fit literature. This framework suggests that, depending on the trait, dissimilar traits may offer or reduced opportunities within groups (Tett et al., 2021).

Understanding trait distribution is pivotal for discerning the relevance of homogeneous or heterogeneous group-fit hypothesis (Cable & Edwards, 2004; De Dreu & Weingart, 2003; Jackson et al., 2019). Diverse group members positively influence outcomes enabled by complementary skills, fostering interaction (Bekele & Menzel, 2005; Moreno et al., 2012; Seong & Hong, 2020; Van Dijk et al., 2017). Conversely, homogeneous groups may foster better learning experiences (den Hartog, 2019; Wilson et al., 2016). However, the prevalent formation of heterogeneous groups (84 primary Studies; Borges et al., 2018) has resulted in inconclusive or incomplete conclusions regarding trait distributions. Homogeneous fit, mostly used as a baseline, compromises the investigation of its independent effects (Apfelbaum et al., 2014).

1.3.1. Group Formation using the Distribution of Extraversion

Individual extraversion significantly shapes behaviors and responses in social contexts (Huang & Wu, 2020), relating to positive feelings, and impacts social interactions and organizational citizenship behavior within groups (Mattila et al., 2011; Moon et al., 2008; Wilt et al., 2012). Additionally, extraversion aligns with social network size, cognitive performance, and emotional experiences at an individual-level (LeMonda et al., 2015; Longua et al., 2009; Pollet et al., 2011). Extraversion significantly shapes group dynamics, influencing social interactions and the establishment of hierarchical or non-hierarchical group structures (Taggar et al., 2006; Wilmot et al., 2019). Literature on group formation underscores the profound impact of extraversion levels among group members on group functioning (McCabe & Fleeson, 2012; Wilmot et al., 2019). Higher levels of individual extraversion correlate with leadership behaviors such as initiating discussions and offering support to other group members (Judge et al., 2002; Porter et al., 2003). In hierarchical group structures, a leader with high extraversion typically leads, while others assume follower-roles, highlighting the role of extraversion in establishing group hierarchies

(Kramer et al., 2014). Moreover, the intricate relationship between group extraversion and effectiveness is intertwined with communication patterns and leadership behaviors (Rothstein & Goffin, 2006).

The effect of extraversion distribution on group outcomes remains a subject of debate. While some Studies advocate for the benefits of heterogeneous extraversion distribution (French & Kottke, 2013; Humphrey et al., 2007; 2011) in enhancing student performance, others propose the superiority of homogeneous distribution for outcomes like innovative thinking and negotiation (den Hartog, 2019; Wilson et al., 2016). However, limited research has experimentally explored the impact of trait distribution on group-formation outcomes (Borges et al., 2018). Exploring hierarchical group structures reveals their varied impact on group effectiveness. Stable hierarchies foster cooperative communication, role clarity, and facilitate decision-making (Roney et al., 2012; Woolley et al., 2022). Conversely, unstable hierarchies, membership instability, and skill differentiation can lead to disruptive communication patterns and conflicts, detrimentally affecting group performance (Greer et al., 2018; Woolley et al., 2022). The lack of clarity in this area necessitates further research to better understand the role of extraversion distribution in diverse outcomes. The dominance complementarity theory, emphasizing balanced dominance, assertiveness, compliance, and submissiveness (Kiesler, 1983), aligns with extraversion's ability to enhance social experiences. This may elucidate why extraverts experience greater subjective well-being, leisure satisfaction, and happiness (Harris et al., 2017; Lu & Hu, 2005). Extraversion's positive aspects, like sociability and assertiveness, contribute to the emergence of transformational leaders (Bono & Judge, 2004). However, while extraverted leadership might enhance group performance in passive situations, it might have counterproductive effects in proactive group settings (Tiedens & Fragale, 2003).

1.5. Research Questions: Group Formation Research and the Role of Extraversion

To address previous inconsistencies in the literature and gain deeper insights (Maqtary et al., 2019; Odo et al., 2019) we employ an experimental design using an algorithm for group formation that manipulates individual differences in extraversion. The distribution of extraversion within groups can be conceptualized as a group characteristic with either homogeneous or heterogeneous appearance (Deckers et al., 2022). This approach enables more reliable conclusions about the causal relationship between group formation and the variables of interest.

The influence of extraversion operates differently at individual and group-levels, creating a complex interplay (Turban et al., 2009). Extraversion's influence on positive affect at the individual-level is firmly established (Wilt et al., 2012). However, the group-level significantly shapes this relationship, impacting sociable behavior within a group context (Mattila et al., 2011; Moon et al., 2008). The distribution of extraversion within a group influences the emergence of leaders and high-status individuals within these groups (Alam et al., 2022). The collective level of extraversion within a group impacts group satisfaction and political participation, reflecting the intricate influence of other extraverts within the group (French & Kottke, 2013; Huber et al., 2021). Research Studies have indicated potential advantages in both heterogeneous and homogeneous distributions of extraversion (den Hartog, 2019; French & Kottke, 2013; Humphrey et al., 2007; 2011; Wilson et al., 2016). However, each distribution model presents unique advantages and challenges. While heterogeneous distributions might facilitate task delegation and conflict management (Humphrey et al., 2007; Tekleap & Quigley, 2014), homogeneous distributions could foster innovation and negotiation skills (den Hartog et al., 2019; Wilson et al., 2016). Additionally, while existing research has predominantly focused on performance as the primary outcome, it is imperative to expand our scope to include other relevant measures (Cachia et al., 2018; Rogat et al., 2022). Furthermore, we explore whether group outcomes are shaped by the interplay among members' constellations rather than solely by their individual attributes (Hitt et al., 2007; Kozlowski & Bell, 2013).

Our research addresses these gaps by experimentally manipulating extraversion in group formation, thereby providing robust evidence to understand the causal relationships between extraversion, group formation strategies, and outcomes. First, we hypothesize that collective dynamics within a group are expected to play a pivotal role in shaping overall outcomes. Hypothesis 1 is as follows:

H1: Group-level characteristics, particularly the distribution of extraversion within groups, will significantly influence satisfaction (a), time investment (b), and performance (c) outcomes compared to individual-level factors.

As a second hypothesis posed, we further anticipate that groups with algorithmically established mixed distributions of extraversion are expected to positively contribute to collaborative dynamics. Hypothesis 2 states:

H2: Groups with mixed (heterogenous) distributions of extraversion are expected to be superior concerning satisfaction (a), time investment (b), and performance (c) outcomes, compared to groups with a similar (homogeneous) distribution.

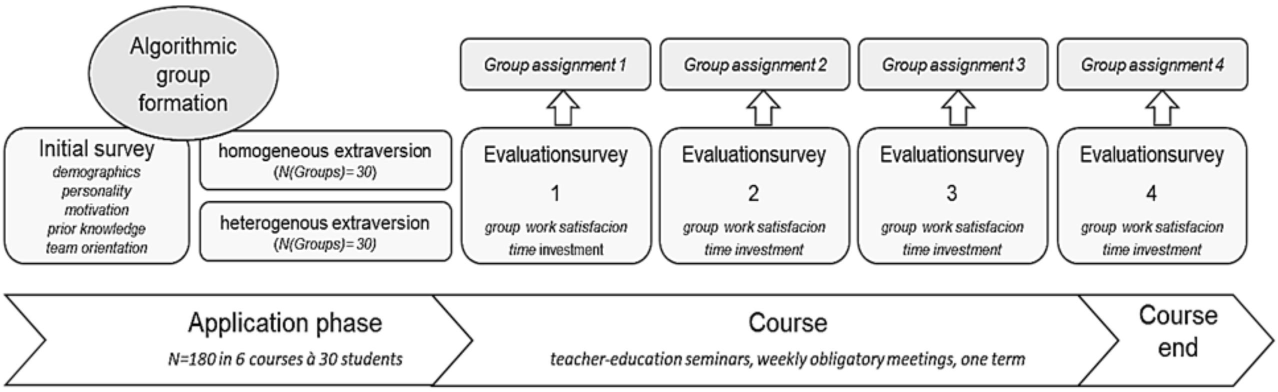
2. Methods and Materials

2.1. Sample and Design

An experimental study design with one factor (extraversion) was manipulated in two conditions (homogeneous and heterogeneous) across students and courses. The total sample included 180 teacher-education students (female $n = 115$; Age $M = 22.2$, $SD = 1.65$) from a German university, all of whom had participated in the seminar as part of an academic module. The seminars were arranged in six courses, each with approximately 30 participating students.

In the initial survey, which was conducted before the formation of the experimental groups, demographic data of the participants, such as age and gender, as well as all personality traits and group orientation, were collected to identify possible differences between the experimental groups before the study onset. Once informed consent and the initial survey were completed by all participants within each course, the algorithm GroupAL initially formed groups of three, for a total of 60 groups, of which members of 30 groups were homogeneously distributed and members of the other 30 groups were heterogeneously distributed in extraversion. After the algorithm GroupAL formed the groups, we informed the students about their group membership and instructed them to sit and work together throughout the entire term. Course meetings were conducted weekly. In addition, group assignments were constructed to stimulate groupwork and turned in weekly. During the course of the study, the outcome variables were regularly recorded in four evaluation surveys. To provide a comprehensive overview of the research method and process of the study, the study's timeline is displayed in Fig. 1.

Figure 1. *Study Timeline*



Note. Study Process divided in the three phases: Application Phase (initial survey, algorithmic group formation, established groups), the course (with evaluations 1-4, group assignments 1-4), and end of the course after one term.

2.2. Group Formation Algorithm to Obtain Experimentally Distributed Groups

The university where we conducted this study utilized Moodle's online learning management system (LMS). To support the formation of learning groups, we developed a Moodle-plugin called MoodlePeers. This plugin provides a user interface for teachers to set up group formation within a course, as well as the administration of questionnaires and an overview of the status of group formation (e.g., not yet started, open for answers, and groups formed). In addition, the plugin includes an implementation of the optimization algorithm GroupAL. The algorithm GroupAL was used to determine how to group the participants and apply a different set of matching criteria to each part. By minimizing (or maximizing) the distance between all three group members simultaneously, it maintains the same fitness level (i.e., prior knowledge and motivation) in the matched overall groups (Konert et al., 2016).

2.2.1. *Experimentally Distributed Groups*

The research interest here lies in how the variance in extraversion [or other traits] within a group, as well as between different groups, influences group dynamics and outcomes. To investigate the role of group formation by extraversion, we carefully designed two types of experimental groups by the trait expression of extraversion: homogeneous and heterogeneous. This study design enables us to investigate the significance of both the overall formation of the group and within each group (Bellhauser et al., 2018; Müller et al., 2022).

Homogeneous Groups. Homogeneous groups comprised members with similar extraversion levels, providing an appearance of uniformity. In other words, they display comparable behaviors and tendencies related to extraversion, such as the need for social contact, communication, and the desire to engage in conversation (AbuSeileek, 2007, 2012; Barrick & Mount, 1991; Hogan et al., 1994). Methodologically, for homogeneous groups, the algorithm aimed to maximize the similarity among their members in terms of extraversion, while still ensuring variation in the trait levels among different homogeneous groups. Consequently, even if the persons within these groups can be classified as similar at first glance, due to the names of these groups, they encompass a wide spectrum of extraversion levels, including high, medium, and low expressions, depending on the sample's trait values, such as range and mean. This approach allows us to assess the similarity (homophily) between group members, not their absolute levels of trait expression.

Heterogeneous Groups. In a heterogeneous group, each member exhibits a different level of extraversion, ranging from high to low expression. The algorithm strives to create a high standard deviation in this trait within the group, indicating the significant differences in the level of extraversion between members of the same group. In contrast to homogeneous groups, where the algorithm maintains a high standard deviation between the groups by keeping trait expressions consistent within each group.

For homogeneous groups, the algorithm aimed to create maximum similarity among members concerning extraversion, while ensuring diversity in expression levels. In contrast, heterogeneous groups were designed to have the same mean extraversion level, but significant variations within the group. This method was aimed at evaluating the effect of standard deviation within and between groups.

2.3. Instruments and Measures

2.3.1. Data Collected before Group Formation

The initial survey included questions on personality traits, motivation, prior knowledge, and group orientation. Participants answered all the questions online using a rating scale ranging from 1 (‘not true’) to 6 (‘true’).

Personality. We used the short version of the BFI-K questionnaire to measure Big Five personality traits (BFI-K; Rammstedt & John, 2005). The questionnaire shows robust reliabilities in this setting (extraversion: eight items, e.g., 'I am talkative like, I like to chat,' α =.88, conscientiousness: nine items, e.g.,' I work reliably and conscientiously,' α =.85, openness: five items, e.g., 'I have an active imagination, I am creative,' α =.69, neuroticism: four items, e.g., 'I worry a lot,' α =.70, agreeableness: four items, e.g., 'I easily trust others, I believe in the good in people,' α = .66).

Motivation. We used the expectancy-value-cost-scale (EVC) developed by Kosovich et al. (2015), to measure motivation. EVC is a self-report survey designed to measure student motivation within four subscales: Expectancy (four items, α =.76), value (five items, α =.86), cost (six items, α =.75), and interest (six items, α = .85).

Self-assessed Prior Knowledge. We measured prior knowledge as an average subjective rating (self-estimation) per knowledge topic of course content ('How do you judge your knowledge about the course content regarding [specific topic]') ranging from 0 to 100 points.

Group orientation. We measured group orientation using three items (e.g., 'If I have a choice, I would rather work in a group than alone') and demonstrated robust reliability (α = .85).

2.3.2. Outcome Variables

Group members completed online questionnaires four times to assess their experience with groupwork. To this end, we constructed an evaluation questionnaire that included questions about (1) satisfaction and belief that group work was an appropriate work method (Group Work Satisfaction), (2) time investment, and (3) number of participating members. In addition to (4), group performance was determined based on the four group assignments rated by the course tutors. Consequently, we collected data for multiple dependent variables over time. These outcome variables were divided into three main headings: satisfaction, time investment, and performance-related dependent variables.

Group Work Satisfaction. Collaborative Group Work Satisfaction is a dependent variable in this study and measures participants' satisfaction with their collaborative groupwork experience, including their perceptions of learning outcomes, skill-building, and social interactions. The construct consists of 5 items, which participants rate on a 6-point Likert scale (ranging from '1 = does not apply' to '6 = applies'). The items include statements such as "I learned more through the group work than I would have learned alone", "I improved my social skills through the group work", and "I improved my project management skills through group work". Participants are also asked to indicate whether group work was better suited than individual work for the specific practice tasks (e.g., ‘Did you learn more through group work than you would have learned alone?'). To ensure the internal consistency of the construct it has been evaluated using Cronbach's alpha, indicating high reliability (α = .89).

Time investment. The construct of time investment was assessed as a means of understanding the level of communication between group members. To subjectively evaluate participants' time investment, they were asked about the frequency of communication in their respective groups using a rating scale of '1 = never' to '6 = very often'. While the construct assessed the mean value of subjective time investment, the questions from which this construct was based did not have to be directly related, as this was a collective indication. Nevertheless, we report this construct for the sake of completeness and note that its internal consistency was good (α = .79).

Performance. We measured group performance based on the points on each of the four group assignments, rated by tutors based on a previously established criteria catalogue (ranging from 0 to 100 points).

2.4. Data Exclusion

We scanned the questionnaire for traces of careless responses and eliminated participants when there were obvious fraud-cases, as well as cases with incomplete or missing data (Meade & Craig, 2012). Overall, we deleted these three cases.

2.5. Data Analysis Procedure

We conducted data analysis and tested two hypotheses using statistical software, SPSS 23.2, and R version 4.0.0 (R Core Team, 2014). Hypothesis 1 aimed to determine whether most variances in our dependent variables (satisfaction with group work, time investment, and performance) could be explained at the group- or individual-level across all four measurement time points. Hypothesis 2 examined the effect of grouping by extraversion (homogeneous or heterogeneous) on these dependent variables.

To address data dependencies, we utilized multilevel modeling (MLM) with the package lme4 in R (Bates et al., 2020). The random intercept model (Geiser et al., 2010) served as the simplest example to consider possible dependencies. We used maximum likelihood estimation for parameter estimation, incorporating a penalty term to account for excessively large random effects, and aimed to identify groups with consistently positive outcomes for almost all members.

Random effects (group effects) were modified by introducing different predictors. In a sequence of interconnected models, the addition of predictors reduced the variance in random effects (Hox et al., 2017). Predictors at both student and group-levels helped explain some differences between groups. We assumed that consideration of predictors influenced group variances. Building on the earlier analysis steps, we tested hypothesis 2 to examine the estimated abilities of students (individuals) at the group-level, focusing on inter-individual differences between groups.

To assess the extent to which group-level factors contributed to explaining differences between groups, we employed the intraclass correlation coefficient (ICC) as a statistical measure (Bliese, 2000). Additionally, we used model-fit indicators AIC and BIC. We isolated the added value of groups from possible influencing factors and calculated MLMs. To understand how much variance could be explained and how the model fit changed, we introduced the study's experimental variable in each Random Intercept Model and Random Slope Model. The sizes of the coefficients reflected the relative importance of variables in the respective models, as again demonstrated by AIC/BIC.

3. Results

3.1. Baseline Descriptive Measures and Comparison of Experimental Groups

To provide context for our study, we initially examined baseline descriptive statistics from the first wave of data collection. These statistics allowed us to compare the mean values of various variables in the two experimental groups before our study began. We conducted t-tests on variables such as age, gender, personality traits, motivation, and prior knowledge of the course, all of which were grand mean-centered prior to analysis. The aim was to identify any significant differences in the distribution of these factors across the two experimental conditions, to ensure that any observed differences in outcomes were a result of our experimental manipulation rather than pre-existing group disparities that might impact the internal validity of our study.

The results showed no significant differences between the groups, matched by a heterogeneous or homogeneous contribution of extraversion in terms of gender ($t(164) = -0.11$, $p = .83$), age ($t(164) = -0.07$, $p = .53$), prior knowledge ($t(164) = -0.41$, $p = .82$), group orientation ($t(164) = -1.26$, $p = .31$), or any other personality traits. To provide additional information, we have included the mean-level-changes in Table 1.

Table 1. *Descriptive Measures and Mean-level Changes between Groups in First Wave Data*

Variables	Heterogeneous *N= 84*		Homogeneous *N=82*					CI	
	M	***SD***	***M***	***SD***	***t***	***df***	***p***	***Lower***	***Upper***
Gender	1.32	0.47	1.33	0.47	-0.11	164	0.83	-0.15	0.14
Age	22.19	1.82	22.21	1.46	-0.07	164	0.53	-0.52	0.50
Openness	4.35	0.83	4.45	0.89	-0.72	164	0.41	-0.36	0.17
Extraversion	4.21	0.96	4.27	0.87	-0.39	164	0.27	-0.34	0.23
Conscientiousness	4.37	0.84	4.39	0.80	-0.17	164	0.97	-0.27	0.23
Agreeableness	4.15	1.04	4.08	1.00	0.45	164	0.51	-0.24	0.38
Neuroticism	2.96	1.00	3.14	1.00	-1.15	164	0.29	-0.48	0.12
Prior knowledge	37.48	13.72	38.37	14.25	-0.41	164	0.82	-5.18	3.40
Team orientation	3.75	1.23	4.00	1.36	-1.26	164	0.31	-0.65	-0.14

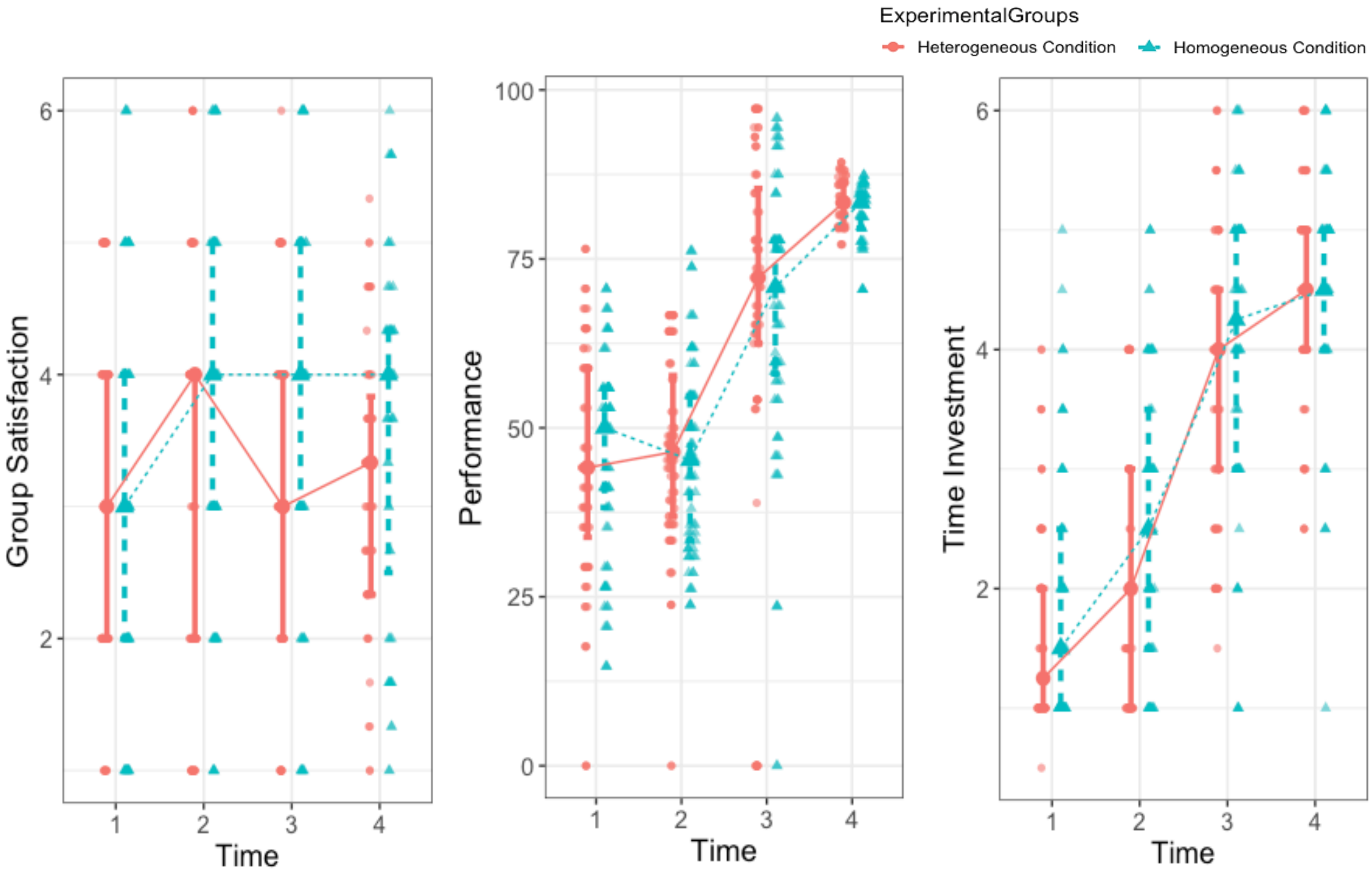
Note. N = 166; CI = Confidence interval. 1 = heterogeneous group formation; 2 = homogeneous group formation. M and SD are represented as the number of observations and standard deviation, respectively.

3.1.1. Visualizing Key Outcomes Over Time

Our primary research question concerns the influence of individual differences in extraversion on group work dynamics and outcomes. To investigate this, we present visualizations of mean-level changes for the three outcome variables - performance, group-work-satisfaction, and time investment - across four measurement-time points, separately for the two experimental conditions based on extraversion distribution. Figure 2 illustrates the results for these dependent variables in the order of group satisfaction, performance, and participation for both heterogeneous and homogeneous groups over four time points.

The figure suggests that the level of extraversion within a group may have varying effects on different outcomes. Differences in satisfaction and time investment have implications for group collaboration, while the similarity in performance may indicate that diverse perspectives and skill sets do not necessarily impact this outcome. However, further analysis, including accounting for the multi-level structure of the data, is necessary to fully comprehend these findings.

Figure 2. *Group Satisfaction, Performance and Time Investment over Time*



Note. Figure 2 presents the outcomes for both groups over the four time points. Results (left to right): First time point heterogeneous groups slightly higher in satisfaction, similarly at the second time point, at the third and fourth time points, homogeneous groups with higher satisfaction, heterogeneous group with lower standard deviation. Performance by heterogeneous groups slightly better. More time investment by homogeneous groups for all measurement time points, with the similar standard deviations.

3.2. Testing Hypothesis 1: Contribution of Variance by Individual- and Group-Level over Time

This section presents multilevel models that explain the variance contribution of parameters. To enhance interpretability, we present predictors, that have been grand mean centered. The Intraclass Correlation Coefficient (ICC), also referred to as the Variance Partition Coefficient (VPC), serves as a valuable metric for assessing variation at each level, especially in models featuring random intercepts. For clarity, we have included all nontrivial ICCs, defined as those exceeding .05, to evaluate the significance of higher levels as the ICC increases (Hox et al., 2017). Table 1 showcases the ICCs for the dependent variables in the Intercept-Only-Model at both individual- and group-levels. However, the ICC for the dependent variable 'performance' was negligible at the individual-level (.001) and substantial at the group-level (.080) and is therefore not included in Table 2.

Table 2. *Results of Random Intercept-Only Model per Individual- and Group-Level*

Dependent Variables	**Individual-Level**	**Group-Level**
Satisfaction Group Work	39%	44%
Time Investment	33%	57%

Note. Indication of percentages, residual variance remains open.

3.3. Considering Outcomes by the Distribution of Extraversion at the Group-Level

Group Work Satisfaction. To investigate the impact of extraversion distribution on group work satisfaction, we utilized the Random Slope Model. Our findings substantiate Hypothesis 1a. Additionally, we incorporated the experimental condition (extraversion distribution) in both the Random Intercept Model and the Random Slope Model, to gauge its explanatory power. Extraversion exhibited significant variance in both models, with the Intercept Model offering a slightly better fit. Contrary to Hypothesis 2a, our results indicate that groups homogeneously formed in extraversion were more satisfied with groupwork than those with heterogeneous distribution. The summarized results are presented in Table 3.

Table 3. *Individual-level and Group-level Predictors of Group-Work-Satisfaction*

	Random Intercept Model	Random Slope Model
Experimental Condition	0.31	0.38*
	(0.22)	(0.19)
Constant	3.21**	3.23**
	(0.15)	(0.13)
Observations	404	404
Log Likelihood	−682.24	−682.69
AIC	1,376.48	1,381.39
BIC	1,400.49	1,413.40

Note. Experimental Condition: Criterion Extraversion heterogeneous = 0 homogeneous = 1. Unstandardized coefficients were reported. Robust standard errors are reported in parentheses. Missing data handled with case deletion. *p = .05. **p = .01.

Time Investment. For the variable of time investment, the model for intercepts per individual did not yield the best fit. We compared a model for the Random Intercept of groups to the Intercept-Only Model to determine, if a group-level model was a better fit. The best-fitting model for time-investment was the intercept-per-group model, affirming Hypothesis 1b. Like group satisfaction, we evaluated whether the experimental variable of extraversion distribution could explain the variance. Unfortunately, our results did not support Hypothesis 2b, suggesting that groups with a heterogeneous distribution of extraversion did not outperform those with a homogeneous distribution. These detailed findings are provided in Table 4.

Table 4. *Individual-level and Group-level Predictors of Time Investment*

	Random Intercept Model	Random Slope Model
Experimental Condition	0.13	0.13
	(0.18)	(0.18)
Constant	3.00**	2.99**
	(0.12)	(0.12)
Observations	523	523
Log Likelihood	−943.94	−943.85
AIC	1,899.88	1,903.70
BIC	1,925.43	1,937.78

Note. Experimental Condition: Criterion Extraversion heterogeneous = 0 homogeneous = 1. Unstandardized coefficients were reported. Robust standard errors are reported in parentheses. Missing data handled with case deletion. *p = .05; **p = .01.

Performance. The best-fit model for performance was the model for the intercept of individuals per group, representing a three-level model structure. As with previous analyses, we introduced the experimental variable of extraversion distribution to both the Random Intercept Model and the Random Slope Model. Our results substantiate Hypothesis 1c, indicating that group-level factors significantly influence performance outcomes. Hypothesis 2c, positing that heterogeneously distributed groups would outperform homogeneously distributed groups, was not supported. Detailed results are provided in Table 5.

Table 5. *Individual-level and Group-level Predictors of Performance*

	Random Intercept Model	**Random Slope Model**
Experimental Condition	−0.15	−1.16
	(1.61)	(1.77)
Constant	26.92**	27.12**
	(1.14)	(1.25)
Observations	619	487
Log Likelihood	−2,677.81	−2,117.74
AIC	5,367.62	4,251.47
BIC	5,394.18	4,284.98

Note. Experimental Condition: Criterion Extraversion heterogeneous = 0; homogeneous = 1. Unstandardized coefficients were reported. Robust standard errors are reported in parentheses. Missing data handled with case deletion. *p = .05. **p = .01.

4. Discussion

Based on the research question of the impact of extraversion variance on group formation, this study aimed to investigate the role of group formation (Hypothesis 1) by experimentally manipulating the distribution of extraversion in groups and examining its effect on outcomes (Hypothesis 2). To achieve this, an experiment was conducted, in which participants were randomly assigned to either heterogeneously or homogeneously distributed groups based on their level of extraversion. The outcomes assessed in this study were group work satisfaction (a), time investment (b), and performance (c). In line with Hypothesis 1, we provide evidence of the relevance of group formation as a crucial factor in group work, while the rejection of Hypothesis 2 raises questions about the assumption of the benefits of extraverted individuals as leaders in collaborative learning.

4.1. Interpretation of Results based on Group-Roles and Hierarchies

The results of the study are aligning with Hypothesis 1, which states that group-level variance plays a more important role in explaining outcome variables than individual-level variance. This finding is in line with previous research on the effects of group formation (Horwitz & Horwitz, 2007; Loignon et al., 2018; Mannix & Neale, 2016; Voltmer et al., 2022), highlighting the importance of considering group formation as a relevant factor in group work. However, further analysis of other outcome measures is warranted.

