Team Dynamics and the Marshmallow Challenge: studying team performance and personal satisfaction with a focus on verbal interactions

Hanna Daoudy and Michel Verstraeten

The present study analyses the impacts of verbal interactions as well as the team’s international diversity on team performance and on team members’ satisfaction during a game called the Marshmallow Challenge. Ninety-one students from a business school participated in the game, forming twenty-three teams. The purpose was to construct the highest freestanding structure with 20 sticks of spaghettis and a marshmallow on top. Participants only had eighteen minutes to achieve this goal. The variables were measured through observations and through individual questionnaires. Results show that verbal interactions played a critical role on both performance and satisfaction. Teams where some of the members spoke more than others were more likely to achieve higher performance. Members in these teams were also more satisfied regarding the team outcome. Furthermore, open discussions in teams decreased the members’ communication process satisfaction. Finally interesting results appeared in international teams. For instance, the average level of anger and frustration was highest in these teams. This in turn had an impact on personal satisfaction. More specifically, the team’s international diversity affected negatively the members’ communication process satisfaction. Taken together, these findings show that communication strongly affected performance and satisfaction and it significantly influenced members’ willingness to remain in the same team. Despite these observations, the current study presents some limitations that will be discussed and that should be taken into account for further research.

Keywords: team performance, team members’ satisfaction, verbal interactions
Team Dynamics and the Marshmallow Challenge: studying team performance and personal satisfaction with a focus on verbal interactions

Hanna Daoudy
Michel Verstraeten

Author Note

Hanna Daoudy, Solvay Brussels School of Economics and Management, Université Libre de Bruxelles
Michel Verstraeten, Solvay Brussels School of Economics and Management, Université Libre de Bruxelles

Correspondence concerning this article should be addressed to Michel Verstraeten, Av. F. Roosevelt 50 – CP 114/3, 1050 Bruxelles,
Contact : mverstra@ulb.ac.be
Abstract

The present study analyses the impacts of verbal interactions as well as the team’s international diversity on team performance and on team members’ satisfaction during a game called the Marshmallow Challenge. Ninety-one students from a business school participated in the game, forming twenty-three teams. The purpose was to construct the highest freestanding structure with 20 sticks of spaghetti and a marshmallow on top. Participants only had eighteen minutes to achieve this goal. The variables were measured through observations and through individual questionnaires. Results show that verbal interactions played a critical role on both performance and satisfaction. Teams where some of the members spoke more than others were more likely to achieve higher performance. Members in these teams were also more satisfied regarding the team outcome. Furthermore, open discussions in teams decreased the members’ communication process satisfaction. Finally interesting results appeared in international teams. For instance, the average level of anger and frustration was highest in these teams. This in turn had an impact on personal satisfaction. More specifically, the team’s international diversity affected negatively the members’ communication process satisfaction. Taken together, these findings show that communication strongly affected performance and satisfaction and it significantly influenced members’ willingness to remain in the same team. Despite these observations, the current study presents some limitations that will be discussed and that should be taken into account for further research.

Keywords: team performance, team members’ satisfaction, verbal interactions,
Team Dynamics and the Marshmallow Challenge: studying team performance and personal satisfaction with a focus on verbal interactions

Individuals in society are more and more subject to teamwork in a challenging environment that requires collaboration with the team members. Several issues arise when people are gathered together, which could then have an impact on the quality of the outcome on the tasks they are asked to do. Many studies and analyses explored the dimension of team outcome with the aim of understanding what team processes occurred and enhanced team success (Kozlowski & Bell; Sanna & Parks in Balkundi & Harisson, 2006). Teams often fail to achieve their potential due to faulty processes such as coordination and motivation losses (Steiner in Brown, 2000).

When it comes to studying teams, more attention has been paid to the evaluation of team outcomes and results than to the interactions that produced them (Keyton & Beck, 2008). Despite a prevalent level of research in this area, a lack of empirical evidence remains when it comes to assessing whether and how communication and team diversity are related to team outcomes. Important gaps prevail in the understanding of these relationships (Kearney, Gebert & Voelpel, 2009). Therefore, the present study sheds light on these team processes that may affect team outcomes. More specifically, it examines the impact of verbal interactions and the team’s international diversity on team performance and personal satisfaction. Finally, it emphasises the high time pressure to which members are exposed. The dimension of time has indeed been strongly neglected in the research of teams (Kozlowski & Bell in Mohammed & Nadkarni, 2011), and should be kept in mind as the level of verbal interactions may play a significant role under this pressure.
We based our research on a team experiment – the Marshmallow Challenge – that was created and introduced in 2002 by Peter Skillman and that has been tested worldwide since then by an award-winning innovator called Tom Wujec. Inspired by Peter Skillman, Tom Wujec aimed back then to understand what made a team more performing than another based on team composition. Therefore he tested the challenge with different categories of teams such as teams of CEOs, teams of architects, teams of engineers, teams of business students and teams of kindergarten children.

The game consists in building, in teams of four, the highest freestanding structure with 20 sticks of spaghettis, one yard of tape, one yard of string and a marshmallow on top in eighteen minutes. The challenge exposes team members to a design situation where the degree of uncertainty and time pressure is high.

In the present study, the Marshmallow Challenge has been tested with twenty-three teams of four, composed of master and exchange students from the Solvay Brussels School of Economics and Management. Ninety-one students participated in the challenge.

**Theoretical framework and hypotheses**

As mentioned, the aim of the present study is to analyse the impact of verbal interactions and team diversity on two team outcomes that are the team performance and the team member satisfaction. This section introduces and defines the main variables used in this paper. In addition, it highlights some prevailing gaps that exist between past theories. Finally it formulates new hypotheses based on these theories and empirical research.

With this structure in mind, we start this introduction by turning to a well-known approach in the study of teamwork that is the input-process-output model (IPO) of performance. It is one of the most common approaches that have been adopted in the research of teams (Ilgen et.al. 2005 in Nijstad 2009). According to the IPO model, the performance is
an outcome that depends first on design factors put in place such as group composition, organizational contexts and task designs. Second it depends on team processes.

