Does International Status Affect Competence Ratings in Newly Formed Multinational Teams? The Role of Psychological Safety Growth and Verbal Behavior

Samuel Farley[1] , **Jeremy Dawson[2], Thomas Greenaway[2], Keelan Meade[2], and Daría Hernández Ibar[2]**

Abstract

National status has been found to influence how people are perceived in multinational teams. Team members from an international background are often perceived as less competent than those from the local context. Studies mainly focus on language differences to explain this phenomenon, but in this study, we offer a different theoretical explanation. We propose that national status can affect psychological safety and its development within teams, which in turn affects verbal behavior and competence ratings. To test this notion, we examine differences in psychological safety growth, verbal behavior and competence ratings among home country nationals based in the United Kingdom (UK) and international members of newly formed multinational teams. In a sample of 519 team members (101 teams), results showed that internationals, compared to home country nationals, have lower initial psychological safety, as well as slower development in psychological safety over time. Furthermore, the relationship between national status and competence ratings was partially mediated by psychological safety growth and verbal behavior. These results were fully replicated on a separate sample of 538 team members (90 teams) in a second study using an identical research design. However, exploratory analyses indicated that the pattern of findings were not consistent across team members from Africa, Asia, and Europe. The psychological safety of home nationals only started and grew more quickly than that of Asians, while only African and Asian team members spoke less and were rated as less competent. Together these results have implications for managers of newly formed multinational teams.

In this study we seek to understand how national status influences competence ratings in newly formed multinational teams. Multinational teams tend to adopt a common language, typically

[1]University of Leeds, UK
[2]University of Sheffield, UK

*Thomas Greenaway is now affiliated to University of Warwick, UK; Keelan Meade is now affiliated to Digital Poverty Alliance, UK.
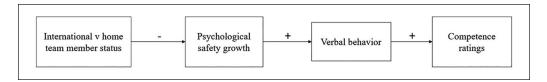
Corresponding Author:
Samuel Farley, Leeds University Business School, Maurice Keyworth Building, Leeds LS2 9JT, UK.
Email: s.j.farley@leeds.ac.uk

Figure 1. Hypothesized model.

English (Paunova, 2017). As a result, most studies addressing competence within multinational environments have focussed on features of language ability, such as proficiency, fluency, and accent (Li et al., 2019; Neeley, 2013; Russo et al., 2017).

Although language plays an important role in determining competence ratings, the influence of psychological safety is understudied in this process. We therefore examine how national status affects psychological safety, defined as the extent to which people feel safe to engage in interpersonal risk-taking, and speak up with relevant ideas, questions, or concerns (Edmondson, 1999). Psychological safety captures team members perceptions about the interpersonal team climate (Edmondson & Lei, 2014). It is important within multinational teams because international team members can offer unique perspectives on entering new markets or streamlining operations in a particular cultural context, but only when they feel safe speaking up (Tröster & van Knippenberg, 2012). Yet, we suggest that in the early phases of multinational team formation, international team members experience slower development of psychological safety compared to home national team members. In turn, we argue that this causes internationals to speak less than home nationals, which reduces how competently they are perceived (see Figure 1). We test our predictions across two studies that use the same research design with two different samples of undergraduate student teams.

Our study seeks to extend understanding on multinational teams and psychological safety in three main ways. First, we present a novel account of the mechanisms that influence the relationship between national status and competence ratings in multinational teams. Prior studies have examined how status develops within multinational teams by examining such factors as language proficiency (Li et al., 2019; Neeley, 2013), core-self evaluations (Paunova, 2017), advice centrality (Salk & Brannen, 2000), and power sources, including hierarchical position and expertise (Tenzer & Pudelko, 2017). In contrast, we use insights from expectation states theory (EST; Berger et al., 1980) to offer an alternative perspective on how competence ratings develop, which highlights the mechanistic roles of psychological safety and verbal behavior.

Second, we extend research on the individual differences antecedents of psychological safety by examining how it is influenced by national status. Little is known about how psychological safety is affected by cultural differences (Woodson, 2020), but it is crucial to develop a more culturally nuanced perspective on how it unfolds, as different development trajectories may suppress the potential of those from minority backgrounds. Indeed, Woodson (2020) states "*We cannot designate a space or performance as conducive to psychological safety without attention to cultural nuance within the strategies used to communicate vulnerability, resolve conflict, or resist domination*" (p. 576). Previously, Ulusoy et al. (2016) found that Turkish employees working in Germany experienced lower psychological safety than German nationals. However, this research examined psychological safety across workplaces, rather than within workplaces. We therefore extend understanding by examining whether internationals and home nationals working in the same team experience different levels of psychological safety.

Finally, our study contributes to knowledge on psychological safety by presenting a dynamic picture of how it unfolds in the early stages of multinational team development. One limitation of the psychological safety literature concerns the dominance of cross-sectional studies, which

can only provide a static snapshot at a given moment in time. We therefore address calls for research that uses growth curve modeling to highlight inter-individual and team variability in psychological safety development (Newman et al., 2017).

Cultural Diversity Within Teams

Stahl et al. (2010) note that cultural diversity influences teams in contrasting ways. Similarity-attraction theory (Byrne, 1971) argues that mutual attraction is established when people share similar characteristics. Similarly, self-categorization theory (Tajfel, 1982) proposes that people classify themselves into in-groups and differentiate themselves from out-groups on the basis of demographically observable differences. Both theories highlight that homogenous team members are more likely to experience mutual attraction which can facilitate positive team integration and communication, whereas the cultural diversity among members of multicultural teams can hinder these outcomes (e.g., Pauliene et al., 2019). However, a differing perspective is offered by information/decision-making theory (van Knippenberg et al., 2004), which argues that cultural diversity can positively benefit teams. This perspective states that when team members have different cultural backgrounds, members can access a broader range of perspectives and networks, which offers opportunities for greater creativity, enhanced problem solving, and resilience. Meta-analytic data supports both perspectives as team cultural diversity was linked to process losses, through task conflict and reduced social integration, however it was also linked to process gains in the form of greater satisfaction and creativity (Stahl et al., 2010).

In this paper, we seek to understand how national status impacts psychological safety, verbal behavior, and competence ratings in temporally short-lived groups during the first 5 days of team functioning. Short-lived, self-managed teams are increasingly being used by organizations (Bakker, 2010) and have clear managerial importance (Kalish & Luria, 2016), yet we know little about how team dynamics unfold in these types of multinational teams. By following newly formed teams over a 5-day period, we are able to examine how national status influences the way that team members experience psychological safety, and in turn how this influences the competence evaluations of team members as team dynamics unfold and settle. This is imperative as the status hierarchies established in the early stages of group formation are often reinforced over time (Ridgeway, 2001).

To theorize how national status impacts competence ratings in newly formed multinational teams, we adopt expectation states theory (EST; Berger et al., 1980). EST is well suited to explain the formation of competence ratings, as unlike other diversity theories it was developed to understand why certain members of groups without a formal hierarchy quickly adopt higher status roles, which include more opportunities to talk, shape decisions, and display competence (Berger et al., 1980). The main assumption is that members of groups develop stable conceptions of each other's competence using status characteristics, defined as "*attributes on which people differ (e.g., gender, computer expertise) and for which there are widely held beliefs in the culture associating greater social worthiness and competence with one category of the attribute (men, computer expert) than another (women, computer novice)*" (Correll & Ridgeway, 2006, p. 32). Therefore, one member may be perceived as more capable than another because they hold salient status characteristics that members perceive and associate with value (Ridgeway & Erickson, 2000).

Two types of status characteristic are differentiated. Specific status characteristics are those that carry cultural expectations for competence at limited, defined tasks, and they *only impact competence expectations in those tasks* (e.g., computer ability; Correll & Ridgeway, 2006). Comparatively, diffuse status characteristics are those used to make *general assumptions* about another member's capability. Therefore, ethnicity, gender, national status, and physical attractiveness all sometimes operate as diffuse status characteristics. Diffuse status characteristics are particularly important in newly formed performance-oriented groups as they influence how members come to form expectations about the likely quality of each team member's contribution

to the task (Ridgeway, 2001). In this paper, we argue that national status acts as a diffuse status characteristic that signals each member's expected performance.

