VIRGINIA SANCHINI*
*UNIMI; IEO; IFOM*

SARAH SONGHORIAN
*UNISR*

PIER PAOLO DI FIORE
*IEO; UNIMI; IFOM*

DAVIDE DISALVATORE
*IFOM*

PAOLO SPADA
*Southampton University*

# DELIBERATION AND PUBLIC BIOETHICS: A TEST CASE IN REPRODUCTIVE GENETICS[1]

*abstract*

*Background. Since the nineties, policymakers and theorists working across several disciplines started to entertain the notion of directly engaging the public in matters of public concern. Accompanying this was the attempt to resort to deliberative democracy in order to make such an involvement effective. Seizing on its intrinsic dialogic nature, some scholars proposed the application of public deliberation to the realm of public bioethics. Drawing upon the theory and practice of deliberative public bioethics, the purpose of this paper is to shed light onto the figure of the bioethical expert and its role in public bioethics, and relatedly, to investigate how deliberation may be implemented in a public bioethics context.*
*Methods. We set up a large-scale experiment to investigate whether, and to what extent, different moderation styles impact on participants' moral preferences. The study combines a survey of a representative sample of the general population with a laboratory experiment based on a random sample of students that ex-ante has identical attitudes to the general population.*
*Results. Findings show that: i) different moderation styles can significantly influence deliberative outcomes; iii) the effects of deliberation are not necessarily immediate, but may be revealed after the end of deliberative session; iii) participants tend to better appreciate a bioethical expert acting as "passive moderator", namely as someone who acts in order to ensure non-domination and non-interference, thus allowing the creation of basic conditions for equality within the deliberative setting.*
*Conclusions. Our experiment represents an example of how deliberation can be employed in public bioethics.*

*keywords*

*public bioethics, deliberation, genetic testing, bioethical expertise, moderator*

1 **Acknowledgments**. Many are those who contributed to the design, organization, and conduct of the experiment. As for the interventional study, the authors wish to acknowledge Marco Annoni, Federico Boem, Luca Chiapperino, Giuseppe D'Agostino, Giulia Ferretti, Ilaria Galasso, Regina Gregori Grgic, Maria Nicolaci, Mario Sanchini for having devoted their time and expertise in the observation and moderation of the participants within deliberative sessions; to Emanuela Orlando and all the technicians of the Istituto Firc di Oncologia Molecolare (IFOM) for having helped us in the huge organizational work propaedeutic to the interventional study itself. As for the preliminary experiments, we wish to acknowledge all the professors of the University of Milan, faculty of Medicine, Literature, Philosophy and Political Science, as well as those of the Politecnico di Milano, who allowed us to collect eighty-hundred questionnaires in their classes and during their lectures. A special thanks goes to David Teira for his precious comments on the several versions of this work, and to Luca Marelli for his deep conceptual insights, his thought-provoking comments and suggestions.

Phenomenology and Mind, n. 19 - 2020, pp. 150-193
DOI: https://doi.org/10.17454/pam-1911
Web: http://www.rosenbergesellier.it/elenco-libri?aaidriv=14



ISSN 2280-7853 (print) - ISSN 2239-4028 (on line)

## 1. Introduction

Since the nineties, policymakers and theorists working across several disciplines started to entertain the notion of directly engaging the public in matters of public concern (e.g., Fishkin 1991; Bohman 1998; Moore 2010; Landemore 2012; Neblo 2015). This was accompanied by the attempt to resort to deliberative democracy (henceforth DD) to make such an involvement effective (Gutmann and Thompson 1996; Gutmann and Thompson 2004). Geared towards bringing the core tenets of DD into different contexts of the public sphere, the main aim of these deliberative processes – generally defined under the common, albeit variously interpreted (Blacksher et al. 2012), label of "public deliberation" – is that of eliciting citizens' opinions, while also, in some cases, informing policy-making (Abelson et al. 2012).
Leveraging upon its intrinsic dialogic nature, some scholars have proposed the integration of deliberation into the realm of public bioethics, which is considered to be a field of bioethics dominated by "value conflict and high pressure for decision and regulation" (Moore 2010, p. 715). In particular, seizing on its capacity to constructively deal with (value) conflicts, deliberation was considered, by some, to be a promising tool for addressing moral disagreements of public relevance (Crawshaw et al. 1985; Bowling, Jacobson, and Southgate 1993; Bowie, Richardson, and Sykes 1995; Gutmann and Thompson 1997; MacLean and Burgess 2010; King et al. 2010; Meagher and Lee 2016).
In parallel, the *theoretical* proposal of a *deliberative public bioethics* has been accompanied by the attempt to empirically test deliberation in the context of public dialogue over ethical issues (Abelson et al. 2003a; Abelson et al. 2003b; Abelson et al. 2012; Abelson et al. 2013). These deliberative experiments have been implemented in different forms – from citizens' juries to national issue forums, from deliberative opinion polls to participatory budgeting (Abelson et al. 2003a; Goold et al. 2012, p. 24) – and focus on a wide range of ethically sensitive issues, ranging from priority setting in healthcare to the ethics and regulation of (healthcare) technologies (Abelson et al. 2013).