The IPO model needs however to be treated carefully as it presents specific limitations. For instance it does not consider interdependencies between variables. In addition it has been argued that certain dimensions of the model called “processes” should be rather defined as features of the team that arise as a result of interactions between the members involved. Finally, the reverse causal relationship between the variables is not necessarily taken into account (Forsyth, 2009). Despite these limitations, the IPO model is still worth mentioning as it provides a basic framework for the analysis of Team Outcomes. Nevertheless, the current research goes beyond this model by considering the interdependency between the variables (such as interactions and team diversity).

**Team Performance (TP)**

This paper defines “team performance” (TP) as the successful achievement of a task in a certain period of time. TP can be referred to an “interactionist phenomenon” (Davis J.H.; Laughlin in Stangor: 184, 2004), as it is a function of both the team members’ skills and the way they combine these competencies in the team itself (Davis J.H.; Laughlin in Stangor, 2004).

Human interactions are all composed of two essential components: the content and the process (Hanson in Biech, 2008). While the first is concerned with the task itself, the second deals with the members’ behaviour. It deals with how individuals are interacting together in the team. In that context, dimensions such as influence, participation, conflict and leadership emerge in team processes (Hanson in Biech, 2008). Furthermore, team processes are defined as being the “interdependent acts” taken by the members to transform inputs into outputs (Marks, Mathieu & Zaccaro in Nijstad: 167, 2009). These processes consist in cognitive, behavioural and verbal actions with the aim of getting organized in order to
achieve the required goals (Marks et al. in Nijstad 2009). In this context our first aim will be to study the association between verbal interactions and TP.

When individuals are gathered together in teams, they tend to interact together through both verbal and non-verbal communication with the aim of achieving a common goal (Forsyth, 2009). Robert F. Bales (1950, 1999) concluded that there are two types of interactions: relationship and task interactions (Bales in Forsyth 2009). His investigations and development of the interaction process analysis (Forsyth, 2009) go beyond the scope of the present paper. However, it is important to point out that these interactions have an influence on the members’ actions. In this matter, communication plays a critical role in team performance as it enables the exchange of information between team members (Pinto & Pinto in Hoegl & Gemuenden, 2001). One dimension of communication that is emphasized in this research is the number of verbal interactions. The number of verbal interactions is defined as “the number of times each member makes a verbal contribution” (Gray & von Broembsen, Doreian in Freeman, White & Romney: 157, 1992).

Based on Tom Wujec’s findings, one key to success of the Marshmallow Challenge was prototyping: he gave the examples of kindergarten children who immediately took action by making prototypes. Moreover, he claimed that one issue that arose amongst business graduates was that they spent too much time planning and organizing during the first minutes of the game. The time pressure is critical, and members do not have much time to spend during the decision-making process. Based on these ideas, it can be argued that the level of verbal interactions may affect the level of performance, especially during such a short period of time. In addition, a difference in the number of verbal interactions between the beginning of the game and the end of the game may affect the final score.

Considering these theories and findings, three first hypotheses are tested in the present study:
Hypothesis 1: “The total number of verbal interactions within the team is negatively correlated with team performance”.

Hypothesis 2: “The percentage of verbal interactions within the team during the first six minutes of the game\(^1\) is negatively associated with team performance”.

Furthermore, being part of a team requires the involvement of every member in the process (Biech, 2008), especially in the case of the Marshmallow Challenge where not only verbal contributions are required but also physical contributions. As Tom Wujec said, the challenge demands that all team members collaborate quickly (Wujec, 2010). In that matter, both participation and feelings of belonging (membership) are believed to enhance TP. Moreover, a good balance of the members’ contribution has an impact on the quality of the task (Hoegl & Gemuenden, 2001).

It can be argued that one member’s willingness to participate in the game depends on how the other members of the team include him/her in the interactions. Similarly, the feeling of membership is concerned with the degree of acceptance and inclusion in the team (Forsyth, 2009). Thus if the member has a sense of belonging and membership to the team, he/she should be more willing to put extra effort into the task (Hanson in Biech, 2008). Additional studies have found a positive correlation between group cohesion and team performance (Stogdill in Nijstad, 2009). However, this association is assumed to be only positive when all the members have agreed upon how they are going to accomplish the task (Podsakoff, MacKenzie & Ahearne in Nijstad, 2009). In that case, they will be more motivated to achieve the final goals.

Those ideas of feeling of membership and inclusion lead to the following hypothesis:

Hypothesis 3: “The percentage of verbal interactions directed to more than one member (no one-to-one communication) is positively associated with team performance”.

\(^{1}\) The higher the percentage, the more they talk during the first minutes of the game compared to the last minutes.
We would be tempted to formulate another hypothesis based on a positive possible link between the balance of individual verbal contributions and TP. But through Tom Wujec’s experiment, he stated that teams composed of CEOs and executive administrators were the most performing teams due to the executives’ skills in managing the process. Based on these results, the presence of one or two members taking the lead in organizing and managing the process could still enhance the performance of the team. This suggests that a concentration of the verbal interactions on one or two members would foster TP. We’re thus facing two contradictory performance factors. However, considering that in this case the time pressure plays an important role, we would consider an hypothesis giving favour to the second factor:

Hypothesis 4: “A significant difference between each team member’s number of verbal interactions\(^2\) is associated with team performance”

After considering the associations between verbal interactions and TP, the present study examines the impact of team diversity on team performance. Despite prevalent investigations, the relationship between both variables is not straightforward. Results related to the effect of diversity on TP are contradictory and still remain unclear (Badke-Schaub, Goldschmidt & Meijer, 2010). The current paper does not only aim to examine the impact of team diversity on TP, but also looks on certain interactions and emotions that may occur in these teams which in turn could affect TP.