Hypotheses Development

National Status

In newly formed multinational teams only a few diffuse status characteristics are observable to other members, including gender, race, and national status. Paunova (2017) states that national status is the primary status-determining characteristic in multinational teams, with gender and race being of secondary importance, as national status activates stereotypes which members use to understand interpersonal differences. Indeed, one's values, cognitive schemas, demeanor, and primary language are all influenced by national status (Hambrick et al., 1998).

National status as a construct is open to various definitions and operationalizations. Hambrick, et al. (1998, p. 183) defined it as "*the country in which an individual spent the majority of his or her formative years.*" However, it can also refer to one's legal status, or the identities of one's parents. Given this variation, national status has been measured in contrasting ways (Castilla, 2008; Paunova, 2017; Tröster & van Knippenberg, 2012). In this study we focus on the distinction between team members from an international background and home nationals from the local context. Using Hambrick et al.'s (1998) definition, home nationals are defined as those who spent the majority of their formative years in the country where the multinational team is based. Comparatively, internationals are defined as those working in a different national environment to the one in which they spent their formative years. Previous research in this domain has used different labels to explore this distinction. For example, Haas (2005) investigated status differences between "locals," who live and work in the country in which the team operates and "cosmopolitans" who have lived and worked in many countries. Other studies use terms such as expatriates and host country nationals (Mol et al., 2005; Tung, 1998), or natives and non-natives (Neeley, 2013; Russo et al., 2017). Regardless of the label used, the distinction is common in research, as it provides a measurable and theoretically relevant way of understanding status differences between people (Au et al., 2017; Castilla, 2008; Tröster et al., 2014).

Psychological Safety

Psychological safety enables people to overcome fear of evaluation by making them feel comfortable to speak out and express themselves (Edmondson, 1999). When people experience psychological safety, they are comfortable saying what they think, they engage in constructive conflict, they respect their colleagues', and feel safe taking risks (Edmondson, 1999). Despite being investigated at the individual, group, and organizational levels, psychological safety has been described as a phenomenon that lives at the group level, as groups tend to hold similar perceptions on whether the team climate is safe for interpersonal risk taking (Edmondson & Lei, 2014). However, we argue that psychological safety perceptions are likely to vary across multinational team members, particularly at the start of multinational teamwork projects. This is because individual differences are likely to have a greater impact earlier in a team's lifecycle when team and organizational factors play a less central role (Edmondson & Mogelof, 2006).

The Impact of National Status on Psychological Safety

EST states that when a goal-oriented group come together for the first time, the pressure to perform the group task leads members to evaluate the likely quality of one another's contributions. The assumptions that individuals make about their own and other member's contributions are

called performance expectations, and they are often formed unconsciously, sometimes in a matter of minutes (Paunova, 2017; Ridgeway, 2001). Status characteristics can influence the development of performance expectations, but only under certain conditions. First, a status characteristic must be salient in order to activate performance expectations (Correll & Ridgeway, 2006). For a status characteristic to be salient, it must either differentiate group members, or be relevant to the group's task. Therefore, a diffuse status characteristic such as gender would not be activated in an all-female group as it would not differentiate members. Second, diffuse status characteristics will inform performance expectations even when the characteristic is not relevant to the task at hand, unless it is shown to be unrelated to the task, or there is no other basis for discrimination between actors (such as specific status characteristics that inform members of a particular person's task ability).

In newly formed multinational teams, the salience of national status as a diffuse status characteristic has been activated, as members are differentiated by national status. Since organizations align their norms, values, and beliefs to the most dominant social group (Alderfer et al., 1980) and home national team members are likely to have values, demeanor, and language that are aligned with organizational and societal norms (Foldy et al., 2009), we contend that home nationals will be afforded higher performance expectations than international team members. This will be the case unless there is no other basis for discrimination between actors, such as task expertise. For example, in a newly formed soccer team, an international member who demonstrates greater soccer ability than home nationals will be afforded a higher performance expectation. In this example, soccer ability would be a specific status characteristic which is relevant to the task at hand.

It is important to note that EST assumes that performance expectations are consensually shared across members (Correll & Ridgeway, 2006). This means that members with lower performance expectations will know and recognize that they are lower in the social hierarchy than those with higher performance expectations. Therefore, when national status becomes relevant as a diffuse status characteristic, international team members will consider their own performance expectation as lower than home nationals. EST states that when an actor holds higher performance evaluations, they are more likely to receive opportunities to act, to take up opportunities to act and to have their actions positively evaluated (Correll & Ridgeway, 2006). We argue that this gives way to differing psychologically safety perceptions, as those with greater performance expectations are more likely to view the team climate as receptive to their suggestions, ideas, and concerns. Indeed, research has shown that higher status individuals experience more psychological safety than lower status individuals (Bienefeld & Grote, 2014; Nembhard & Edmondson, 2006), and that home nationals tend to have higher psychological safety than internationals (Ulusoy et al., 2016).

> Hypothesis 1: International team members will have lower psychological safety than home national team members at the start of multinational teamwork projects.

We also contend that national status will influence the development of psychological safety over time. Psychological safety should increase in a linear fashion for both home and international team members, as Edmondson and Lei (2014) state that it develops through familiarity and positive responses to vulnerability and interpersonally risky actions. However, team members with lower performance expectations may experience slower growth in psychological safety as double standards theory predicts that their contributions to the team will be evaluated according to a stricter standard than members with higher performance expectations (Correll & Ridgeway, 2006). This is because when low status individuals perform the group task well, it is inconsistent with expectations of them, which leads to critical scrutiny. Comparatively, when high status individuals perform well, their performance aligns with expectations, leading to lesser scrutiny. Experiencing a more

lenient level of scrutiny should allow the psychological safety of higher status members to unfold at a faster rate over time, as responses to their interpersonally risky actions are more likely to be met with positive responses. Comparatively, it will take lower status members longer to develop psychological safety, as their input will be subjected to greater critical scrutiny.

Hypothesis 2: Over the first 5 days of multinational team development, psychological safety will increase at a slower rate for international team members.

Psychological Safety Growth and Verbal Behavior

A body of research has examined how psychological safety relates to voice behavior, which involves speaking up with constructive information that seeks to improve the status quo (Edmondson & Lei, 2014; Walumbwa & Schaubroeck, 2009). One meta-analysis found that a moderate positive relationship exists between the constructs (Frazier et al., 2017), as employees only feel comfortable engaging in voice behavior when the climate is perceived as safe for interpersonal risk taking.

In this study, we examine how growth in psychological safety affects engagement in verbal behavior during team meetings. Verbal behavior is broader than voice behavior as it encompasses all speech acts. Within teams, opportunities to engage in verbal behavior are limited, as only one person can speak at a time. As a result, the speaking behavior within teams can be dominated by one or two individuals (Sherf et al., 2018). We argue that growth in psychological safety positively relates to engagement in verbal behavior, as those who increasingly believe that their ideas will be accepted are more likely to put their ideas, thoughts, and suggestions forward in the form of verbal behavior.

Hypothesis 3: Psychological safety growth is positively related to verbal behavior (i.e., those who experience lower growth in psychological safety will engage in less verbal behavior).

Verbal Behavior and Competence Ratings

Competence perceptions involve subjective impressions of another person's competence level (Li et al., 2019). Within expectation states theory, there are a variety of cues that inform competence perceptions, including upright posture and speaking up in a confident manner (Correll & Ridgeway, 2006). Speech cues have been found to be particularly important indicators of competence. For example, Li et al. (2019) found that the number of times people spoke up in team meetings was positively related to other people's perceptions of their competence. Engaging in higher levels of verbal behavior signals competence because it suggests to others that the speaker has more information to offer the team. EST states that over time, there is a bi-directional relationship between speaking behavior and competence perceptions, such that those who speak more are recognized as more competent, which affords them more opportunities to speak.

Hypothesis 4: Verbal behavior is positively related to competence ratings.