* Virginia Sanchini[1,2,3], Davide Disalvatore[3], Sarah Songhorian[4], Paolo Spada[5], Pier Paolo Di Fiore [2,1,3]

Authors' affiliations: [1]Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy; [2]IEO, European Institute of Oncology IRCCS, Milan, Italy; [3]IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Milan, Italy; [4]Faculty of Philosophy, Vita-Salute San Raffaele University, Milan, Italy; [5]Department of Politics and International Relations, Southampton University, Southampton, UK.

Corresponding Author: Virginia Sanchini, virginia.sanchini@unimi.it

Drawing upon the theory and practice of deliberative public bioethics, this paper presents *the overarching goals* to shed light on the figure of the bioethical expert and its role in public bioethics, and – relatedly – to investigate how DD may be implemented in a public bioethics context, assuming the principles of DD at face value[1].
The paper sets out from the dominant view–according to which public bioethics can be considered an exemplary case for practically embedding the deliberative democratic ideals (Moore 2010; Rei et al. 2009) –to contend that, in fact, it is deliberation in its role of managing (value) conflicts, that can be considered a useful tool to be employed in public bioethics' settings. To this end, we devised a "validated laboratory experiment" (in line with Abelson et al. 2003a, p. 98), designed according to results obtained through two preliminary surveys[2].
Specifically, *the endpoint* of our study was to empirically investigate whether, and to what extent, different moderation styles – implemented by different figures (active moderator, passive moderator[3], and observer) – would impact upon participants' individual moral preferences and, in particular, on preference shifts. The idea of devising and testing a 'supervisor' – a well-established figure within deliberative experiments, albeit underestimated in its impact (Moore 2010), something that is mostly absent in bioethics literature – represents the most relevant novel element in this experiment. As it will be explained herein, we did not intend to measure the mere shift of preferences (Stewart, Kendall, and Coote 1994; Coote and Lenaghan 1997; Fung 2006; Fishkin 2011), but rather the shift towards what Engelhardt defined as *"the Principle of Permission" (or forbearance)* (Engelhardt 1996) – see section below "Theoretical Background".

The paper is structured as follows. First, we explain the theoretical background upon which the paper builds. Then, we present the rationale, design, and metrics of the study. Next, we present the study results. Finally, we analyse and discuss our findings in relation to four phenomena: i) the apparent lack of impact of information; ii) the role of time in unmasking the effects, iii) the better appreciation of passive moderation over active moderation; iv) the impact of our findings on the bioethical debate regarding bioethics expertise and public bioethics.

### *1.1 Theoretical background*

The "Principle of Permission" (henceforth PoP) is a non-substantive negative principle interpretable as "non-interference". In Engelhardt's view, the principle of permission represents the most fitting principle for a "secular bioethics", i.e. the contemporary bioethical reflection characterised by deep moral disagreement, inhabited by "moral strangers", i.e. individuals endorsing different moral views.
By adopting this view, we measured, as the end point of our study, a shift towards PoP[4]. The

1 We are not claiming that DD is *always* the best tools to address moral disagreements, as this, we do believe, would require a case-by-case discussion that we cannot carry out within the scope of this paper. Rather, we aim to investigate whether it could be a useful tool to address moral disagreements, and its impact on participants' preferences.
2 The preliminary surveys consisted of two pilot studies aimed at defining the topic to be addressed in the main experiment through the analysis of the similarities between the Italian general population and the study population, so as to increase the external validity of the experiment. In this paper we do not report and/or discuss the preliminary surveys.
3 The expression "active" and "passive moderators" is drawn from Farrar and colleagues' paper (Farrar et al. 2009). However, despite the same wording, and for the reasons that will be extensively covered in the discussion section, our connotation of the terms is different from theirs.
4 As already noted, the outcome of our study was not the mere shift of preferences after deliberative session. Following Smith (Smith 2009). Democratic Innovations: Designing Institutions for Citizen Participation. Cambridge: Cambridge University Press), who claims that "in itself, opinion change tells nothing about whether judgments represent *enlightened preferences*" (Ibid: 95), we believed that the mere opinion change after deliberation cannot be considered proof of the success or failure of different interventions, since transformation of preferences *per se* does not prove that the purposes of deliberation have been met.

main justification lying behind this proposal is the following[5]: we argue that this shift is consistent with the purposes that proponents of DD have attributed to deliberation itself: *pluralism awareness*, i.e. to make people aware that the public arena is a domain dominated by moral pluralism – namely, of what Rawls defines as "the fact of pluralism" (Rawls 1993); *pluralism recognition*, i.e. the fact that moral pluralism characterises the public arena should be recognised and endorsed, and that pluralism-oriented discussions and decision-making strategies to cope with it should be put in place.

Once recognized that deliberation intends to fulfill these purposes, to quantify the shift of participants' preferences towards pluralism awareness and recognition, we considered the shift to be in place when participants replied in line with PoP. Indeed, since PoP, as non-interference, is the condition enabling the simultaneous co-existence of a wider spectrum of substantive positions, this appears to be the principle that better serves deliberative purposes. So conceived, PoP is the descendant of the Rawlsian liberal tradition, which draws on the acceptance of moral pluralism as an undisputed fact, and on the consideration that this is the principle that maximises it the most. Accordingly, secular bioethics cannot be regulated by the principle of autonomy, since the latter represents a substantive principle, bound to substantive moral doctrines. Only a non-substantive principle, PoP, safeguards the simultaneous coexistence of different substantive moral positions dominating the public arena.