Based on our results, Hypothesis 2 needed to be rejected. However, the result here was contrary to our initial expectations, as it revealed that groups formed homogeneously in extraversion showed significantly higher levels of group work satisfaction. While this finding contradicts some previous Studies (French & Kottke, 2013; Humphrey et al., 2007; 2011), it is consistent with Wilson et al.'s (2016) study, which demonstrated that groups composed of members with similar levels of agreeableness and extraversion tend to express more positive emotional displays during negotiation, leading to faster agreements, less relationship conflict, and more positive impressions of their negotiation partners, regardless of whether the groups were similarly high, average, or low in these two traits. Likewise, Jackson et al. (2019) observed a tendency for participants to form groups based on similarity, indicating an unconscious or conscious bias toward homophily over time. Additionally, our results find support in Shemla et al.'s (2016) review, which indicated that perceived group heterogeneity can yield both positive and negative outcomes. This underscores the variability in research on this topic, not only in terms of definitions and concepts, but also in methodological approaches to measuring heterogeneity. Furthermore, our study's results challenge the assumption that collaborative learning benefits from the dynamics of an extroverted leader and less extroverted followers. Although a hierarchical group structure has the potential to benefit group effectiveness by increasing coordination and improving communication patterns, it can also create conflicts that harm group effectiveness, especially when aspects of the group structure and hierarchy itself create conflicts (Greer et al., 2018; Tiedens & Fragale, 2003; Woolley et al., 2022). Rather than relying on a hierarchical structure, our results suggest that group work can be a suitable working method for groups homogeneously formed in extraversion, as such a distribution was experienced as significantly more satisfying by its members.

The finding of significantly higher levels of group work satisfaction based on the homogeneous distribution of extraversion is important because positive experiences with group work can shape students' perceptions and attitudes toward working collaboratively, as an intricate link of group effectiveness to the satisfaction of its members (Harris et al., 2017; Lu & Hu, 2005; Mohrman et al., 1995). Therefore, higher education institutions must prioritize student satisfaction in their group-work-activities to ensure a more positive and successful long-term experience (Fazal-e-Hasan et al., 2021). Previous research has demonstrated a strong relationship between students' past and present group experiences, highlighting the importance of individual satisfaction as a critical factor for future collaborative work (Peeters et al., 2006). A negative group work experience can lead to negative anticipation of future group work exposure, while a positive experience can lead to more anticipating and socially skilled group work behavior in the long term and reinforce the benefits of learning in small groups (Hillyard et al., 2010).

Research on the different effects of group formation considering the heterogeneous and homogeneous distribution of criteria is limited. A meta-analytic integration of previous research on the effects of homogeneous and heterogeneous groups found relatively small, combined effect sizes in favor of heterogeneous groups (Bowers et al., 2000), suggesting that differences may not be substantial. However, researchers may not always be able to establish a convincing causal relationship between the chosen research design for group formation and observed results. Here, not considering criteria related to personality traits before the group formation led to lower results (Revelo-Sanchez et al., 2021). This is particularly true when neither an experimental research design nor theoretically validated group formation criteria are utilized, which leads to the strengths and limitations of the respective study. Overall, in line with previous findings (Blasco-Arcas et al., 2013), our results emphasize the role of active collaborative learning, while integrating technologies to improve students' learning performance. In line with others, the relevance of personality traits must be further experimentally investigated (Maqtary et al., 2019; Revelo-Sanchez et al., 2021). Factors such as gender and specialty play a substantial role in shaping group work dynamics (Kucukozer-Cavdar & Taskaya-Temizel, 2016), challenging conventional beliefs.

4.2. Strengths and Limitations of the Study

This study contributes to the existing body of research on group formation by using an experimental design to investigate the role of extraversion distribution. The experimental design allowed causal assumptions to be made about the results, adding to the understanding of the relationship between extraversion and group formation. As such, the strength of this study is the use of an experimental design, which ensures that the effect of extraversion on group formation can be isolated and attributed to the independent variable rather than other extraneous factors. Furthermore, the use of a university-seminar-structure in the research setting provides several benefits for studying group formation and its outcome variables. This setting allows for the examination of groupwork over a longer period, providing valuable insights into group development and dynamics. The use of longitudinal data and consistency of the groups working on tasks over the course of the study enhanced the generalizability of results. Additionally, the study utilized a robust methodology that minimized the potential for errors and biases in the results. An algorithm that considers individual differences provides a cheap and economical means for group formation in various contexts such as education, business and the public sector.

Despite these strengths, some limitations of this study should be considered when interpreting its findings. One major limitation is the small sample size, which may have affected the statistical power of the study and its ability to detect significant effects. This was particularly relevant given the high standard deviation of extraversion scores among participants, making it challenging to detect the effect of different grouping extremes due to statistical power. A general limitation concerns the underlying field of experimental conditions, in which one cannot control or account for all potentially confounding variables. Another limitation is the potential bias of the self-reported data. Participants were asked to report their thoughts, behaviors, and experiences, which may not always be an accurate representation of their true experiences. In line with this, subjective responses in homogeneous groups might be considered less objective or accurate than responses in heterogeneous groups (Apfelbaum et al., 2014), which provides the potential for misleading interpretations of the effects found. Additionally, as personal information is required to form groups, it is essential to consider the willingness of individuals in a particular setting to take a personality test as part of the group formation process, which may impact the generalizability of these findings to other contexts. Replication issues may arise with the construct used in the study 'Group work satisfaction scale.' It is necessary to replicate Studies with the construct and adhere to current reporting standards, including measures of internal consistency and evidence of convergent and discriminant validity. However, further research is needed to fully address these concerns.

Additionally, the results of this study are limited in their generalizability to other populations outside higher education. While the sample was representative of the population being studied, it included a preselected set of people, such as students majoring in teacher education, in a specific age range in the given set of the course structure. Therefore, these findings may not necessarily be applicable in other populations or settings. This emphasizes the need for further research to better understand the effects of homogeneity on group outcomes.

4.3. Contributions and Implications for the Research Field and Educational Practice

The present study has made valuable contributions to understanding the relationship between group formation, distribution of extraversion, and several group-work-outcomes. However, given the strengths and limitations of this study, several key issues must be addressed in future research to deepen the understanding of the underlying mechanisms and implications.

First, the results of this study challenge previous research that supports the superiority of a heterogeneous distribution of extraversion, as homogeneously distributed extraversion was found to enhance group-work-satisfaction. However, such conflicting findings highlight the need for further research into the makeup of groups' extraversion. When replicating the study, researchers should focus on enhancing the application of experimental research. Here, the methodological difficulties of researching groups can usually be solved only through appropriate research designs, methods, and reviews based on experimental study results, to contribute to theory building in such an important area (Shemla et al., 2016).

Secondly, some methodological implications should be considered. The experimental design for future research must be stressed. Additionally, increasing the sample size enables the examination of context-specific effects and increases the generalizability of the findings. Incorporating control groups, for example, with no group formation, and using other objective outcome measures or group-process data, such as video analysis and qualitative data, is recommended to improve the validity of the results and reduce the influence of extraneous factors. An exemplary outcome measure is the observational rubric for assessing collaborative disciplinary engagement in groups (Rogat et al., 2022). This approach takes advantage of observational methods and provides a rubric for quality assessments that enable efficient qualitative analysis of larger samples.

Third, besides replicating previous findings, research should examine the impact of extraversion under different conditions (e.g., school, workplace, or private study groups), settings (e.g., short-term vs. long-term, voluntary vs. mandatory, and present vs. virtual), and populations (e.g., students from other disciplines, schoolchildren, adults, persons with special needs, and educational requirements). Replicating and extending the findings across different populations, settings, and conditions would allow for greater generalizability and enhance our understanding of the mechanisms underlying group-work hierarchies and enable us to better predict group-outcomes.

Future research should consider additional criteria for group formation, such as more indirect components of group work, like the attributes of communication skills, fluency in using computers, and group-work-attitude (Acharya & Sinha, 2018). As Chen and Kuo (2019) state, while it is essential to acknowledge the benefits of systematic group formation, further research is needed to explore the implications of diverse group roles and the potential influence of factors such as gender diversity.

Finally, some practical implications for education can be derived from this research, including a stronger focus on systematic group formation, evaluation of group work satisfaction in addition to performance, and potential for data-driven approaches (e.g., algorithmic group formation) in the educational sector. However, ethical considerations such as informed consent, proper authorization for psychometric tests, and privacy must be considered when implementing such strategies. In line with others, we recommend the utilization of genetic algorithms for group formation in collaborative learning scenarios (Ani et al., 2010) and propose to further explore personality traits as a grouping criterion, in line with recent experimental research employing validated constructs (Revelo-Sanchez et al., 2021). In university settings with small groups within selective work environments, where students are often grouped without the option to decline applications, our methodology offers clear advantages. Dealing with a predetermined population and the need for effective grouping of all students makes an algorithmic approach, accounting for individual differences like extraversion, essential for optimizing group outcomes (Shemla et al., 2016).

While our initial hypotheses on the superiority of heterogeneous extraversion in groups lacked support, noteworthy groups homogeneously distributed in extraversion significantly impacted satisfaction. In our study, the group performance score showed overall little variance, limiting the potential to differ between experimental groups. However, group performance is closely tied to member satisfaction, with dissatisfied members known to hinder overall performance (Mohrman et al., 1995). Moreover, individual satisfaction within group contexts has broader implications for future collaborative work (Peeters et al., 2006). Recognizing and nurturing individual satisfaction within group settings could lead to downstream effects, potentially reducing absenteeism (Makary et al., 2006), considering its significance in university student life. Despite the well-established association between extraversion and positive affect (Wilt et al., 2012), this finding underscores the crucial role of satisfaction in group dynamics.

Given our predominantly female sample, it's imperative not to overlook gender dynamics, aligning with Woolley et al.'s (2022) findings suggesting that a higher proportion of women in a group can enhance overall group performance. Hence, understanding the influence of gender on group interactions is critical, guiding group formation strategies in university contexts. Additionally, the significance of interaction quality and quantity in collaborative learning outcomes should be considered in group work research, as highlighted by Yücel and Usluel (2016), aiding in better understanding the dynamics of group work settings.

In professional work environments, where there's typically more control over group member selection, the applicability of our approach may vary. Tailoring the approach according to project goals becomes essential, particularly in tasks involving diverse responsibilities, where factors like extraversion and other traits might significantly shape the hierarchical structure of the group. Moreover, when dealing with culturally diverse groups, acknowledging the influence of the cultural context is imperative. While our study focused on a Western academic population, recognizing that different cultures may have distinct preferences regarding extraversion is crucial. Future research should explore how extraversion impacts groupwork within diverse cultural contexts, enriching our comprehension of how cultural factors interact with personality traits.

Distinct student populations may exhibit specific differences in group formation outcomes (Acharya & Sinha, 2018). In initial programming courses, considering personality traits notably improved collaborative performance, especially in software development activities (Ani et al., 2010; Revelo-Sanchez et al., 2021). Besides these differences, various factors such as group composition, setting, and the relevance of the group-work-situation should be carefully considered in each unique case to craft nuanced and tailored group-formation solutions.

In the realm of human-AI collaboration and learning in CSCL-environments, our research implies potential extensions. Our study sheds light on the roles that AI-based systems can undertake within a group. Similar to human group members, AI can assume varied roles such as a tutor, evaluator, peer, or manager (Seeber et al., 2018). The significance lies in carefully considering these roles when forming AI-human groups for specific tasks in learning environments. The findings underscore the importance of distributing tasks meaningfully, leveraging the complementary capabilities of humans and AI, rather than having AI take over every role within a group.

The research on AI-based group-roles emphasizes the context-specific nature of choosing the role for an AI-system. While AI can enhance human cognition in decision-making tasks

(Jarrahi, 2018) or act as an idea evaluator (Maher & Fisher, 2012), this role-assignment should be deliberate and aligned with the specific learning objectives and context in CSCL settings. For instance, an AI-based system might serve as an emotional support agent (Hofeditz et al., 2022) or a peer (Elshan & Ebel, 2020), depending on the learning goals. Moreover, our research contributes to the understanding of the role of group formations. Similar to human-only groups, the effectiveness of AI-human groups hinges on contextual factors, the nature of tasks, and specific objectives. It is crucial to assemble groups that harness each other's skills and capabilities, leading to higher performance and superior outcomes.

4.4. Conclusion

In summary, our study highlights the impact of extraversion-based group formation on group work outcomes in a higher education setting. Specifically, groups with a homogeneous extraversion distribution tend to report higher satisfaction in their groupwork than those with heterogeneous distribution. These findings stress the importance of considering both individual personalities and the collective extraversion of traits within groups to enhance overall outcomes. Challenging existing assumptions, these outcomes highlight the intricate nature of group dynamics, urging further research to devise effective strategies, adaptable across various contexts. Our study contributes to comprehending the framework for successful group work, particularly in educational and professional domains, laying the groundwork for future research by utilizing algorithmic group formation. However, to fully comprehend broader implications and limitations, additional research remains imperative. We advocate for continued exploration of the makeup of group members' personality traits as pivotal elements in forming highly effective groups.

6.3.1 Summary of Study 3 and Motivation for Study 4

Study 3, conducted over a term in an educational science course involving face-to-face group work, aimed to investigate the impact of extraversion distribution on group-work-outcomes. The results revealed that members in groups with homogeneously distributed extraversion exhibited significantly higher satisfaction with groupwork.

Encouraged by study 3 findings, and in response to the challenges faced in earlier Studies, in study 4, advanced evaluation questionnaires were administered. This strategic adjustment aimed to provide a more comprehensive understanding of face-to-face-group-work-dynamics. In this subsequent investigation, the goal was to replicate and generalize the findings by exploring the effect of extraversion distribution on groupwork outcomes in a broader context. Therefore, the sample of the study was expanded to include teacher education students from the University of Mainz and international business students from the University of Reutlingen. The results affirmed the homogeneous distribution of extraversion to enhance performance and member participation. This continuous effort to refine research methodologies highlights the commitment to advance group research in various educational settings.

6.4 Manuscript 4: Müller, A., Goeddeke, A., Kneip, P., Konert, J., Röpke, R., & Bellhäuser, H. (under review). How the Experimental Manipulation of Extraversion Distribution Impacts Group-Work-Outcomes in Technology-Assisted Education. *Computers and Education Open.*

Experiment on Extraversion Distribution in Groups Through a Group-Formation-Algorithm

Abstract

Advancements in technology have sparked a surge of interest in systematic group formation in educational contexts. This study aims to investigate the impact of group formation based on extraversion distribution, which is expected to influence group hierarchy, on group work outcomes. To address this, we employed an algorithm to create groups with either consistent or varied levels of extraversion. Over the course of a semester, a total of N = 114 students participated in multiple parallel seminars and were allocated into groups using an algorithmic group formation tool, that resulted in two experimental conditions: one with heterogeneous group members in terms of extraversion levels and another with homogeneous group members. Group formation through group-level extraversion was found to significantly influence performance and participation. Specifically, a homogeneous distribution of extraversion had a positive impact on group performance, as evidenced by improved grades in course-group assignments and increased active participation in group meetings. These findings underscore the importance of considering personality traits on the group-level during group formation, to enhance the success of group work in higher education.

1. Introduction

Working in groups is a widely recognized technique for promoting learning progress through social interaction (Ertmer & Newby, 2013; Johnson et al., 2000). It is commonly used in various educational and professional settings (National Education Association, 2014). However, despite the benefits of group work (Kyndt et al., 2013), Studies have also highlighted potential downsides (Chang & Brickman, 2018; Chiriac, 2014). The underlying group-dynamic processes and factors, that lead to either success or failure in group work, are not fully understood (Mathieu et al., 2017). Therefore, research has focused on developing effective strategies for forming, shaping, and actively managing groups to ensure positive outcomes (Graesser et al., 2018; LePine et al., 2011).

Optimizing group formation is essential for promoting group work in higher education, which is critical for improving students' learning experiences and outcomes, but still inconclusive in research results (Borges et al., 2018). In previous research, experiments were conducted on group formation, primarily within the context of online group work. In these Studies, additional criteria were chosen for experimental group formation, and their effects on group work outcomes were investigated. However, it's important to note that these experiments were conducted in online settings, whereas the current study is centered on face-to-face interactions. Notably, the present study features successful longitudinal data collection with minimal dropout, setting it apart from other group-work-research, including some of our own work (Bellhauser et al., 2018; Müller et al., 2022). The unique experimental study design represents a distinctive feature, yielding valuable insights into the dynamics of group formation.

Building on this foundation, our research aims to delve specifically into the implications of extraversion distribution within groups. As noted by Mohammed and Angell (2003), extraverted individuals, with their assertiveness and confidence, tend to dominate discussions and inadvertently establish hierarchical structures. This phenomenon has been further elucidated by Studies conducted by Taggar et al. (2006) and Wilmot et al. (2019), indicating that a heterogeneous extraversion distribution fosters leadership by those with higher extraversion, ultimately leading to the formation of a group hierarchy. Consequently, the present study endeavors to expand upon these observations by conducting a thorough examination of how varying degrees of extraversion influence the dynamics and functioning of heterogeneous groups in face-to-face interactions.

1.1. From Random to Strategic: The Didactic of Group Formation

To form groups effectively, teachers require adequate support to manage logistics and ensure timely execution. A systematic approach to group formation requires prior knowledge of the characteristics of the individuals to be assigned to groups, as well as the techniques to perform group formation, including technical support required (Maqtary et al., 2019). However, this is not trivial, as the application of group formation techniques in collaborative learning contexts is a complex process and combinatorial challenge (Hwang et al., 2008) that is influenced by various factors.

With the advent of technology-supported learning environments, algorithmic tools have gained popularity for their potential to facilitate objective and criteria-driven group formation (Bellhauser et al., 2018; Müller et al., 2022). The use of algorithmic group formation tools allows for experimental research to investigate relevant group formation criteria and their constellation, driven by the goal of providing the best possible group-learning-experience for each learner (Maina et al., 2017; Maqtary et al., 2019; Sun & Chen, 2013). Additionally, algorithms can efficiently handle large datasets and intricate combinations of criteria, overcoming the limitations of manual allocation methods (Wilkinson et al., 2010). However, it's important to acknowledge that the effectiveness of algorithms depends on the quality and relevance of the criteria used for grouping, and they may not account for nuanced contextual factors that teachers can consider (Dwork et al., 2012). Given the positive contribution of systematic, algorithmic group formation to group work outcomes, as shown in previous research (Borges et al., 2018; Liang et al., 2021; Mujkanovic et al., 2019; Odo et al., 2019), the question now is, how to implement it for these positive group work outcomes. For this, more research is needed to examine the criteria for group formation. This research is essential for developing evidence-based best practices in group formation (Patrício & Franco, 2022). In this regard, the role of individual differences, especially personality traits, could be further investigated as potential criteria for group formation.

1.2. Interplay of Individual Differences and Group-Level-Resources in Group Formation

Building on the concept of individual and group-level factors, it's essential to consider various individual differences during the group formation process. These differences encompass demographics, personality traits, attitudes, and cognitive preconditions of group members (Hübscher, 2010; Müller et al., 2022; Revelo-Sánchez et al., 2020; 2021). The choice between homogeneous or heterogeneous distribution of these characteristics within learning groups can depend on the specific criteria used (Apfelbaum et al., 2014; Bell et al., 2018; Bowers et al., 2000). Furthermore, it is essential to test and distinguish the independent effects of group formation and individual traits. Model-fit analysis reveals whether group- or individual-level factors play a more significant role in explaining group differences in outcomes within a specific context, assessing the relative impact of both group- and individual-level factors (Hitt et al., 2007).

Importantly, the configuration of trait expressions of specific characteristics among group members is a crucial factor in determining outcomes. It often holds more influence than the isolated trait expressions of individual members. This configuration implicitly triggers group dynamics, such as the formation of hierarchies, which, in turn, impact outcomes over time (Blanco-Fernández et al., 2023). This interplay raises the question of whether group formation exerts an independent influence on outcomes, distinct from individual characteristics.

In the broader context, existing literature highlights the significance of distinguishing between group- and individual-level factors to gain a comprehensive understanding of group behavior (Kozlowski & Bell, 2013). Previous Studies consistently demonstrate the discernible effects of group formation at the group-level, influencing decision-making, problem-solving, and creativity (Mannix & Neale, 2016; Van Knippenberg & Schippers, 2007; Voltmer et al., 2022). Leadership dynamics and communication patterns, inherently linked to the constellation of individuals' personality traits, significantly impact group performance (Gawande et al., 2003; Zennouche et al., 2014). In essence, understanding the factors responsible for the variability in group work outcomes is paramount in the field of group formation research. It becomes increasingly evident that the outcomes of groups are significantly shaped by the dynamic interplay between individual- and group-level factors over time (Blanco-Fernández et al., 2023), leading us to question whether it is reasonable to exclusively attribute group outcomes to an individual's isolated trait configuration. This emphasizes the necessity of adopting a holistic perspective that acknowledges the joint influence of the contextual environment in which groups operate (LePine et al., 2011).

1.3. Understanding the Big Five Personality Traits

The "Big Five" personality traits, also referred to as the Five-Factor-Model, represent a well-established and widely utilized psychological framework (Hough & Oswald, 2000, 2005). This model provides a standardized and valid framework for comprehending and predicting personality (Saucier & Ostendorf, 1999; Stewart, 1999). It encompasses five fundamental personality dimensions: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Costa

& McCrae, 1992, 2008). The Big Five-model is employed to depict an individual's personality based on these traits, each characterizing general personality features at varying levels of abstraction. Various tests and questionnaires, such as the NEO Personality Inventory-Revised (NEO-PI-R), the Sixteen Personality Factor Fifth Edition (16PF-5), and the Big Five Questionnaire (BFQ), have been developed to assess an individual's standing on each of the Big Five dimensions (Saucier & Ostendorf, 1999; Stewart, 1999).

The Big Five traits, regarded as a construct-oriented approach, have been extensively studied, both within the context of individuals and groups, with the potential to influence thoughts, emotions, behaviors, and social interactions (Bell, 2007; Costa & McCrae, 1992, 2008; MacDonnell et al., 2009; Mohammed & Angell, 2003; Peeters et al., 2006; Wang & Hsu, 2012). Numerous Studies consistently highlight their significant impact on individual behavior and group outcomes (Fleeson & Gallagher, 2009). Furthermore, the examination of specific trait constellations in group formation can have a positive influence on group hierarchy, power dynamics, and overall group functionality (Marks et al., 2001; Mujkanovic & Bollin, 2019; Perry, 2021). Nonetheless, comprehending the relationship between the personalities of individual group members and group outcomes remains an ongoing challenge (Borges et al., 2018; Prewett et al., 2018). While the Big Five model provides a standardized and effective framework for describing personality, achieving consensus on how to employ these traits as criteria for group formation and how their distribution among group members impacts overall group performance and social behavior continues to be a subject of ongoing discussion (Bell, 2007; Driskell et al., 2006; Lykourentzou et al., 2016).

1.3.1. Group Formation considering the Expression of Extraversion

Extraversion is characterized by sociability, assertiveness, and a tendency to seek out and enjoy social interactions. This trait indicates how individuals may exhibit friendliness, approachability, talkativeness, and activity (Barrick et al., 1998; Hough, 1992; Judge et al., 2002; Neuman et al., 1999). Individuals with high scores in extraversion tend to be sociable, enjoy interactions with others, and possess an easy-going nature. Moreover, those with high levels of extraversion often demonstrate leadership behaviors, maintain a positive attitude toward group interactions, and prefer engaging in social activities. Conversely, individuals with low scores in extraversion may exhibit more reserved behaviors and may lean towards solitary work (Costa & McCrae, 1992). Due to its interpersonal expression and direct association with social behavior, extraversion plays a pivotal role in understanding an individual's interactions within groups (Barrick et al., 1998; Hough, 1992; Judge et al., 2002; Neuman et al., 1999) and is related to the quality of group interaction (Wilmot et al., 2019). Further, its association with leadership behavior and improved group communication makes it an interesting trait to consider when forming groups (Bell et al., 2018; Judge et al., 2002).

Literature on group formation suggests that the level of extraversion among group members can have a significant impact on group functioning (McCabe & Fleeson, 2012; Wilmot et al., 2019), particularly social interactions and group role formation, such as the formation of hierarchical or non-hierarchical group structures (Taggar et al., 2006; Wilmot et al., 2019). In a group, people with high extraversion initiate discussions (Judge et al., 2002; Mohammed & Angell, 2003) and provide support to other group members (Porter et al., 2003), and because of that, they often take over the leadership role, while less extraverted individuals are more likely to follow their lead (Nonaka et al., 2016; Taggar et al., 2006). Thus, one can infer that a heterogeneous distribution in extraversion implicitly establishes a group hierarchy (Kramer et al., 2014).

However, it's essential to note that while extraverted leadership can enhance group performance in situations where employees are passive, it may have a reverse effect when employees are proactive (Tiedens & Fragale, 2003), with extraverts contributing more original solutions and comments than introverts (Yellen et al., 1995). This phenomenon aligns with the dominance-complementarity-theory, which suggests that interactions thrive when dominance and assertiveness are balanced by compliance and submissiveness (Kiesler, 1983). In line with this, extraversion's ability to enhance social experiences might explain extraverts' greater subjective well–being (Harris et al., 2017), and higher leisure-satisfaction and happiness (Lu & Hu, 2005). This underscores the need for a nuanced understanding of how varying degrees of extraversion influence group dynamics and functioning in face-to-face interactions. The literature provides mixed findings regarding the benefits of heterogeneous versus homogeneous distributions of extraversion for group functioning. Some Studies have found that heterogeneity in extraversion is beneficial for task delegation, and that heterogeneity in dominance, related to extraversion, enhances the management of relationship conflicts (Humphrey et al., 2007; Tekleap & Quigley, 2014). Conversely, others suggest that a homogeneous distribution of extraversion may lead to higher team innovation (den Hartog et al., 2019).

Despite this mixed evidence, there is a general assumption that a heterogeneous distribution of extraversion is superior to a homogeneous distribution for group functioning (Bell, 2007; Roney et al., 2012; Thanh & Gillies, 2010) as it can promote cooperative communication, clarify roles and responsibilities, and facilitate coordination and decision-making (Roney et al., 2012; Woolley et al., 2022). However, research shows inconclusive results on several group formation outcomes, and research concerning specific, relevant outcomes influenced by distributions of extraversion has not been thoroughly studied, leaving questions about the effectiveness of group formation strategies (Maqtary et al., 2019). Our research question aims to fill this gap by experimentally exploring group formation based on the distribution of extraversion, with the hope of generating more knowledge about its systematic use for group formation. Our research aims to fill this gap by experimentally investigating group formation by the distribution of extraversion, hoping to generate more knowledge about its use for systematic group formation.

1.4. Research Goals and Justification for Group Formation by Extraversion

The use of personality as a criterion for group formation has received relatively little attention in research, underscoring the necessity for further investigation (Borges et al., 2018). The existing research on group formation reveals notable gaps, often characterized by a reliance on correlational designs that cannot establish causal relationships or the use of invalid criteria for group formation (Kirschner, 2017; Klein et al., 2009). Additionally, there is limited research on the impact of various group formation criteria on group outcomes (Eckhaus & Davidovitch, 2019; Eckhaus et al., 2017; Filade et al., 2019; Zheng & Pinkwart, 2014). These limitations present significant challenges in ascertaining the most effective methods for selecting and weighing group formation criteria. Moreover, efforts to address the group formation problem in diverse contexts have yielded incomplete solutions, leaving significant voids in the existing body of literature (Maqtary et al., 2019). Furthermore, comprehensive evaluations of current group formation strategies are conspicuously absent, and the assessments of implemented algorithms exhibit notable variations. In certain cases, the algorithmic tools employed for group formation lack clear justification (Odo et al., 2019). Consequently, our understanding of this subject remains fragmented and inconclusive.

Previous research has predominantly focused on performance as the primary indicator of group success or failure (Mehar & Kaur, 2020; Shaw, 2013). However, for a more comprehensive understanding of successful group work, especially in higher education settings, it is crucial to consider additional outcome measures. Given the evidence suggesting that emotions can exert a notable impact on performance, the inclusion of measures assessing group emotions appears to be a valuable addition for a more profound understanding of group success (Putwain et al., 2018). Consequently, it becomes essential to explore other relevant outcome measures that can serve as indicators of successful group work, with a specific emphasis on factors such as group members' satisfaction or willingness to continue group work (Peeters et al., 2006). This broader approach leads to a more holistic and thorough understanding of group success.

Moreover, personality is an essential factor in group functioning and member satisfaction (Müller et al., 2022). Regardless of how a person's personality trait level or distribution interacts with those of other group members, it inevitably affects the group's work process and outcome through the contextual setting in which the group operates (Prewett et al., 2018; Stipelman et al., 2019). Driven by group-social-capital-theory (Oh et al., 2004), we predict that the associations between the distribution of extraversion and group functioning can be conceptualized as group-level resources related to outcomes. The outcome of group formation by the distribution of extraversion should be explained by the structure of the groups and not by the composition of the individuals within these groups (Loignon et al., 2018). This aligns with previous research suggesting that diverse group members can positively affect group outcomes, as it emphasizes the importance of considering various group-level resources in the context of group formation (Van Dijk et al., 2017). The intricate relationship between individual characteristics and group-level dynamics further underscores the need to explore the effects of extraversion distribution on group-work-outcomes, with a focus on the group as a whole. Therefore, we expect more variance to be explained at the group-level than at the individual-level for all outcome measures, and formulate the first hypothesis as follows:

H1: In the models that include experimental manipulation of extraversion, the group-level will explain more variance in the outcome's satisfaction with group work (H1a), member participation (H1b), and performance (H1c) than the individual-level (e.g., lower model fit indicators).

Research suggests that while heterogeneous group characteristics complement each other (Bekele & Menzel, 2005; Moore, 2011; Seong & Hong, 2020), a homogeneous distribution leads to increased comfort and motivation to work together (De Dreu & Weingart, 2003; McPherson et al., 2001). Here, factors such as social homophily (Bradley & Hebert, 1997; Martin & Paredes, 2004) and the similarity-attraction paradigm (Byrne, 1971) influence group similarity preferences, even when similarities may not significantly affect group outcomes (Jackson et al., 2019).

In this study, our aim is to investigate the impact of group formation on various outcome variables by manipulating the distribution of extraversion within groups, categorizing them as either homogeneous or heterogeneous. To offer a more comprehensive understanding, our study will experimentally test the hypothesis that a heterogeneous distribution of extraversion within a group is more advantageous than a homogeneous one, with inconsistent results reported in past research (den Hartog et al., 2019; Müller et al., 2022). Within heterogeneous groups, individuals with higher extraversion levels, characterized by assertiveness and confidence, often dominate discussions, inadvertently establishing hierarchical group structures and fostering leadership emergence (Mohammed & Angell, 2003; Taggar et al., 2006; Wilmot et al., 2019). Building upon previous research that suggests the superiority of groups with heterogeneously distributed extraversion (French & Kottke, 2013; Humphrey et al., 2007; 2011), we aim to experimentally test this research hypothesis. In doing so, we will utilize a range of outcome measures, encompassing both subjective and objective indicators such as satisfaction, participation, and group performance.

