The Meetology® Laboratory 2012

Behavioural Research Results

Undertaken by the research division of

the meetology® group

exploring the science behind effective meetings

In association with

WORLDWIDE EXHIBITIONS FOR INCENTIVE TRAVEL, MEETINGS & EVENTS

imex GROUP

Because the world never stands still

barbican

مكتب دبي للمؤتمرات

DUBAI CONVENTION BUREAU

pcma Education Foundation

February 2013

INTRODUCTION

This report presents the results of the Meetology® Laboratory research project conducted at the IMEX 2012 Frankfurt exhibition and supported by the Barbican, PCMA Education Foundation, the IMEX Group and the Dubai Convention Bureau.

The Meetology® Laboratory was a live behavioural research project on the show floor of the IMEX exhibition in Frankfurt and was designed to answer the following question: "Does meeting face to face improve creativity in meetings compared to virtual meetings?".

The research was undertaken by the research division of the Meetology® Group whose psychologists conducting an experiment in which pairs of participants worked together on a creative problem solving task. Participants were divided into three separate conditions, face-to-face, video link and voice only. We acquired data on the ideas generated as well as their own psychological experience of the task. Overall it was found that the number of ideas generated, the quality of ideas and the variety of ideas seemed to improve in the face to face condition over the virtual meeting condition. Interestingly, the research indicated that these increases in creativity were not as a result of a more positive or motivating experience, therefore face to face seems to improve creative problem solving, despite no benefit in terms of social experience.

Why creativity?

Most meetings involve the discussion of ideas. These may be focused on thinking of new ways to solve existing problems, trying to generate ideas of new business opportunities or ventures, ideas to overcome a challenge or trying to think of all the possible problems and issues that may arise from a forthcoming event or product launch. It is unusual for meetings not to involve some discussion of people's thoughts and ideas on a topic. Therefore understanding factors that improve or hinder the generation of ideas and the quality of ideas generated is essential for organisations.

Why different environments?

Examining the impact of different environments or meeting types have on creative thinking is essential. As businesses aim to reduce costs, or simply to improve communication, more organizations are moving towards virtual meetings using either voice or video platforms. However, the impact that these different technologies have on the success of the meeting is not fully understood. This research project focused on three of the most popular ways people meet – i.e. face to face, via voice (phone) or via video.

AIMS

The research project aimed to investigate whether the way in which a meeting is conducted (e.g. virtually or face to face) impacts upon creative problem solving outcomes. This research examines three different means of communication and their impact on creative problem solving. It was predicted that

creative problem solving would be improved in face-to-face meetings over virtual meetings. The following research questions are examined.

Research Question 1
*Does interacting face to face improve the **number** of creative ideas you have over virtual meetings?*

Research Question 2
*Does interacting face to face improve the **quality** of the ideas you have over virtual meetings?*

Research Question 3
*Does interacting face to face improve the **variety** of ideas you have over virtual meetings?*

Research Question 4
*Does interacting face to face improve the **social experience of interaction** over virtual meetings?*

METHOD

When conducting research it is important to understand the tension between conducting a research project that is as "real" as possible, whilst controlling for as many extraneous variables as is necessary. For research to be of practical and applied benefit it should reflect real applications in organizations and businesses. As such it should be conducted in a way that simulates or reflects real world scenarios rather than completely contrived and unusual laboratory surroundings. If the research laboratory did not mimic real life situations then the results would be difficult to generalise to the applied setting. However, there is also a tension when trying to make the scenario as real as possible. Real life is often messy and complicated. For example if you merely asked someone to state whether they thought face to face were better than virtual meetings the results would also include differences in preference for face to face and virtual meetings (some may prefer one style over the other), differences in experience of these meetings and also differences in actually understanding what "better" means in this case - for example, does it mean more productive, more enjoyable or more efficient? In a research setting it is important that as much as possible the 'messiness' of real life is controlled as well as clear objective measurement carried out. This then allows the researcher to ensure that if differences between groups or conditions arise then they can rule out alternative explanations and ensure that the results are reliable and valid as clear objective measures have been used.

In the Meetology® Laboratory at IMEX 2012 it was decided to run three groups, which mimicked the three key ways that people meet. These are face to face, via video conference and via voice (phone). Furthermore, it was important to choose tasks that individuals would not have experience in as to ensure that differences were not accounted for by differences in experience or expertise.