To summarise, once recognised that the purposes of deliberation are the awareness and recognition of pluralism, we identified PoP as the best means to serve these purposes and, therefore, that the shift towards this principle represents a sign of the success of the deliberation. Accordingly, for the authors of this contribution, claiming that a deliberation is successful if it promotes PoP does not equal the endorsement of any substantive moral stance[6].

## 2. Methods

In what follows, we will discuss materials and methods of the validated laboratory experiment.

### *2.1 Rationale*

Our experiment consisted of a randomized controlled trial (RCT) design based on a template of a laboratory experiment (Figure 1). The overall aim was to observe what happened to participants' preferences after the deliberative intervention and to observe whether the deliberative approach devised here led to positions compliant with the purposes of

5 Actually, there is another reason for adopting a value-laden analysis of the outcome, which is connected with the potential artifact of "inconclusive addition" of the mere shift of preferences. To explain this effect, let us imagine a two-question questionnaire. One individual in the control group (the observed group, in our case) changes her preferences, after the intervention. On one question, she moves one notch towards PoP; on the other, one notch away from it. Her total score would be 2. In another arm of the study, one individual changes preferences moving one notch towards PoP, on both questions. Her score would also be 2. In this "inconclusive addition" the "real" effect, i.e., that one individual moved towards PoP, while the other did not, would be lost. In many political science studies, this caveat would not apply, since the choices are binary (agree/do not agree or in favour/against) and the mere shift suffices. In our study, the adoption of a value-laden 5-point scale allowed this possibility (hence our choice to assign a "direction" to the shift). As an experimental proof of this notion, when we reanalyzed our data without considering the direction, all effects were nullified (data not shown).

6 In the specific context from which we draw our analysis – genetic testing employed in reproduction – it is nevertheless important to point out that, for some specific techniques – namely Non Invasive Prenatal Testing (NIPT) – PoP is compatible with different substantive moral views, spanning those supporting the principle of procreative liberty in its different connotations (Robertson 1983; 1985) to those granting the foetus with an unconditional intrinsic moral value. Indeed, performing NIPS may be considered ethically legitimate by supporters of both these views, as the former may draw from very different intentions and lead to very different outcomes. As an example, the decision to perform NIPS can be grounded in the intention of verifying the health of the foetus to decide whether to interrupt the pregnancy or to gain more knowledge in order to be prepared to properly welcome the newborn, even in cases of genetic disorders. In any case, it is not our intention to endorse a specific moral viewpoint in the paper, especially insofar as the different authors of this piece would endorse very different moral viewpoints.

deliberation, in this specific context to PoP (§1/1.1.). The RCT was designed to accomplish the aim of challenging various modalities of supervision in the deliberative setting applied to bioethics. To this end, the RCT comprised three arms: i) *Observed*, in which a supervisor was present but silent and did not intervene in the discussion; ii) *Passively Moderated*, with a supervisor acting as a promoter of some "negative" deliberative values[7]; and iii) *Actively Moderated*, with a supervisor acting as a promoter of positive and negative deliberative values. The roles and the rules of engagement of these three figures were precisely defined and the supervisors were extensively trained before the sessions (Appendix 1A). In addition, precise lists of "DOs" and "DON'Ts" were provided (Appendix 1A).

*2.2 Experimental design*

274 students were enrolled[8]. The study was implemented through the administration of a questionnaire (Q, Appendix 1B) centred on "*Genetic testing related to reproductive choices*", selected through the Pilot Studies. The questionnaire comprised 10 statements that could be answered using a 5-point Likert scale. The same questionnaire was administered at various time points (Q1 – Q4 corresponding to time points T1 – T4, respectively) (Figure 1).
At Time 1 (T1), the initial preferences of the students were recorded in Q1. Students then received the "Informative Material"[9], consisting of simple, yet accurate, information on the scientific aspects of genetic testing related to reproductive choices (Appendix 1C)[10]. They were then asked to complete Q2 (T2) to measure the impact of information on their preferences. Q2 also included 5 questions for evaluating their comprehension of the Informative Material (Appendix 1D).
Next, students were randomized, through a double randomization process, into the three aforementioned arms[11]. They then attended the deliberative sessions that lasted 90 minutes.

7 By "negative deliberative values" we refer to those values that try to prevent from some group's dynamics to occur (e.g. interference, domination). Accordingly, these values may be defined as "negative" since, rather than promoting some actions/behaviours, they try to limit and/or impede actions and/or behaviours (e.g. try to limit domination dynamics within the group). Differently, by "positive deliberative values" we refer to those values that, rather than simply limiting some dynamics, try to promote certain additional behaviours which should enable group's dynamics (e.g., promote mutual respect, promote equal consideration, etc.).