National Status and Competence Ratings: A Serial Mediation Model

We now lay out a full model that seeks to explain how national status influences competence ratings. We argue that national status acts as a diffuse status characteristic which influences the development of performance expectations upon team formation. Home national team members are likely to be afforded higher performance expectations than international team members because they have norms, values, and beliefs that are more aligned with the organizational and

societal context. Those with higher performance expectations will be asked more questions and will have their suggestions positively evaluated, whereas those with lower performance expectations are more likely to accept the decisions of others (Ridgeway, 2001), Therefore, home nationals will experience greater psychological safety, both initially and over time. In turn, home nationals are more likely to speak up with suggestions, ideas, and concerns, whereas internationals are likely to speak less. Verbal behavior is positively related to competence perceptions (Li et al., 2019), therefore by speaking more often, home nationals will be evaluated as more competent than internationals. As a result, the performance expectations formed in the early stages of a team's lifecycle, create a more psychologically safe space for home nationals, who consequently speak more often and are in turn judged as more competent (see Figure 1).

> Hypothesis 5: The relationship between national status and competence ratings is mediated sequentially through psychological safety growth and verbal behavior.

Methodology

Two studies were conducted to test the hypotheses. The hypotheses were initially tested on a sample of first year undergraduate students at a large UK university. A second study was then conducted using the same research design with a sample of second year students to determine whether the results were replicable. Both studies received full ethical approval.

Research Context

The studies were conducted using a sample of engineering students as they participated in a week-long teamwork module. During the module, students work in four to six person teams to conduct a research project that aims to provide a solution to a real-world engineering problem. Students were organized into teams by the module leader, who ensured that there was at least one international team member on each team. Team members did not have prior experience of working with one another.

The research projects were adapted from the Engineering Without Borders (EWB) Design Challenge, a series of team challenges based upon real world problems faced by the communities in which EWB works. Between Monday and Friday, the teams work in hubs that accommodate several teams. They participate in daily meetings, attend workshops, and meet with university alumni. On the last day of the challenge, the teams submit a report and give a presentation on their research findings.

Study 1

Data were collected from 519 first year undergraduate students, who were organized into 101 teams. Of the participants, 127 (24.5%) were female (391 males, 73.3%, with one missing value) and 165 (31.8%) were from an international background (354 were from the local context, 68.2%). Of the 165 international members, 88 (53.3%) were from Asian countries, 48 (29.1%) were from European countries, 27 (16.4%) were from African countries, 1 (0.6%) was from the Americas and 1 (0.6%) was from Oceania. The students were mostly organized into five-person teams ($n = 83$, 82.2%), although 18 six-person (17.8%) teams participated in the study.

Measures. National status: participants were categorized as international ($n = 165$), or home nationals ($n = 354$) based on the domicile reported in their university records. Those from United Kingdom nations were categorized as home national team members, whereas those from all other countries were categorized as international. Domicile was chosen as the focal

variable rather than nationality, as it was more closely aligned with Hambrick et al.'s (1998) definition as "the country in which an individual spent the majority of his or her formative years." However, it should be noted that there was a very high degree of overlap between nationality and domicile (95.8%).

Psychological safety was measured using Edmondson's (1999) 7-item scale, which participants completed at the end of all five challenge days. The reliability of scale, as measured by Cronbach's alpha, varied between .64 and .86 across the challenge days. ICC(1) values showed that no more than 15% of the variance in psychological safety could be explained at the team level on each of the 5 days, indicating that it was appropriate to analyze psychological safety at the individual-level. The proportion of variance attributable to the team on each challenge day was as follows: day 1 = 3%, day 2 = 3%, day 3 = 11%, day 4 = 15%, and day 5 = 5%.

Verbal behavior was measured using a method known as Behavior Analysis (BA), which involves an independent "observer" sitting in on team meetings to code verbal behavior in real-time (Farley et al., 2018). Similar to other interaction process analysis methodologies, observers code distinct "sense units" that reflect conversational statements. During BA, an observer records each time a team member makes a verbal statement using a custom iPad app. Accordingly, 111 observers were trained to collect data on the teams participating in the study. On day 4 of the challenge, observers recorded data from two meetings (one in the afternoon and one in the morning). The morning meeting lasted roughly 2 hours while the afternoon meeting lasted roughly 3 hours. Data from the morning and afternoon meetings were aggregated to create a verbal behavior variable for the entire day. This variable was created by dividing each individual's total verbal statements by the team's total verbal statements. Verbal behavior was therefore operationalized as the proportion of an individual's verbal statements, relative to their team's overall verbal statements. Therefore, the mean verbal behavior score of 0.20 in study 1 indicates that in a five-person team, each team-member would be responsible for 20% of the team's verbal statements on average. To assess observer interrater reliability, 15 meetings were selected at random in which two different observers applied codes at the same time in the same meeting. The ICC values were obtained by a mean-rating, absolute-agreement, two-way random effects model [ICC (2,k)] with the pairs of observers coding 15 teams (85 participants). The ICC for verbal behavior across team participants was .90, indicating excellent reliability.

Competence ratings were collected on day 5 of the challenge. Each team member was asked to rate themselves and their teammates along two dimensions: (1) their contribution to the project deliverables; and (2) their contribution to the process of team working. These ratings took place on an online peer-moderated marking system. The rating process involved each individual having 100 points to distribute to everyone in their team for both dimensions. Therefore, in a five-person team, each team member would have 100 points to distribute across five team members. As such, if an individual perceived that all team members had contributed equally, they would give each member 20 points. If every member received 20 points from their teammates, their WebPA score would be 100. The WebPA score that each team member received was an aggregate of all team member ratings and this score was also an average of the two dimensions they were rated on. To ensure that the WebPA scores were comparable across teams of different sizes, the WebPA scores were group-mean centered by subtracting the team's WebPA average from each individual's score. Therefore, a competence rating above 0 indicates that the individual was generally perceived as more competent than others in the team, a rating less than 0 indicates that the individual was generally perceived as less competent than others in the team and a score of 0 indicates an average score. Finally, it should be noted that the WebPA variable includes the participants rating of themselves. However, we also examined WebPA ratings without each members' self-rating and found identical results. Therefore, the original WebPA ratings were used in the analysis.

Results

Analysis Strategy

Growth curve modeling was used to test hypotheses 1 and 2 using Mplus Version 8. Growth curve modeling assesses repeated measures of a variable to identify the initial starting point (intercept) and its rate of change across time points (slope). In a growth curve model, the intercept represents the variable's starting value, whereas the slope indicates the change in that variable per time point (Burant, 2016). Growth curve modeling also facilitates the analysis of independent variables associated with the trajectory of a dependent variable. It was therefore possible to model the starting values and temporal change of participants' psychological safety based upon their national status. Furthermore, conditional growth curve models allow for the inclusion of covariates that theoretically affect the rate of change (Bollen & Curran, 2006). This allows an examination of how covariates affect the mean intercept and mean slope. We therefore controlled for gender and team size in the analyses, as these variables can influence psychological safety (Edmondson, 2004). We also controlled for the students' average grade to ensure that any differences were due to national status, rather than academic ability. It should also be noted that all statistical tests were performed without the control variables and the findings remained the same without them.

We fitted a linear growth trajectory, as psychological safety is expected to increase in an upwards trajectory in the early phases of team development (Edmondson & Lei, 2014). Psychological safety measured across all 5 days were included in the growth model. Some missing data was apparent across the 5 days, as the number of complete psychological safety cases was 474, 484, 476, 456, and 397. For all statistical tests, we handled missing data using full information maximum likelihood (FIML) estimation, which makes use of the available data to estimate the model. We also accounted for the multi-level nature of the data by using the cluster function in Mplus which models the non-independence of observations, as the participants were nested within teams. In order to minimize model complexity and maximize statistical power, the psychological safety items were averaged at each time point. Furthermore, models were estimated using a robust maximum likelihood (MLR) estimator, which is an Mplus requirement for adopting the cluster function.

To test hypotheses 1 and 2, a model contrasting approach was adopted, starting with a model that fixed the paths from national status to the intercept and slope of psychological safety. A series of models were then specified that allowed the intercept and slope to vary according to national status and the control variables (see Table 1). In model 1, the paths from national status and the control variables to the intercept and slope of psychological safety were constrained at zero. In model 2, the path from national status to the intercept of psychological safety was freed to examine whether the starting values of psychological safety varied by national status. In model 3, the path from national status to the slope of psychological safety was also freed to determine whether growth in psychological safety over time differed according to national status. Slope loadings were fixed at 0, 1, 2, 3, and 4. Finally, in model 4, the paths from the covariates to the intercept and slope were freed to determine whether national status explained variation in the intercept and slope after accounting for the control variables.[1] We assessed the fit of the different models using model fit indices, including the chi-square (χ^2) statistic, the comparative fit index (CFI), the Tucker-Lewis index (TFI), the root mean squared error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the Akaike information criterion (AIC). Hu and Bentler (1999) suggest that CFI and TLI values between 0.90 and 0.95 represent acceptable fit, whereas values above 0.95 indicate good fit. Similarly, models with an RMSEA and SRMR of less than 0.08 have acceptable fit, with RMSEA values below 0.05 indicating good fit. Finally, lower chi-square and AIC values indicate better fit, although the former is influenced by sample size.