Participants

One hundred and four participants took part in the study. The participants were all attendees at the IMEX 2012 exhibition, which took place over three days in July in Frankfurt, Germany. Anyone at the exhibition had the opportunity to take part. Participation was not limited to exhibitors or visitors. Participants were invited to take part whilst they passed the laboratory and were asked on an opportunistic basis. The laboratory and study was also advertised via twitter and through other social media aimed at the delegates at the event. Participants received no incentive to take part apart from the fact that this was a research project examining the impact of different types of communication. No financial incentive was involved.

It is likely that the participants themselves were a self-selected group who may be more motivated to take part or engage in this type of activity. However, as participants were randomly allocated to any one of the conditions it is unlikely that this had any detrimental impact on the results as they were just as likely to be in a virtual meeting condition as a face to face condition.

Procedure

Participants were invited to take part in the study and assured that this would take no longer than 10mins. Participants were selected in pairs (either known or unknown) to work together on the task. The participants were given the following instructions:

> *The research involves taking part in a problem solving exercise with another participant. You will work together to generate solutions to a problem. Participation in this research is voluntary and anonymous. Personal details will only be kept to allow you to withdraw your data; no data will be published in a way that could identify individual participants. You may withdraw at any time or omit the answers to any questions you do not wish to answer.*

Participants were then given the opportunity to consent to take part, or decide not to. If the participants agreed to take part they were then given the following information.

> *In a moment you will be given a problem or issue, which you are required to generate some solutions or possible issues that arise from the problem. There is no correct answer. This study involves generating as many ideas as you can. You will be scored on the number of ideas that you produce as well as the innovativeness of those ideas, not whether they are right or wrong.*

This demonstrates that all participants were encouraged to produce as many as possible and as innovative as possible ideas. They were then given the following instructions.

> *You will be required to work together during the task to help produce your ideas. You will have THREE MINUTES in which to work together to generate as many creative ideas as possible. We will*

be monitoring and making notes during this time, but will not be able to answer questions. Please do not be distracted by us and try and focus on the task.

You will be required to write down all the ideas and solutions you have. We will collect these at the end. We will tell you when your time will begin and end.

The participants were randomly allocated to one of three conditions. These were as follows:

1: VIDEO CONDITION
In the video condition participants sat in front of a computer screen with an attached headset. The participants had a barrier between them and could only see the other participants via the video screen and could only hear them via the headset. They were given the following instructions.

You will be working with your partner over a video screen on a monitor in front of you. We will ensure that this works and is set up prior to beginning the study. You will have headphones and a microphone to hear and talk to each other. Please try and focus on this task and ignore any other background noise you may hear. You will not be required to use the computer keyboard for anything but should just write down the ideas and answers that you have.

Participants were required to write down on paper the ideas that they had generated in the task.

2: PHONE CONDITION
In the phone condition participants sat in front of a blank computer screen with an attached headset. The participants had a barrier between them and could not see the other participants and could only hear them via the headset. They were given the following instructions.

You will be working with your partner over a headphone set via computer and only hear the voice of your partner. We will ensure that this works and is set up prior to beginning the study. You will have headphones and a microphone to hear and talk to each other. Please try and focus on this task and ignore any other background noise you may hear. You will not be required to use the computer keyboard for anything but should just write down the ideas and answers that you have.

3: FACE TO FACE CONDITION
In the face to face condition participants sat face to face either side of a desk. They were given the following instructions.

You will be working with your partner over a desk. Please try and focus on this task and ignore any other background noise you may hear. You should write down the ideas and answers that you have.

When the participants had been allocated to their condition they were then randomly allocated one of three creativity tasks. The three tasks were chosen to cover a range of creative thinking types, just

suppose, unusual uses and product improvement. These were drawn from the validity Torrance Tests of Creative Thinking tasks. All these involve the generation of innovative ideas. Each is described below.

Task

1: JUST SUPPOSE – CLOUDS WITH STRINGS
In this condition participants were given the following instructions.