8 In line with the results of the two preliminary surveys, students were first- or second-year undergraduates from the University of Milan, selected on a voluntary basis. At enrollment, students were only informed of the "bioethics nature of the experiment". Upon arrival, they were received by the experimenter in charge of the study (V.S.), who provided participants with the following information: 1) that the experiment consisted of anonymously recording their preferences on the issue of genetic testing in reproduction; 2) that the preferences would be recorded repeatedly, upon completion of a number of phases – the nature of these phases was not disclosed; c) no information was provided as to the rationale, the aim, and the structure of the experiment; d) participants were instructed to be truthful in their opinions and behaviors, since there was no expected outcome, no right/wrong answers, in order to exclude and/or minimize 'the expectancy effect' (McDermott 2002). Students enrolled received credits; however, credits were not linked to any mandatory course, in order to prevent undue influence.

9 Participants received the Informative Material only during the course of the experiment and not beforehand. This was intended to prevent participants from looking for further information and/or discussing it with others, in order to achieve uniformity in the access to background information and a cleaner measure of the impact of information on the subsequent expression of preference. All the study material was prepared by the researcher responsible for the trial (V.S.), and was then corrected/modified/integrated by different experts – namely, two physicians (one specialized in genetics); five PhD students in philosophy (three of them specialized in bioethics); two statisticians; five professors of philosophy (one of political philosophy, two of philosophy of science; one of political science); two psychologists; a sample of high school students at their last year (therefore of a similar age to the sample). After this, DOXA, a leading Italian polling organization also checked the material (information sheet and questionnaires) to provide a final professional verification.

10 Students had one hour to read the material. During this time, they were not allowed to interact with each other or to use other sources of information.

11 Immediately after T2, students were randomized (without communicating to them the groups to which they had been assigned) into three equal groups. Students within each group were then further randomized into subgroups of 4 or 5 students (Verba 1961; Karpowitz et al. 2012). The study was run over 10 days with a total of 59 subgroups: due to several no-shows on a specific day, one subgroup - in the observed arm - was not formed.

At the end of the deliberative session T3, students were asked to complete Q3, which also included a series of questions aimed at qualitatively evaluating the procedure (Appendix 1E-G). Finally, all students were recalled one month after the deliberative sessions to fill in Q4, to evaluate the impact of time.

*2.3 Metrics*

As mentioned, distinct from similar studies, the endpoint of our study was not the mere change of preferences (Stewart et al. 1994; Coote and Lenaghan 1997; Fung 2006; Fishkin 2011), but the shift towards PoP. To quantitatively estimate the shift, we developed a 5-point scale. A score of 5 was attributed to answers closest to PoP as non-interference (scoring matrix is in Appendix 1H). The quantitative outcome of the RCT was the mean individual change (MIC) towards (or away from) a perspective in line with PoP with respect to the use and implications of genetic testing in the context of reproduction at the time points T2, T3 and T4, relative to the baseline, T1. We calculated the transformation for each student. The effect of the intervention was measured as the difference in pairwise comparisons between the three groups. To calculate the minimum observable difference of MIC between two groups, we used a two-sided t-test. Assuming: i) an enrolment of at least 100 students per group, ii) a significance level of 1%, iii) a statistical power of 80%, and iv) a variance of MIC in each group between 5 and 100, the minimum observable MIC was calculated to be 1.08 and 4.83 for a variance of 5 and 100, respectively. We considered this range in the minimum MIC as a reasonably observable one.
By analysing the MIC between T1 and T2 and by applying Tukey's interquartile rule for outliers to identify poor quality data (Tukey 1977), we identified 31 (11%) students as outliers, defined as external to the median range ± (1.5 x the interquartile range), i.e., score ≤ – 6 or score ≥ 6. These students were excluded from further consideration and all outcomes were calculated on the remaining 243 students (Appendix 1I-J).

## 3. Results

*3.1 Quantitative analysis*

The distribution of answers to questionnaires Q1-Q4 are reported in Appendix 1K. The quantitative analysis of the preferences at the various time points of the study (Table 1), revealed that:

- There were no significant differences at T1 between the three arms of the intervention (Observed, Passively Moderated, Actively Moderated), indicating that the randomization was appropriately conducted (Table 2).
- There were no significant differences in all pairwise comparisons between the three arms in the analyses T2 *vs.* T1 and T3 *vs.* T1. This finding indicates that there were no immediate effects of "information" (T2 *vs.* T1) or of "deliberation" (T3 *vs.* T1) (Table 3, top).
- There was a significant difference at T4 *vs.* T1 in the pairwise comparison Passively Moderated *vs.* Observed (p=0.0019) towards acceptance of freedom in reproductive choices (Table 3, top). No significant differences were evidenced in the comparison T4 *vs.* T1, Actively Moderated *vs.* Observed (see §4).
- All of the above results were confirmed in an independent analysis, where we did not adjust for covariates (Table 3, bottom).