Table 1. Fit Indices of Competing Growth Curve Models For Study 1.

Model	χ^2 (df)	CFI	TLI	RMSEA	SRMR	AIC
(1) Fixed paths from national status and controls to i and s.	94.99* (34)	0.93	0.94	0.06	0.14	2,477.63
(2) Free path from national status to i, but fixed to s. Fixed paths from controls to i and s.	87.41* (33)	0.94	0.94	0.06	0.12	2,469.52
(3) Free paths from national status to i and s. Fixed paths from controls to i and s.	81.99* (32)	0.94	0.94	0.06	0.13	2,464.88
(4) Free paths from national status and controls to i and s.	74.25* (26)	0.94	0.93	0.06	0.12	2,472.44

Note. i = intercept; s = slope.
*p < .05.

Table 1 shows the fit indices for the different models. Based on the chi-square statistic, model 4 appears to be the best fitting model. The procedures for conducting chi-square difference tests for MLR models were used to compare model 1 to model 4. Model 1 was chosen as the null model, as the intercept and slope were constrained in this model. The difference test showed that model 4 had significantly better fit than model 1 (χ^2 (7) = 20.31, $p < .01$) and the overall fit of model 4 is acceptable ($\chi^2 = 74.25$ (26), $p < .05$; CFI = 0.94, TLI = 0.93, RMSEA = 0.06, SRMR = 0.12, AIC = 2,472.44), although the SRMR was above the recommended cut-off of .08. However, Wu et al. (2009) state that the CFI, TFI, and RMSEA have shown good performance in evaluating the fit of growth curve models. In contrast, they state that researchers should be cautious about using the standardized root mean square residual (SRMR) to evaluate growth curve models, as it has low power to detect misspecification in the mean structure. We therefore relied on the chi-square, CFI, TFI, and RMSEA in the analysis of model fit.

Consultation of the model 4 coefficients showed that national status significantly impacted the psychological safety intercept (estimate = −0.10, $p \leq .05$). The estimated starting level of psychological safety for home national team members was 4.25, whereas for international team members it was 4.14. Therefore, international team members started the team working project with lower levels of psychological safety, providing support for hypothesis 1. National status also significantly influenced the psychological safety slope (estimate = −0.04, $p < .01$). The estimated slope of home national team members was .09, while for internationals it was .05. This demonstrates that international team members' psychological safety grew at a significantly slower rate over time, thus supporting hypothesis 2. Finally, neither the intercept (I) nor slope (S) of the control variables were significant.

To test hypotheses 3, 4, and 5 we adopted serial mediation analysis, whereby national status was specified as the independent variable, psychological safety growth was specified as the primary mediator, verbal behavior was specified as the secondary mediator, and competence ratings were specified as the dependent variable. Psychological safety growth was measured using the slope of psychological safety across the five study days, which was estimated in the manner used to test hypothesis 2 (the intercept was fixed at zero). We made the analytical decision of using the psychological safety data from all 5 days of the challenge, as it is recommended that at least five time point variables are used for growth curve modeling. However, since verbal behavior was measured on day 4, we also ensured that the results were equivalent when only the first 4 days of psychological safety data was used to form the growth variable.

The means, standard deviations, correlations, and reliabilities of the study variables are shown in Table 2. Most variables conformed to a normal distribution, although the time 3, 4, and 5

Table 2. Means, Standard Deviations, Correlations, and Reliabilities of Study 1 Variables.

Variables	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Nat. status (0=hom., 1=int.)	0.32	0.47											
2. Sex (0=female, 1=male)	0.75	0.43	−.03										
3. Team size	5.19	0.39	−.07	.02									
4. Average grade	58.00	16.50	−.08	−.06	.01								
5. Meeting attendance	−0.14	0.50	.02	−.08	.06	.15**							
6. Psychological safety (T1)	4.14	0.50	−.11*	.02	−.03	.09	−.07	(.64)					
7. Psychological safety (T2)	4.16	0.56	−.11*	.04	−.05	.08	.02	.60**	(.75)				
8. Psychological safety (T3)	4.29	0.58	−.14**	.05	.01	.01	.01	.47**	.71**	(.77)			
9. Psychological safety (T4)	4.36	0.58	−.16**	.03	−.03	.06	−.05	.48**	.71**	.75**	(.82)		
10. Psychological safety (T5)	4.39	0.65	−.20**	−.03	−.04	.14**	.01	.48**	.70**	.73**	.86**	(.86)	
11. Verbal behavior	0.20	0.10	−.38**	.06	−.08	.14**	.08	.12**	.22**	.20**	.23**	.25**	
12. Competence rating	0.00	0.15	−.28**	.01	.00	.15**	.31**	.14**	.14**	.15**	.17**	.21**	.55**

Note. T=time; reliabilities (Cronbach's α) are in parentheses on the diagonal; N=519.
*p<.05. **p<.01

Table 3. Mean and Standard Deviations (*SD*) For Daily Psychological Safety Across the Subgroups.

National and continental groups	Day 1 psychological safety	Day 2 psychological safety	Day 3 psychological safety	Day 4 psychological safety	Day 5 psychological safety
UK (home national)	4.17 (0.50)	4.20 (0.54)	4.34 (0.57)	4.43 (0.57)	4.48 (0.61)
Asian	3.97 (0.51)	3.98 (0.62)	4.08 (0.65)	4.11 (0.65)	4.05 (0.81)
African	4.11 (0.40)	4.28 (0.42)	4.26 (0.49)	4.29 (0.52)	4.36 (0.36)
European	4.23 (0.45)	4.18 (0.56)	4.32 (0.50)	4.39 (0.43)	4.45 (0.47)

psychological safety variables were negatively skewed. However, the mediation models were fitted using bootstrapping which correctly estimates effects in non-normal data (Finney & DiStefano, 2006). Similar to the procedures used to test hypotheses 1 and 2, we accounted for the multi-level nature of the data by using the type=complex cluster function in Mplus to account for the nesting of participants within teams. When the type=complex cluster function is used alongside bootstrapping, Mplus requires a weight variable to account for the non-independence of observations. A weight variable is a value for each observation in the dataset, which in this case was 1 for all observations.

We conducted serial mediation analyses using Mplus and ran bootstrap analyses (10,000 iterations) to obtain stable regression coefficients. Unlike the analyses of hypotheses 1 and 2, team size was not included as a control variable, as it was accounted for in the adjustment of the competence rating variable. Furthermore, meeting attendance was used as a control variable, alongside sex and grade, as it had a significant impact on competence ratings ($r = .31$, $p < .01$). Attendance was recorded by subtracting the number of challenge days the student missed from zero. Therefore, the positive correlation between competence rating and attendance indicates that students who attended more days were rated as more competent.

In support of hypothesis 3, the serial mediation results showed that the path between psychological safety growth and verbal behavior was positive and significant (standardized beta=.19, $p < .01$).[2] Hypothesis 4 was also supported, as the path between verbal behavior and competence ratings was also positive and significant (standardized beta=.48, $p < .01$). In addition, the results showed a significant indirect effect from national status to competence ratings through psychological safety and verbal behavior (point estimate=−0.006, lower confidence interval (LCI)=−0.011, upper confidence interval (UCI)=−0.003, $p < .01$). The indirect effect from national status to competence ratings via verbal behavior was also significant (point estimate=−0.05, LCI=−0.07, UCI=−0.03, $p < .001$). However, the indirect effect from national status to competence ratings via psychological safety growth was non-significant (point estimate=−0.003, LCI=−0.012, UCI=0.003, $p = .387$). Finally, the direct effect between national status and competence ratings was significant (point estimate=−0.088, LCI=−0.122, UCI=−0.056, $p < .001$), indicating that psychological safety and verbal behavior partially mediated the relationship between national status and competence ratings. These findings provide partial support for hypothesis 5.