> *In your imagination, just suppose that the following situation were to happen. Think of all of the other things that would happen because of it. In other words, what would be the consequences? Make as many guesses as you can.*

The improbable situation – JUST SUPPOSE clouds had strings attached to them which hang down to earth. What would happen? List your ideas and guesses on the blank page. ***Think about as many possible implications as you can.***

2: UNUSUAL USES - CARDBOARD
In this condition participants were given the following instructions.

> *Most people throw their empty cardboard boxes away, but they have lots of interesting and unusual uses. On the blank page list as many of these interesting and unusual uses as you can think of. Do not limit yourself to any one size of box. You may use as many boxes as you like. Do not limit yourself to the uses you have seen or heard about;* ***think about as many possible new uses as you can.***

3: PRODUCT IMPROVEMENT – STUFFED ELEPHANT
In this condition participants were given the following instructions.

> *Imagine a stuffed toy elephant of the kind you can buy in most children's shops for about £5. It is about the size of two fists on top of each other. On the blank page, list as many interesting and unusual ways you can think of for changing this toy elephant so that children will have more fun playing with it. Do not worry about how much the change would cost. Think only about what would make it more fun to play with as a toy.* ***Think about as many improvements as you can.***

The proportion of people in each condition and allocated to each task is shown in Table 1 below. Please note that a smaller number took part in the voice only toy condition. A χ^2 analysis was conducted to examine the distribution of participants to the conditions and tasks. This would demonstrate whether there was a significant uneven distribution. The results demonstrated no significant differences between the number of people in each condition and each task ($\chi^2 = 5.318$, df = 4, p = .256).

Table 1. Percent of participants in each condition and task				
		Task		
		Toy	Cardboard	Clouds
Condition	Face-to-face	17.3%	15.4%	9.6%
	Video-Voice	11.5%	7.7%	7.7%
	Voice only	5.8%	7.7%	17.3%
Total		34.6%	30.8%	34.6%

After participants had completed the 3 minute exercise they were required to complete a short questionnaire (described below), thanked for their time and given information on what the study was about and what happens next with the data.

Measures

Creativity

The Torrance Test of Creative Thinking tasks can assess a number of aspects of creative thinking. These are all important to fully understand the impact that the environment may have on creative thinking. These can be categorised broadly as testing the *number of ideas generated* and the *quality of ideas generated.* Below each criteria is described, these cover *fluency* (the total number of ideas generated), *originality* (the innovativeness of the idea) and *flexibility* (the number of different types of ideas generated).

Number of ideas

Fluency assesses the total number of ideas generated. This is the sum total of the number of ideas generated irrespective of the quality of the ideas.

Quality of ideas

In order to assess the quality of ideas each idea generated needed to be categorised. A total of 545 ideas were generated by the participants (Cardboard responses = 211, Toy responses = 178, Cloud responses = 158).

These total responses were then coded. Coding involves the categorisation of the responses into meaningful categories of similar responses. For example in the cardboard task the responses “burning for fuel” or “setting light to for a fire” would be categorised as similar using the category “burning/heat”. In the toy task the responses of “enlarge trunk”, “make trunk bendable”, “stretch” or “extend trunk” would be categorised as “change shape or size”. The full list of the items and the coded categories are available on request.

Once each item had been categorised, these were checked by an independent researcher for consistency and accuracy.

Originality

Originality is measured by the number of statistically infrequent ideas. Ideas that are relevant (i.e. an appropriate response to the task) but few people generate will be scored as high on originality, conversely common ideas are those which will be scored low on originality. For example if the question was name a grey animal, and the response elephant was very common this would obtain a low originality score, but if the response Koala or African Grey Parrot was uncommon, these would demonstrate high originality. This demonstrates the ability of an individual, pair or team to produce uncommon or unique responses. The scoring procedure counts the number of responses produced for each category (as defined above). The most common responses get the code of 1 and then each successive response gets ranked one higher (e.g. 2, 3, etc.). This ranking continues until the most uncommon response receives the highest score. In the case of a number of categories all receiving equal numbers of responses, they will get the average of their ranked score.

These ranked scores are then divided by the total number of different responses so that the highest score achieved is 1 (i.e. if there were 15 different categories of items and only one item was unique that would be scored 15/15 = 1). This is done so that scoring is comparable across tasks. Some tasks produce more ideas than others and if the standardization (dividing by the total number of different responses) was not conducted then the tasks with the largest variety of responses (high flexibility) would be related to the originality scores due to the scoring being related.