Consequently, we formulate the second hypothesis as follows:

H2: Groups with a heterogeneous distribution of extraversion will report greater satisfaction with their group work (H2a), show a higher degree of group member participation (H2b), and achieve better results (H2c) than groups with a homogeneous distribution of extraversion.

2. Method

2.1. Sampling

We recruited participants from two undergraduate classes at two public universities in Germany: Microeconomics at University Reutlingen and Educational Studies at University Mainz. Participation in the courses and groupwork was mandatory. We obtained written consent from all participants and matched international business students ($N = 65$) from University Reutlingen and teacher education students ($N = 58$) from University Mainz into groups of three ($N(groups) = 38$). We excluded a total of 29 students from our initial sample, who had previously attended the seminar and might biases the outcomes due to prior course knowledge. Among these, 16 students belonged to University Reutlingen, while 8 students were from University Mainz. Additionally, students responding carelessly in the questionnaire for experimental group formation were additional excluded ($N = 5$). Those students were excluded in random groups and, consequently, not included in the final sample for analysis. Table 1 displays sample and respective group conditions.

Table 1. *Sample Characteristics and Group Conditions by University after Dropout*

	Total *N*	**Group *k***
Overall	114	38
Homogeneous in Extraversion	60	20
Heterogeneous in Extraversion	54	18
By University		
University X	54	18
Homogeneous in Extraversion	30	10
Heterogeneous in Extraversion	24	8
University Y	65	20
Homogeneous in Extraversion	30	10
Heterogeneous in Extraversion	30	10
Exclusion Criteria	University Reutlingen	University Mainz
Prior Knowledge (Previous Attended Seminar)	16	8
Careless Response in Questionnaires	1	4

Note. Number of participants and groups by university and conditions, after the initial survey. Exclusion Criteria Prior Knowledge (Attended Seminar) N = 29, University X N = 16 before group formation, University Y N = 8. Exclusion Criteria Careless Response in Questionnaires N = 5.

In Condition 1, we established 20 groups with a heterogeneous distribution of extraversion, and in Condition 2, we established 18 groups with a homogeneous distribution of extraversion. We ensured equal motivation and prior knowledge levels across all groups by employing a rigorous randomization and group formation procedure. This meticulous approach ensured that motivation and prior knowledge levels remained constant across all groups. After the algorithm had assigned individuals to groups, the students worked on problem sets in face-to-face-groups throughout the term, completing assignments and evaluations on the quality of their groupwork. Three assignments received grades, and every student completed three evaluations.

The study at University Reutlingen was conducted in an introductory Microeconomics-class during the first term of the freshman year. Students were assigned to two groups (A and B) to allow for an engaging and supportive learning environment. Each class had approximately 35 students, who were taught for a total of 180 minutes weekly. To pass the course, students had to sit a final exam (scoring a minimum of 51 out of 100) and the opportunity to collect up to 20 bonus points throughout the term. The teacher also awarded up to five bonus points for class participation and up to 15 for three group work submissions, making group work mandatory for passing the course.

At University Mainz, participants were teacher education students in the bachelor's program for the teaching profession. Like University Reutlingen, the course was mandatory for graduation and required participation (allowance for a maximum of two missed dates), preparation of meetings at home (literature, slides, podcasts), active participation in discussions, and group work, including turning in three group assignments. The teachers evaluated the group assignments three times, according to previously defined evaluating criteria, with each assignment stimulating group work.

2.1.1. Exclusion Criteria and Data Elimination Procedures

To ensure data quality, we implemented exclusion criteria. We excluded students retaking the course due to their prior knowledge and eliminated participants' data when detecting careless responses or incomplete or missing data, using case deletion. We also excluded data from participants who had not filled in the questionnaire before experimental grouping, resulting in the algorithm placing them in random groups with other participants with missing data. In instances, where participants forgot their codename or misspelled it in the posttest, we were unable to match data over time, leading to the exclusion of these participants from the analyses. To further ensure data quality, we scanned questionnaire data for traces of careless responses and eliminated any data that was deemed unreliable. Finally, the number of students excluded using each exclusion criterion was $N = 29$, for University Reutlingen $N = 16$ and University Mainz $N = 8$, additionally $N = 5$ students were grouped randomly and therefore not included in the final sample.

2.1.2. Justification of Sample Size

Following suggestions by Lakens (2021), researchers needed to consider the resources available to conduct a study. The limiting factor in our study was the availability of students. We recruited as many students as possible, but resources limited our sample size. We also performed an a priori statistical power analysis for sample size estimation, based on data from a previous study (*N(groups)* = 60) (Glimmpse software[1]). The effect size in this study was for outcome variable *performance d* = .03, *participation d* =.07, and *satisfaction d* = .04, considered extremely small using Cohen's (1988) criteria (Müller et al., 2022). With α = .05 and power *d* = 0.90, the projected required sample size approximately reflects outcome measures: *performance N(groups)* = 24, *participation N(groups)* = 36, and *satisfaction N(groups)* = 44 to perform a between-group comparison. Thus, our proposed sample size of *N(groups)* = 38 was adequate. It should also allow for expected attrition and our additional objectives of controlling for possible mediating or moderating factors and interpreting results.

2.2. Study Design and Experimental Conditions

The study employed a longitudinal experimental design, with a factor (extraversion) at two levels (heterogeneous and homogeneous) with three evaluation time points. The Universities, where we conducted the study, utilized Moodle's online learning management system (LMS). The Moodle[2] platform played a central role, as it enabled the implementation of the grouping plugin MoodlePeers, developed for this purpose, which adjusted ensure equal or unequal distribution of mean values for selected criteria across groups.

---

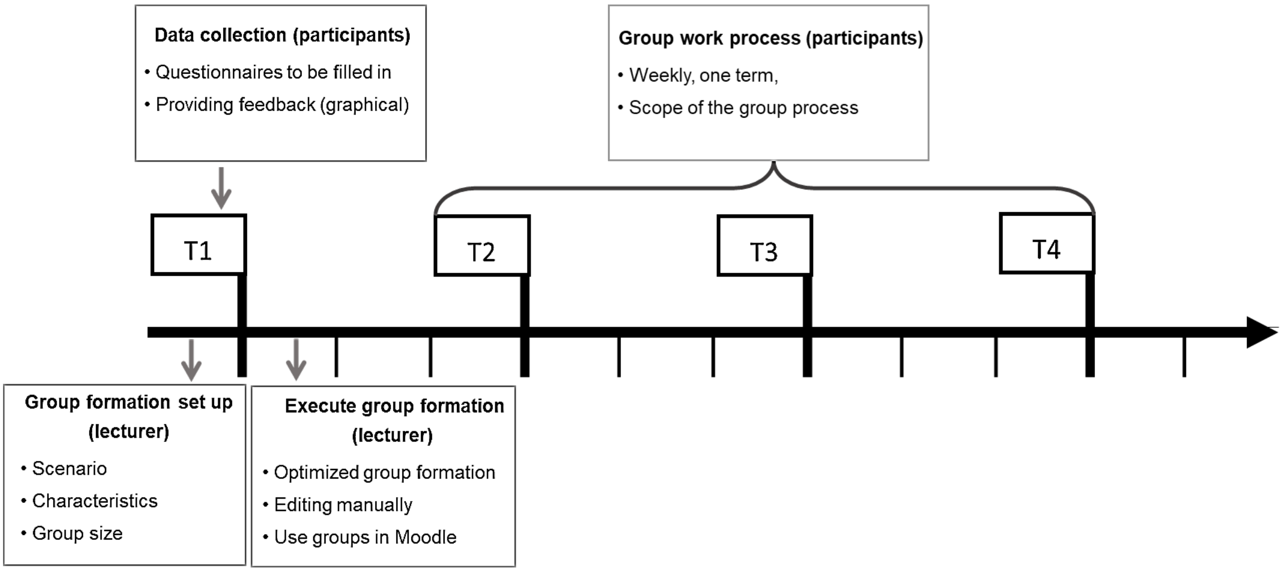
[1] Glimmpse 3.0.0 (samplesizeshop.org) accessed on June 26, 2022.

Students initially completed the T1 questionnaire directly in the Moodle-Platform. Afterwards, the MoodlePeers plugin randomly divided the entire sample into two comparable halves and ensured a balanced distribution of the chosen criteria by testing whether both halves had equal mean values in the group-formation-criterion extraversion and in the control-variables prior knowledge and motivation. Afterwards, the algorithm performed the group formation in both halves separately from each other.

In one half, the experimental group, the algorithm was programmed to create groups that were heterogeneous in extraversion, while trying to achieve a similar extent of heterogeneity in all groups in this experimental condition. In the other half, the control group, the algorithm created groups being homogeneous in extraversion, again aiming to achieve the same extent of homogeneity in all groups in this experimental condition. To control for the impact of prior knowledge and motivation, in both halves the algorithm tried to create groups that were heterogeneous in the control variables simultaneously with the respective manipulation of the distribution of extraversion. Thereby, all resulting groups in both halves have similar standard deviations in prior knowledge and motivation, which necessarily also leads to similar mean values in both control variables. With respect to the experimental condition, all groups in the experimental group had similarly high standard deviations in extraversion, while all groups in the control condition had similarly low standard deviations in extraversion.

Following the group-formation-process, students engaged in coursework tailored to their respective seminars throughout the semester, culminating in three graded group-work-submissions, each complemented by evaluation-questionnaires (T2-T4). At the end of the semester, students completed the T4-final-evaluation-questionnaire. Fig. 1 offers an illustrative representation of the steps involved in the group formation process as well as data collection for the study. The methodology utilized an experimental, longitudinal design, with distinct measurement time points, represented as T1-T4.

Figure 1. *Group Formation and Coursework*



Note. Overview of Study Process by T1 = experimental group formation, T2-T4 group work progress thereafter.

2.3. Measurement Instruments

2.3.1 Control variables

Algorithmic Support. To form groups, we developed a Moodle-Plugin called MoodlePeers for optimizing group formation. This plugin provides a user interface for instructors to set up group formation within a course, as well as the administration of questionnaires and an overview of the status of group formation (e.g., not yet started, open for answers, and groups formed). In addition, the plugin includes an implementation of the optimization algorithm GroupAL. The algorithm was used to determine how to group the participants and apply a different set of matching criteria to each part. By minimizing (or maximizing) the distance between all three group members simultaneously, it maintains the same fitness level (i.e., prior knowledge and motivation) in the matched overall groups (Konert et al., 2014, 2016; Müller et al., 2022). GroupAL was specifically designed to optimize the formation of learning-groups within Moodle. It emphasizes achieving balanced group composition, in terms of relevant criteria chosen. The algorithm endeavors to ensure an equally good fit between all possible pairs within a group, considering both homogeneous and heterogeneous optimization criteria.

In our study, the criteria for homogeneous distribution were student motivation and self-assessed initial prior knowledge level. The algorithm employed the distribution of extraversion, either homogeneously or heterogeneously within groups, as another set of optimization criteria. It is important to clarify that "homogeneous criteria" refer to characteristics that the algorithm aims to make ideally equal within a group, while "heterogeneous criteria" refer to characteristics that the algorithm aims to make ideally different within a group. For further details, please refer to the provided references (Konert et al., 2014, 2016).

Personality. We measured extraversion using the German short version of the Big Five Inventory (BFI-K) (Rammstedt & John, 2005). The personality questionnaire had robust reliabilities in this setting (extraversion: eight items, $\alpha = .89$, conscientiousness: nine items, $\alpha = .83$, openness: five items, $\alpha = .70$, neuroticism: four items, $\alpha = .79$, agreeableness: four items, $\alpha = .64$). We measured prior knowledge as an average subjective rating (“How do you judge your knowledge about the course content,” etc.) ranging from 0 to 100 points.

Motivation. We used four subscales to measure motivation: expectations (four items, e.g.,” I know that I can learn the contents of the preliminary course,” $\alpha = .86$), use (five items, e.g.,” I understand how important the preliminary course is for my future,” $\alpha = .78$), cost (six items, e.g.,” The time required for the preliminary course seems great to me,” $\alpha = .83$), and interest (seven items, e.g.,” I’m looking forward to the preliminary course,” $\alpha = .80$). Reliabilities of the motivation scales were high.

Team Orientation. We measured attitudes toward teamwork (team orientation), using three questions (e.g., “If I have a choice, I’d rather work in a team than alone,” $\alpha = .86$). Reliabilities of the scale were high. Participants rated all questions online, using a scale from 1 (“does not apply”) to 6 (“does completely apply”).

2.3.2. Outcome Measures

We collected data on dependent variables by administering online evaluation surveys at regular intervals, three times throughout the semester, following the submission of each group assignment. The surveys were conducted at the same intervals in both universities. The dependent variables can be broadly categorized into satisfaction, participation of members, and performance-related variables.

Satisfaction. Within the short evaluation questionnaire, participants self-rated their satisfaction regarding groupwork based on 4 items (e.g.,” *I am satisfied how we work together as a group*”). Participants rated questions on a six-point Likert-scale ranging from “1 = *does not apply*” to “6 = *applies*”. Satisfaction comprises the average of the items, whereby higher values indicate higher satisfaction. Reliability analysis showed a good internal consistency ($\alpha = .93$).

Participation. Participation in group work was assessed through self-report of the attendee number using an online evaluation questionnaire. The questionnaire was distributed to participants immediately after each group work session, which was mandatory for all group members and took place on the same day. Participants were asked to report the number of group members actively participating in solving the week's homework, including themselves ("*How many people in your group have actively participated in solving this week's homework?*"). Since each group comprised three members, the possible response options ranged from 1 (= only me) to 3 (= everyone). If no member answered, participation was rated as 0. This method allowed for the individual assessment of participation and collaboration quality, capturing the perspectives of all group members.

Performance. Serving as an indicator for the outcome performance, three homework assignments were graded by the tutors. To transfer the values of the performance scores to be merged for the different universities in a consistent manner, percentages were calculated based on the maximum score. The scale ranged from 0 to 100, with higher scores indicating better performance.

2.4. Transparency and Openness

Before commencing the study, we informed participants about the general experiment, including the group formation process and obtained written consent prior to participation, to maintain transparency and ethical standards. However, we intentionally withheld specific details such as the criterion (extraversion) for group formation and the study's hypotheses to prevent potential bias in participant behavior and evaluation. To address the sensitive nature of the questions asked in the survey, participants were assured of the confidentiality of their responses, and were informed that their raw data would not be shared with any third parties. Data were analyzed using R, version 4.0.0 (R Core Team, 2020) and the package nlme (v3.1-152; Pinheiro et al., 2021). We describe our sampling plan, all data exclusions (if any), all manipulations, and all measures in the study.

2.5. Statistical Models

The manipulated categorical independent variable was personality trait extraversion (either homogeneous or heterogeneous within each group). We measured the dependent variable based on the study-questionnaires, the submitted homework, and the grades obtained for the project. We conducted a multilevel analysis (MLM) due to the research design structure and accounting for the nested data. Thereby, we could test *hypothesis 1*, if group-level or individual-level predictors explained most of the variances. By applying random intercept and random slope models separately and in combination, the model-fit indicators AIC and BIC will answer the question of level. A model with a lower AIC and BIC provides a reasonable fit (Burnham & Anderson, 2016). If the group-level could explain the variance, we could test, if extraversion manipulation supported *hypothesis 2* for the outcome measures, *satisfaction (H2a), participation (H2b),* and *performance (H2c)*.

3. Results

3.1. Data Preparation and Analysis

We ran analyses using SPSS 23.2 and R. Before the analysis, an Epsilon-Test was performed to assess whether the data of the two universities were equivalent and could be combined. The Epsilon Test is a widely used procedure in multilevel modeling that evaluates the comparability of data from different levels, such as the data from the two universities in this study. This test examines the similarity of the variance-covariance matrices of the different levels, and if the result is not significant, it implies that the data can be pooled and analyzed together. Since the analyses revealed that the data sets from the two universities were comparable, and based on that, the necessary conditions were met to proceed with the data analysis of the merged data, which was then carried out. All variables were grand mean-centered, to allow a more straightforward interpretation of results. We used the standard $p = .05$ criterion to determine, if a predictor on any given level explained variance. The post hoc test suggests that the results are significantly different from those expected if the null hypothesis is retained.

3.2. Distribution of Individual Differences across Groups

3.2.1. Results of t-test to Compare Experimental Groups Divided by Extraversion Level

Before conducting the main analyses, we ensured the comparability of the experimental conditions. To achieve this, we performed a t-test to compare the experimental groups divided by their level of extraversion in the grand mean values of the initial survey, which included gender, age, team orientation, and personality traits. We found that there were no significant differences between the groups in terms of gender, $t(110) = 0.30$, $p = .77$, or age, $t(110) = -0.28$, $p = .78$. However, we did find a significant difference in team orientation between the scores for the heterogeneous ($M = -0.29$, $SD = 1.49$) and homogeneous ($M = 0.32$, $SD = 1.08$) conditions, with $t(111) = -2.48$, $p = .02$. In addition, the t-test validated the intended purpose of the implemented algorithm, because the standard deviation of extraversion within groups in the scores for heterogeneous ($M = 0.92$, $SD = 0.58$) and homogeneous ($M = 0.47$, $SD = 0.40$) conditions differ significantly ($t(109) = 4.74$, $p < .01$).

3.2.2. Descriptive Statistics in Outcomes between Experimental-Group-Conditions

To provide an initial overview of the data, we first present descriptive analyses of the main outcome measures *satisfaction, participation, performance* for the two experimental groups, divided by extraversion in the heterogeneously (1) and homogeneously (2) structured group conditions. The results are presented in Table 2.

Table 2. *Descriptive Measures of Main Dependent Variables, as Divided by the Algorithm*

Dependent Variable	Experimental Group	**1**		**2**		**3**	
		N	***M (SD)***	***N***	***M (SD)***	***N***	***M (SD)***
Performance (Assignment)	Heterogeneous	58	42.96 (23.36)	46	56.20 (17.76)	22	70 (18.09)
	Homogeneous	54	44.97 (28.85)	45	67.39 (17.62)	30	69 (10.21)
Satisfaction	Heterogeneous	58	3.59 (1.69)	53	3.33 (1.58)	51	3.34 (1.60)
	Homogeneous	54	3.40 (1.78)	45	3.76 (1.67)	50	3.66 (1.84)
Participation	Heterogeneous	60	2.14 (0.54)	60	2.21 (0.46)	60	1.43 (0.99)
	Homogeneous	54	2.37 (0.49)	54	2.34 (0.55)	54	1.44 (1.22)

Note. N, M, and SD represent the number of observations, the mean, and the standard deviation, respectively. 1, 2, 3 = Time (Measurement time points)

3.3. Randomizing at the Individual-Level for Improved Precision to Test for the Effects of Group-Formation

To prepare for the testing of hypotheses 1 and 2, we constructed a three-level MLM with time (level 1) nested in individuals (level 2) and individuals nested within groups (level 3). We use the term ‘nested’ as each student only learns in one group, and each group is doing evaluations and group tasks over time (three events to be nested). MLMs offer information about which of the levels should be used for randomized experimental conditions. In terms of statistical precision and power, it is usually best to randomize at the lowest level possible, in our case level 2 (Van Landeghem et al., 2005). Therefore, we chose the individual-level (level 2) of a three-level MLM.

MLM analysis also identifies the unexplained variance at each model level. For example, in the case of the influence of other personality traits at the individual-level, we can assume that some elements not considered in any given questionnaire represent unexplained variances at any level. Specifically, by not including information about the groups, we may miss important variables at the group-level that might explain performance at the individual-level. Therefore, we also developed an incorrect model to understand the outcome variable of interest beyond the known problem with underestimating standard errors. In the context of MLMs, including variables at each level is relatively simple, as are interactions among variables at different levels. To sum up, the greater the model complexity, the greater the possibility of understanding the phenomenon of interest (Hox et al., 2017). The empty or null-level-model, initially set up without any explanatory variables, describes the partition (ICC) of variance between the student- and group-levels. Since it was assumed that all 3 group members should have the same score in *participation* and in *performance*, was no variance on the individual-level expected. The resulting ICCs for both individual-level and group-level are listed in Table 3. All outcome measures were found to provide more explanation at the group-level.

Table 3: *Results of Intercept-Only-Model per Individual- and Group-Level*

	Individual-Level	Group-Level
Satisfaction	11%	53%
Participation	0 %	63 %
Performance	0 %	64 %

Note. The intraclass correlation coefficient (ICC) explained variance at the individual- and group-level.

The ICCs do not add up to 100% because they are a measure of the proportion of total variance that is explained by the group-level or individual-level factors. It represents the percentage of the total variability in the outcome that is due to the grouping variable (e.g., group membership) or individual-level factors (e.g., personality traits, motivation, other individual traits). However, there may still be other sources of variability that are not accounted for by these factors (i.e., residual variance) and contribute to the remaining percentage of variance.

3.3.1. Impact of Distribution of Extraversion on Outcomes

We initially established different models to test them against each other before choosing the best model fit. Following the initial step of building an intercept-only model or null-model, we included the variations among participants in the model. Next, we created the models to compare variances across groups alone and individuals within these groups. We tested each model against the null model. Afterward, we added our experimental condition, criterion extraversion—responsible for the structure of extraversion within groups - to the model to determine, if it could explain the amount of variance. In doing so, we explicitly included the influence of experimental group-level manipulation (distribution of extraversion within the group, heterogeneous or homogeneous). To assess potential multicollinearity within the dataset, we expanded our analysis by employing additional models. These models introduced further predictors at the individual-level, including an array of personality traits, namely extraversion, conscientiousness, neuroticism, agreeableness, openness, and team orientation, as independent variables (Model 3) as well as the inclusion of those predictors, namely personality traits and team orientation, in their squared form (Model 4). We describe the procedure in detail for different dependent variables in the following.

Satisfaction. We specified a model with the dependent variable *Satisfaction*. In the model, extraversion did not explain significant variance for the group-level, thereby rejecting hypothesis 1a. Furthermore, the model did not show a significant effect of the experimental grouping, thus rejecting hypothesis 2a. In Model 3, none of the additional variables showed statistical significance in this context. Model 4 specifically focused on the quadratic effects of personality traits and team orientation on participants' satisfaction levels, revealing a significant negative curvilinear effect for extraversion^2. This indicates that satisfaction is highest at moderate levels of extraversion. In other words, students with moderate levels of extraversion reported the highest levels of satisfaction, with satisfaction decreasing for students with both lower and higher levels of extraversion. Additionally, there was a positive curvilinear effect for the quadratic term of team-orientation on satisfaction. This implies that satisfaction was highest at moderate levels of team orientation. Table 4 displays the model results for the dependent variable satisfaction, including the experimental condition variable of the study.

Table 4. *Individual-level and Group-level Predictors of Dependent Variable Satisfaction*

	Dependent variable: Satisfaction			
	(1)	**(2)**	**(3)**	**(4)**
Constant	3.643***	3.644***	3.744***	3.827***
	(0.323)	(0.209)	(0.276)	(0.348)
Experimental Condition	-0.197	-0.190	-0.307	-0.425
	(0.442)	(0.290)	(0.304)	(0.301)
Extraversion			-0.053	-0.199
			(0.216)	(0.222)
Consciousness			0.026	-0.046
			(0.228)	(0.234)
Neuroticism			-0.021	-0.047
			(0.191)	(0.204)
Agreeableness			0.056	0.348
			(0.179)	(0.280)
Openness			-0.001	0.073
			(0.169)	(0.183)
Team orientation			-0.116	0.017
			(0.128)	(0.143)
Extraversion2				-0.232*
				(0.124)
Conscientiousness2				-0.331
				(0.246)
Agreeability2				-0.012
				(0.136)
Neuroticism2				0.155
				(0.118)
Openess2				0.118
				(0.144)
Teamorientation2				0.143**
				(0.072)
Observations	322	322	319	319
Log Likelihood	-497.049	-530.803	-524.224	-518.997
Akaike Inf. Crit.	1,006.097	1,077.605	1,076.448	1,077.994
Bayesian Inf. Crit.	1,028.744	1,107.802	1,129.160	1,153.298

Note. Treatment Effect: Extraversion homogeneous = 0, heterogeneous = 1. Unstandardized coefficients are reported. Robust standard errors are in parentheses. Missing data handled with case deletion. *p**p***p<0.01

Participation. For the dependent variable *participation*, we established an intercept-only model and compared it to the random-intercept-model of individuals to determine the best fit. Next, we compared a model of random intercepts of groups to the intercept-only model and found it to be the best fit. The best-fitting model for *participation* was the model for the intercept of individuals per group, and therefore hypothesis 1b can be accepted, which confirms that the *participation* is explained at the group-level. Considering hypothesis 2b, we added the study's experimental variable in the Random Intercept Model and Random Slope Model to see the variance explained. The distribution of extraversion has a significant effect on the *participation* of group members. But contrary to hypothesis 2, homogeneous groups formed higher member *participation* than heterogeneous in extraversion-formed groups. Therefore, we must reject hypothesis 2b. Model 3 revealed a significant, positive effect for consciousness. This finding suggests that higher levels of consciousness were positively associated with increased participation in group activities. In contrast, among the additional variables considered, none exhibited statistical significance. In Model 4, there were no observable curvilinear effects to report. This implies that participation was not significantly influenced by the quadratic effects of the included variables. It suggests that the linear effects alone are sufficient to explain the relationship between these traits and participation. Detailed results can be found in Table 5.

Table 5. *Individual-level and Group-level Predictors of Dependent Variable Participation*

	Dependent variable: **Participation**			
	(1)	**(2)**	**(3)**	**(4)**
Constant	2.317***	2.675***	2.695***	2.776***
	(0.094)	(0.044)	(0.057)	(0.074)
Experimental Condition	-0.111	-0.160**	-0.147**	-0.155**
	(0.129)	(0.062)	(0.063)	(0.064)
Extraversion			0.008	-0.002
			(0.044)	(0.047)
Consciousness			0.082*	0.066
			(0.047)	(0.050)
Neuroticism			-0.056	-0.057
			(0.040)	(0.044)
Agreeableness			0.009	0.011
			(0.037)	(0.060)
Openness			0.042	0.037
			(0.035)	(0.039)
Team orientation			0.008	-0.005
			(0.026)	(0.031)
Extraversion2				-0.015
				(0.026)
Conscientiousness2				-0.065
				(0.052)
Agreeability2				0.001
				(0.029)
Neuroticism2				0.0005
				(0.025)
Openess2				-0.016
				(0.031)
Teamorientation2				-0.010
				(0.015)
Observations	321	321	318	318
Log Likelihood	-330.613	-304.417	-298.424	-296.559
Akaike Inf. Crit.	673.225	624.835	624.848	633.119
Bayesian Inf. Crit.	695.854	655.006	677.517	708.360

Note. Treatment Effect: Extraversion homogeneous = 0, heterogeneous = 1. Unstandardized coefficients are reported. Robust standard errors are in parentheses. Missing data handled with case deletion. *p**p***p<0.01

Performance. For the dependent variable *performance*, we established an intercept-only model and compared it to the random-intercept-model of individuals to determine the best fit. The random intercepts of groups proved the best fit, thus hypothesis 1c can be accepted, which confirms that the *performance* is explained by the group-level. We added the study's experimental variable to the models to determine whether it could explain the variance, and it significantly did.

Hereby, groups with the homogeneous distribution in extraversion outperformed the heterogeneous distributed ones, thus not confirming Hypothesis 2c, despite the significant effect of extraversion distribution on member *performance*. In Model 3, the results showed that none of the variables had a significant effect on performance. However, in Model 4, a significant negative curvilinear effect for neuroticism on performance was observed. This implies, that performance reaches its peak at moderate levels of neuroticism and decreases for students with both lower and higher levels of this trait. Furthermore, agreeableness was found to have a negative effect on performance in a linear manner, but the quadratic effect of agreeableness (agreeableness^2) was not significant, suggesting a more straightforward linear relationship between agreeableness and performance. The results are shown in Table 6.