*3.2 Additional analyses*

During the course of the study, participants were asked to answer additional questions at various time points.
At T2, students were asked to answer questions (Appendix 1D) to verify their comprehension of the Informative Material. Setting the threshold of understanding at 3 correct answers out of 5, only 10 participants out of 274 (3.6%) failed to meet it. Raising the threshold to 4 out of 5, 16.1% of participants did not meet the conditions of comprehension. Thus, since 83,9% of participants answered at least 4 out of 5 questions correctly, we concluded the material was comprehensible (p=0.0001**)** (Appendix 1F).

At T3, the additional questions (Appendix 1E) aimed at analysing interactional aspects of the experiment, such as the behaviour of participants, the general tendency of deliberative sessions, the implicit or explicit consensus reached between participants, and so on. In detail, 82% of the subjects found the questionnaire clear or very clear. Furthermore, 90% did not feel at all manipulated. Similarly, 95% felt highly or very highly free to express their preferences within deliberative sessions. Thus, there was almost no perception of any kind of manipulation.
Three questions dealt with the topics of respect, consensus, and transformation of preferences, broadly addressing issues of perceived legitimacy[12]. On the question asking whether the discussion promoted an attitude of higher respect towards the preferences of others, 88% answered "High/Very High".
In addition, on the question exploring whether the discussion was perceived as designed to strive towards consensus, 69% answered "High/Very High" ($p=0.00005$, calculated on the entire distribution with respect to a null hypothesis of equal distribution among the categories); despite the fact that i) consensus-reaching was not the aim of the deliberative sessions, and ii) no indication in this direction was given to participants, the latter, nevertheless, appeared to conceive their task as an attempt to reach a consensus.
Finally, concerning the question related to the transformation of preferences, the majority of subjects did not perceive that they had changed their minds significantly from T1 to T3: 63% answered "Not at all/Small degree". This result is in line with the fact that at T3 no significant differences in the MIC were present. These data suggest that perceived legitimacy was in line with real legitimacy.
To conclude, having adopted a rather stringent criterion to define significance ($p\text{-value}<0.01$), we did not observe significant differences between the three experimental arms, regarding the aforementioned questions. However, in the question regarding manipulation (Q5), a trend towards the actively moderated group ($p=0.011$) was noticed. And yet, if we consider the answers "Not at all" and "Small degree" as different, albeit comparable, proofs of absence of substantial manipulation by the side of the supervising figure, we can consider such a result not significant ($p=0.53$).

## 4. Discussion

The primary endpoint of the study was to describe the effects of deliberation on individual moral preferences in a wide sample of undergraduates representative of the general Italian population. In particular, we wanted to investigate whether deliberation might have led participants to adopt practices more oriented to the awareness and recognition of pluralism than the ones they initially expressed. The secondary endpoint was to investigate whether this shift was emphasized or downsized by different moderation styles.
In the discussion of results, we will focus on the four major study outcomes: the apparent lack of impact of information, the role of time in unmasking the effects, the better appreciation of passive over active moderation, and the impact of study results over the debate around bioethical expertise and public bioethics accounts.

### *4.1 Is there an impact of information?*

The first outcome of the RCT is that providing informative material demonstrated no effect, as illustrated by the lack of change at T2 (Table 3). One could argue that, since it is impossible to

12 Here we draw a distinction between *perceived* legitimacy – i.e. what participants declare as a consequence of their perception –, and *real* legitimacy – i.e. what they declare after having investigated, reflected upon, and rationalized what they have perceived. This distinction rests upon the idea that what participants claim through their preferences does not always correspond to their *considered* preferences; i.e., the preferences that participants would have expressed if they had had enough time and information to reflect upon them (Parkinson 2006).

distinguish between the impact of information and deliberation at T3 and at T4, the fact that an effect was observed at T4 in the passively moderated group (*vs*. the observed one) could be due to some additive effect (over time) of information plus deliberation. In principle, this is a reasonable objection, given that information might have laid the foundations for a less biased approach to the subsequent deliberation, therefore representing an "enabling condition" for the subsequent shift of preferences. However, this potential 'enablement' was unmasked only in the moderated arm (see below), while one might have reasonably expected it in all arms. It remains indisputable that there was no effect of information *in the short term* (from T1 to T2) and this raises a number of methodological questions and caveats. First, we did not measure the impact of information over time *per se*. This may be relevant since time was an important factor in influencing opinions. Second, the time that participants were given to read and to comprehend the information might have been too short (1 hour), even though the comprehension test revealed a satisfactory level of understanding. Third, we did not include sessions with experts, which are frequently part of traditional deliberative studies (Fiorino 1990; Fishkin, Luskin, & Jowell 2000; Abelson 2003b; Fishkin 2011), thereby reducing exposure to information. However, at least the last two modifications were intentionally introduced to test an approach that was less idealized than the ones currently present in the literature and that can hardly be applied in real-life settings. Our design, while retaining the rigor of a laboratory setting to obtain causal inferences from the adopted interventions, was conceived with the intent of moving closer to practice[13].
To conclude, although more work is needed, our hypothesis is that, at least in the close-to-practice setting that we enacted and limited to bioethical issues, deliberation is more effective than information in promoting pluralism.