Generally speaking, team diversity (TDIV) is defined as the variation among team members on any characteristic that can be emphasized in order to differentiate individuals (Larson JR., 2010). For instance, teams can be diverse due to differences between members in gender, in national background, in age, in disciplines and more. Thus, TDIV can be

\(^2\) Teams in which all members do not speak in the same amount
interpreted in several different ways and these differences are likely to have different impacts on the overall performance (Larson JR., 2010).

As teams are becoming more and more diverse in workplaces, it is important to examine the association between TDIV and performance (Shin, Kim, Lee & Bian, 2012). Team diversity has been described in the literature as a “two-edged sword” (Milliken & Martins in Canella Jr., Park & Lee, 2008) due to empirical research that predicts both positive as well as negative impacts of TDIV on TP (Canella Jr. et al., 2008). This dual aspect (positive or negative) will depend on the context. In addition, previous studies show differences in the associations between TDIV and TP depending on whether TDIV is referred to task-oriented diversity (ex. Education) or relations-oriented diversity (ex. gender, race/ethnicity) (Joshi & Roh, 2009).

The current paper investigates the impact of the team’s international diversity (TIIntDIV) on TP. In that context, it takes into account the members’ nationalities by measuring the total number of different nationalities present in the team. It is the only dimension of TDIV referred to in this study.

The concept of “national culture” is critical in the study of diversity. National culture has been at the heart of attention in the study of Human Sciences, especially since the results of Hofstede’s dominant work in this field in 1980 (Leung, Bhagat, Buchan, Erez & Gibson, 2005). Generally speaking, national culture refers to “values, beliefs, norms and behavioural patterns of a national group” (Leung et al., 2005). The deep study of culture and its variance through different nationalities goes beyond the scope of this paper. However it is still worth mentioning as the latter contributes to the explanation of differences in national background. Thus, in line with these reflexions, it can be argued that individuals with a different national background (and as a result, with a different national culture) may behave in different ways, which may in turn affect the team’s performance. For instance, empirical studies have found
differences in commitment and communication behaviour in teams due to differences in cultural values (Kirkman & Shapiro, 2001). TIntDIV could be both beneficial as well as detrimental to the team performance. On the one hand, it can be beneficial as the team gathers individuals with different point of views (due to their differences in background). Therefore, it could provide a wider range of bright ideas and insights (Jehn; Jehn, Nortcraft & Neale in Badke-Schaub, Goldschmidt & Meijer, 2010). On the other hand, it can be detrimental because people with different cultures and languages may face difficulties in communicating and interacting with each other. Once again, we are facing two opposite factors and, again, we consider the context of high time pressure, under which communication obstacles are more likely to occur. This leads to the next hypothesis:

Hypothesis 5: “The number of different nationalities within the team impacts negatively team performance”.

To conclude this section, it is important to highlight that there are many difficulties in assessing team performance. For instance it can be questioned whether a failure in achieving high performance is due to one or more of the members involved, thus the unit of analysis can be problematic to evaluate (Brannick, Salas & Prince 1997). Furthermore, there are always some doubts regarding the validity and reliability of measures that are used (Dickinson, McIntyre in Brannick et al. 1997). These elements will be taken into account in the limitation section of this paper.

Team Members’ Satisfaction (TMS)

The second variable that is studied in this paper is the team members’ satisfaction (TMS) throughout the game. According to the IPO model of performance mentioned in the previous section, TMS is described as a team outcome (Ilgen et al. in Nijstad, 2009). It is an affective outcome that arises at the level of the individual (Hackman in Nijstad, 2009). Moreover, some empirical investigations have demonstrated that TMS leads to additional
outcomes such as higher levels of commitment and more importantly higher levels of performance (Mathieu & Zajac in Kirkman & Shapiro, 2001). TMS is considered as an important dimension of team effectiveness (Gladstein & Hackman in Stark & Bierly 2009). Thus, personal satisfaction should not be neglected in the research of team dynamics.

The present study emphasizes three dimensions of TMS: team outcome satisfaction (TOS), team communication process satisfaction (TCPS) and the members’ willingness to remain in the same team again in the future (REMAIN). While the first dimension (TOS) concerns the degree of the members’ satisfaction related to the outcome of the teamwork, the second (TCPS) focuses on members’ satisfaction related to verbal and non-verbal interactions between members during the game. Finally, the third dimension (REMAIN) can be linked at some level to the team’s cohesion in the sense that the latter indicates the degree to which an individual is willing to remain in the same team (Cartwright in Hoegl & Gemuenden, 2001).

Taking into account verbal interactions, it can be argued that TMS may depend on how team members interact with each other. Several empirical investigations have demonstrated the existence of a strong association between cooperation, communication and team members’ satisfaction (Pinto et al. in Hoegl & Gemuenden, 2001). Research shows that TMS is influenced by teamwork quality due to gains in knowledge and new skills (Hoegl & Gemuenden, 2001). Teamwork quality can be conceptualized as the quality of both task-related and social interactions between members (Hoegl & Gemuenden, 2001). These theories and reflexions prove that the dimension of satisfaction is a complex concept to study due to its strong dependency on both team- as well as individual-specific characteristics. Therefore, this paper aims to bring some contribution to previous literature by analysing the impact of specifically verbal interactions and the team’s international diversity on TMS. We analyse first verbal interactions.
Open communication and frequent face-to-face interactions are related to feelings of shared identity and responsibility (Guzzo & Shea in Stark & Bierly, 2009). This provides a sense of purpose and belonging to the team members. In addition, goal interdependence can be related to the idea of shared responsibility among team members (Guzzo & Shea in Stark & Bierly, 2009). This in turn, may enhance their motivation to participate and contribute to the final goal. Furthermore, empirical research found that paying attention to the other team members and being sensitive to others’ needs are related to TMS (Henderson, 2008). Therefore, open communication and a good balance of the team members’ contribution in terms of verbal interactions could be related to TMS. Two additional hypotheses are tested:

**Hypothesis 6:** “The percentage of verbal interactions directed to more than one member (No one-to-one communication) is positively associated with (a) TOS, (b) TCPS and with (c) REMAIN.”