Exploratory Analyses

Since those in the international status samples held a variety of national backgrounds, one question arising from the results concerned whether psychological safety, verbal behavior, and competence ratings differed according to continental background. We therefore conducted further analyses to explore whether there were differences in the main study variables based on whether participants were from the UK, Africa, Asia, or Europe.[3] Descriptive analyses were undertaken to examine the psychological safety of each subgroup on the challenge days using the mean and
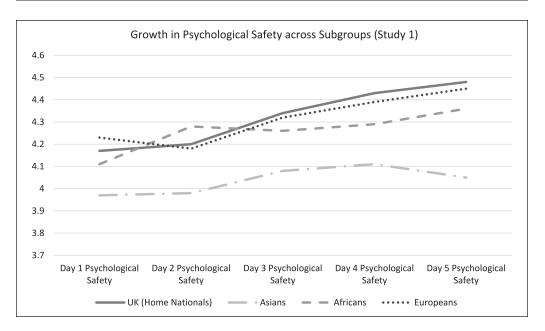
Figure 2. Growth in psychological safety across subgroups over time (study 1)

standard deviation (SD). Analysis of Variance (ANOVAs) were then conducted to determine whether differences existed in relation to verbal behavior and competence ratings.

Table 3 and Figure 2 indicate that Europeans had the highest psychological safety upon group formation, followed by UK home nationals, Africans, and Europeans. The analysis was descriptive; therefore, it is not possible to determine whether psychological safety differs significantly across the subgroups. However, the data show that UK home nationals have the greatest growth in psychological safety from day 1 to day 5, followed by Africans, Europeans, and Asians respectively. Moreover, on day 5, UK home nationals have the highest psychological safety, followed by Europeans, Africans, and Asians.

A one-way analysis of variance (ANOVA) was conducted to determine whether participants from the four subgroups differed according to how much verbal behavior they enacted. There was a significant effect of subgroup on verbal behavior $F(3, 499) = 30.06$, $p < .001$ and follow-up Bonferroni post hoc tests revealed that UK home nationals (Mean $(M) = 0.23$, $SD = 0.09$) spoke significantly more than Asian team members ($M = 0.13$, $SD = 0.09$) and African team members ($M = 0.14$, $SD = 0.06$). However, there was no significant difference in verbal behavior between UK nationals and European team members ($M = 0.19$, $SD = 0.09$).

A further one-way ANOVA found that there were significant differences between the subgroups in relation to competence ratings $F(3, 506) = 18.96$, $p < .001$. A Bonferroni post-hoc comparison revealed that UK home nationals ($M = 0.03$, $SD = 0.12$) received significantly higher competence ratings than Asian ($M = -0.08$, $SD = 0.16$) and African team members ($M = -0.07$, $SD = 0.21$). However, they did not receive higher ratings than European team members ($M = 0.01$, $SD = 0.14$).

Study 2

During study 1, the hypotheses were tested on a sample of first year students who would have been new to the host country context, and it is likely that many of these students would not have worked in multinational teams previously. This replicates the context faced by many international members of multinational teams who move abroad to secure work for the first time.

Table 4. Fit Indices of Competing Growth Curve Models For Study 2.

Model	χ^2 (*df*)	CFI	TLI	RMSEA	SRMR	AIC
(1) Fixed paths from national status and controls to *i* and *s*.	118.09* (34)	0.94	0.95	0.07	0.16	2,114.69
(2) Free path from national status to *i*, but fixed to *s*. Fixed paths from controls to *i* and *s*.	91.59* (33)	0.96	0.96	0.06	0.14	2,116.77
(3) Free paths from national status to *i* and *s*. Fixed paths from controls to *i* and *s*.	85.68* (32)	0.96	0.97	0.06	0.14	2,113.17
(4) Free paths from national status and controls to *i* and *s*.	72.22* (26)	0.97	0.96	0.06	0.13	2,115.80

Note. *i* = intercept; *s* = slope.
*$p < .05$.

However, some members of multinational teams are familiar with the cultural context in which they work and have previous experience of working in multinational teams. As a result, we wanted to determine whether the results of study 1 were replicable on a sample of second year students who were completing a similar teamwork module at the same university.

In addition, the exploratory analyses showed few differences between UK home nationals and Europeans with regards to psychological safety, verbal behavior, and competence ratings. Moreover, both subgroups spoke more and received higher competence ratings than African and Asian team members. This indicates that the home/international dichotomy may not fully explain the observed findings, and that a more culturally nuanced theoretical explanation may be required. Indeed, EST outlines how information associated with several characteristics is used to form performance expectations (Correll & Ridgeway, 2006). Therefore, whilst international status signals to home nationals that a team member may hold different values, demeanor, language, and knowledge of the organizational context, the specific nationality of an international team member may provide further information relevant to performance expectation formation. For example, international team members from more developed nations may be ascribed higher performance expectations than those from less developed nations (Paunova, 2017). Cultural distance dimensions may also be relevant in this process, particularly individualism, defined as "*the relationship between the individual and the collectivity that prevails in a given society*" (Hofstede, 2001, p. 209). Research finds that the individualism/collectivism distinction influences information sharing (Sarker, 2005) and self-enhancement (Sedikides et al., 2005), which may influence both verbal behavior and competence ratings. Therefore, this new sample offered a more rigorous test of the hypotheses, as well as an opportunity to conduct more in-depth exploratory work.

Procedure. Identical procedures, measures, and analyses were adopted to replicate study 1 on this new sample. The data were collected from 538 second year undergraduates, who were grouped into 90 teams. The participants included 94 (17.5%) females and 434 (80.7%) males (with 10 missing values). Of these, 353 (65.6%) were from UK countries and 175 (32.5%) were from international countries (10 missing values). Of the 175 international team members, 104 (59.4%) were from Asian countries, 38 (21.7%) were from European countries, 28 (16%) were from African countries, 4 (2.6%) were from American countries, and 1 (0.7%) was from Oceania. The students were predominantly grouped into six-person teams (*n* = 87, 95.6%), although two five-person (2.2%) and one four-person (1.1%) team participated (with one missing value).

Data screening. The reliabilities of the psychological safety variables varied between .65 and .86 (all alphas are reported in Table 4). ICC(1) values showed that no more than 6% of the variance in psychological safety on any 1 day could be explained by team membership, indicating the appropriateness of individual-level analyses. The study variables were normally distributed, with the exception of psychological safety on days 3, 4, and 5, which were negatively skewed. The number of complete cases for psychological safety across the 5 days were 474, 459, 436, 441, and 376.

Growth curve modeling was again used to test hypotheses 1 and 2. We used the model contrasting approach adopted in study 1 to determine whether there was a linear increase in participants' psychological safety. Table 4 shows the fit indices of the different models. Similar to study 1, model 4 (involving unconstrained paths from national status and the control variables to psychological safety) is the best fitting model, displaying good fit overall ($\chi^2 = 72.22$ (26), $p < .05$; CFI = 0.97, TLI = 0.96, RMSEA = 0.06, SRMR = 0.13, AIC = 2,115.80).

It was not possible to correctly estimate the intercept or slope for model 4 due to a problem when estimating a team size parameter on the slope. We therefore removed team size from this model, which altered the fit statistics ($\chi^2 = 57.83$ (23), $p < .05$; CFI = 0.97, TLI = 0.97, RMSEA = 0.06, SRMR = 0.13, AIC = 2,117.19). Further analysis was then conducted on the adjusted model. An MLR chi-square difference test was used to compare model 1 with model 4, which showed that the latter held significantly better fit (χ^2 (10) = 49.43, $p < .001$). Furthermore, national status significantly influenced the psychological safety intercept (estimate = −0.24, $p < .001$). Home team members estimated starting level of psychological safety was 4.04, compared to 3.80 for international team members, which provides additional support for hypothesis 1. Support for hypothesis 2 was also replicated, as national status significantly influenced the psychological safety slope (estimate = −0.03, $p < .001$). Home team members' slope was estimated at .06, whereas international team members' slope was estimated at .03. This provides strong support to indicate that international team members' psychological safety grows more slowly over time in newly formed multinational teams. Neither the intercepts nor slopes of the control variables were significant.