Table 6. *Individual-level and Group-level Predictors of Performance*

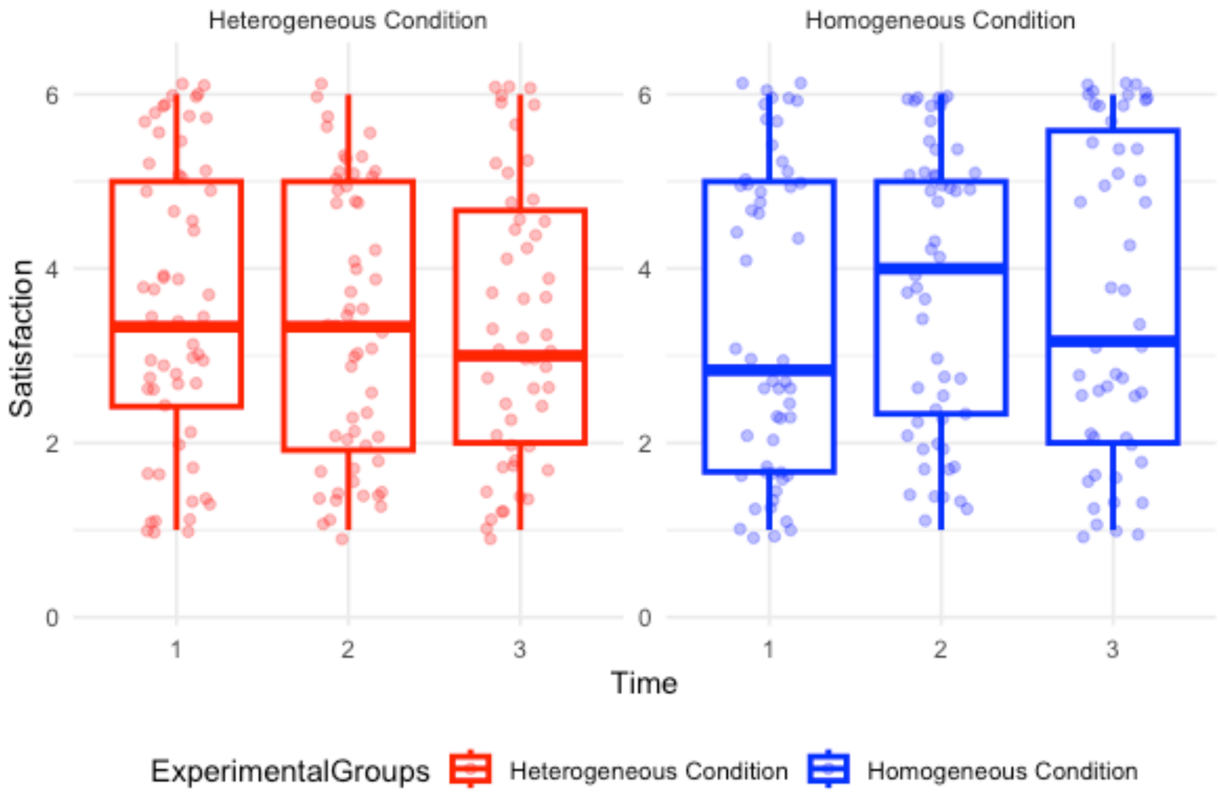
	Dependent variable: Assignment			
	(1)	**(2)**	**(3)**	**(4)**
Constant	57.051***	62.821***	63.036***	61.727***
	(3.432)	(1.800)	(2.443)	(3.053)
Experimental Condition	-5.096	-7.872***	-8.825***	-7.255***
	(4.794)	(2.630)	(2.688)	(2.657)
Extraversion			-2.088	-0.754
			(1.860)	(1.954)
Consciousness			1.913	2.457
			(2.019)	(2.047)
Neuroticism			1.521	1.107
			(1.753)	(1.876)
Agreeableness			-0.494	-6.126**
			(1.639)	(2.601)
Openness			0.467	0.085
			(1.517)	(1.625)
Team orientation			-1.198	-1.637
			(1.154)	(1.215)
Extraversion2				1.641
				(1.078)
Conscientiousness2				1.384
				(2.053)
Agreeability2				-0.755
				(1.246)
Neuroticism2				-2.992***
				(1.090)
Openess2				0.528
				(1.286)
Teamorientation2				-0.562
				(0.662)
Observations	235	235	233	233
Log Likelihood	-1,037.717	-1,069.331	-1,058.326	-1,053.269
Akaike Inf. Crit.	2,087.434	2,154.661	2,144.652	2,146.539
Bayesian Inf. Crit.	2,108.191	2,182.338	2,192.967	2,215.560

Note. Treatment Effect: Extraversion homogeneous = 0, heterogeneous = 1. Unstandardized coefficients are reported. Robust standard errors are in parentheses. Missing data handled with case deletion. *p**p***p<0.01

We could not create a model with an advantageous fit for the other dependent variables, including criterion extraversion. Additionally, it is important to highlight that we calculated the same models for a homogeneous fit of extraversion in groups, showing the same positively significant values as seen in the tables above.

Additionally, the results are summarized graphically below, with each figure displaying one of the three dependent variables (*satisfaction, participation,* and *performance)* and highlighting the group formation (heterogeneous = red colored, or homogeneous = blue colored in extraversion) over three measurement time points. This enables us to identify, which group condition is more beneficial for each outcome. The results are displayed in Figure 2, 3 and 4.

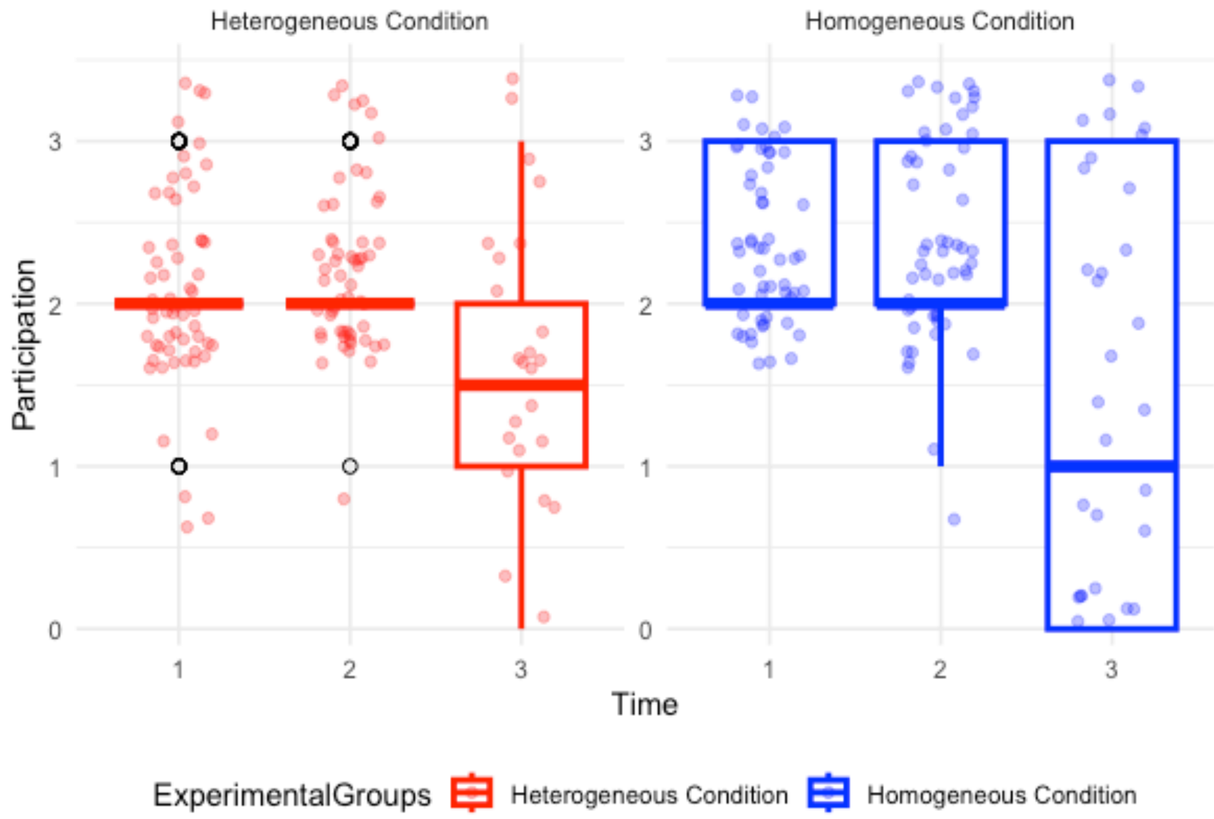
Figure 2. *Boxplots of Satisfaction over Time by Experimental Group Conditions*



Note. box = interquartile range (IQR), whiskers = min/ max range, middle line = median

Figure 2 shows boxplots of the satisfaction outcome over three time points, divided by the experimental group conditions. The boxplots display minor changes. No uniform results can be reported. However, the mean values for the homogeneous groups are more fluctuating than for the heterogeneous groups.

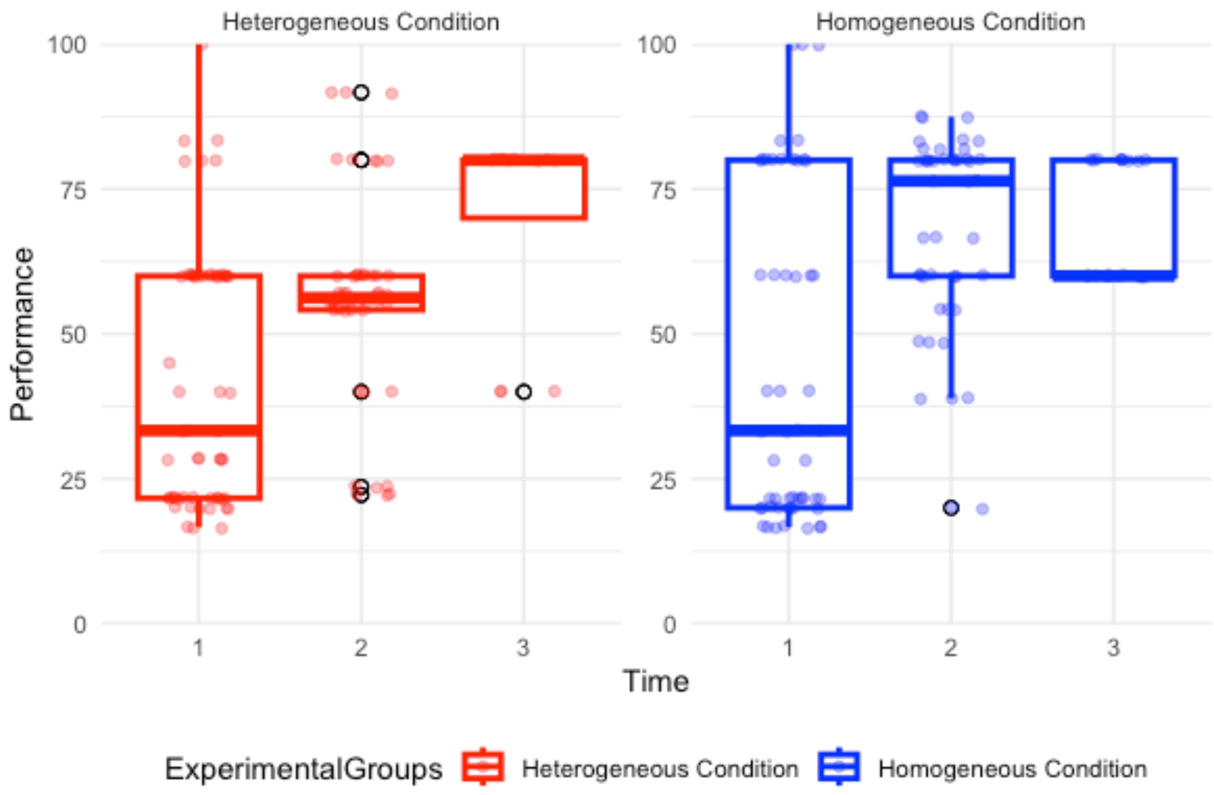
Figure 3. *Boxplots of Participation over Time by Experimental-Group-Conditions*



Note. box = interquartile range (IQR), whiskers = min/ max range, middle line = median

Figure 3 boxplots show that the mean values for both groups are similar, with an overall high variance. Specifically, the mean values for the heterogeneous groups are higher in time 1, while the homogeneous group has higher mean values in time 2.

Figure 4. *Boxplots of Performance over Time by Experimental Group Conditions*



Note. box = interquartile range (IQR), whiskers = min/ max range, middle line = median

Figure 4 boxplots show higher values for the homogeneous condition as well as more variance, suggesting that the homogeneous group formation may be more effective in promoting performance outcomes over time.

4. Discussion

The purpose of our research was to evaluate the impact of group formation while experimentally examining the distribution of extraversion as a group formation criterion to consider. The algorithm systematically generated homogeneous and heterogeneous groups, which we then analyzed to determine their effect on outcome variables, accounting at the same time for the nested data structure. As a prerequisite for analyzing the effect of specific distributions of personality traits, we first tested whether the outcomes could be attributed to group-level effects.

Therefore, we want to address the question of the respective model fit and the significance of the random slopes, followed by a discussion of the limitations and strengths of this study. Based on the previous, we will provide a research conclusion.

4.1. The Premise of Analyzing the Level of Outcomes with Extraversion

This discussion part highlights the respective level fit of the data by the group-formation-experiment, with a focus on the relationship between the experimental groups, manipulated by group-level in extraversion-distributions, as a criterion, and outcome measures such as *satisfaction*, *participation,* and *performance*.

Our first hypothesis posited that the variance in outcome variables is primarily influenced by group-level effects rather than individual-level effects. This was confirmed for the dependent variables of performance and participation. However, when considering model-fit indices AIC and BIC while accounting for the experimental manipulation of extraversion within groups, this did not hold true for satisfaction. Notably, for outcome satisfaction in the model based on AIC/BIC, the variability can be better explained by the individual-level. Here the variable did not vary between the groups and was not significantly affected by them.

We attribute this finding mainly to a ceiling effect, as all participants reported being very satisfied with their groups, which led to high overall satisfaction scores. The study's use of algorithmic group formation was uncommon in higher education settings, where students typically self-select or are allocated to groups. As a result, the study's approach drew attention to the composition of the groups, which may have contributed to the high levels of satisfaction among the students. Although the participants were unaware of the criteria used to form the groups and the underlying hypotheses, we believe that they expected the algorithm to choose the perfect group members for them. This positive expectancy might then have contributed to positive interactions between group members, which in turn lead to better group outcomes—even in the control group. Therefore, we assume that the algorithmic group formation and participation in the research project led to high levels of satisfaction among the students, in line with the research participation effect (Mccambridge et al., 2014).

4.2. Impact of Extraversion Distribution on Group Work Outcomes

In the following section, we focus on the effect of extraversion distribution on outcome variables. Here, the second hypothesis posed that groups with a heterogeneous distribution of extraversion will report greater *satisfaction (H2a)* with group formation and group work, show higher group member *participation (H2b)*, and achieve better *performance (H2c)* results than groups with a homogeneous distribution of extraversion. In contrast to expectations, the results from our overall analysis suggest that a homogeneous distribution of the trait extraversion was significantly superior to the outcome variable's *performance* and *participation*. Moreover, it is essential to acknowledge the inability to confirm our second hypothesis, which posited the benefits of trait heterogeneity in groups, based on existing literature suggesting its advantages (Nijstad & De Dreu, 2002; Thanh & Gillies, 2010; Zheng & Pinkwart, 2014). These results imply that similarity in extraversion levels among group members is beneficial for group-work-success. This finding resonates with social capital theory (Oh et al., 2004), emphasizing the value of social cohesion and homophilia (Rienties & Héliot, 2018). Research in older adults found that social capital provides a viable explanation for the association between extraversion and volunteering, such as participation in organizations and contact with friends. It is important to note that our results challenge previous literature that suggested the benefits of a heterogeneous distribution of extraversion to be beneficial to performance-results (French & Kottke, 2013; Humphrey et al., 2007; 2011)

Still, we found that in our sample team-orientation was significantly higher in homogeneously organized groups. Homogeneous groups give themselves significantly higher scores in response to evaluation questions and indicate that they and their team members learn more through groupwork than alone. Groups, whose members differ in the degree of extraversion, also vary significantly in their team orientation, compared to homogeneously extraverted groups. However, multilevel analyses that controlled for the levels of respective variance supported the significant effect of homogeneous distribution in extraversion on group performance and member participation. In addition, we controlled for other variables, including team orientation and still found a significant effect for the experimental group condition. We argue that the nature of both outcomes was not affected by team orientation since they relate more to individual and group characteristics, based on theoretical assumptions. Therefore, we can assume that the experimental distribution of extraversion had a significant impact on the success of group work, independent of the effect of team orientation. In summary, our results suggest that a homogeneous distribution of the trait extraversion has a positive effect on the outcome variables of *performance* and *participation*, but not on *satisfaction*.

4.2.1. Exploring the Benefits of Random Slope Models in Group Research: An Analysis of Extraversion and Group Formation

The following theoretical rationale for the application of the random slope model, coupled with the conceptualization of groups as complex adaptive systems (Ramos-Villagrasa et al., 2018), offers a robust framework for interpreting the outcomes of group formation based on extraversion distribution. This approach contributes to a nuanced understanding of how individual differences dynamically interact within groups, shaping overall group outcomes over time.

To justify the adoption of the random slope model in our study, we delve into both statistical and practical considerations. One of the distinctive features of the random slope model is its capacity to capture how extraversion influences outcome variables differentially across varied groups. While our study employed group extraversion heterogeneity as an experimental variable for grouping, the nuances introduced by individual variations within each group, be it heterogeneous or homogeneous, play a crucial role. This nuanced variation in extraversion distribution within groups significantly shapes the observed outcomes, as revealed by the random slope model. Our findings emphasize the importance of considering the structural interplay of individual differences within groups. The concept of group-level resources underscores that the configuration of trait expressions among group members holds greater significance in influencing outcomes than the isolated trait expressions of individual members. To discern the independent influence of group formation on outcomes over time, separate from individual characteristics, the evaluation of model-fit-indicators is crucial (Hitt et al., 2007). This perspective aligns with the idea that specific group dynamics, including the emergence of group hierarchies, are implicitly influenced by these configurations of trait expressions, impacting outcomes over time, as supported by our results and Blanco-Fernández (2023).

The impact of group-formation-outcomes, rooted in the distribution of extraversion, is better understood when considering the structural interplay among individual differences within the group. It emphasizes the configuration of trait expressions across the group, aligning with the concept that 'the whole equals more than the sum of its parts' (Loignon et al., 2018). Rather than focusing solely on isolated traits of individual group members, our research underscores the importance of these trait configurations at the group-level. This perspective highlights that specific group dynamics, including the emergence of hierarchies, are influenced by the collective trait expressions of group members. Our findings suggest that these configurations significantly influence outcomes over time, a notion consistent with similar studies (Blanco-Fernández et al., 2023). Importantly, this perspective shifts the emphasis from the individual's isolated traits to the holistic influence of trait configurations within the group, contributing to a deeper understanding of the outcomes of group formation based on extraversion distribution.

Existing literature emphasizes the importance of distinguishing between group- and individual-level factors, to gain a comprehensive understanding of the intricacies of group behavior (Kozlowski & Bell, 2013). Moreover, previous Studies have consistently shown that group formation can distinctly affect group-level processes, influencing decision-making, problem-solving, and creativity (Mannix & Neale, 2016; Van Knippenberg & Schippers, 2007; Voltmer et al., 2022). Leadership dynamics and communication patterns, which are inherently linked to the constellation of individuals' personality traits, also significantly impact group performance (Gawande et al., 2003; Zennouche et al., 2014), which the results of our research can confirm.

4.3 Implications

The findings of our study lead to several theoretical, methodological, and practical implications. When developing a theoretical framework of factors that are relevant for the success of groupwork, the study highlights the importance of considering the degree of similarity or difference in personality traits when forming groups. Extraversion is considered particularly relevant for group formation and composition (Humphrey et al., 2007) and, according to our results, should be included in such a framework.

Methodologically, our study highlights the fact that effectively enhancing group formation remains a challenge that can only be tackled through interdisciplinary research (Den Hartog et al., 2019; Odo et al., 2019). Conducting randomized-field-experiments with manipulation of group distributions is only possible by combining the diagnostic expertise from fields like psychology together with the algorithmic expertise of computer-scientists. Also, we want to emphasize the methodological decision to analyze curvilinear patterns (Curşeu et al., 2019) in our data. Even though in our case we did not find such non-linear effects, the idea of "too-much-of-a-good-thing" is a statistical method that can be applied to many contexts.

From a practical perspective, algorithmic group formation might be a promising option for many educational settings, as soon as feasible grouping criteria will be empirically established through more research. When learners have very little information about their potential group members, is it difficult for them to form groups on their own. Limited social interactions between group members make group difficult, especially in distance education. The algorithmic selection of group members holds the promise of providing the necessary skills for this kind of group work in every single group—not only in one group in which the strongest students gathered based on their homophily. Further, implementing algorithmic group formation saves educators time and effort by eliminating the need for manual group formation, allowing them to focus on other critical aspects of teaching.

The result of a homogeneous distribution of extraversion being beneficial for student groups performance aligns with the new concept of agile group work. Agile teams favor a flat, non-hierarchical structure, where people are given the autonomy to work independently and organize themselves (Junker et al., 2022). In the context of extraversion, a homogeneous distribution means that all members of the group have similar levels in this trait. This can lead to a more balanced and harmonious group dynamic with less conflict or imbalance.

However, ethical considerations such as informed consent, proper authorization for psychometric tests, and privacy must be considered when implementing such strategies. Overall, in line with previous findings (Blasco-Arcas et al., 2013), the results of our study emphasize the role of active collaborative learning, while integrating new technologies to improve students' learning performance.

4.4. Strengths and Limitations of the Study

Our study represents one of the few attempts to conduct a randomized field experiment on the effects of group formation in a real educational setting. We collected data from two very different populations of students, namely Microeconomics and Educational Studies. from two different universities, which makes it easier to generalize the findings. Also, we analyzed the outcomes of their groupwork over several weeks, further increasing the validity of the setting. Due to the course structure, there was almost no drop-out, resulting in a higher power overall measuring points despite a low number of participants. When creating the groups, our algorithm made sure that in both experimental conditions two relevant predictors of group success—prior knowledge and motivation—were distributed equally, while simultaneously manipulating our grouping criterion of interest, extraversion. We also did not rely solely on subjective outcome measures, but also included objective data in our study.

However, several limitations of our approach should be considered when interpreting the findings. First, the sample size was rather small, particularly given the fact that the main research question was situated on the group-level rather than the individual-level. This is a common problem in group-research, as larger samples are difficult to recruit for empirical studies. For this reason, we opted to form small groups, resulting in a larger number of groups. However, the limited number of students within the groups leads to increased variance at the group-level compared to that at the individual-level, making first-type-errors more probable.

When it comes to our outcome measures, we want to point out that the quality of collaborative participation was assessed through self-report. While an objective measure would have been desirable, it was not possible to collect behavioral data from students or ratings by the teachers in the setting of our study. Also, self-reports are commonly used in comparable Studies (Fryer & Dinsmore, 2020; Virk et al., 2020) and our measurement instruments have undergone rigorous testing for reliability and validity. We provided participants with a general overview of the experiment, without specific information of the hypotheses posed, to ensure transparency. This approach aimed to maintain natural interactions within the groups and avoid inducing artificial behavior or perceptions. While this transparency may have influenced participants' awareness of the experiment, it aligns with ethical research practices. Still, this may have influenced participant behavior and could thereby introduce limitations to the outcomes of the study. Additionally, collecting and utilizing personal data for algorithmic group formation may raise privacy concerns. Institutions must establish clear guidelines for handling sensitive information and maintain transparency. As a last concern, algorithmic group formation is not inherently superior to teacher-based methods but offers specific advantages. It reduces bias and subjectivity by relying on predefined criteria and mathematical rules. This objectivity is valuable in educational settings, where fairness is paramount. Additionally, it enhances scalability and efficiency, particularly in contexts with many students.

4.5 Future Directions

Future research on group formation based on personality traits should address the identified limitations and strive for improvements. This involves increasing the sample size, controlling for additional variables (e.g., self-regulation, emotional intelligence) that may impact group outcomes, and exploring the impact of personality traits as group formation criteria in a variety of settings. Apart from extraversion, there are several other possible conditioning factors to consider, such as the optimal group size, the type of task to be performed, or the group's goal. While this study focused on experimentally manipulating extraversion, it represents only one facet of the broader picture. To advance our understanding further, it is recommended to include control groups, for example, with randomized group formation, and to employ other objective outcome measures or group-process-data, such as video analysis and logfile data. Utilizing well-established outcome measures, like the observational rubric for assessing collaborative disciplinary engagement in groups (Rogat et al., 2022), can improve the validity of results and reduce the influence of extraneous factors.

While contributing to the current understanding of group dynamics, this research also points towards a future path for optimizing collaborative learning experiences, particularly in higher education settings. The development of specialized training programs and tools, including apps, can aid individuals in mastering group work and conflict resolution based on their unique characteristics, thereby enhancing the overall group-learning-experience. Better knowledge of group formation driven by research could enable learners to make better decisions in situations, where they are in charge of group formation by themselves (Magpili & Pazos, 2018).

Looking ahead, we recognize the evolving landscape of human-AI collaboration, where AI-based systems become integral group members in various scenarios (Mirbabaie et al., 2021; Seeber et al., 2020; Siemon et al., 2018). This transformation necessitates a reevaluation of established theories on group phenomena and processes, considering how human-AI collaboration impacts the way groups work together (Krämer et al., 2011; Seeber et al., 2020). While there are similarities between human-human and human-machine interactions, many aspects of human-AI-Collaboration still require further investigation. Besides well-explored topics like trust in AI, reciprocity in human-AI collaboration, and anthropomorphism, we propose a focus on group formation and, specifically, the potential roles of AI-based systems within a group.

4.5. Conclusion

In conclusion, the proposed solution to student group formation using support algorithms and extraversion levels as criteria is a useful contribution to collaborative learning and automates the group formation process efficiently and effectively. Although our experiment on the distribution of extraversion within groups did not provide a full understanding of its impact, we found that groups with a more similar distribution of extraversion perform better.

The study utilized a unique approach to determining the efficacy of grouping students by incorporating individual traits, such as extraversion, in the group formation process with the help of an algorithm. However, it is essential to consider ethical considerations when implementing algorithmic approaches, such as informed consent, proper authorization for psychometric testing, and privacy acts. Overall, our findings highlight the importance of considering personality traits in group formation and call for more attention to this aspect in future research.

Part III: General Discussion

7. Discussion

The dissertation investigated specific group-formation-criteria that emerged as beneficial for predicting group outcomes in different collaborative environments, with consideration to groups underlying structures. Following the research question of how group formation can facilitate group-work-experiences, four studies were conducted implementing the group formation algorithm to align members within groups according to the set of proposed criteria.

The criteria under investigation were personality traits extraversion and conscientiousness, as well as prior knowledge. Group formation outcomes performance, time spent, participation and satisfaction were evaluated to examine how group formation mechanisms operate in both short-term online (Studies 1, 2) and long-term face-to-face (Studies 3, 4) collaborative environments. The studies focused on the influence of group-level extraversion on the social-structural configuration of groups. Groups with heterogeneous extraversion (Studies 1, 2, 3, 4) and homogeneous conscientiousness (Study 1) or prior knowledge (Study 2) were hypothesized to have better outcomes. The subsequent sections will discuss the overall findings, implications, and comparative analysis, considering the initial hypotheses.

7.1 Collective Research Results of Studies Underlying the Dissertation

Examining the combined research findings of studies aimed to uncover connections, patterns, and relationships. Contrary to expectations, the central hypothesis that heterogeneously distributed extraversion leads to better group-work-outcomes was not supported by the studies. In particular, the superior results found for its homogeneous distribution in face-to-face groupwork (Studies 3 & 4) warrant attention, highlighting the need for a reassessment of existing assumptions. Additionally, interaction effects resulting from the experimental manipulation of two traits, as explicated by the research design applied to online group work (Studies 1 & 2), were further addressed.

In study 1, with no significant main effects to be reported, the expected benefits of experimental grouping by extraversion and conscientiousness distribution on outcomes did not accrue. While no interaction effect was found for outcomes like communication frequency or participation, significant interaction effects prevailed for outcome satisfaction: Satisfaction, initially low at the first evaluation, increased from the second to the third. Groups with heterogeneous conscientiousness and extraversion reported the highest satisfaction, while groups with heterogeneous extraversion and homogeneous conscientiousness reported the lowest scores. Moreover, team orientation was found to have a negative curvilinear effect on satisfaction, with peak satisfaction found at moderate levels of team orientation. As such, findings align with those implicating a joint or moderating impact of the different factors involved in group formation (French & Kottke, 2013). Consistent with findings in study 1, no significant main effects of experimental grouping were observed in study 2. Still, interaction effects indicated prolonged group stability for a heterogeneous distribution of extraversion and a homogeneous distribution of prior knowledge. Nonetheless, both studies 1 and 2 encountered methodological challenges, limiting interpretability of the reported findings.

In face-to-face-studies, the results implied the superiority of homogeneously extraverted groups for most outcomes (satisfaction in study 3, participation, and performance in study 4). Beyond the shift in the research setting, the findings emphasize the importance of considering time to gain a comprehensive understanding of how individual traits might change through the influence of time in terms of the unfolding nature of individual traits within group processes in different settings. Understanding social contexts requires acknowledging the intricate dynamics of the social structure, where observations at different levels are interconnected, allowing for inferences to be made about the population. This approach highlights the significance of examining individual behavior within a group setting rather than in isolation, which enables the analysis of a variable that varies at the group-level (between groups) and a variable that varies for each individual (within groups).

Results revealed that the group-level had a higher explanatory value for most of the dependent variables. Thus, group membership, and hence experimental group formation, impacted the nature of outcomes, without implying any direction of effect, e.g., better or worse. Yet, some outcomes accounted for a greater proportion of variance at the individual level. Taking outcome satisfaction as an example, the variation in satisfaction levels could be better explained by individual differences, e.g., individual-level predictors (e.g., neuroticism in study 1 and team orientation in study 4), rather than being attributed to any group-level predictor. This could suggest that individuals may enjoy groupwork in relation to their inherent predispositions rather than their affiliation with a particular group.

Overall, to ensure accurate predictions and insightful analyses in research, it is crucial to comprehend the interconnected relations among variables underlying social structures. The evident association between the research settings, manifested by the patterns in result consistency, adds to the reliability of research findings. Yet, the divergent results identified could point to the impact of further external contextual factors that were not explicitly controlled or assessed in the research design. Hence, it is imperative to regard these findings as valuable insights rather than absolute truths.

7.2 Results Considering Nonlinear Assumptions of Variables

Following related research (Bowers et al., 2000), studies 1 and 4 acknowledge the presence of curvilinear patterns in the relationship between variables, which can be appropriately modelled using non-linear functions. Such non-linear relationships of variables build on the notion that the positive effect of one variable may decrease as the value of the other variable increases, resulting in a U-shaped (or inverted U-shaped) relationship in which changes in one variable affect changes in the other - often referred to as the 'too much of a good thing' effect (Chen et al., 2018).

Certain curvilinear variables offered insights into aspects of the social behavior and interaction present within groups. For instance, conscientiousness and neuroticism were associated with levels of communication, implying the importance of personality trait level besides distribution to further examine group communication patterns. By including this curvilinear assumption, further exploratory insights could be gained. For example, in terms of homework, higher levels of agreeableness and a quadratic variable of extraversion (indicating increased extraversion up to a certain threshold) predicted its effectiveness. While this implies that a certain level of extraversion is beneficial, excessive levels may have diminishing returns. Thus, individuals with high levels of extraversion may contribute positively to group work due to their diverse interests. However, extremely high levels of extraversion may lead to eccentric behavior, thereby negating the positive effects and being associated with narcissistic tendencies (Zajenkowski & Szymaniak, 2021). In turn, narcissistic individuals were found to be frequently perceived and chosen as group leaders by others, which negatively affects group dynamics, leading to conflicts and undermining cohesion and effectiveness (Brunell et al., 2008).

In sum, the consideration of non-linear interactions also opens the possibility of addressing the assumption that different characteristics (e.g., agreeableness vs. conscientiousness) can account for different outcomes (e.g., group performance vs. social processes). Nevertheless, the interpretation of these analyses should be approached conservatively, due to the restructuring of the data sets required for curvilinear analysis.