**Hypothesis 7:** “A significant difference between each team member’s number of verbal interactions is negatively associated with (a) TOS, with (b) TCPS and with (c) REMAIN.”

The last dimension that is studied is team diversity (TDIV), and more precisely the number of different nationalities present in each team (TIntDIV), and its impact on TMS. As discussed earlier, TDIV impacts the interactions between members due to differences in cultural values and languages (Kirkman & Shapiro, 2001). Tdiv can thus be detrimental because people with different cultures and languages may face difficulties in communicating and interacting with each other. Therefore, international teams could face more views, interests and values clashes (Badke-Schaub, Goldschmidt & Meijer, 2010). Thus TIntDIV can be subject to internal breakdowns and interpersonal conflicts (Greening & Johnson; Hambrick & D’Aveni; O’Reilly, Snyder & Boothe in Canella Jr., Park & Lee, 2008). As a result, this may lead to higher degrees of frustration and anger among team members. Therefore it would be interesting to measure the impact of TIntDIV on the three dimensions
of TMS, but also on a fourth dimension which is the level of anger and frustration during the game.

Based on these reflexions, the following hypothesis is proposed:

Hypothesis 8: “The number of different nationalities within the team impacts (a) TOS, (b) TCPS, (c) REMAIN and (d) the team’s average level of anger and frustration”

To close this section on TMS, the present paper analyses whether the three main levels of satisfaction (TOS, TCPS and REMAIN) are interrelated. More precisely, the aim is to understand which dimension of satisfaction (TOS, TCPS) plays a critical role on the degree to which the member is willing to remain in the same team (REMAIN). After analysing whether a significant association exists between the three levels of satisfaction, the current study raises a last question:

Question 1: “which dimension of satisfaction (TOS or TCPS) has the most impact on the members’ willingness to remain on the same team for future projects (REMAIN)?”

The figure below summarizes the proposed model of the present study:

**Figure 1- Proposed Model**
Method

Sample and procedure

As mentioned earlier, the experiment in the present paper consisted of a team game called the Marshmallow Challenge. The participants were master students and exchange students from the Solvay Brussels School of Economics and Management (SBS-EM). Business students have been identified as the worst performing team category in Tom Wujec’s experiment. In total, 91 students participated in the challenge. In addition, cameramen and one photographer were present that day to film and take pictures of the participants during the game. The teams were composed of four members and one observer except one team that was only composed of three members and one observer. Thus 23 teams (N=23) were competing against each other.

The whole experiment took place during an “Organizational Behaviour” lecture. A short introduction was given in class before the start of the game, explaining the goals and the rules of the challenge. After the instructions delivery, 114 students received a team number and role (team member or observer) by drawing lots. Thus, the teams were randomly composed. While the team members went down in the Atrium where the game took place, the twenty-three observers remained ten minutes longer in class in order to carefully explain to them what indicators they had to observe. Therefore, team members had no idea what the observers were measuring and observing during the game.

The participants were sitting on the floor where a place was specifically attributed to each member of the team. Each team was sitting in a circle enabling all members to communicate with each other. Scotch tape was used to form squares on the floor that would then be assigned to each member. The “seats” were placed at an equal distance from each
other. The observer sat on the floor as well but outside of the circle. The team members were not allowed to talk to the observer.

As explained earlier, the Marshmallow Challenge lasted eighteen minutes, in that matter a video projector was used to display the countdown time. The organizer whistled each six minutes to inform the students of the time left. At the end of the game, all the structures were measured with a measuring tape, while all the participants went back to class to fill out an individual questionnaire. Finally, the winning teams were announced in class. No reward was given.

Each team had received a bag containing twenty sticks of uncooked spaghettis, 1.45 meters of string, one tape role, one piece of marshmallow and a pair of scissors. The goal of the game was for each team to construct the tallest freestanding structure and to place the entire marshmallow on top of it, using all the equipment mentioned above. The participants were allowed to use as much or as little of the material, and were free to break the spaghettis, the string and the tape, if needed. They were not allowed to talk to members of other teams. As said previously, they all had 18 minutes to achieve the goal. Once the time ran out, no one was allowed to touch the structure anymore.

Four sources of data were collected after the game:

- Individual questionnaires (filled out by the participants)
- Observers’ sheets (with the coding of verbal interactions)
- Ratings of the final structures

Measures

The Pearson correlation coefficient was used in order to measure the strength and direction of the associations between the variables (Stangor, 2004). For simplification reasons, the analysis consisted in two-tailed tests (Agresti & Agresti, 1979).
Team performance (TP) was the first team outcome to analyse. In the present study, the performance is defined by the height of the structure expressed in centimetres. It was measured with a measuring tape from the surface of the floor to the top of the marshmallow. The taller the structure, the better the performance of the teams. The unit of analysis for this dimension was the team. Thus TP was measured for the 23 teams (Number of observations, N=23).

The team member satisfaction (TMS) was the second team outcome to analyse. This variable was measured through individual questionnaires that were distributed to the students right after the game. In the present study, the satisfaction consists of three mains dimensions:

The participant’s satisfaction regarding the team outcome (TOS)
The participant’s satisfaction regarding the team communication process (TCPS)
The extent to which the participant was willing to remain in the same team in the future (REMAIN)

As mentioned previously, we also measured Teams’ average level of anger and frustration as a fourth dimension of team satisfaction. This dimension specifically intervenes in the study of some impacts of team diversity.

For each dimension, the participant was required to set a score between 1 (very low/very unlikely) and 7 (very high/very likely). The unit of analysis for the three first dimensions of satisfaction was the participant. Thus TMS was measured for each of the 91 participants (N=91). However, regarding the fourth dimension, the team average level was computed in order to measure the overall level of anger/frustration. Thus this variable was not considered at the individual level, but instead at the team level.

Teams’ average level of anger and frustration - This variable was measured through the individual questionnaires. The members were asked to rank their “level of anger/frustration during the game” on a scale of 1 (Not at all) to 7 (Very much so). After
obtaining this level for the four members of each team, the team average level was computed in order to measure the overall level of anger/frustration. Thus this variable was not considered at the individual level, but instead at the team level.