The same variables and analytical approach were used to identify whether support for hypotheses 3, 4, and 5 would be replicated. The means (*M*), standard deviations (*SD*s), correlations, and reliabilities of the variables are shown in Table 5. Results of the serial mediation showed that the path between psychological safety growth and verbal behavior was positive and significant (standardized beta = .17, $p < .001$), which provides additional support for hypothesis 3. Similarly, the path between verbal behavior and competence ratings was also positive and significant (standardized beta = .41, $p < .001$), which replicates support for hypothesis 4. The results also showed that psychological safety growth and verbal behavior partially mediated the relationship between national status and competence ratings. The indirect effect from national status to competence ratings via psychological safety growth and verbal behavior was significant (point estimate = −0.005, LCI = −0.010, UCI = −0.002, $p < .05$). Moreover, the indirect effect between national status and competence ratings was significant via verbal behavior (point estimate = −0.03, LCI = −0.04, UCI = −0.02, $p < .001$), but not psychological safety growth (point estimate = −0.004, LCI = −0.012, UCI = 0.001, $p = .198$). Finally, the direct effect was significant (point estimate = −0.079, LCI = −0.111, UCI = −0.053, $p < .001$), meaning that the partial support for hypothesis 5 was replicated in study 2.[4]

Exploratory Analyses

By collecting additional data in study 2 we were able to conduct more detailed exploratory analysis on whether the results were consistent across the continental subgroups. We combined the data from the two studies, which increased the sample size within each subgroup: home nationals = 707, Africans = 55, Asians = 192, and Europeans = 86. This allowed us to conduct growth

Table 5. Means, Standard Deviations, Correlations, and Reliabilities of Study 2 Variables.

Variables	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Nat. status (0 = hom., 1 = int.)	0.33	0.47											
2. Sex (0 = female, 1 = male)	0.82	0.38	−.02										
3. Team size	5.97	0.22	.00	.04									
4. Average grade	60.05	13.68	−.06	−.07	.02								
5. Meeting attendance	−0.15	0.60	.11*	.04	.04	.03							
6. Psychological safety (T1)	4.13	0.51	−.20**	.06	.06	.09*	.01	(.65)					
7. Psychological safety (T2)	4.15	0.57	−.24**	.03	.09*	.03	−.04	.74**	(.74)				
8. Psychological safety (T3)	4.22	0.62	−.19**	.03	.03	.05	.01	.65**	.75**	(.82)			
9. Psychological safety (T4)	4.26	0.66	−.23**	−.02	.08	.07	−.01	.64**	.76**	.80**	(.85)		
10. Psychological safety (T5)	4.31	0.64	−.24**	−.01	.13*	.12*	.04	.57**	.70**	.79**	.85**	(.86)	
11. Verbal behavior	0.17	0.09	−.32**	.01	−.08	.07	.13**	.18**	.21**	.21**	.22**	.25**	
12. Competence rating	0.00	0.13	−.28**	.04	.00	.13**	.33**	.17**	.14**	.17**	.20**	.26**	.50**

Note. T = time; reliabilities (Cronbach α) are in parentheses on the diagonal; N = 538.
*p < .05. **p < .01.

curve modeling to determine whether hypotheses 1 and 2 were supported when the home nationals were compared with the African, Asian, and European subgroups separately. In these growth curve models we controlled for study year group to ensure that any differences were not attributable to differences between first- and second-year students. These analyses showed that the psychological safety intercept of UK home nationals significantly differed from Asians (estimate$=-0.28$, $p < .001$), indicating that Asian team members started with significantly lower levels of psychological safety. However, the intercept of UK home nationals did not significantly differ from those of Africans (estimate$=-0.06$, $p=.42$) or Europeans (estimate$=-0.01$, $p=.91$). Similarly, the psychological safety slope of home nationals differed significantly from Asians (estimate$=-0.04$, $p < .01$), but not from Africans (estimate$=-0.02$, $p=.24$) or Europeans (estimate$=-0.02$, $p < .16$). However, the subgroup sample sizes may have affected the pattern of significant results, as the European and African subgroups were much smaller than the Asian subgroup.

One-way ANOVAs showed significant differences regarding verbal behavior (F(3, 1,007)$=62.85$, $p < .001$) and competence ratings (F(3, 1,025)$=37.93$, $p < .001$) across the subgroups. Bonferroni post hoc tests showed that home nationals ($M=0.21$, $SD=0.09$) spoke significantly more than Africans ($M=0.14$, $SD=0.07$) and Asians ($M=0.11$, $SD=0.09$), but not more than Europeans ($M=0.19$, $SD=0.09$). Europeans spoke significantly more than Africans and Asians, while there were no significant differences between Africans and Asians. Post hoc tests also indicated that home nationals ($M=0.03$, $SD=0.11$) received significantly higher competence ratings than Africans ($M=-0.05$, $SD=0.17$) and Asians ($M=-0.08$, $SD=0.16$), but there were no significant differences between the competence ratings of home nationals and Europeans ($M=0.01$, $SD=0.12$). Moreover, Europeans received significantly higher competence ratings than Africans and Asians, while no significant differences existed between African and Asian competence ratings.

Discussion

Using EST, we argued that international team members would be afforded lower performance expectations than home nationals when multinational teams are formed. We therefore hypothesized that international team members would start multinational teamwork projects with lower psychological safety and that they would experience slower development in psychological safety over time. We found support for these predictions across the two studies, which both showed that psychological safety started lower and grew more slowly for international team members. However, the exploratory analyses conducted on the combined dataset showed that the psychological safety of home nationals only started and grew more quickly compared Asians, with the differences being non-significant for Africans and Europeans.

In turn, we argued that the slower development in psychological safety among international team members would be associated with lower verbal behavior and lower competence ratings. Findings replicated across both studies partially supported these predictions, however exploratory analyses showed that only African and Asian team members spoke significantly less and received significantly lower competence ratings than home nationals. This suggests that an international team member's specific nationality is a more important predictor than national status in determining how competence perceptions unfold in newly formed national teams. EST states that members of newly formed groups use various information to develop performance expectations about the likely contributions of members (Correll & Ridgeway, 2006). Therefore, whilst national status can highlight whether an individual is familiar with the host country and organization's customs, it does not provide the same breadth of information as nationality, which better signals factors that may be relevant when performance expectations are formed (e.g., cultural distance, language differences, and economic development; Paunova, 2017).

This study took place in the UK context which may help to explain the pattern of exploratory results observed. Previously, Paunova (2017) found that members of multinational teams from more developed nations were more likely to be considered team leaders than individuals from less developed nations. She argued that "*members of multinational teams possess implicit knowledge about the global order that allows them to, perhaps unintentionally, estimate a nation's relative level of development, and thereby their partners' potential worth*" (Paunova, 2017, p. 889). The UK is a developed country, and whilst there is considerable variation across continents, European nations are generally placed higher on the human development index than Asian and African countries (United Nations Development Programme, 2020). Therefore, implicit believes about national development may have driven the assignment of performance expectations. However, Asian team members had the lowest psychological safety, verbal behavior, and competence ratings, which suggests that other cultural factors are relevant. Indeed, within multinational team contexts the distinction between individualism and collectivism may have a particularly strong influence on how team members assign value and behave. In individualistic cultures, such as the UK and other western nations, individuals seek to demonstrate value by highlighting their personal efficacy and social dominance, whereas individuals from collectivist nations self-enhance by emphasizing collectivist behaviors, including avoiding conflict, or supporting group decisions (Sedikides et al., 2005). Therefore, individuals from individualistic countries tend to share more knowledge in multinational team contexts than individuals from collectivist countries (Sarker, 2005). When this occurs within an individualistic host country, westerners are more likely to be perceived as competent as their behavior aligns with societal and organizational expectations, whereas collectivist forms of self-enhancement are less likely to be valued. This may explain why UK nationals and Europeans spoke more and received higher competence ratings than their Asian and African counterparts. Nevertheless, this does not necessarily explain why Asian team members started multinational teamwork projects with particularly low levels of psychological safety. In cultures characterized by higher collectivism, power distance and uncertainty avoidance, engaging in voice behavior and promoting new ideas comes with the risk of losing face, or potential ostracism from the group (Newman et al., 2017). Therefore, Asian team members may have been reluctant to speak up and may have been anxious about doing so in a different language. Given our analyses were exploratory and only considered continental rather than national differences, further research is needed to determine how nationality informs performance expectations in multinational teams. Future studies should also seek to establish whether the same results would be obtained if the study was conducted in a different national context. For example, if the study was conducted in a collectivist context, would westerners have lower psychological safety, verbal behavior, and competence ratings than home nationals?