7.3 Interactions between Student Characteristics in Online Settings

This section discusses the challenges of online CL and relevant research in the context of the online studies of this dissertation. The specific context of the STEM-student body is noted, while also considering the various factors influencing student performance in mathematics (Springer et al., 1999). In the pursuit of forming high-performing groups, it is suggested that the optimal leader may exhibit a combination of moderate levels of extraversion and conscientiousness, aimed at motivating and initiating group interactions. It was observed that groups designated as heterogeneous had a higher probability of including an individual with adequate levels of both extraversion and conscientiousness. This aligns with other research indicating combined effects of trait levels, for instance, enhanced group performance arising from group heterogeneity in conscientiousness as well as group heterogeneity in extraversion (Kickul & Neuman, 2000).

In study 2, conscientiousness was replaced by prior knowledge. Nevertheless, exploratory findings revealed that extraversion and conscientiousness remained more influential than demographics, gender, and age. All performance-related outcomes identified conscientiousness as a significant predictor in the exploratory analysis. Given the multitude of individual- and group-level variables influencing CSCL-processes, studies 1 and 2 could only incorporate a selection of factors deemed potentially relevant to group work. The chosen outcome variables were centered on the assessment of group work processes, particularly concerning the challenges inherent in CSCL. In essence, the exploratory results indicated a mere tendency in the data for the online group-work-setting.

Notably, related research has identified significant variations in the quality of interaction and learning outcomes in online group work (Strijbos et al., 2004), likely attributed to differences in technology utilized, units of analysis, research methods, and outcome measures employed (Lipponen, 2002; Odo et al., 2019). In general, while providing technology can facilitate communication and social interaction, particularly in an online setting, its mere provision does not guarantee their occurrence (Kreijns et al., 2003). Even though groupwork occurred in an online setting, where individuals could have opted to work alone, was the variance in results attributed to the group-level. However, students exhibited a decreased inclination to actively participate in collaborative learning, interactions, and discussions within the online environment (Studies 1 & 2) compared to face-to-face settings (Studies 3 & 4), aligning with findings from prior research (Dumford & Miller, 2018). The lack of willingness to maintain groupwork was a significant contributor to both dropout rates and social loafing (Tsovaltzi et al., 2019; Williams et al., 2006). Understanding the specific circumstances in which groupwork was conducted could help to explain differences in behaviors or results within a group.

7.4 The Distribution of Extraversion in Long-Term-Group-Work

To address the shortcomings in acquisition identified in studies 1 and 2, studies 3 and 4 were conducted. Deviating from the initial focus on remote, online, short-term group-work-interactions, the subsequent studies 3 and 4 shifted the focus of the research objective to another setting of obligatory, prolonged face-to-face groupwork within a more conventional student-setting. The shifted research focus towards longer-term, face-to-face group work, anticipating that the sustained nature of group interactions would heighten the expression of traits, chosen as group formation criteria, as well as rendering the effects of variables for group formation more pertinent over time and consequently resulting in distinct group-process-outcomes. Both studies scrutinized group structure and centered on the research question, prioritizing hierarchical group formation, with either heterogenous or homogeneous distribution of extraversion as the criterion for experimental group formation.

While contextualizing the divergent findings of studies 3 and 4, it is imperative to contrast them with existing research. Contrary to the initial hypothesis, the results indicated that rather homogeneity in group extraversion outperformed heterogeneous distribution for some of the main outcome variables. Based on the findings, there is evidence in favor of a non-hierarchical group structure in face-to-face groupwork, possibly linked to greater member negotiation in homogeneous groups (Wilson et al., 2016). Contrary, research by Humphrey et al. (2007, 2011) revealed that a heterogeneous distribution of extraversion leads to better outcomes. Kramer et al. (2014) agreed to this notion by highlighting the benefits of group hierarchy and structured leadership in enhancing overall group performance. Moreover, French and Kottke (2013) conducted a related study on undergraduates in a mandatory, long-term groupwork setting. They assessed teamwork interest and the mean and variance of team extraversion in correlation to outcome satisfaction. Satisfaction was not predicted by either teamwork-interest or mean extraversion. Instead, low satisfaction was predicted by greater extraversion dispersion, particularly when individuals were more interested in teamwork, related to the results of study 3 and 4. For instance, study 4 showed higher performance and participation scores for homogeneously distributed groups. This corresponds with their findings of a significant interaction effect between teamwork interest and extraversion dispersion. With increasing teamwork interest, extraversion dispersion had a greater effect on individual satisfaction with the group, indicating that the satisfaction of individuals more interested in group work depended on the similarity of group members' extraversion (French & Kottke, 2013).

Although research has shown that individual differences in group formation are important for predicting performance (Bradley & Herbert, 1997), it is worth considering that the specific actions of each individual can also play a crucial role in shaping the overall social dynamics within a group and thus potentially shaping outcomes (Chiu, 2000). It is also possible that outcomes may be more dependent on the processes of group work, referring to contextual factors and temporal aspects of group processes that were determined by the course structure and academic discipline in the studies of the dissertation, as prevailed by leadership roles change over time (Bendersky & Shah, 2013). The extended duration of group work also contributed to the outcomes of group work, given that the best model fit was found for models accounting for group and time. Thus, differences in subject matter may explain differences in findings across studies. Likewise, variations in other factors may have had a greater impact in explaining the findings, such as the specific research setting and the nature of the (group-)work task or assignment to be performed.

7.5 The Influence of Student Characteristics and Group-Setting on Outcomes

To integrate and advance the understanding of the complexities involved in group formation, this section seeks to re-evaluate assumptions about the impact of personality traits on group work dynamics. Given the variation in outcomes observed in online (Studies 1 & 2) and face-to-face (Studies 3 & 4) groupwork studies, it is imperative to challenge the prevailing notions of uniformity in group dynamics across settings, as suggested by previous research (Goñi et al., 2020). Indeed, the intricate interplay of individual characteristics and situational factors still constitutes a central focus for further discussion.

Exemplifying extraversion's reduced salience in online group work (Wilson et al., 2021), supported by studies 1 and 2, highlights the trait distribution's dependency on the underlying setting. Extraversion's less salient role in online group work, supported by results from both studies, aligns with the assumption that situational factors might be more crucial in online settings. Linking this to the results of the online groupwork studies, the high variance observed at the group-level might be explained by potentially uncollected group-level variables differing between groups.

Based on the exploratory results of study 1, extraversion and neuroticism were found to partially explain the frequency of communication within groups. In online settings, shy students may feel more confident to participate and express their opinions and ideas, due to the less intimidating computerized environment and lack of anxiety caused by direct interference from others (AbuSeileek, 2007; 2012). This interpretation is in line with research indicating that groups having a heterogeneously distribution of personality traits can lead to more successful learning outcomes in certain situations and for certain traits (Roberge & van Dick, 2010). However, interpreting personality trait distributions in online settings is intricate, with unexplored factors that complicate interpretations. These factors add to the methodological shortcomings discussed in the section on research limitations.

Studies 3 and 4 challenged the presumed benefits of a heterogeneous distribution in group extraversion. However, in other studies, the benefits of a mix of different personality trait distributions improved problem-solving and decision-making (Bradley & Herbert, 1997; Chen et al., 2015; Chiu, 2000; De Dreu & Nijstad, 2008). Given the differences in research results, it can be assumed that the ideal distribution of personality traits within a group may be determined by various factors. While research suggests that both homogeneous and heterogeneous groups can succeed or fail, the potential importance of other predictors such as performance goals, behavioral engagement (Giel et al., 2020), gender, and domain expertise appear to be more relevant for successful group work (Kucukozer-Cavdar & Taskaya-Temizel, 2016). Reconsidering the underlying assumptions about the merits of personality traits as criteria for group formation, some literature has argued for, or rather advocated, a task-dependent approach (Bowers et al., 2000; Wax et al., 2017). In addition, relations between personality traits and outcomes may vary across performance categories (Zell & Lesick, 2022), highlighting the need for a consistent approach to group effectiveness when assessing group formation. Regardless of the chosen criteria for group formation, it remains possible for groups to contain individuals, who are disruptive to group work and thus contribute to group failure. However, such individuals may exhibit extreme characteristics, reinforcing the potential of the curvilinearity approach.

7.5.1 A Guiding Model to Uncover the Complexity of Group-Formation Processes

To comprehensively understand social group behavior, considering situational factors, social context, past experiences, cultural values, and societal norms is necessary. The question remains as to whether a single profile, regardless of its nature, can fit all groups, or if the effectiveness of a group depends on individual members or specific member constellations that may vary from group to group. Notably, the results for conscientiousness (Study 1) and extraversion (Studies 3 & 4) contradict previous research, hinting at a partial lack of knowledge regarding their distribution on group-work-outcomes, as supported by other studies (Maqtary et al., 2019; Odo et al., 2020). An interaction of individual characteristics with situational factors appears to be a likely model for further research into group formation and processes.

To guide future research on group formation, the modified model of Gladstein (1984) and the proposed taxonomy by Maqtary et al. (2019) can be used in conjunction with algorithmic support (see Figures 1 & 2). The taxonomy displays the decisions to take in a group formation process, which include member attributes such as personality traits and prior knowledge, as well as group attributes such as group structure, the setting underlying group work, and outcomes. Recognizing and integrating all these attributes into research methodologies is essential, to incorporate valuable insights for future research and its review. Understanding these challenges can assist researchers and educators in designing, applying and evaluating effective group formation techniques.

By highlighting conflicting research findings, this dissertation attempts to illustrate how group formation can be influenced by a range of attributes, and in turn requires the definition and consideration of both the individual-level (characteristics used) and the group-level factors (input structure, e.g., homogeneous/heterogeneous). Additionally, the situation (online/face-to-face/blended learning), task, duration of group work, setting, and motivation can all influence group work outcomes, from satisfaction to performance, differently. Special needs for working in groups may arise from individual situational circumstances or as a combined result of group-level behavior. Therefore, it is essential to consider the situations and needs of each member to shape optimal groups for the goal of enhancing individual learning in various fields.

8. Research Strengths

The advantages of the studies underlying this dissertation include the experimental design implemented in an authentic research environment, which encouraged continuous group work. This approach not only nurtured the natural behavior of students, but also implied high external validity, enhancing the generalizability of the findings. The reasons for this were the large sample size in studies 1 and 2, before dropout, as well as the constant, long-term group work of studies 3 and 4. Since both settings - online and face-to-face - were tested, differences in the settings can be identified, which is a clear strength of this work. In addition, the studies were based on various research projects conducted by researchers from different fields, including psychology, computer science, and mathematics. It is important to highlight the interdisciplinary environment of research as an outstanding advantage. By showing that the standard deviation of extraversion within groups in the scores for heterogeneous and homogeneous conditions differed significantly, the functionality and purpose of the implemented group-formation-algorithm could be confirmed across all studies.

In terms of data analysis, the advantage of multilevel modeling must be mentioned as another strength of research, making it possible to analyze data independent of time (Level 1), individuals (Level 2), and individuals within groups (Level 3). The consideration of outcome variables at both the individual and group-level, reveals a spectrum of outcome variables and promotes the ability to differentiate statements about the experimental variables used for grouping. In study 2, homework completion was an objective measure, for which dropout was considered as being absent during data analysis. In addition, evaluations were a prerequisite for submitting homework assignments and were completed by students more frequently than in study 1. Studies 3 and 4 had almost no dropouts due to the obligatory course structure. It was possible to keep the groups going until the end, resulting in greater validity of all measurement points. Complete data collection of consistent groups working together over a long period could be achieved, and thereby longitudinal data analyzed regarding personality-diversity-interaction over time and its influence on group members' perceptions and performance. Inevitably, thanks to this weekly labor-intensive seeding, personality traits within the group could develop and as intended, the influence of the experimentally manipulated group structure could be revealed.

9. Research Limitations

While the studies incorporated in this dissertation provide insights and assistance with group formation, it is crucial to acknowledge research limitations. The initial limitation pertains to the selection of research-participants. Across all four studies, the target group consisted of university students. Therefore, results cannot be transferred to other groups, e.g., in schools, workplace environments, or in free-time-group activities. Furthermore, in study 1 and study 2, participants were prospective students in MINT-subjects, according to which there is most likely no variation in cognitive ability, as seen in study 2 by an overall low value of prior knowledge. In study 3, participants were pre-service teachers enrolled in a specific course of educational psychology, which limits the generalizability of findings to other subjects, majors, or populations.

Participants with more diverse backgrounds enrolled in study 4, which improved the generalizability of findings. However, the matching process, by which students apply to universities, and universities, in turn, admit students, selects sets of attributes that constrain the population in particular ways. These constraints may include factors such as income, ability, class, motivation, drive, and narrow age ranges. Additionally, students tend to self-select in majors and fields of study, which can further limit heterogeneity on some, and often many dimensions. Therefore, in terms of age, cognitive ability, achievement motivation, and most likely conscientiousness, a selection effect can be assumed. As a result, each study's sample probably tended to be driven by respective selection effects (e.g., western, educated, industrialized, wealthy, and democratic individuals; Henrich et al., 2010), which can introduce bias in research results.

The first two studies investigated groupwork in an online setting over a short period. The overall low intensity and limited time frame may have diminished the effects of the respective group formation: Homogeneity and heterogeneity might have played a less important role, as the members did not reach a level of familiarity with each other. Considering that groups are often formed, intending to perform tasks in the same constellation over a longer period, a format of four weeks proved to be insufficient, especially due to dropout. Due to the reduced group sample size after dropout, a lack of consistency with previous findings due to several reasons may be related to the limitations of the study design, compromising external validity. Voluntary participation in studies 1 and 2, coupled with the previously mentioned short period, could have hindered the development of a group dynamic, as students' traits may not have had time to solidify and allow them to identify and compare similarities or differences between each other. Additionally, high dropout surely disrupted the group-process and most of the groups simply no longer existed. No longitudinal data could be obtained, meaning that effects could not be measured over time. Furthermore, one could speculate that only highly conscientious students attended until study completion, implying additional self-selection effects.

The small sample size of studies 3 and 4 obviously limited statistical power. Group research must address this pervasive problem, as calculations are done at the group-level. These considerations usually favor small group sizes, as in our case, to increase the number of available groups. Thus, the small group size and the large number of groups might lead to an artificially high group-level ICC (Hox et al., 2017), which is beneficial for group comparisons, but worth considering when interpreting the results (Gelman & Hill, 2007).

Some limitations in research designs need to be addressed for all studies, as no control groups (with no experimental manipulation or group work) were created besides the two experimental group conditions because of the small sample size and the generation and analysis of group data. Although field studies increase the external validity of findings compared to highly controlled laboratory experiments, they face several threats to internal validity (McMillan, 2007). Field studies not only address experimental conditions, but also contend with differential attrition when students are randomly assigned to them. In study 4, a limitation in this regard was identified, as unintentionally team orientation was observed to be higher in homogeneously in extraversion formed groups. Higher team orientation typically correlates with a preference for working in groups. Therefore, satisfaction values might have been mediated by interaction effects that included higher team orientation. Nevertheless, since team orientation is known to correlate primarily with the assumedly affected outcome satisfaction, we do not further explore these differences in team orientation, due to the non-significant roles of both extraversion and team orientation on outcome satisfaction. The last point concludes that for all studies, the group-formation-criteria in use, as well as most of the research-outcomes, relied exclusively on self-report questionnaire formats, and only an objective outcome measure for performance was integrated. Subsequently, ideas for further research are derived based on the above-mentioned research limitations and discussion points.

10. Ideas for Further Research

Ongoing research is imperative for achieving a broader understanding of the complex interdisciplinary challenge posed by group formation. The ambiguous results obtained underscore the necessity of a meticulous examination of the multifaceted nature of groups. The selection and weighting of relevant criteria for group formation, coupled with the exploration of other prerequisites for effective group work, continue to be demanding interdisciplinary areas of research. The application of a global research model is necessary to assist in the systematic collection and review of findings, thereby contributing to enhancing the overall understanding of group dynamics. Establishing such a research model is closely related to creating a common understanding of what constitutes *effectiveness* in the context of group work outcomes. Reference has already been made to several dimensions of *effectiveness* in various studies (see section 4 for more details). Besides, as variables such as satisfaction are not necessarily related to variables such as performance, it is essential to note that outcomes of group work should not be limited to a single measure. By incorporating different outcome measures and introducing additional objective measures, researchers can obtain valuable insights into the multifaceted nature of group work outcomes and foster the comparability of research.

In a review of studies, even when focusing on groups in which the distribution of extraversion was experimentally manipulated (aligned with the criteria of studies 1 & 2), the full extent of its role in groups remained elusive. Therefore, additional research on extraversion and its distribution, along with conscientiousness and prior knowledge, conducted with larger numbers of participants over extended periods of time, is warranted. While some studies emphasize the relevance of personality traits (Roberge & van Dick, 2010), others argue for the significance of factors like gender and area of expertise (Kucukozer-Cavdar & Taskaya-Temizel, 2016). The findings from studies 3 and 4 demonstrate benefits for homogeneous groups, indicating the potential for a more equal and less hierarchical relationship between leaders and followers, which aligns with the concept of agile work transformation. Agile principles prioritize collaboration, adaptability, and team-interaction, with the goal of replacing traditional hierarchical structures with agile frameworks that are based on self-governance and consensus-building (Bundtzen & Hinrichs, 2021). However, implementing flat hierarchies may face obstacles due to societal expectations, resulting in limited acceptance despite their potential for productivity and innovation.

Research findings on face-to-face groups have led to a focus on social perceptions of leadership personality and behavior (Latu et al., 2013; Woolley et al., 2010). The demonstrated influence of characteristics potentially ascribed to women in social contexts on group work highlights the advantages that accrue with an increasing number of women in groups (Eagly & Carli, 2003). Notably, groups with more women achieve significantly higher performance scores, surpassing even groups with more highly intelligent members ("The more women, the better"; Woolley & Malone, 2011). Accordingly, potential gender-specific characteristics and associated social conditioning should be explored as factors in future research, and as such, embraced by both genders in a transforming world of work. This development is exemplified by current trends in leadership research, which deconstruct personality into agency and warmth (Dubois et al., 2016). Plausibly, attributes such as the ability to recognize the emotions of others could act as underlying predictors of the respective influence of gender on group success.

The creation of experimental groups based on other characteristics and their properties or constellations is a topic that certainly warrants exploration in future research. Other relevant characteristics of interest include factors such as students' prior technical knowledge and social sensitivity. Emotional ability, specifically the understanding of emotional expressions in others (e.g., Theory of Mind (Perner, 1999)), is particularly intriguing, due to its potential for objective measurement. Additionally, individual factors such as social status and popularity goals have varying effects on academic outcomes (Jones & Cooke, 2021), and thus be considered in further research. Drawing from the methodological challenges and findings from study 2, it is advisable for future research to consider replacing prior knowledge with technological proficiency. This adjustment is particularly relevant, given that technological proficiency, previously presumed to be mediated by neuroticism in study 1, can provide a more nuanced understanding of learners' competencies and abilities within digital environments.

The interplay between different group-formation-criteria and their interactions should be considered in future research. For instance, study 1 revealed interaction effects of extraversion and conscientiousness on group outcomes, supported by prior research (Bowers et al., 2000; Brunell et al., 2008). The absence of significant results in satisfaction in study 4 may indicate that satisfaction is influenced by multiple factors beyond trait distribution, such as communication, work effort, and individual motivations. Considering alternative explanations, the observed results could have been influenced even by characteristics not measured in the studies. The impact of individuals with unique personality trait constellations on group outcomes further raises questions about the role of specific individuals in shaping group dynamics.

In CSCL, sustained, coherent and high-quality discourse may be lacking (Lipponen et al., 2002, 2003), along with students' tendency to focus misleadingly on technical support rather than group interaction, resulting in limited overall communication. (Strømsø et al., 2007). To overcome the challenges posed by technology and enhance the virtual learning experience (Lim & Newby, 2020), researchers have advocated for the promotion and development of online group activities (Wessner & Pfister, 2001). Conversely, issues related to a lack of tools and the need for a more systematic, technology-assisted approach to support group formation have been identified (Maqtary et al., 2019). Particularly in situations of social distance, group formation holds significant potential for creating social interaction and a sense of belonging among students. This is especially relevant for first-year-students, who may not be acquainted with each other, as seen in studies 1 and 2. Employing communication tools may offer a viable solution to foster and ensure social interaction. Beyond communication tools, further research should explore and set situation-specific strategies to support students engaged in online group work in a pedagogically meaningful way (Harrer et al., 2006; Inaba et al., 2000; Strijbos et al., 2004). This entails promoting collaboration and communication among students through the establishment of guidelines for groupwork and encouraging a sense of ownership in their learning process.

The dissertation focuses on examining face-to-face and online groupwork separately, to allow for a more compliant focus on the research question posed. However, it is worth noting that a comparison of multiple settings of group work is a fruitful future research topic. Further research could deepen the understanding of the relationship between group-formation-inputs and outcomes. For instance, examining the moderating effect of group size could be interesting. Kirschner (2017) found that heterogeneity had a stronger negative impact on group performance, but only in smaller groups. Examining the impact of heterogeneity on group performance over time could be valuable. Rispens et al. (2021) stated that group members may become more cohesive or accustomed to working with a diverse group when groupwork occurred over time. Additionally, investigating the effect of different levels of heterogeneity (e.g., low, moderate, high) as well as homogeneity on group outcomes could provide further insights. Finally, it might be worth researching the effects of training or interventions aimed at increasing group cohesion in a range of diverse groups.

In sum, future research on group formation should meticulously consider the multifaceted aspects of group work, including diverse student characteristics, varied settings, and the intricate interplay among these factors. It is imperative to systematically integrate these elements when synthesizing them for review or constructing research models.

11. Practical Implications

Exploring the potential impact of additional member features on group work in both online and face-to-face settings remains a relevant research question. This dissertation emphasizes the significance of intricate group formation in defining roles, goals, and activities for each learner before initiating group work. To facilitate effective group formation, it is essential to consider a nuanced approach to the underlying factors, their direct and indirect features, and their interaction. For instance, determining the extent of homogeneity or heterogeneity in selected characteristics among group members demands careful consideration. As each approach has advantages and disadvantages, the optimal solution will vary depending on individuals' specific needs, the availability of a diverse spectrum of potential group-formation-criteria to be selected, and the respective setting and task to be performed.

The approach to group formation described in this dissertation can (and should) be used in conjunction with other approaches, including interventions in different group-process-stages (Tuckman, 1965), and previously described guidelines to increase the benefits of group learning by offering structured information that facilitates CL in a variety of settings. The goal may be to provide social and learning support to all members of the cohort, as well as to assist individual students in developing collaboration skills. Continuation of research might help to develop nuanced guidance on grouping strategies. Such further research may help in the development of resources or training programs that can assist trainers in making informed decisions about grouping, considering the specific context and objectives of the learning task. As no definite instructions can be given at this stage, such courses could be of an informative or technical nature. Similarly, student modules would be intended to introduce learners to the potential benefits and challenges of different group structures and help them to adapt their collaborative skills accordingly for a more effective and harmonious group-work-experience.

From a practical perspective, gaining insights into how online learning environments could be hypothetically designed to facilitate group work, considering individual differences between students structured by group formation, is important (Study 1 & 2). In connection with this, practitioners could benefit from insights into the potential design of homogeneous groups of students based on extraversion, as suggested in studies 3 and 4. The findings tentatively suggest that forming groups based on homogeneity in extraversion may lead to better results for face-to-face group work, while a heterogeneous distribution of extraversion may be beneficial for online group work.

Studies 1 and 2 underscore the challenges associated with fostering social interaction and cultivating a sense of belonging in CSCL-environments. Given the high dropout rates documented in online settings, as evidenced in both these Studies and existing literature (e.g., Lim & Newby, 2020; Strømsø et al., 2007), the practical implications of these findings for enhancing online group-work-scenarios are limited. This assumption shifts the focus not only to the consideration of different individual characteristics, but also on how to shape more complex interaction patterns experimentally. In study 1, I consider the potential of having one person responsible for either the success or failure of group work to be like having a group leader, in line with others (Kickul &

Neuman, 2000; Putro et al., 2020). Furthermore, the results under the curvilinear condition suggested that a person with a particular personality type should be found and allocated in each group, as the group's leader as a potential solution to group formation, by which the singular person having more impact on the group's outcome variables than the combined traits of all group members. However, this might be true for some trait constellations but not for others.

The benefits of using algorithms for group formation as a didactic-application-tool should be emphasized. This development allows for on-the-fly grouping according to different criteria, with a user-friendly interface, implementation possibilities, and accessibility as OpenSource software. Therefore, it represents an economical and practical tool for research and application. The use of algorithms for grouping students has several implications in education. These include the potential for personalized instruction and support, early identification of struggling students, the creation of balanced and diverse groups, and adaptive e-learning systems. Grouping algorithms can be used to create personalized and adaptive e-learning systems. This is achieved by grouping students based on their performance development and providing customized learning paths and resources. Online learning portfolios can also be used to provide help on demand in a student-to-student case of group formation or by respective learning goals. The practical implications of systematic group formation could extend beyond the university setting and be applied in schools, workplace training programs, and community organizations.

12. Conclusion

The way groups are formed can have a major impact on the social behavior and overall experience of students in collaborative learning. To achieve the best possible outcomes for a group in any given situation, it is important to consider the unique characteristics of the individuals involved, as well as other factors that may influence the group as a whole. This dissertation presents an approach for investigating specific group formation criteria that encourage desired interaction processes for designing more engaging and effective learning environments. Gaps in the literature indicate the necessity for more comprehensive research considering the multifaceted aspects of group dynamics. Although results do not give a definitive answer on the best way to form groups, studies 1-4 employed experimental design representing a potential methodological approach to explore group formation criteria and their influence on relevant outcomes. Considering alternative explanations, the role of specific individuals in shaping group dynamics, as well as the influence of specific characteristics, that were not measured in the studies, illuminate.

To fully support CL, it is important for grouping methods to take into account critical elements that impact learner-interaction. The dissertation aimed to contribute to the initial development of an integrated concept for group formation research. This involves carefully designing specific formations that encourage expected interactions during collaborative learning activities. Future research can explore potential mediators and moderators to establish flexible guidelines that effectively shape group formation based on individual needs.

Erklärung über bisherige Promotionsversuche

gemäß § 6 Absatz 2 f) der Promotionsordnung der Fachbereiche 02, 05, 06, 07, 09 und 10 vom 04. April 2016

NAME: MÜLLER

VORNAME: ADRIENNE MARA AIMÉE

Haben Sie sich bereits früher einem Promotionsverfahren unterzogen?

- [x] Nein
- [ ] Ja

- [ ] Wenn ja, an welcher Universität? ____________________
- [ ] zu welchem Zeitpunkt? ____________________
- [ ] und mit welchem Erfolg? ____________________

Haben Sie die vorgelegte schriftliche Prüfungsleistung oder Teile daraus in einem anderen Verfahren zur Erlangung des Doktorgrades oder eines sonstigen akademischen Grades oder einer anderen Prüfung eingereicht bzw. bereits früher eingereicht?

- [x] Nein
- [ ] Ja

______________________ ______________________

DATUM UNTERSCHRIFT

Eigenständigkeitserklärung

gemäß § 6 Absatz 2 g) und gemäß § 6 Absatz 2 h) der Promotionsordnung der Fachbereiche 02, 05, 06, 07, 09 und 10 vom 04. April 2016

NAME: MÜLLER

VORNAME: ADRIENNE MARA AIMÉE

Hiermit erkläre ich, dass ich die eingereichte Dissertation selbständig, ohne fremde Hilfe verfasst und mit keinen anderen als den darin angegebenen Hilfsmitteln angefertigt habe, dass die wörtlichen oder dem Inhalt nach aus fremden Arbeiten entnommenen Stellen, Zeichnungen, Skizzen, bildlichen Darstellungen und dergleichen als solche genau kenntlich gemacht sind.
Von der Ordnung zur Sicherung guter wissenschaftlicher Praxis in Forschung und Lehre und zum Verfahren zum Umgang mit wissenschaftlichem Fehlverhalten habe ich Kenntnis genommen. Meine Erklärung bezieht sich auf Schriften, die ich als alleiniger Autor bzw. Autorin eingereicht habe oder bei Ko-Autorenschaft auf jene Teile, für die ich mich verantwortlich zeichne. Ich habe keine Hilfe von kommerziellen Promotionsberatern in Anspruch genommen.