Each participant (or pair of participants in this case) then obtain an originality score for each of the items they have produced. There are then two further ways of dealing with these scores. The average originality score is the total of all the originality scores for all of the pair's items, divided by the number of items they generated. This gives us the average amount of originality of the items these participants generated. This means that individuals who produce few, but original items will still score higher than those who produce many, but fairly unoriginal items. The average originality score therefore separates the originality of the items with the number of items produced (fluency).

However, it is also useful to look at the total originality score. This is sum of the originality scores. This then produces a single score for each pair which reflects the total amount of originality of the items they generated in the creativity task. This is useful as some pairs may have generated quite a few quite original items, whereas other pairs may have produced one very original item. In the case of the mean score the participants with one highly original item will score higher than the participants who produced many more fairly original items.

Flexibility

Flexibility measures the number of *different* ideas generated. This criterion demonstrates the ability of an individual, pair or team to produce many different ideas, rather than ideas that are all rather similar. Whereas fluency examines the number of ideas as a whole and originality measures the originality of

each idea, it is useful also to examine the flexibility of the ideas. When many ideas are produced it can be the case that these are produced by individuals or pairs just elaborating on a theme. For example in the clouds task a pair may generate items such as “it will change the amount of rain”, “it will change the amount of snow”, “people may be able to change the weather”, etc. As you can see these are all about weather related ideas. These would be less flexible than a pair of individuals who generated the ideas “the weather may be affected”, “kids could use the strings as swings” and “we could use the strings as a transport system”. In this second case the same number (3) of ideas are generated but this latter set is much more *flexible* in that they cover a wider variety of issues. Therefore the flexibility score reflects the number of different categories of items that were included. The categories were used from the originality scoring. The participants were then scored regarding the number of different categories they had included in their responses.

Other means of coding responses include: Elaboration (the ability to continually add to an idea, demonstrating an ability to develop and elaborate upon ideas), Abstractness (the degree to which individuals go beyond the labelling of an idea) and Resistance to Premature Closure (the degree to which individuals continue to keep an open mind and continually produce new relevant ideas). These were not included in the current study as ideas were not required to be elaborated on and the task had a very limited time frame so premature closure was hardly ever demonstrated.

Post task measures

After participants completed the task they were required to complete a short questionnaire which included a number of measures about their experience of the task. These were as follows.

Vitality

This measured how energised and enthusiastic they felt at the end of the task. This scale included items such “as I feel energized right now” and “At this moment, I feel alive and vital”. This scale demonstrated acceptable levels of reliability $\alpha = .845$.

Intrinsic Enjoyment

This measured how much they had enjoyed the activity. This included items such as “I found the task very interesting” and “Doing the task was fun”. This scale demonstrated good levels of reliability $\alpha = .74$.

Perceived competence

This measured how competent they felt in terms of their abilities in the task. This scale included items such as “I think I am pretty good at this task”, and “After working at this task for a while, I felt pretty competent”. This scale demonstrated good levels of reliability $\alpha = .74$.

Pressure Tension

This measured how much pressure and tension they felt when doing the task. This scale included items such “as I was anxious while doing the task” and “I felt pressured while doing the task”. This scale demonstrated good levels of reliability $\alpha = .78$.

Relatedness

This measured how they felt towards the person they worked with on the task. This scale included items such as "It is likely that this person and I could become friends if we interacted a lot" and "I feel close to this person". This scale demonstrated good levels of reliability $\alpha = .75$.

RESULTS

To examine whether any of the specific creativity tasks resulted in higher levels of creativity a multivariate analysis of variance (MANOVA) was conducted. This demonstrated no significant group difference on any of the key outcome variables (fluency, flexibility, originality), Wilks $\lambda = .77$, $F(5, 88) = 1.231$, $p = .283$. This demonstrates that the three separate creativity tasks resulted in similar responses overall and can therefore be collapsed.

Research Question 1:

Does interacting face to face improve the number of creative ideas you have over virtual meetings?

We investigated the above research question through examining the total number of ideas that the pairs had during the creativity tasks. First we analysed whether there were any differences between the two non-face to face conditions (video and voice). There were no significant differences between these two conditions, $t(28)=-.619$, $p = 5.41$, with both conditions on average generating around 10 ideas (video mean = 10, sd = 3.68, voice only mean = 10.81, sd = 3.51). As there were no differences between these we combined these into a "virtual" group.