Our analyses support the predictions of Edmondson and Lei (2014) who commented that psychological safety is likely to increase in a linear fashion as team members become more familiar with each other. Indeed, it is interesting to note that psychological safety grew quite substantially over a very short period of time. This may have been a function of the amount of time that team members spent together during the challenge days, as well as the intensity and level of interdependence involved in the task. Our results confirm that as psychological safety grows, individuals are more likely to engage in verbal behavior. This is crucial to the success of multinational teams, as the performance benefits arising from broader knowledge pools can primarily be achieved through knowledge sharing (van Knippenberg et al., 2004). Indeed, having a broad range of perspectives can enhance creativity, learning, and decision-making (Guillaume et al., 2013), but only when team members feel psychological safe to share their insights. In this respect, Hajro et al. (2017, p. 353) found that multinational team knowledge exchange is most effective in engagement-focussed diversity climates, where diversity is used "*to inform and enhance work processes based on an assumption that cultural differences give rise to different insights.*" Our research found that international team members started with lower levels of

psychological safety than home nationals, and it was consistently found to develop more slowly for them. This underscores the importance of promoting engagement-focussed diversity climates where the value of different cultural perspectives is continuously emphasized throughout the team's lifecycle.

Our findings also showed that first-year international team members started with higher psychological safety and demonstrated quicker growth than second year international team members. This provides further support of EST, which states that performance expectations formed in one encounter are carried over to new situations and groups (Correll & Ridgeway, 2006). Therefore, whilst it may have been expected that international team members in study 2 should have had greater levels of psychological safety, due to being in the country for a year longer, the opposite was observed. This can be explained by home and international team members forming performance expectations in their first-year teamwork projects, which were transferred into the second-year teamwork project. This suggests that although international team members who have lived for longer in the host country context may have a better understanding of the language and organizational context, they may still have to overcome stereotypes regarding their national background.

Practical Implications

We have argued that the disparity in psychological safety, verbal behavior, and competence ratings between home and international team members arises because national status is a diffuse status characteristic used to form competence evaluations in the absence of other relevant information. To minimize the extent to which national status is used as a diffuse characteristic and to maximize the benefits of multinational teams, it is important to highlight the value of diversity within multinational team contexts. Hajro et al. (2017) suggest that assertive knowledge exchange between team members can increase effectiveness, which can be achieved by encouraging members to express the characteristics that make them unique. Therefore, when international teams first come together, each team member could be encouraged to outline their background, skills, and expertise. By doing so, more relevant status characteristics may be used to form competence evaluations.

At the organizational level, Singh et al. (2013) found that minorities experienced higher psychological safety in supportive diversity climates. Diversity climates are defined as "*employees' perceptions regarding the extent to which an organization values and integrates diversity and supports it through fair employment practices*" (Singh et al., 2013, p. 243). Organizations can signal their commitment to diversity by increasing minority representation at senior organizational levels (Singh et al., 2013), by highlighting the successes and achievements of international employees, and by training home employees to recognize the difficulties that international employees face in unfamiliar cultural contexts (Li et al., 2019).

Limitations and Future Directions

Some limitations should be acknowledged. First, we examined student members of multinational teams as they participated in a short-term teamwork project. Future research should therefore determine whether the results can be replicated with a representative sample of working teams. Indeed, it would be particularly interesting to examine the antecedents of competence ratings in established multinational teams to determine the extent to which they are facilitated by psychological safety and verbal behavior. Researchers have found that psychological safety tends to be lower for international employees compared to home nationals (Ulusoy et al., 2016), which suggests that the mechanisms highlighted in this study may prevail in more established teams.

Second, the use of a student sample may affect the generalizability of our findings; however, this should be balanced against the use of rigorous multi-source and multi-time methods to test our hypotheses. Psychological safety was measured consistently across 5 days, verbal behavior was measured using observational methods and competence ratings were obtained from direct ratings. Together, these sources of data allowed us to examine the relationships more precisely, in a way that limited bias (Kozlowski, 2015). We also conducted a replication study on a separate sample to further evidence the internal validity of our findings.

Conclusion

To release the potential of multinational teams, organizations need to create conditions where international team members feel comfortable voicing their ideas, suggestions, and opinions. However, this study suggests that international team members experience slower development of psychological safety in newly formed multinational teams, which may inhibit verbal behavior and in turn reduce competence perceptions. These insights should be used to ensure that international team members feel valued throughout the team's lifecycle by creating conditions that highlight the benefits of cultural diversity.

Declaration of Conflicting Interests

Funding

ORCID iD

Samuel Farley https://orcid.org/0000-0003-0782-4750

Supplemental Material

Supplemental material for this article is available online.

Notes

1. The Mplus syntax is attached as a Supplemental Material.
2. Path diagrams of the mediation models specified in studies 1 and 2 are available in a Supplemental Material.
3. It was not possible to include participants from the Americas or Oceania in these analyses as the sample size was insufficient in these subgroups.
4. According to our theoretical framework, national status should affect the level and trajectory of psychological safety. Although we examined the trajectory in the mediation tests, equivalent results were also obtained when we assessed the level of psychological safety by taking an average across days 1 to 3.

References

Alderfer, C. P., Alderfer, C. J., Tucker, L., & Tucker, R. (1980). Diagnosing race relations in management. *The Journal of Applied Behavioral Science*, *16*(2), 135–166. https://doi.org/10.1177/002188638001600202

Au, T. K. F., Kwok, A. F. P., Tong, L. C. P., Cheng, L., Tse, H. M. Y., & Jun, S. A. (2017). The social costs in communication hiccups between native and nonnative speakers. *Journal of Cross-Cultural Psychology*, *48*(3), 369–383. https://doi.org/10.1177/0022022116687852

Bakker, R. M. (2010). Taking stock of temporary organizational forms: A systematic review and research agenda. *International Journal of Management Reviews*, *12*(4), 466–486. https://doi.org/10.1111/j.1468-2370.2010.00281.x

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

Bienefeld, N., & Grote, G. (2014). Speaking up in ad hoc multiteam systems: Individual-level effects of psychological safety, status, and leadership within and across teams. *European Journal of Work and Organizational Psychology*, *23*(6), 930–945. https://doi.org/10.1080/1359432X.2013.808398

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Wiley.

Burant, C. J. (2016). Latent growth curve models: Tracking changes over time. *The International Journal of Aging and Human Development*, *82*(4), 336–350. https://doi.org/10.1177/0091415016641692

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

Castilla, E. J. (2008). Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, *113*(6), 1479–1526. https://doi.org/10.1086/588738

Correll, S. J., & Ridgeway, C. L. (2006). Expectation states theory. In J. Delamater (Ed.), *Handbook of social psychology* (pp. 29–51). Springer.

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383. https://doi.org/10.2307/2666999

Edmondson, A. C. (2004). Psychological safety, trust, and learning in organizations: A group-level lens. In R. M. Kramer & K. S. Cook (Eds.), *Trust and distrust in organizations: Dilemmas and approaches* (pp. 239–72). Russell Sage.

Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 23–43. https://doi.org/10.1146/annurev-orgpsych-031413-091305

Edmondson, A. C., & Mogelof, J. P. (2006). Explaining psychological safety in innovation teams: Organizational culture, team dynamics, or personality? In L. Thompson & H. Choi (Eds.), *Creativity and innovation in organizational teams* (pp. 129–156). Psychology Press.

Farley, S., Evison, R., Rackham, N., Nicolson, R., & Dawson, J. (2018). The behavior analysis coding system: An applied, real-time approach for measuring and improving interactive skills. In E. Brauner, M. Boos, & M. Kolbe (Eds.), *Cambridge handbook of group interaction analysis* (pp. 584–593). Cambridge University Press.

Finney, S. J., & DiStefano, C. (2006). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Information Age Publishing.