*Verbal Interactions* - The unit of analysis of the verbal interaction to be coded was a sentence composed of minimum two words. When the subject said only one word, the observer was required to code the letter with a bullet point under it. These coding rules had as a purpose to increase the reliability of the measurements by making replication possible. Reliability occurs when the information provided by the indicator (in this case the observer) will not vary depending on the indicator’s attributes (Neuman 1997).

*Total number of verbal interactions* - The observers were required to code the verbal interactions between team members on their sheets of paper. A letter (A-B-C-D) was attributed to each member in order to enable the observer to code the source and the direction of the verbal interactions. On the paper sheet, a row corresponded to each subject. In theses rows (A-B-C-D), the observer was required to code the letter of the member to which the initial member was talking. Thus the row represented the source and the letter that was coded in the row represented the direction of the verbal interaction. For instance, if the subject A was talking to the subject B, the observer had to code in the row “A” the letter “B”.

*Total number of verbal interactions during the three periods of time* - The time of the game was divided into three periods of six minutes. Each observer had received three sheets of paper, each sheet corresponding to six minutes of the games. The instructor whistled four times during the game in order to mark the three periods of time. Thanks to the observers’ task, it was possible to compute the total number of interactions during these periods of time and then to express, for each period, this level of verbal interactions as a percentage of the total number of verbal interactions during the eighteen minutes. This had as a purpose to
determine if the members were talking more at the beginning or at the end of the game, and then to analyse if it had an impact on TP.

Percentage of verbal interactions directed to more than one member - When one person was talking to more than one member, the observer had to code the verbal interaction as a “+”. Thus the percentage of verbal interactions directed to more than one member equals the total number of “+” divided by the total number of verbal interactions and finally multiplied by 100 in order to obtain the percentage. This procedure was aimed to understand how many participants were involved in the discussion and there on analyse this impact on TP.

Difference between each team member’s numbers of verbal interactions (VI) - This variable was obtained by measuring for each team, the variance of the members’ verbal contributions. The members’ verbal contributions correspond to their number of VI during the game. Thus, this total number of VI was computed for each of the team members on the observers’ sheets. Then the variance of these four numbers of VI (linked to the four members) was computed in order to determine if there was a considerable difference in the number of interactions between the four members. Thus the higher the variance, the less were the verbal interactions equally split between the members.

Number of different nationalities - The respondents had to mention their nationality in the questionnaires. As a result, it was possible to compute the number of different nationalities present in each team. This number could be between 1 (All members had the same nationality) and 4 (None of the members had the same nationality). The variable was therefore measured at the team level.

Results

Team Performance
As explained, TP was measured at the level of the team. Thus the team was the unit of analysis for the following results; therefore the number of observations (N) was 23. Fourteen teams out of twenty-three succeeded to build a structure in eighteen minutes. The team performance scores (expressed in centimetres) are as follows: 89; 88; 87; 84; 71; 70.5; 67.5; 63; 51; 47; 46; 39; 25; 0; 0; 0; 0; 0; 0; 0; 0.

Regarding the two first hypotheses, no significant correlations were found. Indeed, the positive correlation (r=.236) between team performance and the total number of verbal interactions was not significant (p=.277). Similarly there was no significant correlation between TP and the percentage of verbal interactions during the first six minutes although the signs of the Pearson correlation coefficient was in the hypothesized direction (r=-.087; p=.692). Hypotheses 1 and 2 are therefore rejected.

In hypothesis 3, it was predicted that there would be a positive association between the percentage of verbal interactions directed to more than one member and team performance. Results failed to support this assumption and showed instead a non-significant correlation between these variables (r=.011; p=.961).

Hypothesis 4 examined the impact of the difference between each team member’s numbers of verbal interactions on team performance. Results showed a significant correlation between these two variables. Thus this hypothesis was consistent with this finding due to the positive association between them (r=.43; P<.05). This shows a positive impact on TP when one or two team members speak more than the others.

Closing this section on TP, hypothesis 5 studied the association between the number of different nationalities within the team (TIntDIV) and TP. The findings could only partially
support this hypothesis that predicted a relationship between the number of different nationalities within the team and TP, due to the low level of significance (r=-.233; p=.285).

**Team Members’ Satisfaction (TMS)**

The 6th hypothesis tested the effect of the percentage of verbal interactions directed to more than one member on TMS. The impact on the first dimension (TOS) was not significant. The correlation was r=-.86 (p=.417). As a result, hypothesis 6a could not be supported. Moreover, hypothesis 6b predicted a positive association between this variable and TCPS. Despite the significant correlation found between these variables, the association was surprisingly negative. Thus the hypothesis was rejected due to the sign of the correlation (r= -.251; p<.05). Finally, the negative effects of the percentage of verbal interactions directed to more than one member on REMAIN was not significant and was equal to r=-.078 (p=.460). Hypothesis 6c was therefore rejected as well.

Furthermore, a difference between each team member’s numbers of verbal interactions had a positive and significant impact on TOS (r=.259; p<.05). Hypothesis 7a was therefore rejected, as it would predict an opposite direction in the association between these variables. In addition, the results showed that the positive effect of this variable on TCPS was not significant. This positive correlation was r=.070 (p=.507). The results showed that the positive effect (r=.080) of this variable on REMAIN was insignificant (p=.450). As a result, hypotheses 7b and 7c were rejected.

Hypothesis 8 examined the relationship between the number of different nationalities within the team (TIntDIV) and the three main levels of TMS and the average level of anger and frustration. This variable had nearly no effect on TOS, and the negligible correlation was insignificant (r=-.092; p=.385). The hypothesis 8a could not be verified. Hypothesis 8b was consistent with the results, as TIntDIV was associated with TCPS. The correlation was negative and significant at the 0.05 level (r=-.206; P=.050). TIntDIV had a small negative
effect on REMAIN. This negative correlation was however insignificant and amounted to \( r = -0.082 \) (\( p = 0.441 \)). Finally, in contrast, the average level of anger and frustration within the team was significantly correlated with TIntDIV (\( r = 0.424 \), \( p < 0.044 \)). This result was in agreement with hypothesis 8d.