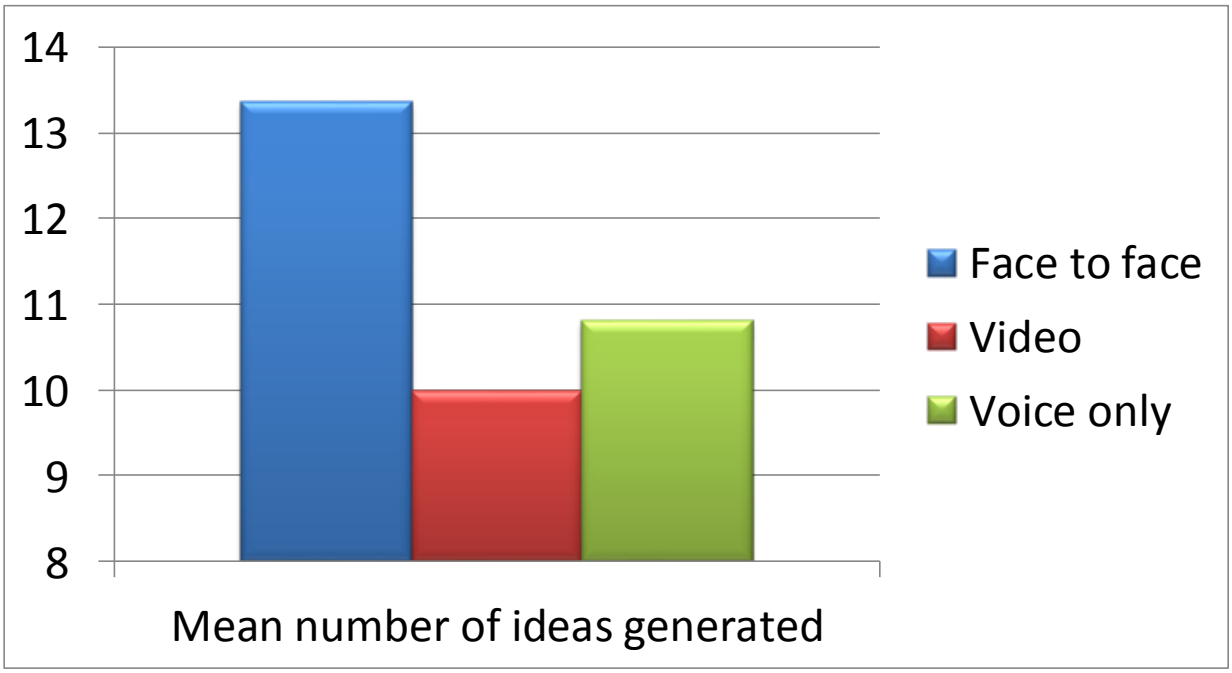
To examine whether face-to-face meetings improved the number of ideas generated we compared the number of ideas generated in the face to face condition with those generated in the virtual condition. We found that on average more ideas were generated in the face to face group than the virtual groups, $t(31.72) = 2.05$, $p = .049$[1]. On average those pairs in the face to face condition generated 13.36 ideas (sd = 5.98) compared to those in the virtual group (mean = 10.43, sd = 3.55). Therefore on average the face to face pairs generated 30% more ideas than virtual pairs. Cohen's D effect size indicates and effect of .6, which is regarded as medium or moderate size effect of face to face in improving number of ideas generated over virtual meetings.

These results were also apparent in the maximum scores (the maximum number of creative ideas generated by a pair). In the face to face condition the highest number of ideas generated by a pair was 29, this is was more than 50% more than the highest number in the voice only condition (18) and 70% more than the highest number of solutions in the video condition (16). Furthermore the differences between the groups are not accounted for by a small number of high scoring pairs in the face-to-face condition (which would result in a non-normal distribution of scores). Both the median and modal number of ideas in the face-to-face group was 13 (compared to a median of 9.5 and mode of 7 in the

[1] Note as the Levene's test for the equality of variances was violated equality of variances not assumed corrected scores are presented.

video condition and a median of 10 and mode of 9 in the voice condition), demonstrating a clear indication that the face to face group did generate more ideas on average.