Foldy, E. G., Rivard, P., & Buckley, T. R. (2009). Power, safety, and learning in racially diverse groups. *Academy of Management Learning & Education*, *8*(1), 25–41. https://doi.org/10.5465/amle.2009.37012177

Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A., & Vracheva, V. (2017). Psychological safety: A meta-analytic review and extension. *Personnel Psychology*, *70*(1), 113–165. https://doi.org/10.1111/peps.12183

Guillaume, Y. R., Dawson, J. F., Woods, S. A., Sacramento, C. A., & West, M. A. (2013). Getting diversity at work to work: What we know and what we still don't know. *Journal of Occupational and Organizational Psychology*, *86*(2), 123–141. https://doi.org/10.1111/joop.12009

Haas, M. R. (2005). Cosmopolitans and locals: Status rivalries, deference, and knowledge in international teams. *Research on managing groups and teams: Status and groups*, *7*, 201–227. https://doi.org/10.1016/S1534-0856(05)07009-X

Hajro, A., Gibson, C. B., & Pudelko, M. (2017). Knowledge exchange processes in multicultural teams: Linking organizational diversity climates to teams' effectiveness. *Academy of Management Journal*, *60*(1), 345–372. https://doi.org/10.5465/amj.2014.0442

Hambrick, D. C., Davison, S. C., Snell, S. A., & Snow, C. C. (1998). When groups consist of multiple nationalities: Towards a new understanding of the implications. *Organization Studies*, *19*(2), 181–205. https://doi.org/10.1177/017084069801900202

Hofstede, G. (2001) *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. SAGE.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kalish, Y., & Luria, G. (2016). Leadership emergence over time in short-lived groups: Integrating expectations states theory with temporal person-perception and self-serving bias. *Journal of Applied Psychology*, *101*(10), 1474–1486. https://doi.org/10.1037/apl0000126

Kozlowski, S. W. (2015). Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations. *Organizational Psychology Review*, *5*(4), 270–299. https://doi.org/10.1177/2041386614533586

Li, H., Yuan, Y. C., Bazarova, N. N., & Bell, B. S. (2019). Talk and let talk: The effects of language proficiency on speaking up and competence perceptions in multinational teams. *Group & Organization Management*, *44*(5), 953–989. https://doi.org/10.1177/1059601118756734

Mol, S. T., Born, M. P., Willemsen, M. E., & Van Der Molen, H. T. (2005). Predicting expatriate job performance for selection purposes: A quantitative review. *Journal of Cross-Cultural Psychology*, *36*(5), 590–620. https://doi.org/10.1177/0022022105278544

Neeley, T. B. (2013). Language matters: Status loss and achieved status distinctions in global organizations. *Organization Science*, *24*(2), 476–497. https://doi.org/10.1287/orsc.1120.0739

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

Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. *Human Resource Management Review*, *27*(3), 521–535. https://doi.org/10.1016/j.hrmr.2017.01.001

Pauliene, R., Diskiene, D., & Matuzeviciute, E. (2019). Complex approach on multicultural teams management & leadership. *Journal of Educational and Social Research*, *9*(2), 8–16. https://doi.org/10.2478/jesr-2019-0008

Paunova, M. (2017). Who gets to lead the multinational team? An updated status characteristics perspective. *Human Relations*, *70*(7), 883–907. https://doi.org/10.1177/0018726716678469

Ridgeway, C. L. (2001). Gender, status, and leadership. *Journal of Social Issues*, *57*(4), 637–655. https://doi.org/10.1111/0022-4537.00233

Ridgeway, C. L., & Erickson, K. G. (2000). Creating and spreading status beliefs. *American Journal of Sociology*, *106*, 579–615. https://doi.org/10.1086/318966

Russo, M., Islam, G., & Koyuncu, B. (2017). Non-native accents and stigma: How self-fulfilling prophesies can affect career outcomes. *Human Resource Management Review*, *27*(3), 507–520. https://doi.org/10.1016/j.hrmr.2016.12.001

Salk, J. E., & Brannen, M. Y. (2000). National culture, networks, and individual influence in a multinational management team. *Academy of Management Journal*, *43*(2), 191–202. https://doi.org/10.5465/1556376

Sarker, S. (2005). Knowledge transfer and collaboration in distributed US-Thai teams. *Journal of Computer-Mediated Communication*, *10*(4), JCMC1043. https://doi.org/10.1111/j.1083-6101.2005.tb00278.x

Sedikides, C., Gaertner, L., & Vevea, J. L. (2005). Pancultural self-enhancement reloaded: A meta-analytic reply to Heine (2005). *Journal of Personality and Social Psychology*, *89*(4), 539–551. https://doi.org/10.1037/0022-3514.89.4.539

Sherf, E. N., Sinha, R., Tangirala, S., & Awasty, N. (2018). Centralization of member voice in teams: Its effects on expertise utilization and team performance. *Journal of Applied Psychology*, *103*(8), 813. https://doi.org/10.1037/apl0000305

Singh, B., Winkel, D. E., & Selvarajan, T. T. (2013). Managing diversity at work: Does psychological safety hold the key to racial differences in employee performance? *Journal of Occupational and Organizational Psychology*, *86*(2), 242–263. https://doi.org/10.1111/joop.12015

Stahl, G. K., Maznevski, M. L., Voigt, A., & Jonsen, K. (2010). Unraveling the effects of cultural diversity in teams: A meta-analysis of research on multicultural work groups. *Journal of International Business Studies*, *41*(4), 690–709. https://doi.org/10.1057/jibs.2009.85

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*(1), 1–39. https://doi.org/10.1146/annurev.ps.33.020182.000245

Tenzer, H., & Pudelko, M. (2017). The influence of language differences on power dynamics in multinational teams. *Journal of World Business*, *52*(1), 45–61. https://doi.org/10.1016/j.jwb.2016.11.002

Tröster, C., Mehra, A., & van Knippenberg, D. (2014). Structuring for team success: The interactive effects of network structure and cultural diversity on team potency and performance. *Organizational Behavior and Human Decision Processes*, *124*(2), 245–255. https://doi.org/10.1016/j.obhdp.2014.04.003

Tröster, C., & van Knippenberg, D. (2012). Leader openness, nationality dissimilarity, and voice in multinational management teams. *Journal of International Business Studies*, *43*(6), 591–613. https://doi.org/10.1057/jibs.2012.15

Tung, R. L. (1998). American expatriates abroad: From neophytes to cosmopolitans. *Journal of World Business*, *33*(2), 125–144. https://doi.org/10.1016/S1090-9516(98)90002-5

Ulusoy, N., Mölders, C., Fischer, S., Bayur, H., Deveci, S., Demiral, Y., & Rössler, W. (2016). A matter of psychological safety: Commitment and mental health in Turkish immigrant employees in Germany. *Journal of Cross-Cultural Psychology*, *47*(4), 626–645. https://doi.org/10.1177/0022022115626513

United Nations Development Programme. (2020). *The next frontier Human development and the Antropocene*. AGS. https://report.hdr.undp.org/

van Knippenberg, D., De Dreu, C. K., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, *89*, 1008–1022. https://doi.org/10.1037/0021-9010.89.6.1008

Walumbwa, F. O., & Schaubroeck, J. (2009). Leader personality traits and employee voice behavior: Mediating roles of ethical leadership and work group psychological safety. *Journal of Applied Psychology*, *94*(5), 1275–1286. https://doi.org/10.1037/a0015848

Woodson, A. N. (2020). Don't let me be misunderstood: Psychological safety, Black girls' speech, and Black feminist perspectives on directness. *Journal of Educational Psychology*, *112*(3), 567–578. https://doi.org/10.1037/edu0000458

Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods*, *14*(3), 183–201. https://doi.org/10.1037/a0015858

Author Biographies

Samuel Farley is an associate professor of organizational psychology at Leeds University Business School. He received his PhD from the University of Sheffield. His research focuses on workplace mistreatment and teamwork.

Jeremy Dawson is a professor of health management at the University of Sheffield. He received his PhD from Aston University and his research focuses on healthcare staff and statistical methods.

Thomas Greenaway is a skills developer at the University of Warwick. He received his PhD from the University of Warwick. His research interests include culture and teamwork in higher education.

Keelan Meade is head of research, evaluation and insights at the Digital Poverty Alliance. He received his PhD from the University of Sheffield. His research interests include environmental psychology and team behavior.

Daría Hernández Ibar is a knowledge exchange associate at the University of Sheffield. She received her PhD from the University of Sheffield and her research interests include emotion regulation, team behavior and leadership.