*4.2 The role of time*

At T3, i.e., immediately after the deliberative sessions, we did not observe any significant effect of the various moderation styles. However, after one month (T4) and in the absence of any other intervention, there was a significant shift towards PoP. The effect was not only highly significant (Passively Moderated *vs*. Observed, p=0.0019), but also of a relevant absolute magnitude. In principle, if each participant had shifted all his/her preferences (i.e., to each single question) one notch towards PoP, we should have seen a MIC difference of 10. In the Passively Moderated group, the change in mean MIC between T1 and T4 was ~1. This means that, on average, with only 90 minutes of passively moderated deliberation followed by one month of "reflection" time, ~10% of all participants embraced a view compliant with PoP on all questions (or that all the participants did so on 10% of the questions).
This is interesting in light of the doubts that have been raised on whether deliberation is a useful learning process, above all when devised as a single event. Chlivers, for instance, reported that those who attended deliberative experiments repeatedly asked to have "enough time [...] to become informed and develop a competent understanding" (Chlivers 2008, p. 174). Similarly, several of our participants expressed the need to extend the time devoted to deliberation in the final Evaluation Questionnaire. Analysing our data, however, rather than being a matter of the number (or length) of deliberative sessions, the issue might be the time that participants need in order to properly digest the deliberative session.
At a minimum, therefore, our results show the need to evaluate the impact of deliberative

13 In real-life settings, procedures will have to be streamlined to ensure citizens' compliance (doing otherwise might introduce severe sampling biases due to the selection of a population with more time available), and to contain costs. Thus, we opted for a series of time-saving and parsimonious approaches, such as providing written balanced material as "information" and limiting the entire RCT to 5 hours per participant. And yet, being aware that time to metabolize what was learned is important, we introduced a final questionnaire (T4) one month after the RCT.

interventions some time after them – which is atypical for experimental settings, while present in mini-publics (e.g., Nabatchi 2010). If relatively brief sessions prove effective, our results encourage devising protocols for the real-life application of deliberative methodologies. We submit here, based on our results, that a limited number of short sessions might be sufficient, provided that the end-points are evaluated after a reasonable amount of time.

*4.3 Passive Moderators vs. Active Moderators*

In political science, the role of the so-called moderators has been recognized as crucial, as they serve the purpose of fostering negative deliberative values, such as non-domination and non-interference, allowing the creation of the basic conditions for political equality (Mansbridge et al. 2006; Smith 2009; Gerber 2011; Moore 2012; Landwehr 2014). Ideally, as public bioethics is a domain dominated by substantive disputes, a proactive figure helping participants to develop their own preferences – both in terms of internal (logical coherence) and external (awareness of the consequences) consistency, and not only someone acting as guarantor of freedom of speech and equal participation – might be advantageous. Accordingly, we devised and tested, not only the traditional figures of Observer and Passive Moderator, but also an Active Moderator, crafted as the promoter of a set of positive values, such as autonomy, critical thinking, critical reasoning, and mutual respect.

However, our results showed a remarkable difference ascribable to the presence of Passive Moderators *vs.* Active Moderators in the discussion groups. While in the passively moderated setting there was a clear (and very significant) shift towards PoP, this did not occur in the actively moderated groups.

The first conclusion is that the presence of a non-corrective figure (the Passive Moderator) was beneficial in promoting the purposes of deliberation, in particular pluralism awareness[14]. From a bioethical perspective, the failure of active moderation might be disappointing, as it seems to hinder the public utility and relevance of bioethics' experts (Weinstein 1994; Steinkamp and Gordijn 2003; Steinkamp, Gordijn, and ten Have 2008; Rasmussen 2005; Varelius 2008; Archard 2011; Gordon 2014; Gesang 2010; Cowley 2012; Schicktanz, Schweda and Wynne 2012). In our design, the Active Moderator was devised as a public bioethicist – namely with the express intention of identifying a potential role for the (public) bioethical expert as *ethical expert*, while not being a moral expert (Sanchini 2015). In the bioethics literature, this distinction refers to the fact that the expert possesses some knowledge and skills that do not legitimate her to decide *for others*, but that enable her to help others to decide *for themselves* – i.e. by fostering the formation of participants' considered preferences (Dryzek 2001, Hendriks 2006).

In principle, there were reasonable theoretical justifications to entertain this idea. In particular, although National Bioethics Commissions (NBCs) are already widespread[15], the role of the bioethical expert in the public arena has not yet been standardized and/or institutionalized. Moreover, because of their composition, NBCs are at risk of being epistocratic, thus failing to sincerely mediate between non-experts' needs and institutional requests (Doods and Thomson 2006; Moore 2010). Finally, although DD ideals are paramount

14 Moore has suggested further reflection upon and possibly standardise the different figures mediating deliberation (Moore 2012, op. cit. note 32). In this context, our identification of DOs and DON'Ts for Passive Moderators (Appendix 5A) might represent a valuable starting point.