DATUM UNTERSCHRIFT

Darstellung des Ausbildungs- und Studienverlaufs

Juli 2010	Allgemeine Hochschulreife an der Eleonorenschule, Darmstadt	
Oktober 2011 - Februar 2015	**Bachelorstudiengang Psychologie**	**Lehrende/Prüfende:**
SoSe 2012	B. Methodenlehre	Malte Persike
	C. Allgemeine Psychologie Basis	Prof. Dr. Daniel Oberfeld-Twistel
	E. Persönlichkeitspsychologie und Diagnostik (Basis)	Univ.-Prof. Dr. Boris Egloff
WiSe 2012/13	Praxis der Neuropsychologie psychischer Störungen	Dr. Armin Scheurich
	A. Biologische Psychologie	Prof. Dr. Stefan Berti
	D. Entwicklungs- und Sozialpsychologie: Basis	Fabian Escher; Maike Rathgeber
	Seminar: Einführung in die entwicklungspsychologische Forschung	Fabian Escher; Maike Rathgeber
	D. Entwicklungs- und Sozialpsychologie: Basis	Univ.-Prof. Dr. Randolf Ochsmann
SoSe 2013	Seminar: Sozialpsychologie	Dr. Paul Schaffner
	H. Persönlichkeitspsychologie und Diagnostik (Aufbau)	Univ.-Prof. Dr. Boris Egloff
WiSe 2013/14	Seminar AOW: Ausgewählte Themen der Personalpsychologie	Prof. Dr. Thomas Rigotti
	N. Klinische Psychologie (Aufbau)	Univ.-Prof. Dr. Michael Witthöft
	Vorlesung: Einführung in die Personalpsychologie	Prof. Dr. Thomas Rigotti
	Arbeits-, Organisations-, Wirtschafts- und Personalpsychologie	Prof. Dr. Thomas Rigotti
	F. Allgemeine Psychologie Aufbau	Prof. Dr. Gerhard Voßel
	K. Klinische Psychologie (Basis)	Univ.-Prof. Wolfgang Hiller
SoSe 2014	M. Gesundheitspsychologie (Aufbau) - Anwendungsperspektiven der Gesundheitspsychologie	Daniela Zahn

	O. BSc. - Projekt und Evaluationsforschung	Univ.-Prof. Dr. Günter Meinhardt
	L. AOW - Psychologie (Aufbau): Personalpsychologie - Personalentwicklung (Human Resource Science)	Dr. Christian Skaletz M.B.A.
	J. Gesundheitspsychologie (Basis) - Kontext und Perspektiven der Gesundheitspsychologie	Prof. Thomas Kubiak
WiSe 2014/15	Bachelorarbeit: "Ausschluss der unkomplizierten Trauer aus der Depressionsdiagnose - Pro und Contra der DSM-5-Neuerung"	Univ.-Prof. Wolfgang Hiller
April 2016 - Juli 2017	Masterstudiengang Psychologie	Lehrende/Prüfende:
SoSe 2016	H. Klinische Bedingungsmodelle, Risikofaktoren und Entscheidung	Mila Domke-Wolf
	I. Differentielle Gesundheitspsychologie	Mario Wenzel
	Persönlichkeitsentwicklung	Tamara Pfeiler
	Angewandte Gesundheitspsychologie: Grundlagen	Prof. Thomas Kubiak
WiSe 2016/17	Studium generale RVL: Heimat heute	Dr. Andreas Hütig; Dr. Edith Struchholz-Andre; Dr. Thomas Vogt
	Psychologische Sachverständigengutachten	Michael Zimmer
	Seminar FV SFP: Psychologie im Familienrecht: Ein interdisziplinäres Seminar für angehende Psychologinnen und Psychologen und Juristinnen und Juristen	Dr. Paul Schaffner
	MSc Projektarbeit	Prof. Dr. Thomas Rigotti
	A. Forschungsmethoden und Diagnostik	Univ.-Prof. Dr. Günter Meinhardt
	Klinische Neuropsychologie	Prof. Dr. Michele Wessa
SoSe 2017	Masterarbeit: "Birds of the same feather learn well together? An experimental Study on the effect of homogeneous and heterogeneous learning group composition on satisfaction and performance"	Dr. Henrik Bellhäuser Maria Theobald

References

AbuSeileek, A. F. (2007). Cooperative vs. individual learning of oral skills in a CALL environment. *Computer Assisted Language Learning, 20*(5), 493–514. https://doi.org/10.1080/09588220701746054

Abuseileek, A. F. (2012). The effect of computer-assisted cooperative learning methods and group size on the EFL learners' achievement in communication skills. *Computers and Education, 58*(1), 231–239. https://doi.org/10.1016/j.compedu.2011.07.011

Acharya, A., & Sinha, D. (2018). A "mixed" strategy for collaborative group formation and its learning outcomes. *Journal of educational technology systems*, *46*(4), 440–462. https://doi.org/10.1177/004723951774924

Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive Learning Environments, 1–13*. https://doi.org/10.1080/10494820.2020.1813180

Alam, A. (2022). Mapping a sustainable future through conceptualization of transformative learning framework, education for sustainable development, critical reflection, and responsible citizenship: an exploration of pedagogies for twenty-first century learning. *ECS Transactions*, *107*(1), 9827. http://doi.org/10.1149/10701.9827ecst

Al-Dujaily, A., & Ryu, H. (2007). Personality and collaborative learning experience. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007),* 619–621. https://doi.org/10.1109/ICALT.2007.199

Altınay, L., & Paraskevas, A. (2007). A computer-supported collaborative learning (CSCL) approach in teaching research methods. *International Journal of Hospitality Management, 26*(3), 623–644. https://doi.org/10.1016/j.ijhm.2006.05.005

Amara, S., Macedo, J., Bendella, F., & Santos, A. (2016). Group formation in mobile computer supported collaborative learning contexts: A systematic literature review. *Journal of Educational Technology & Society, 19(*2), 258–273. http://www.jstor.org/stable/jeductechsoci.19.2.258

Anewalt, K., Polack-Wahl, J. A., Beidler, J., & Smarkusky, D. L. (2003). Group projects across the curriculum. *Journal of Computing Sciences in Colleges*, *19*(2), 232–237. https://doi.org/10.5555/948785.948819

Ani, Z. C., Yasin, A., Husin, M. Z., & Hamid, Z. A. (2010). A method for group formation using genetic algorithm. *International Journal on Computer Science and Engineering*, *2*(9), 3060–3064.

Apfelbaum, E. P., Phillips, K. W., & Richeson, J. A. (2014). Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity? *Perspectives on Psychological Science, 9*(3), 235–244. https://doi.org/10.1177/1745691614527466

Arispe, K. B., & Blake, R. J. (2012). Individual factors and successful learning in a hybrid course. *System, 40*(4), 449–465. https://doi.org/10.1016/j.system.2012.10.013

Atchley, T. W., Wingenbach, G., & Akers, C. (2013). Comparison of course completion and student performance through online and traditional courses. *The International Review of Research in Open and Distributed Learning, 14*(4). https://doi.org/10.19173/irrodl.v14i4.1461

Baldwin, S. J., & Ching, Y. H. (2019). An online course design checklist: development and users' perceptions. *Journal of Computing in Higher Education 31*(3), 156–172. https://doi.org/10.1007/s12528-018-9199-8

Balthazard, P., Potter, R., & Warren, J. (2002). The effects of extraversion and expertise on virtual team interaction and performance. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences,* 10. https://doi.org/10.1109/HICSS.2002.994426

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology, 44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., Neubert, M. J., Mount, M. K., & Stewart, G. L. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*(3), 377–391. https://doi.org/10.1037/0021-9010.83.3.377

Bass, B. M., & Riggio, R. E. (2005). Transformational leadership (2nd ed.). *Psychology Press*. https://doi.org/10.4324/9781410617095

Bates, D., Maechler, M., Bolker, B., Walker, S., Bojesen Christensen, R. H., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., & Fox, J. (2020, December 1). Package 'lme4'. https://cran.r-project.org/web/packages/lme4/lme4.pdf.

Bayeck, R. Y., Hristova, A., Jablokow, K. W., & Bonafini, F. (2018). Exploring the relevance of single-gender group formation: what we learn from a massive open online course (MOOC). *British Journal* *of Educational Technology, 49*(1), 88–100. https://doi.org/10.1111/bjet.12528

Bekele, R. (2006). *Computer-assisted learner group formation based on personality traits* [Ph.D Dissertation, University of Hamburg, Hamburg, Germany]. https://ediss.sub.uni-hamburg.de/handle/ediss/1199

Bekele, R., & Menzel, W. (2005, February 14). A Bayesian approach to predict performance of a student (BAPPS): A case with ethiopian students. *International Conference on Artificial Intelligence and Applications*, *part of the 23rd Multi-Conference on Applied Informatics*, *(AIA–2005*), 189–194. http://www.actapress.com/Abstract.aspx?paperId=18917

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology, 92*(3), 595–615. https://doi.org/10.1037/0021-9010.92.3.595

Bell, S. T., Brown, S. G., Colaneri, A., & Outland, N. (2018). Team composition and the ABCs of teamwork. *American Psychologist, 73*(4), 349–362. https://doi.org/10.1037/amp0000305

Bellhäuser, H., Konert, J., Müller, A., & Röpke, R. (2018). Who is the perfect match? Effects of algorithmic learning group formation using personality traits. *Journal of Interactive Media (i-Com), 17*(1), 65–77. https://doi.org/https://doi.org/10.1515/icom-2018-0004

Benbunan-Fich, R., Hiltz, S. R., & Turoff, M. (2001). A comparative content analysis of face-to-face vs. ALN-mediated teamwork. *Proceedings of the 34th Hawaii International Conference on System Sciences,* 10. https://doi.org/10.1109/HICSS.2001.926189

Bendersky, C., & Shah, N. P. (2013). 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. https://doi.org/10.5465/amj.2011.0316

Blanco-Fernández, D., Leitner, S., & Rausch, A. (2023). Interactions between the individual and the group level in organizations: The case of learning and group turnover. *Central European Journal of Operations Research, 31*(4), 1087–1128. https://doi.org/10.1007/s10100-023-00843-7

Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., & Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Computers and Education, 62*, 102–110. https://doi.org/10.1016/j.compedu.2012.10.019

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein, & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). Jossey-Bass/Wiley.

Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership: a meta-analysis. *Journal of applied psychology*, *89*(5), 901. http://doi.org/10.1037/0021-9010.89.5.901

Borges S., Mizoguchi R., Bittencourt I. I., & Isotani S. (2018). Group formation in CSCL: A review of the state of the art. In A. Cristea, I. Bittencourt & F. Lima (Eds.), *Higher education for all. From challenges to novel technology-enhanced solutions. Communications in Computer and Information Science* (Vol. 832, pp. 71–88). Springer. https://doi.org/10.1007/978-3-319-97934-2_5

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. https://doi.org/10.1177/104649640003100303

Bradley, J. H., & Hebert, F. J. (1997). The effect of personality type on team performance. *Journal of Management Development, 16*(5), 337–353. https://doi.org/10.1108/02621719710174525

Braun, M. T., Kozlowski, S. W. J., Brown, T. A., & DeShon, R. P. (2020). Exploring the dynamic team cohesion–performance and coordination– performance relationships of newly formed teams. *Small Group Research, 51*(5), 551–580. https://doi.org/10.1177/1046496420907157

Bruffee, K. A. (1999). Collaborative learning: Higher education, interdependence, and the authority of knowledge (2nd ed.). *John Hopkins University Press*.

Brunell, A. B., Gentry, W. A., Campbell, W. K., Hoffman, B. J., Kuhnert, K. W., & DeMarree, K. G. (2008). Leader emergence: the case of the narcissistic leader. *Personality and Social Psychology Bulletin, 34*(12), 1663–1676. https://doi.org/10.1177/0146167208324101

Bundtzen, H., & Hinrichs, G. (2021). The link between organizational agility and VUCA – An agile assessment model. *SocioEconomic Challenges, 5*(1), 35–43. https://doi.org/10.21272/sec.5(1).35-43.2021

Burnham, K. P., & Anderson, D. R. (2016). Multimodel inference. *Sociological Methods & Research. 33*(2). https://doi.org/10.1177/0049124104268644

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences, 29*(6), 1057–1068. https://doi.org/10.1016/S0191-8869(99)00253-6

Byrne, D. (1971). The attraction paradigm. *Academic Press*.

Cable, D. M., & Edwards, J. R. (2004). Complementary and supplementary fit: A theoretical and empirical integration. *Journal of Applied Psychology, 89*(5), 822–834. https://doi.org/10.1037/0021-9010.89.5.822

Cachia, M., Lynam, S., & Stock, R. (2018). Academic success: Is it just about the grades? *Higher Education Pedagogies, 3*(1), 434–439. https://doi.org/10.1080/23752696.2018.1462096

Carron, A. V., Colman, M. M., Wheeler, J., & Stevens, D. (2002). Cohesion and performance in sport: A meta analysis. *Journal of Sport and Exercise Psychology, 24*(2), 168–188. https://doi.org/10.1123/jsep.24.2.168

Carter, M. A. (2013). A Study of students' perceptions of the online component of a hybrid postgraduate course. *Procedia - Social and Behavioral Sciences, 84*, 558–568. https://doi.org/10.1016/j.sbspro.2013.06.604

Cera Guy, J. N., Williams, J. M., Bruce M., & Shore, B. M. (2019). High- and Otherwise-Achieving Students'Expectations of Classroom Group Work: An Exploratory Empirical Study. *Roeper Review, 41*(3), 166–184. http://doi.org/10.1080/02783193.2019.1622166

Chahal, J., & Rani, N. (2022). Exploring the acceptance for e-learning among higher education students in India: combining technology acceptance model with external variables. *Journal of Computing in Higher Education 34*(3), 844–867. https://doi.org/10.1007/s12528-022-09327-0

Chalmers, C., & Nason, R. (2005, May). Group metacognition in a computer-supported collaborative learning environment. In C. K. Looi, M. Ikeda, & D. Jonassen (Eds.), *Towards Sustainable and Scalable Educational Innovations Informed by the Learning Sciences.* IOS Press, 35–41.

Chamorro-Premuzic, T., Furnham, A., & Lewis, M. (2007). Personality and approaches to learning predict preference for different teaching methods. *Learning and individual differences, 17*(3), 241–250.

Chang, Y., & Brickman, P. (2018). When group work doesn't work: Insights from students. *CBE life sciences education, 17*(3), ar42. https://doi.org/10.1187/cbe.17-09-0199

Chen, C. M., & Kuo, C. H. (2019). An optimized group formation scheme to promote collaborative problem-based learning. *Computers & Education*, *133*, 94–115. https://doi.org/10.1016/j.compedu.2019.01.011

Chen, C. T. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy sets and systems*, *114*(1), 1–9. https://doi.org/10.1016/S0165-0114(97)00377-1

Chen, Y., Ning, R., Yang, T., Feng, S., & Yang, C. (2018). Is transformational leadership always good for employee task performance? Examining curvilinear and moderated relationships. *Frontiers of Business Research in China, 12(*22). https://doi.org/10.1186/s11782-018-0044-8

Chiriac, E. (2014). Group work as an incentive for learning - students' experiences of group work. *Frontiers in Psychology, 5*, 558. https://doi.org/10.3389/fpsyg.2014.00558

Chiu, M. M. (2000). Group problem-solving processes: social interactions and individual actions. *Journal for the Theory of Social Behaviour*, *30*(1), 26–49. https://doi.org/10.1111/1468-5914.00118

Chou, C. C. (2001). Formative evaluation of synchronous CMC systems for a learner-centered online course. *Journal of interactive learning research, 12*(2), 173–192. Retrieved March 20, 2023. https://www.learntechlib.org/primary/p/8418/

Christodoulopoulos, C. E., & Papanikolaou, K. A. (2007). A group formation tool in an e-learning context. *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, 117–123. https://doi.org/10.1109/ICTAI.2007.155

Clark, D. A. G., Marnewick, A. L., & Marnewick, C. (2019). Virtual team performance factors: a systematic literature review. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2019*, *December 15–18, pp.* 40–44). https://doi.org/10.1109/IEEM44572.2019.8978809

Cleveland-Innes, M., & Campbell, P. (2012). Emotional presence, learning, and the online learning environment. *International Review of Research in Open and Distance Learning, 13*(4), 269–292. https://doi.org/10.19173/irrodl.v13i4.1234

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal*, *32*(1), 99–120. https://doi.org/10.2307/1163215

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). *Lawrence Erlbaum*

Associates.

Conole, G., de Laat, M., Dillon, T., & Darby, J. (2008). "Disruptive technologies", "pedagogical innovation": what's new? Findings from an in-depth Study of students' use and perception of technology. *Computers & Education. 50*(2), 511–524. https://doi.org/10.1016/j.compedu.2007.09.009

Contractor, N. (2013). Some assembly required: Leveraging web science to understand and enable team assembly. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, (371*)1987. https://doi.org/10.1098/rsta.2012.0385

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment, 4*(1), 5–13. http://dx.doi.org/10.1037/1040-3590.4.1.5

Costa, P., & McCrae, R. (2008). The revised NEO personality inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment: Volume 2 — Personality measurement and testing. SAGE Publications Ltd.* https://doi.org/10.4135/9781849200479.n9

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: ten years of experience and results. *American Journal of Physics, 69*(9), 970–977. https://doi.org/10.1119/1.1374249

Crozier, W. R., & Perkins, P. (2002). Shyness as a factor when assessing children. *Educational Psychology in Practice, 18*(3), 239–244. https://doi.org/10.1080/0266736022000010267

Curry, O., & Dunbar, R. I. M. (2013). Do birds of a feather flock together? *Human Nature 24*, 336–347. https://doi.org/10.1007/s12110-013-9174-z

Curşeu, P. L., Ilies, R., Vîrgă, D., Maricuţoiu, L., & Sava, F. A. (2019). Personality characteristics that are valued in teams: Not always "more is better"? *International journal of psychology: Journal international de psychologie*, *54*(5), 638–649. https://doi.org/10.1002/ijop.12511

Curşeu, P. L., Pluut, H., Boroş, S., & Meslec, N. (2015). The magic of collective emotional intelligence in learning groups: No guys needed for the spell! *British Journal of Psychology, 106*(2), 217–234.https://doi.org/10.1111/bjop.12075

Curşeu, P. L., Rusu, A., Maricuţoiu, L. P., Vîrgă, D., & Măgurean, S. (2020). Identified and engaged: A multi-level dynamic model of identification with the group and performance in collaborative learning. *Learning and Individual Differences, 78,* 101838. https://doi.org/10.1016/j.lindif.2020.101838

Damşa, C. I. (2014). The multi-layered nature of small-group learning: Productive interactions in object-oriented collaboration. *International Journal of Computer-Supported Collaborative Learning 9(*3), 247–281. https://doi.org/10.1007/s11412-014-9193-8

Daradoumis, T., Guitert, M., Gimenez, F., Marques J. M., & Lloret, T. (2002). Supporting the composition of effective virtual groups for collaborative learning. *International Conference on Computers in Education, 1,* 332–336. https://doi.org/10.1109/CIE.2002.1185939

De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology, 88*(4), 741–749. https://doi.org/10.1037/0021-9010.88.4.741

De Dreu, C. K., & Nijstad, B. A. (2008). Mental set and creative thought in social conflict: threat rigidity versus motivated focus. *Journal of Personality and Social Psychology*, *95*(3), 648–661. https://doi.org/10.1037/0022-3514.95.3.648

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

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. In *Springer eBooks*. https://doi.org/10.1007/978-1-4899-2271-7

Deckers, M., Altmann, T., & Roth, M. (2020). The role of individual personality in conceptualizing and measuring group characteristics. *Current Psychology, 41*(6), 4054–4065. http://doi.org/10.1007/s12144-020-00919-6

Den Hartog, S. C., Runge, J. M., Reindl, G., & Lang, J. W. B. (2019). Linking personality trait variance in self-managed teams to team innovation. *Small Group Research, 51*(2), 265–295. https://doi.org/10.1177/1046496419865325

Devine, D. J., & Philips, J. L. (2001). Do smarter teams do better: A meta-analysis of cognitive ability and team performance. *Small Group Research, 32*(5), 507–532. https://doi.org/10.1177/104649640103200501

Di Fabio, A., & Busoni, L. (2007). Fluid intelligence, personality traits and scholastic success: Empirical evidence in a sample of Italian high school students. *Personality and Individual Differences, 43*(8), 2095–2104. https://doi.org/10.1016/j.paid.2007.06.025

Diaz, D. P. (2002). Online drop rate revisited. *Extending the Pedagogy of Threaded-Topic Discussions, 2001*(1). Retrieved March 16, 2023. https://www.learntechlib.org/p/96381/

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.) *Collaborative-learning: Cognitive and Computational Approaches*.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61–91). Open University Nederland.

Dincă, M., Luştrea, A., Onițiu, A., Crașovan, M., & Berge, T. (2021). The effects of disciplinary composition on virtual learning group process dynamics: students' perspectives. *Sustainability, 13*(15), 8493. https://doi.org/10.3390/su13158493

Donovan, D. A., Connell, G. L., & Grunspan, D. Z. (2018). Student learning outcomes and attitudes using three methods of group formation in a nonmajors biology class. *CBE Life Sciences Education, 17*(4), 1–14. https://doi.org/10.1187/cbe.17-12-0283

Driskell, J. E., Goodwin, G. F., Salas, E., & O'Shea, P. G. (2006). What makes a good team player? Personality and team effectiveness. *Group Dynamics: Theory, Research, and Practice, 10*(4), 249–271. https://doi.org/10.1037/1089-2699.10.4.249

Druskat, V. U., & Pescosolido, A. T. (2002). The content of effective teamwork mental models in self-managing teams: ownership, learning and heedful interrelating. *Human Relations, 55*(3), 283–314. https://doi.org/10.1177/0018726702553001

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2016). Dynamics of communicator and audience power: the persuasiveness of competence versus warmth. *Journal of Consumer Research, 43*(1), 68–85. https://doi.org/10.1093/jcr/ucw006

Dumford, A. D., & Miller, A. L. (2018). Online learning in higher education: exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education, 30*, 452–465. https://doi.org/10.1007/s12528-018-9179-z

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *Proceedings of the 3rd innovations in theoretical computer science conference* (pp. 214–226). https://arxiv.org/abs/1104.3913

Eagly, A. H., & Carli, L. L. (2003). The female leadership advantage: An evaluation of the evidence. *Leadership Quarterly, 14*(6), 807-834. https://doi.org/10.1016/j.leaqua.2003.09.004

Eckhaus, E., & Davidovitch, N. (2019). Technology-supported teaching: Technological progress or a sham? *European Journal of Educational Research, 8(*3), 697–702. https://doi.org/10.12973/eu-jer.8.3.697

Eckhaus, E., Klein, G., & Kantor, J. (2017). Experiential learning in management education. *Business, Management and Economics Engineering, 15*(1), 42–56. https://doi.org/10.3846/bme.2017.345

Elshan, E., & Ebel, P. (2020). Let's team up: Designing conversational agents as teammates. *Proceedings 41 International Conference on Information Systems, (ICIS 2020, pp.* 2). https://www.alexandria.unisg.ch/handle/20.500.14171/111505

Engel, D., Woolley, A. W., Aggarwal, I., Chabris, C. F., Takahashi, M., Nemoto, K., Kaiser, C., Kim, Y. J., & Malone, T. W. (2015, April). Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures. *Proceedings Conference on Human Factors in Computing Systems*, 3769–3778. https://doi.org/10.1145/2702123.2702259

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. https://doi.org/10.1371/journal.pone.0115212

English, A., Griffith, R. L., & Steelman, L. A. (2004). Team performance: The effect of team conscientiousness and task type. *Small Group Research, 35*(6), 643–665. https://doi.org/10.1177/1046496404266320

Erez, M., Lisak, A., Harush, R., Glikson, E., Nouri, R., & Shokef, E. (2013). Going global: Developing management students' cultural intelligence and global identity in culturally diverse virtual teams. *Academy of Management Learning & Education, 12(*3), 330–355. https://doi.org/10.5465/amle.2012.0200

Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43–

71. https://doi.org/10.1002/piq.21143

Fahmi, F. Z., & Nurjanah, D. (2018). Group formation using multi objectives ant colony system for collaborative learning. *Proceedings 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 696–702. http://doi.org/10.1109/EECSI.2018.8752690

Fazal-e-Hasan, S.M., Mortimer, G., Lings, I., Sekhon, H., & Howell, K. (2021). Managing relationships: Insights from a student gratitude model. *Research in Higher Education 62*(4), 98–119. https://doi.org/10.1007/s11162-019-09576-6

Feichtner, S. B., & Davis, E. C. (1984). Why some groups fail: a survey of students' experiences with learning groups. *Journal of Management Education, 9(*4), 58–73. https://doi.org/10.1177/105256298400900409

Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality, 37*(2), 62–88. https://doi.org/10.1016/S0092-6566(02)00536-6

Fenwick, G. D., & Neal, D. J. (2002). Effect of gender composition on group performance. *Gender, Work & Organization, 8*(2), 205–225.https://doi.org/10.1111/1468-0432.00129

Filade, B. A., Bello, A. A., Uwaoma, C., Anwanane, B. B., & Nwangburuka, K. (2019). Peer group influence on academic performance of undergraduate students in Babcock University, Ogun State, Nigeria. *African Educational Research Journal, 7*(2), 81–87. https://doi.org/10.30918/AERJ.72.19.010

Fiorentino, G., Miranda, P., Nascimento, A., Furtado, A. P., Bellhäuser, H., Gašević, D., & Mello, R. F. (2021). Contrasting automatic and manual group formation: a case study in a software engineering postgraduate course. In I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, & V. Dimitrova (Eds.), *Lecture Notes in Computer Science* (pp. 168–172). https://doi.org/10.1007/978-3-030-78270-2_30

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: fifteen experience-sampling studies and a meta-analysis. *Journal of personality and social psychology*, *97*(6), 1097–1114. https://doi.org/10.1037/a0016786

Forrest, K. D., & Miller, R. L. (2003). Not another group project: Why good teachers should care about bad group experiences. *Teaching of Psychology, 30*(3), 244–246.

Forrester, W. R., & Tashchian, A. (2010). Effects of personality on attitudes toward academic group work. *American Journal of Business Education (AJBE), 3*(3), 39–46. https://doi.org/10.19030/ajbe.v3i3.397

Forsell, J., Forslund Frykedal, K., & Hammar Chiriac, E. (2020). Group work assessment: assessing social skills at group-level. *Small Group Research, 51*, 87–124. https://doi.org/10.1177/104649641987826

French, K. A., & Kottke, J. L. (2013). Teamwork satisfaction: Exploring the multi-level interaction of teamwork interest and group extraversion. *Active Learning in Higher Education, 14*(3), 189–200. https://doi.org/10.1177/1469787413498034

Fryer, L. K., & Dinsmore, D. L. (2020). The promise and pitfalls of self-report: Development, research design and analysis issues, and multiple methods. *Frontline Learning Research, 8*(3), 1–9. https://doi.org/10.14786/flr.v8i3.623

Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, *14*(1), 47–64. https://doi.org/10.1016/j.lindif.2003.08.002

Gawande, A. A., Zinner, M. J., Studdert, D. M., & Brennan, T. A. (2003). Analysis of errors reported by surgeons at three teaching hospitals. *Surgery, 133*(6), 614–621. https://doi.org/10.1067/msy.2003.169

Geiser, C., Eid, M., Nussbeck, F. W., Courvoisier, D. S., & Cole, D. A. (2010). Multitrait-multimethod change modelling. *AStA Advances in Statistical Analysis, 94*(2), 185–201. https://doi.org/10.1007/s10182-010-0127-0

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. *Cambridge University Press.* https://doi.org/10.1017/CBO9780511790942

Giel, L., Noordzij, G., Wijnia, L., Noordegraaf-Eelens, L., & Denktaş, S. (2020). When birds of the same feather fly together: the impact of achievement goal compatibility in collaborative learning. *Educational Psychology*, *41*(1), 79–98. https://doi.org/10.1080/01443410.2020.1787352

Gillen-O'Neel, C. (2021). Sense of belonging and student engagement: a daily study of first- and continuing-generation college students. *Research in Higher Education 62*, 45–71.

https://doi.org/10.1007/s11162-019-09570-y

Gillies, R. M. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction, 14*, 197–213. https://doi.org/10.1016/S0959-4752(03)00068-9

Gladstein, D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative Science Quarterly, 29*(4), 499–517. https://doi.org/10.2307/2392936

Goñi, J., Cortázar, C., Alvares, D., Donoso, U., & Miranda, C. (2020). Is teamwork different online versus face-to-face? A case in engineering education. *Sustainability*, *12*(24), 10444. https://doi.org/10.3390/su122410444

Graesser, A. C., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. W. (2018). Advancing the science of collaborative problem solving. *Psychological Science in the Public Interest, 19*(2), 59–92. https://doi.org/10.1177/1529100618808244

Graf, S., & Bekele, R. (2006). Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In M. Ikeda, K. D. Ashley & T. W. Chan, (Eds.), Intelligent Tutoring Systems. ITS 2006. *Lecture Notes in Computer Science, 4053,* 217–226. https://doi.org/10.1007/11774303_22

Graham, C. R., & Misanchuk, M. (2004). Computer-mediated learning groups: benefits and challenges to using groupwork in online learning environments. In T. Roberts (Ed.), Online Collaborative Learning: Theory and Practice. *IGI Global,* (pp. 181–202). https://doi.org/10.4018/978-1-59140-174-2.ch008

Greer, L. L., de Jong, B. A., Schouten, M. E., & Dannals, J. E. (2018). Why and when hierarchy impacts team effectiveness: A meta-analytic integration. *Journal of Applied Psychology, 103*(6), 591–613. https://doi.org/10.1037/apl0000291

Grønkjær, M., Curtis, T., de Crespigny, C., & Delmar, C. (2011). Analysing group interaction in focus group research: Impact on content and the role of the moderator. *Qualitative Studies, 2*(1), 16–30. https://doi.org/10.7146/qs.v2i1.4273

Hachey, A. C., Wladis, C. W., & Conway, K. M. (2013). Balancing retention and access in online courses: Restricting enrollment… Is it worth the cost? *Journal of College Student Retention: Research, Theory & Practice, 15*(1), 9–36. https://doi.org/10.2190/CS.15.1.b

Hachey, A. C., Conway, K. M., Wladis, C., & Karim, S. (2022). Post-secondary online learning in the U.S.: an integrative review of the literature on undergraduate student characteristics. *Journal of Computing in Higher Education 34*(3), 708–768. https://doi.org/10.1007/s12528-022-09319-0

Hackman, J. R. (1983). *A normative model of work team effectiveness*. Yale School of Organization and Management.

Hailikari, T., Katajavuori, N., & Lindblom-Ylanne, S. (2008). The relevance of prior knowledge in learning and instructional design. *American Journal of Pharmaceutical Education, 72*(5), 1–8: https://pubmed.ncbi.nlm.nih.gov/19214267/

Halfhill, T., Sundstrom, E., Lahner, J., Calderone, W., & Nielsen, T. M. (2005). Group personality composition and group effectiveness: An integrative review of empirical research. *Small Group Research, 36*(1), 83–105. https://doi.org/10.1177/1046496404268538

Hall, D., & Buzwell, S. (2013). The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution. *Active Learning in Higher Education, 14*(1), 37–49. https://doi.org/10.1177/1469787412467123

Haller, C. R., Gallagher, V. J., Weldon, T. L., & Felder, R. M. (2000). Dynamics of peer education in cooperative learning workgroups. *Journal of Engineering Education, 89*(3), 285–293. https://doi.org/10.1002/j.2168-9830.2000.tb00527.x

Harrer, A., McLaren, B. M., Walker, E., Bollen, L., & Sewall, J. (2006). Creating cognitive tutors for collaborative learning: steps toward realization. *User Modeling and User-adapted Interaction, 16*(3–4), 175–209. https://doi.org/10.1007/s11257-006-9007-4

Harris, K., English, T., Harms, P. D., Gross, J. J., & Jackson, J. J. (2017). Why are extraverts more satisfied? Personality, social experiences, and subjective well–being in college. *European Journal of Personality*, *31*(2), 170–186. https://doi.org/10.1002/per.2101

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. *The Academy of Management Journal, 45*(5), 1029–1045. https://doi.org/10.2307/3069328

Harrison, D., Price, K., & Bell, M. (1998). Beyond relational demography: Time and the effects of surface- and deep-level diversity on work group cohesion. *Academy of Management Journal, 41*(1), 96–107.

https://doi.org/10.2307/256901

Harvey, J.-F., Johnson, K. J., Roloff, K. S., & Edmondson, A. C. (2019). From orientation to behavior: The interplay between learning orientation, open-mindedness, and psychological safety in team learning. Human Relations, 72(11), 1726–1751. https://doi.org/10.1177/0018726718817812

Henrich, J., Heine, S. & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. *The Behavioral and brain sciences,* *33*(2–3), 61–135. https://doi.org/10.1017/s0140525x0999152x

Hertz-Lazarowitz, R., & Bar-Natan, I. (2002). Writing development of Arab and Jewish students using cooperative learning (CL) and computer-mediated communication (CMC). *Computers and Education, 39*(1), 19–36. https://doi.org/10.1016/S0360-1315(02)00019-2

Hillyard, C., Gillespie, D., & Littig, P. (2010). University students' attitudes about learning in small groups after frequent participation. *Active Learning in Higher Education, 11*(1), 9–20. https://doi.org/10.1177/1469787409355867

Hitt, M. A., Beamish, P. W., Jackson, S. E., & Mathieu, J. E. (2007). Building theoretical and empirical bridges across levels: Multilevel research in management. *Academy of Management Journal, 50*(6), 1385–1399. https://doi.org/10.5465/amj.2007.28166219

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review.* https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning.