The last aim of the current paper was to study the relationships between the three levels of TMS. Therefore, it first analysed the association between team outcome satisfaction (TOS) and team communication process satisfaction (TCPS). Second, it examined whether TOS and TCPS had an impact on team members’ willingness to work with the same team again in the future (REMAIN). Results showed that TOS and TCPS were strongly and significantly correlated (\( r = 0.373 \), \( P < 0.01 \)).

Furthermore, both the TOS and TCPS had a significant and a positive effect on REMAIN. The respective correlations were \( r = 0.359 \) (\( P < 0.01 \)) and \( r = 0.642 \) (\( P < 0.01 \)). Interestingly, when the variable TCPS was controlled, the association between TOS and REMAIN did not remain significant anymore (partial correlation \( r = 0.169 \), \( P = 0.111 \)). In contrast, the association remained strong and significant between TCPS and REMAIN when TOS was controlled (partial correlation \( r = 0.586 \), \( P < 0.01 \)).

The figure below summarizes the main conclusions.

**Figure 2 -** Tested Model (non-significant results are not represented)
Team Performance

- Total number of verbal interactions
- Percentage of verbal interactions within the team during the first six minutes of the game
- Percentage of verbal interactions directed to more than one member
- Difference between each team member’s number of verbal interactions
- Number of different nationalities within the team

H4 +

TOS

H6b -

TCPS

H8b -

REMAIN

H8d -

Level of anger and frustration
Discussion

The present paper examines the impacts of verbal interactions and the team’s international diversity first on team performance and then on team members’ satisfaction.

Team performance

As mentioned earlier, the results showed that only fourteen teams out of twenty-three succeeded in constructing the structure in eighteen minutes. This could reflect the fact that one of the main difficulties of the challenge is the time pressure the members were subjected to.

As explained earlier, the first hypotheses (1,2,3,4) studied the association between verbal interactions and team performance (TP). The results demonstrated a lack of association between (1) the total number of verbal interactions and team performance (TP), (2) the percentage of verbal interactions during the first six minutes and TP and finally between (3) the percentages of verbal interactions directed to more than one member and TP. Thus, these findings suggest that the level of verbal interactions during the game did neither enhance nor thwart team performance. The non-significant correlations could be explained by the inclusion of both single words and sentences in the verbal interactions. For instance, single words such as “ok”, “great” are less likely to have a significant impact on TP.

Furthermore, all the verbal interactions were not necessarily related to the task itself. Finally, it can be argued that instead of having a direct effect, verbal interactions could have had an indirect impact on TP through non-verbal interactions. The level of non-verbal interactions could indeed have played a significant role as the task was mainly manual and thus required a manual effort from its team members. Furthermore, the prediction that the percentage of verbal interactions directed to more than one member would have an effect on TP was not
supported by the results. Thus the members’ feeling of being included in the team and their willingness to participate in the task could be explained by other factors that may then have had an impact on the performance. In that context, we can consider dyadic communication as the “basic form of communication” (Smith in Douglas: pp.70, 1983). Empirical research has demonstrated that individuals are not capable in engaging with more than one individual in a “genuine dialogue – total reciprocity – in an existential moment” (Smith in Douglas: pp.71, 1983). Therefore, it is argued that dyadic interactions are indispensible to group processes (Smith in Douglas 1983). Thus the formation of subgroups in the team could have occurred during the Marshmallow Challenge, still giving to the members a feeling of belonging and willingness to participate efficiently to the task. Unfortunately, this effect on TP remains uncertain and goes beyond the scope of this paper.

Hypothesis 4 examined the impact of the difference between each team member’s numbers of verbal interactions on team performance. Results supported this hypothesis and showed a positive and significant relationship between these two factors. The presence of some members who spoke much more than other members of the team can translate a high variance in the number of interactions. Different interpretations to this finding can be made. On the one hand, it came in agreement with Tom Wujec’s conclusions regarding the skills needed to manage the task. Tom Wujec stated indeed that teams of CEOs were the most successful when executive administrators were present thanks to their facilitation skills in managing the process. Thus, in the present case, it can be interpreted that, a high variance in the number of verbal interactions could lead to better performance due to the presence of one or two members that spoke more than others and succeeded to direct the team toward the final goal.
On the other hand, the fact that some members spoke less does not necessarily mean that they participated less. Team members could still have been involved and participated through non-verbal interactions.

In that manner, the positive effect of a good balance in members’ contribution on TP (Hoegl & Gemuenden, 2001) lacks of evidence in the present study as it has only been measured through verbal interactions. It is thus subject to further research.

To close this section on TP, the last dimension to analyse was related to team composition: TIntDIV.

As stated earlier, previous theories argued that a difference in nationalities among members could have an effect on performance due to differences in cultural values and in some case due to language obstacles. These effects are believed to impact communication between individuals as well as their emotions, which could in turn affect TP.

The present findings could not support the prediction that TIntDIV would directly impact TP due to the lack of significant association between the number of nationalities present in the team and the performance (hypothesis 5). Therefore, theories related to team diversity’s impact on performance mentioned earlier could not be totally observed hereby. Several explanations come to mind.

National background represents only one dimension of diversity. Some literatures argue that the influence of diversity functions in different ways depending on the criteria that is being observed (Larson, JR., 2010). For instance, assuming that national background and personality traits are related, diversity based on relationship-related traits could impact TP differently than diversity based on task-related traits (Larson, JR., 2010). Therefore, it can be considered that more diversity criteria should be simultaneously taken into account in order to observe direct effects on TP.
Moreover, some theories state that detecting visible and evident differences in cultural values and attitudes is only possible with extended interactions overtime between members (Larson, JR., 2010). Thus, these extended interactions were unlikely to occur during such a short time as in the case of the Marshmallow Challenge.