Figure 1. Mean number of ideas generated by condition.



These results clearly demonstrate that interacting face to face has clear benefits in terms of the number of ideas generated. However, the sheer number of ideas generated does not necessarily mean they are of greater quality. Furthermore, these ideas may just be more elaboration on a single idea. Therefore, when assessing creativity, it is vitally important to assess some indication of the quality of the ideas.

Research Question 2:
Does interacting face to face improve the quality of the ideas you have over virtual meetings?

To examine whether the face to face group had a significantly higher average level of originality an independent sample t-test was conducted comparing face to face with the virtual group. This demonstrated that although the face to face meeting did not reach statistical significance, it approached significance, $t(49) = 1.378$, $p = .08$, one-tailed. However as statistical significance is affected by sample size it is important to look at the size of effect generated. The effect size demonstrated a moderate effect of condition on originality (Cohen's D = .48). This suggests a moderate impact of face-to-face over virtual meetings in terms of the quality of ideas. The results are shown below in Figure 2.

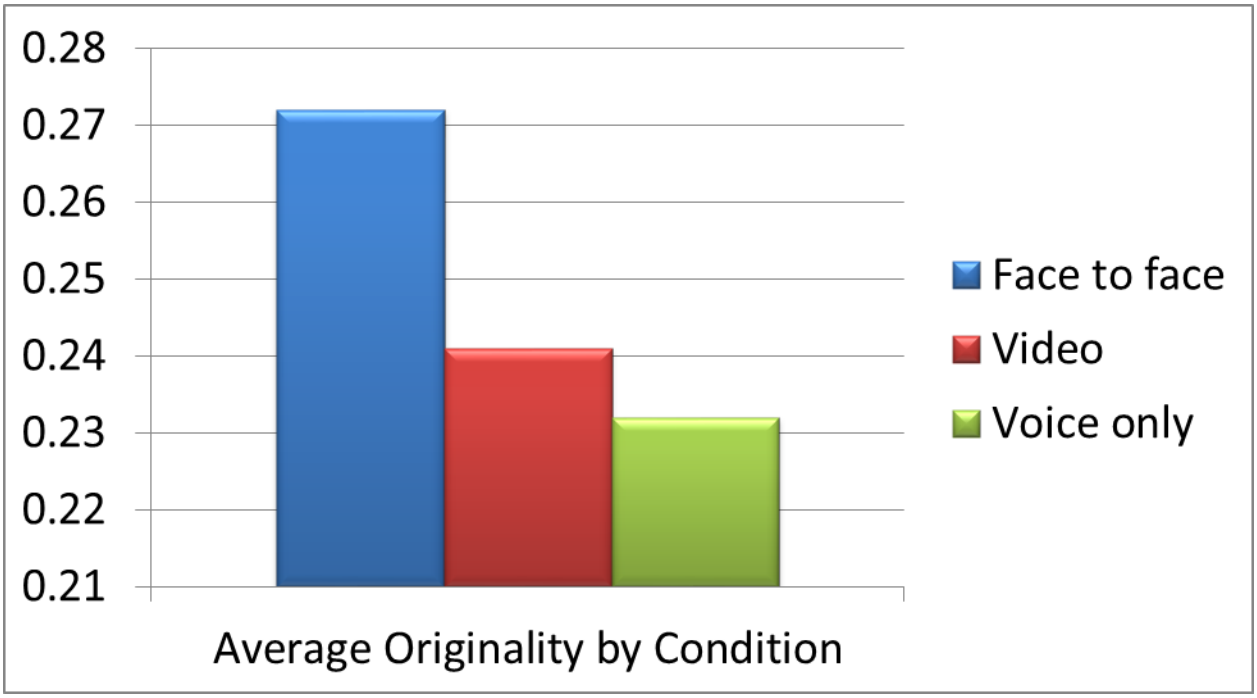


Figure 2. Mean quality of ideas generated by condition.

Research Question 3:

Does interacting face to face improve the variety of ideas you have over virtual meetings?