Hofeditz, L., Clausen, S., Rieß, A., Mirbabaie, M., & Stieglitz, S. (2022). Applying XAI to an AI-based system for candidate management to mitigate bias and discrimination in hiring. *Electronic Markets*, *32*(4), 2207–2233. https://doi.org/10.1007/s12525-022-00600-9

Hogan, R., Curphy, G. J., & Hogan, J. (1994). What we know about leadership: Effectiveness and personality. *American Psychologist, 49*(6), 493–504. https://doi.org/10.1037/0003-066x.49.6.493

Hollenbeck, J. R., DeRue, D. S., & Guzzo, R. (2004). Bridging the gap between I/O research and HR practice: Improving team composition, team training, and team task design. *Human Resource Management, 43*(4), 353–366. https://doi.org/10.1002/hrm.20029

Hollingshead, A. B., McGrath, J. E., & O'Connor, K. M. (1993). Group task performance and communication technology: A longitudinal Study of computer mediated versus face-to-face Work Groups. *Small Group Research, 24*(3), 307–333. https://doi.org/10.1177/1046496493243003

Horton, W. (2000). Designing web-based training: how to teach anyone anything anywhere anytime. Wiley.

Horwitz, S. K. (2005). The compositional impact of team diversity on performance: Theoretical considerations. *Human Resource Development Review, 4*(2), 219–245. https://doi.org/10.1177/1534484305275847

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management, 33*(6), 987–1015. https://doi.org/10.1177/0149206307308587

Hough, L. M. (1992). The "big five" personality variables—construct confusion: Description versus prediction. *Human Performance, 5*(1–2), 139–155. https://doi.org/10.1080/08959285.1992.9667929

Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future - Remembering the past. *Annual Review of Psychology, 51*(1), 631–664. http://dx.doi.org/10.1146/annurev.psych.51.1.631

Hough, L. M., & Oswald, F. L. (2005). They're right, well ... Mostly right: Research evidence and an agenda to rescue personality testing from 1960s insights. *Human Performance, 18(*4), 373–387. https://doi.org/10.1207/s15327043hup1804_4

Houlden, S., & Veletsianos, G. (2022). A synthesis of surveys examining the impacts of COVID-19 and emergency remote learning on students in Canada. *Journal of Computing in Higher Education, 34*(3), 820–843. https://doi.org/10.1007/s12528-022-09323-4

Hox, J. J. (2010). Introduction to multilevel analysis. In J. J. Hox, M. Moerbeek, & R., Van De Schoot, Multilevel analysis: Techniques and applications (2nd ed.), *Routledge eBooks* (pp. 4–46). https://doi.org/10.4324/9780203852279

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). Some important methodological and statistical issues. In *Routledge eBooks* (pp. 41-69). https://doi.org/10.4324/9781315650982

Hox, J. J., Moerbeek, M., van de Schoot, R., Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). Multilevel Factor Models. In *Routledge eBooks* (pp. 269–283). https://doi.org/10.4324/9781315650982-14

Hsiung, C.-M. (2012). The effectiveness of cooperative learning. *Journal of Engineering Education, 101*(1), 119–137. https://doi.org/10.1002/j.2168-9830.2012.tb00044.x

Huang, J.L., Wu, D. (2020). Other-contingent extraversion and satisfaction: The moderating role of implicit theory of personality. *Journal of Individual Differences, 42*(3), 132–139, https://doi.org/10.1027/1614-0001/a000339

Huber, B., Goyanes, M., & Gil de Zúñiga, H. (2021). Linking extraversion to collective and individual forms of political participation: The mediating role of political discussion. *Social Science Quarterly*, *102*(4), 1289–1310. https://doi.org/10.1111/ssqu.12978

Hübscher, R. (2010). Assigning students to groups using general and context-specific criteria. *IEEE Transactions on Learning Technologies, 3*(3), 178–189. https://doi.org/10.1109/TLT.2010.17

Humphrey, S. E., Hollenbeck, J. R., Meyer, C. J., & Ilgen, D. R. (2007). Trait configurations in self-managed groups: A conceptual examination of the use of seeding for maximizing and minimizing trait variance in groups. *Journal of Applied Psychology, 92*(3), 885–892. https://doi.org/10.1037/0021-9010.92.3.885

Humphrey, S. E., Hollenbeck, J. R., Meyer, C. J., & Ilgen, D. R. (2011). Personality configurations in self-managed groups: A natural experiment on the effects of maximizing and minimizing variance in traits. *Journal of Applied Social Psychology, 41*(7), 1701–1732. https://doi.org/10.1111/j.1559-1816.2011.00778.x

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*(6), 869–879. https://doi.org/10.1037/0021-9010.85.6.869

Hwang, G. J., Sung, H. Y., Hung, C. M., & Huang, I. (2013). A learning style perspective to investigate the necessity of developing adaptive learning systems. *Journal of Educational Technology & Society, 16(*2), 188–197. http://www.jstor.org/stable/jeductechsoci.16.2.188

Hwang, G.-J., Yin, P.-Y., Hwang, C.-W., & Tsai, C.-C. (2008). An enhanced genetic approach to composing cooperative learning groups for multiple grouping criteria. *Educational Technology & Society 11*(1), 148–167. https://www.jstor.org/stable/jeductechsoci.11.1.148

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: from input-process-output models to IMOI models. *Annual Review of Psychology, 56*(1), 517–543.

https://doi.org/10.1146/annurev.psych.56.091103.070250

Inaba, A., Supnithi, T., Ikeda, M., Mizoguchi, R., & Toyoda, J. (2000). How can we form effective collaborative learning groups? In *Lecture Notes in Computer Science* (S. 282–291). https://doi.org/10.1007/3-540-45108-0_32

Jackson, C. M., Jackson, J. C., Bilkey, D., Jong, J., & Halberstadt, J. (2019). The dynamic emergence of minimal groups. *Group Processes and Intergroup Relations, 22*(7), 921–929. https://doi.org/10.1177/1368430218802636

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers and Education, 49*(4), 1037–1065. https://doi.org/10.1016/j.compedu.2006.01.004

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business horizons*, *61*(4), 577-586. https://doi.org/10.1016/j.bushor.2018.03.007

Johnson D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory into Practice, 38*(2), 67–73. https://doi.org/10.1080/00405849909543834

Johnson, D. W., & Johnson, R. (2005). New developments in social interdependence theory. *Genetic, Social, and General Psychology Monographs*, *131*(4), 285–358. https://doi.org/10.3200/mono.131.4.285-358

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1991). *Cooperative learning: Increasing college faculty instructional productivity. ASHE-ERIC Higher Education Report No. 4*. ASHE-ERIC Higher Education Reports.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. Cooperative Learning Center at the University of Minnesota Press. Retrieved November 6, 2021. http://www.tablelearning.com/uploads/File/EXHIBIT_B.pdf

Joiner, R. (2004). Supporting collaboration in virtual learning environments. *CyberPsychology & Behavior, 7(2),* 197–200. http://dx.doi.org/10.1089/109493104323024456

Jones, M. H., & Cooke, T. J. (2021). Social status and wanting popularity: different relationships with academic motivation and achievement. *Social Psychology of Education 24*, 1281–1303. https://doi.org/10.1007/s11218-021-09653-8

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*(4), 765–780. https://doi.org/10.1037/0021-9010.87.4.765

Junker, T. L., Bakker, A. B., Gorgievski, M. J., & Derks, D. (2022). Agile work practices and employee proactivity: A multilevel study. *Human Relations, 75*(12), 2189–2217. https://doi.org/10.1177/00187267211030101

Kagan, S. (1994). Cooperative learning. *Kagan Publishing*. http://www.learningtolearn.sa.edu.au/tfel/files/links/3b_cooperative_learning_1.pdf

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*(4), 681–706. https://doi.org/10.1037/0022-3514.65.4.681

Kerr, N. L., & Bruun, S. (1983). The dispensability of member effort and group motivation losses: Free rider effects. *Journal of Personality and Social Psychology, 44*(1), 78–94. https://doi.org/10.1037/0022-3514.44.1.78

Kerr, N. L., Messé, L. A., Park, E. S., & Sambolec, E. J. (2005). Identifiability, performance feedback and the Köhler effect. *Group Processes and Intergroup Relations, 8*(4), 375–390. https://doi.org/10.1177/1368430205056466

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*(1), 623–655. https://doi.org/10.1146/annurev.psych.55.090902.142009

Khalil, H. & Ebner, M. (2014). MOOCs Completion Rates and Possible Methods to Improve Retention - A Literature Review. In J. Viteli, & M. Leikomaa (Eds.), *Proceedings of EdMedia 2014 World Conference on Educational Media and Technology* (pp. 1305-1313). Association for the Advancement of Computing in Education (AACE). Retrieved February 18, 2024 from https://www.learntechlib.org/primary/p/147656/

Kickul, J., & Neuman, G. (2000). Emergent leadership behaviors: The function of personality and cognitive ability in determining teamwork performance and KSAs. *Journal of Business and Psychology 15*, 27–51. https://doi.org/10.1023/A:1007714801558

Kiesler, D. J., (1983). The 1982 interpersonal circle: A taxonomy for complementarity in human transactions. *Psychological review*, *90*(3), 185. http://doi.org/10.1037/0033-295X.90.3.185

Kirchner, K., Razmerita, L., & Sudzina, F. (2009). New forms of interaction and knowledge sharing on Web 2.0. In M. Lytras, E. Damiani, & P. Ordóñez de Pablos (Eds.), Web 2.0., *Springer eBooks,* 1-16. https://doi.org/10.1007/978-0-387-85895-1_2

Kirschner, P. A., (2017). Stop propagating the learning styles myth. *Computers and Education, 106,* 166–171. https://doi.org/10.1016/j.compedu.2016.12.006

Klein, C., DiazGranados, D., Salas, E., Le, H., Burke, C. S., Lyons, R., & Goodwin, G. F. (2009). Does team building work? *Small Group Research, 40*(2), 181–222. https://doi.org/10.1177/1046496408328821

Koh, M. H., & Hill, J. R. (2009). Student perceptions of groupwork in an online course: benefits and challenges. *International Journal of E-Learning & Distance Education: Revue Internationale Du E-Learning Et La Formation à Distance*, *23*(2), 69–92. Retrieved from https://www.ijede.ca/index.php/jde/article/view/477

Konert, J., Bellhäuser, H., Röpke, R., Gallwas, E., & Zucik, A. (2016). MoodlePeers: Factors relevant in learning group formation for improved learning outcomes, satisfaction and commitment in E-Learning scenarios using GroupAL. In K. Verbert, M. Sharples, & T. Klobucar (Eds.), *Adaptive and Adaptable Learning: Proceedings of the 11th European Conference on Technology Enhanced Learning, EC-TEL 2016, Lecture Notes in Computer Science* (vol. 9891, pp. 390–396). https://doi.org/10.1007/978-3-319-45153-4_32

Konert, J., Burlak, D., & Steinmetz, R. (2014). The group formation problem: an algorithmic approach to learning group formation. In *Lecture Notes in Computer Science* (pp. 221–234). https://doi.org/10.1007/978-3-319-11200-8_17

Kosovich, J. J., Hulleman, C. S., Barron, K. E., & Getty, S. (2015). A practical measure of student motivation. *The Journal of Early Adolescence, 35*(5–6), 790–816. https://doi.org/10.1177/0272431614556890

Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds*.), Handbook of psychology: Industrial and Organizational Psychology*, (vol.12, pp. 333–375). https://doi.org/10.1002/0471264385.wei1214

Kozlowski, S. W. J., & Bell, B. S. (2013). Work groups and teams in organizations: Review update. In N. Schmitt, & S. Highhouse (Eds.), *Handbook of psychology. Industrial and organizational psychology* (2nd ed., Vol. 12, pp. 412–469). https://doi.org/10.1002/0471264385.wei1214

Kramer, A., Bhave, D. P., & Johnson, T. D. (2014). Personality and group performance: The importance of personality composition and work tasks. *Personality and Individual Differences, 58*(24), 132–137. https://doi.org/10.1016/j.paid.2013.10.019

Krämer, N. C., Eimler, S., von der Pütten, A., & Payr, S. (2011). Theory of companions: what can theoretical models contribute to applications and understanding of human-robot interaction? *Applied Artificial Intelligence*, *25*(6), 474–502. https://doi.org/10.1080/08839514.2011.587153

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior, 19*(3), 335–353. https://doi.org/10.1016/s0747-5632(02)00057-2

Krejins, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environment. *Journal of Educational Technology & Society, 5*(1), 8–22. http://www.jstor.org/stable/jeductechsoci.5.1.8

Kucukozer-Cavdar, S., & Taşkaya-Temizel, T. (2016). Analyzing the effects of the personality traits on the success of online collaborative groups. *Procedia - Social and Behavioral Sciences, 228*, 383–389. https://doi.org/10.1016/j.sbspro.2016.07.057

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review, 10*, 133–149. https://doi.org/10.1016/j.edurev.2013.02.002

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia—Social and Behavioral Sciences, 31*, 486–490. https://doi.org/10.1016/j.sbspro.2011.12.091

Lakens, D. (2021). Sample size justification. *Collabra*, *8*(1). https://doi.org/10.1525/collabra.33267

Lambić, D., Lazović, B., Djenić, A., & Marić, M. (2018). A novel metaheuristic approach for collaborative learning group formation. *Journal of Computer Assisted Learning, 34*(6), 907–916. https://doi.org/10.1111/jcal.12299

LaRose, R., & Whitten, P. (2000). Re-thinking instructional immediacy for web courses: A social cognitive exploration. *Communication Education, 49*(4), 320–338. https://doi.org/10.1080/03634520009379221

Latu, I. M., Mast, M. S., Lammers, J., & Bombari, D. (2013). Successful female leaders empower women's behavior in leadership tasks. *Journal of Experimental Social Psychology, 49*(3), 444–448. https://doi.org/10.1016/j.jesp.2013.01.003

LeMonda, B. C., Mahoney, J. R., Verghese, J., & Holtzer, R. (2015). The association between high neuroticism-low extraversion and dual-task performance during walking while talking in non-demented older adults. *Journal of the International Neuropsychological Society*, *21*(7), 519–530. https://doi.org/10.1017/S1355617715000570

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. https://doi.org/10.1016/j.hrmr.2010.10.004

Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. *Annual Review of Psychology*, *41*(1), 585–634. https://doi.org/10.1146/annurev.ps.41.020190.003101

Liang, C., Majumdar, R., & Ogata, H. (2021). Learning log-based automatic group formation: system design and classroom implementation study. *Research and Practice in Technology Enhanced Learning, 16*(14). https://doi.org/10.1186/s41039-021-00156-w

Liao, J., Wang, M., Ran, W., & Yang, S. J. H. (2014). Collaborative cloud: A new model for e-learning. *Innovations in Education and Teaching International, 51*(3), 338–351. https://doi.org/10.1080/14703297.2013.791554

Lim, J., & Newby, T. J. (2020). Preservice teachers' Web 2.0 experiences and perceptions on Web 2.0 as a personal learning environment. *Journal of Computing in Higher Education 32*(2), 234–260. https://doi.org/10.1007/s12528-019-09227-w

Lin, L., Mills, L. A., & Ifenthaler, D. (2016). Collaboration, multi-tasking and problem solving performance in shared virtual spaces. *Journal of Computing in Higher Education 28*(3), 344–357. https://doi.org/10.1007/s12528-016-9117-x

Lin, Y. T., Huang, Y. M., & Cheng, S. C. (2010). An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization. *Computers & Education,* *55*(4), 1483–1493. https://doi.org/10.1016/j.compedu.2010.06.014

Lipnack, J., & Stamps, J. (1999). Virtual teams: The new way to work. *Strategy & Leadership, 27*(1), 14–19. https://doi.org/10.1108/eb054625

Lipponen, L. (2002). Exploring foundations for computer-supported collaborative learning. In G. Stahl (Ed.), *Proceedings of Computer Supported Collaborative Learning*, *CSCL '02*, 72–81. https://doi.org/10.3115/1658616.1658627

Lipponen, L., Rahikainen, M., Hakkarainen, K., & Palonen, T. (2002). Effective participation and discourse through a computer network: Investigating elementary students' computer-supported interaction. *Journal of Educational Computing Research, 27*(4), 353–382. https://doi.org/10.2190/MGTW-QG1E-G66E-F3UD

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction, 13*(5), 487–509. https://doi.org/10.1016/S0959-4752(02)00042-7

Lisak, A., & Erez, M. (2009). Leaders and followers in multi-cultural teams: Their effects on team communication, team identity and team effectiveness. In *Proceedings of IWIC 09: International Workshop on Intercultural Collaboration* (pp. 81–88). https://doi.org/10.1145/1499224.1499238

Liu, M., Zou, W., Shi, Y., Pan Z., & Li, C. (2020). What do participants think of today's MOOCs: an updated look at the benefits and challenges of MOOCs designed for working professionals. *Journal of Computing in Higher Education 32*(3), 307–329. https://doi.org/10.1007/s12528-019-09234-x

Loignon, A. C., Woehr, D. J., Loughry, M. L., & Ohland, M. W. (2018). Elaborating on team-member disagreement: Examining patterned dispersion in team-level constructs. *Group & Organization Management, 44*(1), 165 – 210. https://doi.org/10.1177/1059601118776750

Longua, J., DeHart, T., Tennen, H., & Armeli, S. (2009). Personality moderates the interaction between positive and negative daily events predicting negative affect and stress. *Journal of Research in Personality*, *43*(4), 547–555. https://doi.org/10.1016/j.jrp.2009.02.006

Lou, Y., Abrami, P. C., & D'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449–521. https://doi.org/10.3102/00346543071003449

Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Big five personality traits and work drive as predictors of adolescent academic performance. *Personality and Individual Differences, 35*(6), 1231–1239. https://doi.org/10.1016/S0191-8869(02)00330-6

Loya, A., Gopal, A., Shukla, I., Jermann, P., & Tormey, R. (2015). Conscientious behaviour, flexibility and learning in massive open on-line courses. *Procedia - Social and Behavioral Sciences, 191*, 519–525. https://doi.org/10.1016/j.sbspro.2015.04.686

Lu, L., & Hu, C. H. (2005). Personality, leisure experiences and happiness. *Journal of Happiness Studies*, *6*, 325–342. http://doi.org/10.1007/s10902-005-8628-3

Lykourentzou, I., Antoniou, A., Naudet, Y., & Dow, S. P. (2016). Personality Matters: Balancing for Personality Types Leads to Better Outcomes for Crowd Teams. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 260–273). https://doi.org/10.1145/2818048.2819979

MacDonnell, R., O'Neill, T. A., Kline, T. J., & Hambley, L. A. (2009). Bringing group-level personality to the electronic realm: a comparison of face-to-face and virtual contexts. *The Psychologist-Manager Journal, 12*(1), 1–24. https://doi.org/10.1080/10887150802371773

Magnisalis, I., Demetriadis, S., & Karakostas, A. (2011). Adaptive and intelligent systems for collaborative learning support: A review of the field. *IEEE Transactions on Learning Technologies, 4*(1), 5–20. https://doi.org/10.1109/TLT.2011.2

Magpili, N. C., & Pazos, P. (2018). Self-managing team performance: A systematic review of multi-level input factors. *Small Group Research, 49*(1), 3–33. https://doi.org/10.1177/1046496417710500

Maher, M. L., & Fisher, D. H. (2012). Using AI to evaluate creative designs. In *DS 73–1 Proceedings of the 2nd International Conference on Design Creativity* (vol. 1, pp. 45-54).

Maina, E. M., Oboko, R. O., & Waiganjo, P. W. (2017). Using machine learning techniques to support group formation in an online collaborative learning environment. *International Journal of Intelligent Systems and Applications, 9*(3), 26–33. https://doi.org/10.5815/ijisa.2017.03.04

Makary, M. A., Sexton, J. B., Freischlag, J. A., Holzmueller, C. G., Millman, E. A., Rowen, L., & Pronovost, P. J. (2006). Operating room teamwork among physicians and nurses: Teamwork in the eye of the beholder. *Journal of the American College of Surgeons*, *202*(5), 746–752.

https://doi.org/10.1016/j.jamcollsurg.2006.01.017

Mannix, E., & Neale, M. A. (2016). What Differences Make a Difference? The promise and reality of diverse teams in organizations. *Psychological science in the Public Interest, 6*(2), 31–55. https://doi.org/10.1111/j.1529-1006.2005.00022.x

Manske, S., & Hoppe, H. U. (2017). Managing knowledge diversity: towards automatic semantic group formation. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)* (pp. 330–332). http://doi.org/10.1109/ICALT.2017.132

Maqtary, N., Mohsen, A., & Bechkoum, K. (2019). Group formation techniques in computer-supported collaborative learning: A systematic literature review. Technology, *Knowledge and Learning 24*, 169–190. https://doi.org/10.1007/s10758-017-9332-1

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *The Academy of Management Review, 26*(3), 356–376. https://doi.org/10.2307/259182

Martín, E., & Paredes, P. (2004). Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems. In M. Matera & S. Comai (Eds.), Engineering Advanced Web Applications. In *Proceedings of Workshops in Connection with 4th International Conference on Web Engineering (ICWE 2004)* (pp. 188–197). https://repositorio.uam.es/handle/10486/3680

Mathieu, J. E., Hollenbeck, J. R., Van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the Journal of Applied Psychology. *Journal of Applied Psychology, 102*(3), 452–467. http://dx.doi.org/10.1037/apl0000128

Mattila, M., Wass, H., Söderlund, P., Fredriksson, S., Fadjukoff, P., & Kokko, K. (2011). Personality and turnout: results from the finnish longitudinal studies. *Scandinavian Political Studies*, *34*(4), 287–306. https://doi.org/10.1111/j.1467-9477.2011.00273.x

McCabe, K. O., & Fleeson, W. (2012). What is extraversion for? Integrating trait and motivational perspectives and identifying the purpose of extraversion. *Psychological Science, 23*(12), 1498–1505. https://doi.org/10.1177/0956797612444904

Mccambridge, J., Kypri, K., & Elbourne, D. (2014). Research participation effects: A skeleton in the methodological cupboard. *Journal of Clinical Epidemiology, 67*, 845–849.

https://doi.org/10.1016/j.jclinepi.2014.03.002

McCrae, R. R. & Costa, P. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McMillan, J. H. (2007). Classroom assessment: Principles and practice for effective instruction (4th Ed.). Boston: Allyn and Bacon.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Review of Sociology, 27*, 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085

Mehall, S. (2021). Purposeful interpersonal interaction and the point of diminishing returns for graduate learners. *The Internet and Higher Education, 48*(1), 1007744. https://doi.org/10.1016/j.iheduc.2020.100774

Mehar, R., & Kaur, P. (2020). Effect of online collaborative learning strategy on achievement in economics in relation to self-efficacy. *International Journal of Scientific & Technology Research, (9)*2, 6544–6554.

Mehta, S., & Kulshrestha, A. K. (2014). Implementation of cooperative learning in science: a developmental-cum-experimental Study. *Education Research International, 2014*(2), 1–7. https://doi.org/10.1155/2014/431542

Meindl, J. R. (1995). The romance of leadership as a follower-centric theory: A social constructionist approach. *The Leadership Quarterly, 6*(3), 329–341. https://doi.org/10.1016/1048-9843(95)90012-8

Meyer, J., Lüdtke, O., Schmidt, F. T. C., Fleckenstein, J., Trautwein, U., & Köller, O. (2022). Conscientiousness and cognitive ability as predictors of academic achievement: evidence of synergistic effects from integrative data analysis. *European Journal of Personality, 0*(0). https://doi.org/10.1177/08902070221127065

Michalsky, T., Zion, M., & Mevarech, Z. R. (2007). Developing students' metacognitive awareness in asynchronous learning networks in comparison to face-to-face discussion groups. *Journal of Educational Computing Research, 36*(4), 395–424. https://doi.org/10.2190/320V-8H4W-1123-44R6

Mirbabaie, M., Stieglitz, S., Brünker, F., Hofeditz, L., Ross, B., & Frick, N. R. (2021). Understanding collaboration with virtual assistants–the role of social identity and the extended self. *Business & Information Systems Engineering*, *63*, 21–37. https://doi.org/10.1007/s12599-020-00672-x

Moerbeek, M., van Breukelen, G. J., & Berger, M. P. (2003). A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention Studies. *Journal of clinical epidemiology, 56*(4), 341–350. https://doi.org/10.1016/s0895-4356(03)00007-6

Mohammed, S., & Angell, L. C. (2003). Personality heterogeneity in teams: Which differences make a difference for team performance? *Small Group Research, 34*(6), 651–677. https://doi.org/10.1177/1046496403257228

Mohrman, S. A., Cohen, S. G., & Morhman Jr, A. M. (1995). *Designing team-based organizations: New forms for knowledge work*. Jossey-Bass.

Monllaó Olivé, D., Huynh, D. Q., Reynolds, M., Dougiamas, M., & Wiese, D. (2020). A supervised learning framework: using assessment to identify students at risk of dropping out of a MOOC. *Journal of Computing in Higher Education 32*, 9–26. https://doi.org/10.1007/s12528-019-09230-1

Montoya-Weiss, M. M., Massey, A. P., & Song, M. (2001). Getting it together: Temporal coordination and conflict management in global virtual teams. *Academy of Management Journal, 44*(6), 1251–1262. https://doi.org/10.2307/3069399

Moon, H., Hollenbeck, J. R., Marinova, S., & Humphrey, S. E. (2008). Beneath the surface: Uncovering the relationship between extraversion and organizational citizenship behavior through a facet approach. *International Journal of Selection and Assessment*, *16*(2), 143–154. https://doi.org/10.1111/j.1468-2389.2008.00419.x

Moore, R. L. (2011). The effect of group composition on individual student performance in an introductory economics course. *The Journal of Economic Education, 42*(2), 120–135. https://doi.org/10.1080/00220485.2011.555694

Moreno, J., Ovalle, D. A., & Vicari, R. M. (2012). A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. *Computers and Education, 58*(1), 560–569. https://doi.org/10.1016/j.compedu.2011.09.011

Moynihan, L. M., & Peterson, R. S. (2001). A contingent configuration approach to understanding the role of personality in organizational groups. *Research in Organizational Behavior, 23*(7), 327–378. https://doi.org/10.1016/S0191-3085(01)23008-1

Muchinsky, P. M., & Monahan, C. J. (1987). What is person-environment congruence? Supplementary versus complementary models of fit. *Journal of Vocational Behavior, 31*(3), 268–277. https://doi.org/10.1016/0001-8791(87)90043-1

Mujkanovic, A., & Bollin, A. (2019). Personality-Based group formation. In D. Passey, R. Bottino, C. Lewin & E, Sanchez, E. (Eds.), *Empowering Learners for Life in the Digital Age*. IFIP advances in information and communication technology (vol. 524, pp. 207–217). https://doi.org/10.1007/978-3-030-23513-0_21

Müller, A., Bellhäuser, H., Konert, J., & Röpke, R. (2022). Effects of group formation on student satisfaction and performance: a field experiment. *Small Group Research, 53*(2), 244–273. https://doi.org/10.1177/10464964209885

Murphy, P. K., Greene, J. A., Firetto, C. M., Li, M., Lobczowski, N. G., Duke, R. F., Wei, L., & Croninger, R. M. V. (2017). Exploring the influence of homogeneous versus heterogeneous grouping on students' text-based discussions and comprehension. *Contemporary Educational Psychology, 51*, 336–355. https://doi.org/10.1016/j.cedpsych.2017.09.003

Nami, F., & Vaezi, S. (2018). How ready are our students for technology-enhanced learning? Students at a university of technology respond. *Journal of Computing in Higher Education 30*, 510–529. https://doi.org/10.1007/s12528-018-9181-5

Nederveen Pieterse, A., Van Knippenberg, D., & van Ginkel, W. P. (2011). Diversity in goal orientation, team reflexivity, and team performance. *Organizational Behavior and Human Decision Processes, 114*(2), 153–164. https://doi.org/10.1016/j.obhdp.2010.11.003

Nesher Shoshan, H., & Wehrt, W. (2021). Understanding "Zoom fatigue": A mixed-method approach. *Applied Psychology*, *71*(3), 827-852. https://doi.org/10.1111/apps.12360

Neuman, G. A., Wagner, S. H., & Christiansen, N. D. (1999). The Relationship between work-team personality composition and the job performance of teams. *Group & Organization Management,* *24(*1), 28–45. https://doi.org/10.1177 /1059601199241003

Nijstad, B. A., & De Dreu, C. K. W. (2002). Creativity and group innovation. *Applied Psychology, 51*(3), 400–406. https://doi.org/10.1111/1464-0597.00984

Noël, P. H., Lanham, H. J., Palmer, R. F., Leykum, L. K., & Parchman, M. L. (2013). The importance of relational coordination and reciprocal learning for chronic illness care within primary care teams. *Health care management review, 38*(1), 20–28. https://doi.org/10.1097/HMR.0b013e3182497262

Nonaka, I., Hirose, A., Takeda, Y., Friedrich, T. L., Vessey, W. B., Schuelke, M. J., Ruark, G. A., Mumford, M. D., D'innocenzo, L., Mathieu, J. E., Kukenberger, M. R., Wang, D., Waldman, D., Zhang, Z., Banks, G. C., McCauley, K. D., Gardner, W. L., Guler, C. E., Brower, H. H., … Sawyer, K. B. (2016). What type of leadership behaviors are functional in teams? A meta-analysis. *Leadership Quarterly, 17*(3), 288–307. https://doi.org/10.1016/j.leaqua.2006.02.007

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*(5), 971–990. https://doi.org/10.1016/j.paid.2007.03.017

Odo, C., Masthoff, J., & Beacham, N. (2019). Group formation for collaborative learning. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Artificial Intelligence in Education. AIED 2019* (Vol. 2, pp. 206–212). https://doi.org/10.1007/978-3-030-23207-8_39

OECD (2016). Innovating education and educating for innovation: The power of digital technologies and skills. *OECD Publishing*. http://dx.doi.org/10.1787/9789264265097-en

Oh, H., Chung, M. H., & Labianca, G. (2004). Group social capital and group effectiveness: The role of informal socializing ties. *Academy of Management Journal, 47*(6), 860–875. https://doi.org/10.5465/20159627

Okdie, B. M., Guadagno, R. E., Bernieri, F. J., Geers, A. L., & McLarney-Vesotski, A. R. (2011). Getting to know you: Face-to-face versus online interactions. *Computers in Human Behavior, 27*(1), 153–159. https://doi.org/10.1016/j.chb.2010.07.017

Ortiz de Guinea, A. (2012). A meta-analysis of the consequences of virtualness on team functioning. *Information and Management 49*(6), 301–308. https://doi.org/10.1016/j.im.2012.08.003

Pai, H. H., Sears, D. A., & Maeda, Y. (2014). Effects of small-group learning on transfer: A meta-analysis.