Finally, previous literature argues that additional elements should be considered in the study of team diversity. For instance, considerable attention has been paid to the dimension of team interdependency in the research of diversity. For instance, team members can be interdependent because they rely on each other when accomplishing the task due to the task’s requirements (Joshi & Roh, 2009). The final impact of diversity on TP may depend on these levels of team interdependency. Past research suggests that interdependency enhances members’ commitment and gives a sense of common goal, differences between members are put aside in these situations (Gaertner & Dovidio in Joshi & Roh, 2009). To pursue in this thinking, task interdependency during the Marshmallow Challenge may have varied from team to team depending on how they got organized. For instance, some teams may have delegated roles during the challenge, decreasing levels of interdependency. Thus, T

To sum up, these theories show that team diversity is a complex dimension to treat as it is composed of several different criteria. Each of them may have (or not) different effects (direct or indirect) on team performance. Therefore, more light must be shed on this area and requires more research in the field of teamwork.

Team members’ satisfaction (TMS)

The second outcome was the team member satisfaction (TMS). This paper examines first the impacts of verbal interactions on this outcome.
The findings stated that verbal interactions directed to more than one member had only negative effects on TCPS. This contradicted hypothesis 6 that assumed positive impacts on the three levels of TMS. This result could suggest that members seemed more satisfied by the communication process in dyadic conversations instead of open ones. The feeling of shared identity and responsibility (Stark & Bierly, 2009) could however still remain in subgroups and thus enhance higher levels of satisfaction. In addition, dyadic conversations tend to force members to discuss more and give ideas. In this case, two-by-two interactions could have eased the communication process and gived a higher feeling of involvement, and as a result a higher degree of TCPS.

Furthermore, the results didn’t support the assumption that differences between each team member’s numbers of verbal interactions were negatively associated with the three levels of TMS. Thus the findings were not in agreement with hypotheses 7(b) and 7(c). Surprisingly, hypothesis 7(a) was significantly related to TOS but the result showed an opposite direction of association. Different suggestions could be brought to this positive association. First, members could be satisfied with the team outcome despite the difference in verbal contributions between the members. In this scenario, the member that talked the most may have managed the whole process well and at the same time may have given a feeling of involvement to the others, which in turn may have positively affected their satisfaction. Secondly, as discussed before, differences in verbal contributions do not necessarily mean a difference in non-verbal contributions. For instance, a member may contribute less in verbal interactions but may contribute the most in non-verbal interactions. Thus, the positive association between the variance in verbal interactions and TOS may be spurious. In that case, a third omitted variable (non-verbal contribution) could explain the relationship that has been found.
After examining the impacts of verbal interactions on TMS, this paper covers the dimension related to team diversity. In that matter, hypothesis 8 studied the impact of Team International Diversity (TIntDIV) on four dimensions of team member satisfaction: TOS, TCPS, REMAIN and the team’s average level of anger/frustration. Interestingly, the number of different nationalities within the team (TIntDIV) only had an impact on two levels of satisfaction: TCPS and the team’s average level of anger/frustration. Results indeed revealed a negative and significant association between TIntDIV and TCPS. This finding comes in agreement with the literature, which argues that the communication process is influenced by team diversity due to differences in behaviour and cultural values (Kirkman & Shapiro 2001). The negative relationship between these two variables can be explained by the difficulty to interact verbally with members that speak a different language, or that have a different cultural background. These barriers and clashes may lead to team conflicts due to misunderstandings between members and, as a result, may imply a decrease in members’ satisfaction regarding the communication process. These reflections are in line with previous empirical research that showed the negative association between conflicts and TMS (De Dreu & Weingart, 2003).

In addition, the results demonstrate the TIntDIV had a positive impact on the level of anger/frustration. Thus, this level was highest in international teams. In that context, it can be considered that this negative emotion played an intermediary role in the relationship between TIntDIV and TCPS. In that case, this finding supports the suggestion made above that international teams might be subject to team conflict due to communication obstacles. More specifically, affective conflicts can be mentioned in this context as they are linked to socio-emotional and personal issues between members (Jehn, Nortcraft & Neale in Badke-Schaub, Goldschmidt & Meijer, 2010). They create overall tension within the team and decrease the
quality of communication (Amazon & Sapienza in Badke-Schaub, Goldschmidt & Meijer, 2010).

Finally, the results revealed insignificant associations between (1) TIntDIV and TOS and between (2) TIntDIV and REMAIN. Different reasons come to mind. Nationalities constitute only one dimension of team diversity and are thus not enough in observing direct and significant impacts on these two levels of satisfaction. The effects and attributes of diversity are indeed not equal (Joshi & Roh; Williams & O’Reilly in Shin, Kim, Lee & Bian, 2012).

Furthermore, the global concept of diversity may not be strong enough in observing direct impacts on TOS and REMAIN levels. Participants may include several criteria in these evaluations such as the score, the familiarity with the other members, personality traits, task characteristics and more. In other words, the team outcome satisfaction and REMAIN are subjective dimensions. They both cover several elements. Thus, participants may have used different criteria in their evaluation and measurement of satisfaction.

The last aim of the current paper was to study the relationships between the three levels of satisfaction. Then it was to understand which dimension (TOS or TCPS) had (if any) the most impact on the members’ willingness to remain on the same team for future projects (REMAIN) (Question 1). The results demonstrated that the association between TOS and REMAIN did not remain significant when TCPS was controlled; this finding states that the relationship might be spurious as these two variables were both explained by the TCPS. Thus these reflexions lead to consider that the communication process between members played an important role in their overall satisfaction and in their willingness to remain in the same team again in the future.

**Limitations, implications and future research**
Before drawing any conclusions, some important limitations must be kept in mind. This study presents several constraints that may explain some differences between the current results and prior research. The constraints are the following:

**Sample:** The size of the sample (23 teams) covered by this paper limits the significance of the results. However, this paper aimed to bring some additional contributions to previous research. The findings should be replicated in larger samples and in different countries in order to ensure the generalizability of the results.