15 As reported by Doods and Thomson, The World Health Organisation lists about 90 national bioethics committees on its website (Dodds, S., & Thomson, C. (2006). Bioethics and Democracy: Competing Roles of National Bioethics Organisations. Bioethics, 20(6), 326–338, p. 326).

for public bioethics[16], these remain frequently under-expressed[17].
Our data, however, show otherwise. A possible explanation is that people are more willing to consider different perspectives when they come from their peers rather than from a superior figure. In other words, the corrective (albeit non-directive) role of the Active Moderator might have induced a defensive attitude, which, in turn, produced the rejection rather than the acceptance of a deeper consideration of their initial preferences.
We can only speculate about why this would be so. There is evidence that individuals are more prone to accept positions and arguments that are in line with their pre-existing beliefs (Himmelroos and Christensen 2013). Indeed, although from a theoretical viewpoint being exposed to dissimilar views might be beneficial for deliberation (Calhoun 2002; Manin 2005; Mutz 2002), several concerns have been raised regarding its practicability. Evidence from naturally occurring deliberation shows that people prefer to discuss with likeminded people (Mutz 2006). Huckfeldt and colleagues suggest that this human trait may be either ascribed to the human desire of reducing information costs, or to the psychic discomfort that encountering disagreement may produce (Huckfeldt et al. 2004). "Thus, in case of disagreement, people might not necessarily be inclined to confront the dissent with a counter claim, but rather opt for an escape strategy" (Gerber 2011, pp. 4-5). These considerations might explain why the passively moderated group showed a significant effect *vs*. the actively moderated one. The role of the Passive Moderator was simply to prompt equal contribution by encouraging silent participants to speak or by slowing down too dominant ones (Young 2002). The Active Moderator, conversely, may have been disruptive by prompting reflection on expressed preferences and by pointing to different viewpoints and to their likely consequences. The Passive Moderator did not question participants' preferences, while the Active Moderator did. It can, thus, be hypothesized that, for participants of passively moderated groups, it was easier to conform to a viewpoint more in line with PoP, because they were not induced to develop a defensive attitude or to opt for an escape strategy.

### *4.4 From bioethical experts to public bioethics: which role/s for public bioethics*

Let us now consider the final question of whether there is a role for bioethicists in public bioethics, and, if so, what that may be.
The prudent approach is to take our results at face value and to consider our experiment as preliminary evidence that bioethical experts are not beneficial to the promotion of deliberative values in the *discussion phase* of deliberative processes (where a more passive moderation is more effective). This does not imply, however, that the pre-deliberation phase can be constructed in the absence of expertise: the preparation of both informative material and questionnaires requires the knowledge and skills of bioethical experts[18].
Our results are also consistent with some tendencies in the broader debate on the role of public bioethics and NBCs, whether advisory or policy-making (Black 1998; Doods and Thomson 2006; Trotter 2006). Many scholars have claimed that the role for public bioethics in embedding deliberative ideals – a deliberative public bioethics – is that of mediating discussion on

16 Moore claims that public bioethics has a self-understanding that explicitly draws on deliberative ideals (Moore 2010, op. cit. note 1, p. 715). On the same point see also: Moreno 1995; Trotter 2006; Dzur and Levin 2007.
17 Indeed, traditional moderators, as described in political science, are not usually given reasonable latitude of intervention in the execution of their function, whereas an empowered version of the traditional moderator, the Active Moderator, would be able to fulfill the reason-giving requirement and to promote autonomy.
18 Furthermore, bioethicists could be added to an expert panel in the information phase. One might reasonably argue that providing information in the realm of bioethics does not need to be limited to the provision of factual/scientific information and might include, for instance, providing (upon request) scenarios about the moral implications of some practices.

contentious ethical issues of public relevance (Moore 2010). Accordingly, public bioethics plays a *preparatory* role in opening up and facilitating public debate (Doods and Thomson 2006), clarifying, if necessary, moral concepts and facilitating cooperation (Trotter 2006). In this sense, public bioethics should be considered to have primarily an advisory function, where its policy-making function has been articulated in the notion of "regulation as facilitation" (Black 1998). Another, still tenable, position is that by allowing longer deliberative sessions (and/or perhaps by increasing their numbers), the figure of the Active Moderator/bioethicist can still prevail and be superior to the Passive Moderator. This is a testable hypothesis that can be integrated into future studies aimed at verifying the impact and the cost/benefit balance of repeated short interventions employing various moderation styles.

## 5. Conclusion

Herein, we have reported and discussed the results of a large-scale mixed method study exploring moral preferences of undergraduates, subject to different moderation styles, on the topic of genetic testing in the context of reproductive choices. Drawing from the hypothesis that DD can be an instrument tailored for serving the purposes of public bioethics, acting as a tool for addressing and managing moral disagreement occurring in the public sphere, this study has shown that public deliberation with the presence of a passive moderator is an appropriate means of fulfilling the purposes that proponents of deliberation envisaged in theorising it.

Considering this outcome in relation to the current debate over bioethical expertise, we may argue not only that the "moral expert" – someone who interferes with other's own decisions because of his/her superior expertise – but also the less invasive "ethical expert" – someone who actively fosters the formation of considered preferences on a specific topic (corresponding to our Active Moderator) – should be rejected if deliberative ideals are endorsed, and that, conversely, a third figure, the Passive Moderator, appears, to this end, more promising.

This result is also consistent with the interpretation, within public bioethics debate, in favour of a specific connotation of public bioethics, namely as a sphere aimed at opening up and facilitating public debate, and mediating discussion on controversial ethical issues (Hendriks 2006).