Educational Psychology Review, 27(1), 79–102. https://doi.org/10.1007/s10648-014-9260-8

Palloff, R. M., & Pratt, K. (2005). Collaborating online: Learning together in community. *Personnel Psychology, 59*(4), 980–982. https://doi.org/10.1111/j.1744-6570.2006.00060_9.x

Parmelee, D. X., & Hudes, P. (2012). Team-based learning: a relevant strategy in health professionals' education. *Medical Teacher, 34*(5), 411–413. https://doi.org/10.3109/0142159X.2012.643267

Patrício, L., & Franco, M. (2022). A systematic literature review about team diversity and team performance: Future lines of investigation. *Administrative Sciences, 12*(1), 31. https://doi.org/10.3390/admsci12010031

Peeters, M. A. G., Van Tuijl, H. F. J. M., Rutte, C. G., & Reymen, I. M. M. J. (2006). Personality and group performance: A meta-analysis. *European Journal of Personality, 20*(5), 377–396. https://doi.org/10.1002/per.588

Perner, J. (1999). Theory of mind. In M. Bennett, *Developmental psychology: Achievements and prospects* (pp. 205–230). Psychology Press.

Perry, J. L. (2021). Work team diversity: Refocusing through the lens of team power and status. *Social and Personality Psychology Compass, 15*(12), [e12646]. https://doi.org/10.1111/spc3.12646

Pfaff, E., & Huddleston, P. (2003). Does it matter if I hate teamwork? What impacts student attitudes toward teamwork. *Journal of Marketing Education, 25*(1), 37–45. https://doi.org/10.1177/0273475302250571

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2021). nlme: Linear and nonlinear mixed effects models [Computer software]. https://CRAN.R-project.org/package=nlme

Poropat A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin, 135*(2), 322–338. https://doi.org/10.1037/a0014996

Pollet, T. V., Roberts, S. G., & Dunbar, R. I. (2011). Extraverts have larger social network layers. *Journal of Individual Differences, 32*(3), 161–169. https://doi.org/10.1027/1614-0001/a000048

Porter, C. O. L. H., Hollenbeck, J. R., Ilgen, D. R., Ellis, A. P. J., West, B. J., & Moon, H. (2003). Backing up behaviors in groups: The role of personality and legitimacy of need. *Journal of Applied Psychology, 88(*3), 391–403. https://doi.org/10.1037/0021-9010.88.3.391

Powell, A., Piccoli, G., & Ives, B. (2004). Virtual teams: A review of current literature and directions for future research. *Data Base for Advances in Information Systems, 35*(1), 6–36. https://doi.org/10.1145/968464.968467

Power, R. A., & Pluess, M. (2015). Heritability estimates of the big five personality traits based on common genetic variants. *Translational Psychiatry, 5*(7), e604. https://doi.org/10.1038/tp.2015.96

Prewett, M. S., Brown, M. I., Goswami, A., & Christiansen, N. D. (2018). Effects of Team Personality Composition on Member Performance: A Multilevel Perspective. *Group & Organization Management, 43*(2), 316–348. https://doi.org/10.1177/1059601116668633

Prewett, M. S., Walvoord, A. A. G., Stilson, F. R. B., Rossi, M. E., & Brannick, M. T. (2009). The team personality–team performance relationship revisited: the impact of criterion choice, pattern of workflow, and method of aggregation. *Human Performance, 22*(4), 273–296. https://doi.org/10.1080/08959280903120253

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223–232.https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Protsch, P., & Dieckhoff, M. (2011). What matters in the transition from school to vocational training in Germany: Educational credentials, cognitive abilities or personality? *European Societies, 13*(1), 60–91. https://doi.org/10.1080/14616696.2010.540352

Putro, B. N., Pratama, H. G., Prasetyo, W., & Doewes, R. I. (2020). E-learning implementation in physical education department in higher education during COVID-19 pandemic. *Information Technology, Education and Society, 17*(2), 67–80. https://doi.org/10.7459/ites/17.2.05

Putwain, D. W., Becker, S., Symes, W., & Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction, 54*, 73–81. https://doi.org/10.1016/j.learninstruc.2017.08.004

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Rammstedt, B., & Danner, D. (2017). Die Facettenstruktur des Big Five Inventory (BFI): Validierung für die deutsche Adaptation des BFI [The facet structure of the Big Five Inventory (BFI): Validation for the German adaptation of the BFI]. *Diagnostica, 63*(1), 70–84. https://doi.org/10.1026/0012-1924/a000161

Rammstedt, B., & John, O. P. (2005). Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality. *Diagnostica, 51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Ramos-Villagrasa, P. J., Marques-Quinteiro, P., Navarro, J., & Rico, R. (2018). Teams as complex adaptive systems: Reviewing 17 years of research. *Small Group Research, 49*(2), 135–176. https://doi.org/10.1177/1046496417713849

Rapp, T. L., Bachrach, D. G., Rapp, A. A., & Mullins, R. (2014). The role of team goal monitoring in the curvilinear relationship between team efficacy and team performance. *Journal of Applied Psychology, 99*(5), 976–987. https://doi.org/10.1037/a0036978

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review, 19*(1), 65–83. https://doi.org/10.1007/s10648-007-9042-7

Revelo-Sanchez, O., Collazos, C. A., & Redondo, M. A. (2020). Collaborative learning group formation based on personality traits: an empirical study in initial programming courses. In V. Agredo-Delgado, P. H. Ruiz, & K. O. Villalba-Condori (Eds.), Human-Computer Interaction. *Communications in Computer and Information Science* (vol. 1334, pp. 73–84). https://doi.org/10.1007/978-3-030-66919-5_8

Revelo-Sanchez, O., Collazos, C. A., Redondo, M. A., & Pinto, I. I. (2021). Homogeneous group formation in collaborative learning scenarios: An approach based on personality traits and genetic algorithms. *IEEE Transactions on Learning Technologies, (14)*4, 486–499. https://doi.org/10.1109/TLT.2021.3105008

Riazy, S., Simbeck, K., Traeger, M., & Woestenfeld, R. (2021). The effect of prior knowledge on persistence, participation and success in a mathematical MOOC. In H. C. Lane, S. Zvacek, & J. Uhomoibhi, (Eds.), *Computer Supported Education, CSEDU 2020*. https://doi.org/10.1007/978-3-030-86439-2_21

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387.

https://doi.org/10.1037/a0026838

Rienties, B., & Héliot, Y. (2018). Enhancing (in) formal learning ties in interdisciplinary management courses: A quasi-experimental social network study. *Studies in Higher Education*, *43*(3), 437–451. https://doi.org/10.1080/03075079.2016.1174986

Rispens, S., Jehn, K. A., & Steinel, W. (2021). Conflict management style asymmetry in short-term project groups. *Small Group Research, 52*(2), 220–242. https://doi.org/10.1177/1046496419894637

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. https://doi.org/10.1016/j.hrmr.2009.09.002

Röpke, R., Konert, J., Gallwas, E. & Bellhäuser, H. (2016). MoodlePeers: Automatisierte Lerngruppenbildung auf Grundlage psychologischer Merkmalsausprägungen in E-Learning-Systemen. In U. Lucke, A. Schwill & R. Zender (Eds.), *14. E-Learning Fachtagung Informatik DeLFI* (pp. 233–244). https://dblp.uni-trier.de/db/conf/delfi/delfi2016.html#RopkeKGB16

Rogat, T. K., Hmelo-Silver, C. E., Cheng, B. H., Traynor, A., Adeoye, T. F., Gomoll, A., & Downing, B. K. (2022). A multidimensional framework of collaborative groups' disciplinary engagement. *Frontline Learning Research, 10*(2), 1–21. https://doi.org/10.14786/flr.v10i2.863

Roney, R., Greenaway, K, Anicich, E. M., & Galinsky, A. D. (2012). The path to glory is paved with hierarchy: When hierarchical differentiation increases group effectiveness. *Psychological Science, 23*(6), 669–677. https://doi.org/10.1177/0956797611433876

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human resource management review*, *16*(2), 155–180. https://doi.org/10.1016/j.hrmr.2006.03.004

Rourke, L., & Anderson, T. (2002). Using peer teams to lead online discussions. *Journal of Interactive Media in Education, 12*(2), 122–137. http://doi.org/10.5334/2002-1

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14*(2), 201–241. https://doi.org/10.1207/s15327809jls1402_2

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal, 38*(2), 437–460. https://doi.org/10.3102/00028312038002437

Saalbach, H., & Schalk, L. (2011). Preschoolers' novel noun extensions: Shape in spite of knowing better. *Frontiers in Psychology, 2*, 317. https://doi.org/10.3389/fpsyg.2011.00317

Saleh, M., Lazonder, A., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science, 33*(2), 105–119. https://doi.org/10.1007/s11251-004-6405-z

Salovaara, H. (2005). An exploration of students' strategy use in inquiry-based computer-supported collaborative learning. *Journal of Computer Assisted Learning, 21*(1), 39–52. https://doi.org/10.1111/j.1365-2729.2005.00112.x

Saqr, M., Nouri, J., & Jormanainen, I. (2019). A learning analytics study of the effect of group size on social dynamics and performance in online collaborative learning. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), Transforming Learning with Meaningful Technologies. EC-TEL 2019. *Lecture Notes in Computer Science* (vol. 11722, pp. 466–479). Springer, Cham. https://doi.org/10.1007/978-3-030-29736-7_35

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology, 76*(4), 613–627. https://doi.org/10.1037/0022-3514.76.4.613

Schmutte, P. S., & Ryff, C. D. (1997). Personality and well-being: Reexamining methods and meanings. *Journal of Personality and Social Psychology, 73*(3), 549–559. http://dx.doi.org/10.1037/0022-3514.73.3.549

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological bulletin, 143*(6), 565–600. https://doi.org/10.1037/bul0000098

Seeber, I., Bittner, E., Briggs, R. O., DeVreede, G. J., DeVreede, T., Druckenmiller, D., Maier, R., Merz, A.B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2018). Machines as teammates: a collaboration research agenda. In *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS).* http://hdl.handle.net/10125/49942

Seong, J. Y., & Hong, D. S. (2020). Collective personality fit and diversity – how effective are they in predicting relationship conflict? *International Journal of Conflict Management, 31*(2), 201–219. https://doi.org/10.1108/IJCMA-03-2019-0055

Seong, J. Y., Kristof-Brown, A. L., Park, W.-W., Hong, D.-S., & Shin, Y. (2015). Person-Group Fit: Diversity Antecedents, Proximal Outcomes, and Performance at the Group-level. *Journal of Management, 41*(4), 1184–1213. https://doi.org/10.1177/0149206312453738

Shapiro, H. B., Lee, C. H., Wyman Roth, N. E., Li, K., Çetinkaya-Rundel, M., & Canelas, D. A. (2017). Understanding the massive open online course (MOOC) student experience: An examination of attitudes, motivations, and barriers. *Computers and Education, 110*, 35–50. https://doi.org/10.1016/j.compedu.2017.03.003

Shaw, R. S. (2013). The relationships among group size, participation, and performance of programming language learning supported with online forums. *Computers and Education, 62*, 196–207. https://doi.org/10.1016/j.compedu.2012.11.001

Shemla, M., Meyer, B., Greer, L., & Jehn, K. (2016). A review of perceived diversity in teams: Does how members perceive their team's composition affect team processes and outcomes? *Journal of Organizational Behavior, 37*(1), 89–106. https://doi.org/10.1002/job.1957

Shibley, I. A., & Zimmaro, D. M. (2002). The influence of collaborative learning on student attitudes and performance in an introductory chemistry laboratory. *Journal of Chemical Education, 79(6),* 745–748. https://doi.org/10.1021/ed079p745

Sibley, J., & Parmelee, D. X. (2008). Knowledge is no longer enough: enhancing professional education with team-based learning. *New Directions for Teaching and Learning, 116*, 41–53. https://doi.org/10.1002/tl.332

Sidorenkov A.V., Borokhovski E. F., & Kovalenko V. A. (2018). Group size and composition of work groups as precursors of intragroup conflicts. *Psychology Research Behavior Management, 11*, 511–523. https://doi.org/10.2147/PRBM.S178761

Siemon, D. (2022). Elaborating team roles for artificial intelligence-based teammates in Human-AI collaboration. *Group Decision and Negotiation, 31*, 871–912. https://doi.org/10.1007/s10726-022-09792-z

Slavin, R. E. (1987). Cooperative learning and cooperative school. *Educational Leadership, 45,* 7–13.

Smith, K. A., Sheppard, S. D., Johnson, D. W., Johnson, R. T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education, 94*(1), 1–15. https://doi.org/10.1002/j.2168-9830.2005.tb00831.x

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta analysis. *Review of Educational Research, 69*(1), 21–51. https://doi.org/10.3102/00346543069001021

Srba, I., & Bielikova, M. (2015). Dynamic group formation as an approach to collaborative learning support. *IEEE Transactions on Learning Technologies, 8*(2), 173–186. https://doi.org/10.1109/TLT.2014.2373374

Staples, D. S., & Zhao, L. (2006). The effects of cultural diversity in virtual teams versus face-to-face teams. *Group Decision and Negotiation, 15*, 389–406. https://doi.org/10.1007/s10726-006-9042-x

Steiner, I. D. (1972). Group process and productivity. New York: Academic Press.

Stewart, G. L. (1999). Trait bandwidth and stages of job performance: Assessing differential effects for conscientiousness and its subtraits. *Journal of Applied Psychology, 84*(6), 959–968. https://doi.org/10.1037/0021-9010.84.6.959

Stewart, G. L., Fulmer, I. S., & Barrick, M. R. (2005). An exploration of member roles as a multilevel linking mechanism for individual traits and team outcomes. *Personnel Psychology, 58*(2), 343–365. https://doi.org/10.1111/j.1744-6570.2005.00480.x

Stipelman, B. A., Rice, E. L., Vogel, A. L., & Hall, K. L. (2019). The role of team personality in team effectiveness and performance. In K. Hall, A. Vogel & R. Croyle (Eds.), *Strategies for Team Science Success* (pp. 189–196). https://doi.org/10.1007/978-3-030-20992-6_14

Straub, R. (1997). Students' reactions to teacher comments: an exploratory study comments: an exploratory study. Source: *Research in the Teaching of English, 31*(1), 91–119. https://doi.org/10.2307/40171265

Strijbos, J. W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction, 17*(4), 389–393. https://doi.org/10.1016/j.learninstruc.2007.03.004

Strijbos, J. W., Martens, R. L., & Jochems, W. M. G. (2004). Designing for interaction: Six steps to designing computer-supported group-based learning. *Computers and Education, 42*(4), 403–424. https://doi.org/10.1016/j.compedu.2003.10.004

Strømsø, H. I., Grøttum, P., & Lycke, K. H. (2007). Content and processes in problem-based learning: a comparison of computer-mediated and face-to-face communication. *Journal of Computer Assisted Learning, 23*(3), 271–282. https://doi.org/10.1111/j.1365-2729.2007.00221.x

Stürmer, S., Benbow, A. E. F., Siem, B., Barth, M., Bodansky, A. N., & Lotz-Schmitt, K. (2013). Psychological foundations of xenophilia: The role of major personality traits in predicting favorable attitudes toward cross-cultural contact and exploration. *Journal of Personality and Social Psychology, 105*(5), 832–851. https://doi.org/10.1037/a0033488

Sun, G., & Shen, J. (2013). Teamwork as a service: a cloud-based system for enhancing teamwork performance in mobile learning. In *Proceedings of 2013 IEEE 13th International Conference on Advanced Learning Technologies* (pp. 376–378). http://dx.doi.org/10.1109/ICALT.2013.115

Taggar, S., Hackett, R., & Saha, S. (2006). Leadership emergence in autonomous work groups: Antecedents and outcomes. *Personnel Psychology, 52*(4), 899–926. https://doi.org/10.1111/j.1744-6570.1999.tb00184.x

Tajfel, H., & Turner, J. C. (1986). *The social identity theory of intergroup behavior*. In S. Worchel, & W. G. Austin (Eds.), Psychology of intergroup relations (pp. 7–24). Nelson-Hall.

Tekleap, A. G., & Quigley, N. R. (2014). Team deep-level diversity, relationship conflict, and team members' affective reactions: A cross-level investigation. *Journal of Business Research, 67*(3), 394–402. https://doi.org/10.1016/j.jbusres.2012.12.022

Tett, R. P., Toich, M. J., & Ozkum, S. B. (2021). Trait activation theory: A Review of the literature and applications to five lines of personality dynamics research. *Social Science Research Network, 8*, 199–233. http://dx.doi.org/10.1146/annurev-orgpsych-012420-062228

Thanh, P. T. H., & Gillies, R. (2010). Group composition of cooperative learning: does heterogeneous grouping work in asian classrooms? *International Education Studies, 3(*3), 12–19. https://doi.org/10.5539/ies.v3n3p12

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*(3), 558.

http://doi.org/10.1037//0022-3514.84.3.558

Triana, M. D., Kim, K., Byun, S., Delgado, D. M., & Arthur, W. (2020). The relationship between team deep-level diversity and team performance: a meta-analysis of the main effect, moderators, and mediating mechanisms. *Journal of Management Studies, 58(*8), 2137–2179. https://doi.org/10.1111/joms.12670

Tsovaltzi, D., Weinberger, A., Schmitt, L., Bellhäuser, H., Müller, A., Konert, J., Röpke, R., Gijlers, H., Eshuis, E., Eysink, T., Erkens, M., Manske, S., Hoppe, H. U., Bodemer, D., Sankaranarayanan, S., Dashti, C., Bogart, C., Wang, X., Sakr, M., … van Aalst, J. (2019). Group formation in the digital age: Relevant characteristics, their diagnosis, and combination for productive collaboration. In *Computer-Supported Collaborative Learning Conference, CSCL,* 2.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63*(6), 384–399. https://doi.org/10.1037/h0022100

Turban, D. B., Stevens, C. K., & Lee, F. K. (2009). Effects of conscientiousness and extraversion on new labor market entrants' job search: The mediating role of metacognitive activities and positive emotions. *Personnel Psychology, 62(3), 553--573*. https://doi.org/10.1111/j.1744-6570.2009.01148.x

Van Dijk, H., Meyer, B., van Engen, M. L., & Lewin Loyd, D. (2017). Microdynamics in diverse groups: A review and integration of the diversity and stereotyping literatures. *Academy of Management Annals, 11*(1), 517–557. https://doi.org/10.5465/annals.2014.0046

Van Emmerik, I. J. H., & Brenninkmeijer, V. (2009). Deep-level similarity and group social capital: associations with team functioning. *Small Group Research, 40*(6), 650–669. https://doi.org/10.1177/1046496409346550

Van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology, 58(*1), 515–541. http://dx.doi.org/10.1146/annurev.psych.58.110405.085546

Van Landeghem, G., De Fraine, B., & Van Damme, J. (2005). The consequence of ignoring a level of nesting in multilevel analysis: A comment*. Multivariate Behavioral Research, 40*(4), 423–434. https://doi.org/10.1207/s15327906mbr4004_2

Van Merriënboer, J. J. G., & Kirschner, P. A. (2001). Three worlds of instructional design: State of the art and future directions. *Instructional Science, 29*(4), 429–441. https://doi.org/10.1023/A:1011904127543

Varela, O. E., Cater, J. J., & Michel, N. (2012). Online learning in management education: An empirical Study of the role of personality traits. *Journal of Computing in Higher Education, 24*, 209–225. http://dx.doi.org/10.1007/s12528-012-9059-x

Verbeek, M., & Nijman, T. (1996). Incomplete Panels and Selection Bias. In L. Mátyás, & P. Sevestre (Eds.), The Econometrics of Panel Data. *Advanced Studies in Theoretical and Applied Econometrics (vol. 33).* https://doi.org/10.1007/978-94-009-0137-7_18

Virk, A., Joshi, A., Mahajan, R., & Singh, T. (2020). The power of subjectivity in competency-based assessment. *Journal of Postgraduate Medicine, 66*(4), 200–205. https://doi.org/10.4103/jpgm.JPGM_591_20

Voltmer, J. B., Reich-Stiebert, N., Raimann, J., & Stürmer, S. (2022). The role of multi-attributional student diversity in computer-supported collaborative learning. *The Internet and Higher Education, 55*, 100868. https://doi.org/10.1016/j.iheduc.2022.100868

Voorn, R. J. J., & Kommers, P. A. M. (2013). Social media and higher education: Introversion and collaborative learning from the student's perspective. *International Journal of Social Media and Interactive Learning Environments, 1(*1), 59–73. https://doi.org/10.1504/IJSMILE.2013.051650

Vygotsky, L. S., & Cole, M. (1978). Mind in society: Development of higher psychological processes. Harvard university press. https://doi.org/10.2307/j.ctvjf9vz4

Walker, A. (2007). Group work in higher education: are introverted students disadvantaged? *Psychology Learning & Teaching, 6*(1), 20–25. https://doi.org/10.2304/plat.2007.6.1.20

Wang, D. Y., Lin, S. S., & Sun, C. T. (2007). DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior, 23*(4), 1997–2010. https://doi.org/10.1016/j.chb.2006.02.008

Wang, M. L., & Hsu, B. F. (2012). A study to explore the team virtualization level and team effectiveness from the team personality composition. *Journal of the Knowledge Economy, 3*(2), 199–216. https://doi.org/10.1007/s13132-011-0079-y

Wang, X., Michelle Hinn, D., & Kanfer, A. G. (2001). Potential of computer-supported collaborative learning for learners with different learning styles. *Journal of Research on Technology in Education, 34*(1), 75–85. https://doi.org/10.1080/15391523.2001.10782335

Wang, Y.-M., & Elhag, T. (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. *Expert Systems with Applications 31*(2), 309–319. https://doi.org/10.1016/j.eswa.2005.09.040

Wax, A., DeChurch, L. A., & Contractor, N. S. (2017). Self-organizing into winning groups: understanding the mechanisms that drive successful collaborations. *Small Group Research, 48(*6), 665–718. https://doi.org/10.1177/1046496417724209

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction Volume, 17*(4), 416–426. https://doi.org/10.1016/j.learninstruc.2007.03.007

Wessner, M., & Pfister, H. (2001). Group formation in computer-supported collaborative learning. In *ACM International Conference on Supporting Group Work* (pp. 24–31). https://doi.org/10.1145/500286.500293

Wildman, J. L., Nguyen, D. M., Duong, N. S., & Warren, C. (2021). Student teamwork during COVID-19: challenges, changes, and consequences. *Small Group Research, 52(2)* 119–134. https://doi.org/10.1177/1046496420985185

Wilkinson, I. A. G., & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research, 37*(5), 425–447. https://doi.org/10.1016/s0883-0355(03)00014-4

Williams, B., Myerson, J., & Hale, S. (2008). Individual differences, intelligence, and behavior analysis. *Journal of the Experimental Analysis of Behavior, 90*(2), 219–231. https://doi.org/10.1901/jeab.2008.90-219

Williams, E. A., & Castro, S. L. (2010). The effects of teamwork on individual learning and perceptions of team performance: A comparison of face-to-face and online project settings. *Team Performance Management, 16*(3/4), 124–147. https://doi.org/10.1108/13527591011053232

Williams, E. A., Duray, R., & Reddy, V. (2006). Teamwork orientation, group cohesiveness, and student learning: A Study of the use of teams in online distance education. *Journal of Management Education, 30(*4), 592–616. https://doi.org/10.1177/1052562905276740

Wilmot, M. P., Wanberg, C. R., Kammeyer-Mueller, J. D., & Ones, D. S. (2019). Extraversion advantages at work: A quantitative review and synthesis of the meta-analytic evidence. *Journal of Applied Psychology, 104(*12), 1447–1470. https://doi.org/10.1037/apl0000415

Wilson, K. S., DeRue, D. S., Matta, F. K., Howe, M., & Conlon, D. E. (2016). Personality similarity in negotiations: Testing the dyadic effects of similarity in interpersonal traits and the use of emotional displays on negotiation outcomes. *Journal of Applied Psychology, 101(*10), 1405–1421. https://doi.org/10.1037/apl0000132

Wilson, J. M., Fletcher, T. D., Pescosolido, T., & Major, D. A. (2021). Extraversion and leadership emergence: differences in virtual and face-to-face teams. *Small Group Research, 52*(5), 1–30. https://doi.org/10.1177/1046496420986620

Wilt, J., Noftle, E. E., Fleeson, W., & Spain, J. S. (2012). The dynamic role of personality states in mediating the relationship between extraversion and positive affect. *Journal of Personality*, *80*(5), 1205–1236. https://doi.org/10.1111/j.1467-6494.2011.00756.x

Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology, 87*(1), 164 –169. https://doi.org/10.1037/0021-9010.87.1.164

Wittenbaum, G. M., Hollingshead, A. B., Paulus, P. B., Hirokawa, R. Y., Ancona, D. G., Peterson, R. S., Jehn, K. A., & Yoon, K. (2004). The functional perspective as a lens for understanding groups. *Small Group Research, 35*(1), 17–43. https://doi.org/10.1177/1046496403259459

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*(6004), 686–688. https://doi.org/10.1126/science.1193147

Woolley, A. W., Chow, R. M., Mayo, A. T., Riedl, C., & Chang, J. W. (2022). Collective attention and collective intelligence: the role of hierarchy and team gender composition. *Organization Science, 34*(3), 1315–1331. https://doi.org/10.1287/orsc.2022.1602

Woolley, A., & Malone, T. (2011). What makes a team smarter? More women. *Harvard Business Review*, *89*(6), 32-33. https://www.ncbi.nlm.nih.gov/pubmed/21714385

Wu, Y. W., & Wooldridge, P. J. (2005). The impact of centering first-level predictors on individual and contextual effects in multilevel data analysis. *Nursing Research, 54*(3), 212–216. https://doi.org/10.1097/00006199-200505000-00009

Yang, D., Sinha, T., Adamson, D., & Rosé, C. P. (2013). Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *Proceedings of the NIPS Data-driven education workshop* (vol. 11, pp. 1-8).

Yannibelli, V., & Amandi, A. (2011). Forming well-balanced collaborative learning teams according to the roles of their members: An evolutionary approach. In *Proceedings of the IEEE 12th international symposium on computational intelligence and informatics (CINTI)* (pp. 265–270). http://dx.doi.org/10.1109/CINTI.2011.6108511

Yazici, H. J. (2005). A Study of collaborative learning style and team learning performance. *Education & Training, 47*(3), 216–229. https://doi.org/10.1108/00400910510592257

Yellen, R., Winniford, M., & Sanford, C. (1995). Extraversion and introversion in electronically-supported meetings. *Information Manager, (28)*1, 63–74. https://doi.org/10.1016/0378-7206(94)00023-C

Yin, P. Y., Chang, K. C., Hwang, G. J., Hwang, G. H., & Chan, Y. (2006). A particle swarm optimization approach to composing serial test sheets for multiple assessment criteria. *Journal of Educational Technology & Society, 9*(3), 3–15. https://www.jstor.org/stable/jeductechsoci.9.3.3

Yücel, Ü. A., & Usluel, Y. K. (2016). Knowledge building and the quantity, content and quality of the interaction and participation of students in an online collaborative learning environment. *Computers and Education*, *97*, 31–48. https://doi.org/10.1016/j.compedu.2016.02.015

Yulianto, B., Prabowo, H., Kosala, R., & Hapsara, M. (2018). Implementation of learning analytics in MOOC by using artificial intelligence. *Journal of Computer Science*, *14*(3), 317–323. https://doi.org/10.3844/jcssp.2018.317.323

Zaccaro, S. J., Dubrow, S., & Kolze, M. (2018). *Leader traits and attributes*. In J. Antonakis, & D. V. Day (Eds.), The nature of leadership (pp. 29–55). https://doi.org/10.4135/9781506395029.n2

Zajenkowski, M., & Szymaniak, K. (2021). Narcissism between facets and domains. The relationships between two types of narcissism and aspects of the big Five. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 40*(5), 2112–2121. https://doi.org/10.1007/s12144-019-0147-1

Zell, E., & Lesick, T. L. (2022). Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of personality, 90*(4), 559–573. https://doi.org/10.1111/jopy.12683

Zennouche, M., Zhang, J., & Wang, B. (2014). Factors influencing innovation at individual, group and organisational levels: a content analysis. *International Journal of Information Systems and Change Management, 7(*1), 23–42. https://doi.org/10.1504/IJISCM.2014.065052

Zhang, X., Cheng, G., & Xu, B. (2020). The influence of group big-five personality composition on student engagement in online discussion. *International Journal of Information and Education Technology, 10*(10), 744–750. https://doi.org/10.18178/ijiet.2020.10.10.1452

Zheng, Y., Li, C., Liu, S., & Lu, W. (2018). An improved genetic approach for composing optimal collaborative learning groups. *Knowledge-Based Systems, 139*, 214–225. https://doi.org/10.1016/j.knosys.2017.10.022

Zheng, Z., & Pinkwart, N. (2014). A discrete particle swarm optimization approach to compose heterogeneous learning groups. In *Proceedings of the 2014 IEEE Computer Society 14th International Conference on Advanced Learning Technologies* (pp. 49–51). https://doi.org/10.1109/ICALT.2014.24