To examine whether the face to face group had a wider variety of ideas an independent sample t-test was conducted comparing face to face with the virtual group. This demonstrated that the face to face meeting generated a higher variety of ideas than the virtual group, this result again approached statistical significance, $t(49) = 1.519$, $p = .06$, one-tailed. Again the size of effect demonstrated can be considered moderate (Cohen's D = .43). The results are shown in Figure 3 on the next page.

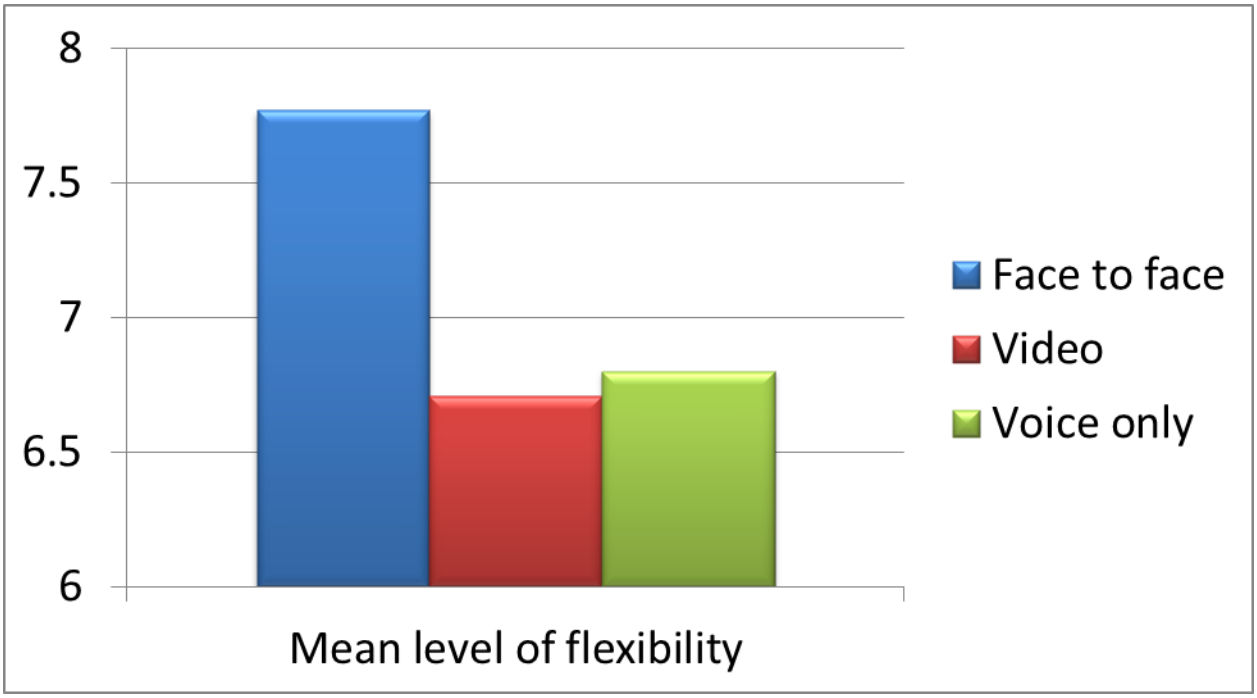


Figure 3. Mean number of different ideas generated by condition.

Interim conclusions

The results so far demonstrate a consistent benefit of face to face meetings on creative problem solving ability. It seems that the face to face meetings improve the number of ideas generated, the quality of those ideas and the variety of those ideas. Furthermore, this seems to be a moderate size of effect. However, it is also important to understand the process through which this increase in ideas operates. In order to begin to examine the process through which face to face meetings might benefit in terms of an increase in creative ideas we examined a number of psychological variables. These were motivation, interpersonal experience and vitality.

Research Question 4:

*Does interacting face to face improve the **social experience of interaction** over virtual meetings?*

4a) Does interacting face to face improve the motivation of the participants?

We assessed three key aspects of motivation, these were whether there were high levels of intrinsic motivation (enjoyment of the task), whether there were higher levels of perceived competency (did they feel better able to do the task) or whether there were differences in the amount of tension (did the task feel pressured or tense). To examine whether there were differences in the conditions between the motivation experienced by the pairs we first again examined whether there were differences between the two virtual conditions on each of the three outcomes (enjoyment, competency, pressure). There were no significant differences between the two virtual conditions on any of these (all p's >.1).