**Verbal Interactions measurement:** the content and the length of the verbal interactions had not been taken into account by the observers. Thus the interactions that had been coded are not necessarily linked and relevant to the decision that had to be made in the team (Stangor 2004).

**Omitted variables:** as discussed earlier, there could be some omitted variables that could have had effects on both team performance and satisfaction leading to spurious correlations between the variables (Glaeser & Scheinkman, 1999). For instance, non-verbal interactions have not been taken into account in the present study. These interactions could have played a critical role in the relationships studied in this research paper.

**Time and Pressure:** Team members had only eighteen minutes to build a structure with a marshmallow on top of it. Thus members had not much time to discuss how they were going to achieve the task. There was too little time to gather all good suggestions and valuable ideas. In addition, the pressure on team members could justify some biases with previous research. The countdown displayed on the projector and the whistling each 6 minutes put the participants under high levels of stress. This environment could have had additional impacts on team performance and TMS.

**Team members:** This study did not take into consideration the team members’ personality traits. These characteristics could have played a role in the associations between
the variables that were studied. For instance, it could have had an influence on members’
degree of satisfaction. In addition, some individuals prefer to work in team than alone, and
therefore would find more satisfaction in interacting with other members (Campion et al.;
Shaw et al. in Stark & Bierly, 2009).

Teams: Previous studies have shown that teams evolve in time and develop through
different stages. Tuckman developed the most dominant model in 1965 (Bonebright, 2010).
According to his model, team members have different behaviours during the different stages
of development (Miller, 2003).

Furthermore, research demonstrates that many group processes, such as decision-making,
cohesion, positive interactions might be time relevant (Miller, 2003). In that manner, team
dynamics could have negative impacts on team outcome during some periods of time but
could be beneficial at other periods of time (Miller, 2003). Related to these researches, an
additional model (the “integrated model of group development”) established by Wheelan
shows how the team members’ behaviour varies over time (Wheelan, 2003). For instance,
this model states that, during the first stage referred to inclusion and dependency, members
are looking for safety and trust (Wheelan, 2003). They rely heavily on dominant members
that could guide them into the task accomplishment. Conflict, open-communication, higher
levels of cooperation occur in later stages (Wheelan, 2003). Thus team processes that occur
in these different stages may have an impact on performance and satisfaction.

As the Marshmallow Challenge was limited in time, the different team development stages
and temporal sequences could not be observed in the study. Hence the differences in team
member behaviour over time could not be analysed.

Therefore, the timing of team processes and its impact on performance and satisfaction
should be considered in future research.
Conclusion

The current paper studied team performance (TP) and members’ personal satisfaction (TMS) during a game called the Marshmallow Challenge. Ninety-one students from a business school participated in the game, forming twenty-three teams. The challenge consisted in building the highest freestanding structure with twenty sticks of spaghettis and one marshmallow on top in eighteen minutes.

In comparison to Tom Wujec’s research, the aim of the present study was to analyse specific team processes related to verbal interactions, and to determine their impacts on both TP and TMS.

The Marshmallow Challenge consists of a very difficult task. High concentration and members’ involvement are required to efficiently build a structure that must stand by itself. As expected, the current results showed that verbal interactions played a critical role in the analysis of performance and satisfaction. They took place between team members and influence the process of the task accomplishment. In addition they had consequences on persons’ behaviour and wellbeing. Verbal interactions constitute the body of this paper as they are interlinked to a wide range of dimensions covered by the present study (ex. diversity).

The current findings showed that teams where some of the members spoke more than others were more likely to construct a stand-alone pyramid in time. Members in these teams were also more satisfied regarding the outcome. Higher levels of organization and pressure management within these teams could explain these findings in addition to the close relationship between performance and team outcome satisfaction. Finally, the team’s average level of anger was highest in the international teams. Nevertheless, there was no significant and direct relationship between the team’s international diversity and team performance.
Analysing the team members’ satisfaction, the current paper first covered the impact of verbal interactions. In that matter, it is important to highlight that verbal interactions had not necessarily the same impacts on team outcome satisfaction (TOS), team communication process satisfaction (TCPS) and members’ willingness to remain in the same team (REMAIN).

For instance, open discussions in teams decreased members’ satisfaction but only regarding the communication process. Furthermore the team’s international diversity had a negative impact on TCPS. One explanation to this effect could be the high levels of anger/frustration within these teams due to obstacles in the communication process.

This demonstrates that communication played an important role and strongly influenced members’ satisfaction and willingness to remain in the same team in future projects.

Interestingly, there were differences among the three levels of satisfaction.

Members satisfied about the team outcome were not necessarily willing to remain in the same team. This shows that performance is not the only dimension that matters, and it is also important to consider members’ satisfaction as this contributes strongly to team cohesion.

Finally, the current paper presented several limitations that should be taken into account in further research. These limitations could explain some biases between previous research and the present results. For instance, a lack of evidence remains regarding the team diversity. While TIntDIV was not the first concern of this paper, it does still merit some attention and emphasis for future research related to team performance and personal satisfaction.

Finally, the study showed that the dimension of time, pressure and context might play an important role in the assessment of team processes. Their influence on TP and TMS should therefore not be taken for granted.

Acknowledgements
Special thanks to all the participants of the Marshmallow challenge and to Professor Patricia Garcia-Prieto Chevalier for authorizing the challenge to take place, for her help in the organization of the game, her support and contribution in the writing of the individual questionnaires, and finally for her useful advice and recommendations.

References


Neuman W.L. (1997), *Social Research Methods: Qualitative and Quantitative Approaches*, 3rd edition, Allyn and Bacon, Boston


Website: http://vimeo.com/37119711 (consulted on the 18th March 2012)
Skillman P. (2007), video: Talk show “Gel conference”.
Website: http://sittingo.com/talk/712 (consulted in April 2012)
Wujec T. (February 2010), video: Talk show “Ted. Ideas worth spreading”
Website: http://marshmallowchallenge.com/Welcome.html (consulted in November 2011, February and March 2012)