We then compared face-to-face pairs with the virtual pairs on their motivation scores. The results demonstrated that there was no differences in intrinsic enjoyment, $t(101) = -.677$, $p = .5$, competency, $t(100) = .396$, $p = .693$ nor pressure, $t(101) = -.267$, $p = .79$. These results demonstrate that there was no observable difference in the experience of the pairs in terms of their enjoyment, perceived competency nor motivation. This demonstrates that it is not due to the experience of the task by the pairs that may influence the increase in ideas. This also seems to demonstrate that the technology required for "virtual" meetings did not affect the participants experience or motivation of the task. This implies that the differences in number of ideas cannot be accounted for by having to use "technology" which may limit competency experiences or de-motivate. On average all participants found the task quite enjoyable (mean=3.8, sd= .69), felt fairly competent (mean = 3.46, sd = .72) and had low levels of pressure (mean = 2.06, sd = .74).

4b) Does interacting face to face improve the interpersonal perception of people?

To examine whether working face to face was a more positive interpersonal experience we examined the *relatedness* of participating pairs. This measure examined the subjective experiences that the pairs had of each other when working together on the task. The results showed that there were no differences between the two virtual conditions, $t(53) = -.227$, $p = .821$ and so were collapsed. Furthermore there were no differences between the face-to-face and virtual conditions, $t(99) = .495$, $p = .622$. This demonstrated that in the type of interaction did not impact on the subjective impressions of the pairs working together. Therefore although the face to face condition generated more ideas, the pairs did not feel more satisfied or close than those in the virtual conditions. This seems to indicate that it is not the positive experience of the other that increases idea generation in face-to-face meetings.

4c) Does meeting face to face improve the subjective experience of individuals working together?

To examine whether the participants who worked face to face felt better and more enthused a measure of *vitality* was used to examine their direct experience after taking part in the study. This showed that there was no difference in vitality in the two virtual conditions, $t(54) = .884$, $p = .380$. There was also no differences in experience of subjective vitality between the face-to-face and virtual meetings conditions, $t(100)=1.05$, $p=.295$. Effect size indicators did demonstrate a small effect (Cohen's D = .211), with the face to face group experiencing slightly more vitality than the virtual group, although this was only marginal (face to face mean = 3.87, sd = .69, virtual mean = 3.72, sd = .728). Therefore these results demonstrate that there was no difference between the groups in terms of their subjective experience of vitality.

CONCLUSION

These results demonstrate an improvement in the number of creative ideas generated through face to face meetings over virtual meetings. They also demonstrated that there seems to be a moderate effect of improvement in the quality and variety of ideas produced in face to face over virtual meetings. From

the Meetology® Laboratory at least, these increases in ideas and improvements in quality were not accompanied by increases in motivation, interpersonal relatedness nor experiences of enthusiasm or vitality. This seems to suggest that face to face meetings increase creativity, but not because participants have a better psychological experience during these meetings, but due to some other process. Reasons behind this benefit should be explored more fully.

Limitations

We need to acknowledge the limitations in this experiment. First the laboratory itself was "unusual" laboratory. We could not control sound, environment etc. to minimise the effects of these. However, we assume the effects were random across each of the tasks and conditions and therefore did not systematically affect one group over another. Moreover the participants were diverse in terms of age, nationality, background and English ability. The last of these may have impacted on some pairs, particularly where they were not known to each other. Again however, these were randomised across conditions and would therefore have only added to the amount of random and not systematic error in the data. A more important issue in terms of the results were the actual tasks themselves. As the data showed, most participants in the study enjoyed the tasks, did not find them stressful and reported fairly high levels of competency. Given the task was novel in the IMEX 2012 environment this may have also contributed to the "positive experience" of the participants. The impact of this may have been that differences in the psychological experience dependent on the type of interaction participants had may have been overshadowed by the novelty and fun nature of the tasks. Therefore the lack of differences found in the subjective experiences of the participants may be accounted for by this "ceiling effect".

Dr Paul Redford, Consultant Psychologist
Dr Jon Sutton, Consultant Psychologist
February 2013

The Meetology® Group
51-53 Church Rd
Hove
East Sussex
BN3 2DB
www.meetology.com
info@meetology.com
+44 (0)1273 903802