### REFERENCES

Abelson, J., Blacksher, E. A., Li, K. K., Boesveld, S. E., & Goold, S. D. (2013). Public Deliberation in Health Policy and Bioethics: Mapping an Emerging, Interdisciplinary Field. Journal of Public Deliberation, 9 (1), 5;

Abelson, J., Eyles, J., McLeod, C. B., Collins, P., McMullan, C. & Forest, P.-G. (2003a). Does Deliberation Make a Difference? Results from a Citizens Panel Study of Health Goals Priority Setting. Health Policy, 66, 95–106;

Abelson, J., Forest, P.-G., Eyles, J., Smith, P., Martin, E. & Gauvin, F.-P. (2003b). Deliberations about Deliberative Methods: Issues in the Design and Evaluation of Public Participation Processes. Social Science & Medicine, 57(2) 239–251;

Abelson, J., Warren, M. E. & Forest, P.-G. (2012). The Future of Public Deliberation on Health Issues. The Hastings Center Report, 42(2), 27–29;

Archard, D. (2011). Why Moral Philosophers Are Not and Should Not Be Moral Experts. Bioethics, 25(3), 119–127;

Black, J. (1998). Regulation as Facilitation: Negotiating the Genetic Revolution. The Modern Law Review, 61(5), 621–660;

Blacksher, E., Diebel, A., Forest, P.-G., Goold, S. D. & Abelson, J. (2012). What Is Public Deliberation? The Hastings Center Report, 42(2), 14–17;

Bohman, J. (1998). Survey Article: The Coming of Age of Deliberative Democracy. Journal of Political Philosophy, 6(4), 400–425;

Bowie, C., Richardson, A. & Sykes, W. (1995). Consulting the Public about Health Service Priorities. British Medical Journal, 311, 1155–1158;
Bowling, A., Jacobson, B. & Southgate, L. (1993). Explorations in Consultation of the Public and Health Professionals on Priority Setting in an Inner London Health District. Social Science and Medicine, 37(7), 851–857;
Calhoun, C. J. (2002). Dictionary of the social sciences. New York: Oxford University Press;
Chilvers, J. (2008). Deliberating Competence Theoretical and Practitioner Perspectives on Effective Participatory Appraisal Practice. Science, Technology & Human Values, 33(2), 155–185;
Coote, A. & Lenaghan, J. (1997). Citizens' Juries: Theory into Practice. London: Institute for Public Policy Research;
Cowley, C. (2012). Expertise, Wisdom and Moral Philosophers: A Response to Gesang. Bioethics, 26(6), 337–342;
Crawshaw, R., Garland, M. J., Hines, B. & Lobitz, C. (1985). Oregon Health Decisions: An Experiment with Informed Community Consent. Journal of the American Medical Association, 254, 3213–3216;
Dodds, S., & Thomson, C. (2006). Bioethics and Democracy: Competing Roles of National Bioethics Organisations. Bioethics, 20(6), 326–338;
Dryzek, J. S. (2001). Legitimacy and Economy in Deliberative Democracy. Political Theory, 29(5), 651–669;
Dzur, A. W., & Levin, D. L. (2007). The Primacy of the Public: In Support of Bioethics Commissions as Deliberative Forums. Kennedy Institute of Ethics Journal, 17(2), 133–142;
Engelhardt, T. H. (1996). *The Foundations of Bioethics*, 2nd ed. New York: Oxford University Press;
Farrar, C., Green, D. P., Green, J. E., Nickerson, D. W., & Shewfelt, S. (2009). Does Discussion Group Composition Affect Policy Preferences? Results from Three Randomized Experiments. Political Psychology, 30(4), 615–647;
Fiorino, D. J. (1990). Citizen Participation and Environmental Risk: A Survey of Institutional Mechanisms. Science, Technology, & Human Values, 15(2), 226–243;
Fishkin, J. S. (1991). Democracy and Deliberation: New Directions for Democratic Reform. New Haven: Yale University Press;
Fishkin, J. S. (2011). When the People Speak: Deliberative Democracy and Public Consultation. Oxford: Oxford University Press;
Fishkin, J. S., Luskin, R. C. & Jowell, R. (2000). Deliberative Polling and Public Consultation. Parliamentary Affairs, 53(4), 657–666;
Fung, A. (2006). Varieties of Participation in Complex Governance. Public Administration Review, 66.s1, 66–75;
Gerber, M. (2011). Who are the Voices of Europe? Evidence from a Pan-European Deliberative Poll. Presented at the ECPR General Conference, Reykjavik, 25–27 August;
Gesang, B. (2010). Are Moral Philosophers Moral Experts? Bioethics, 24(4), 153–159;
Goold, S. D., Neblo, M. A., Kim, S. Y. H., Devries, R., Rowe, G. & Muhlberger, P. (2012). What Is Good Public Deliberation? The Hastings Center Report, 42(2), 24–26;
Gordon, J.-S. (2014). Moral Philosophers Are Moral Experts! A Reply to David Archard. Bioethics, 28(4), 203–206;
Gutmann, A. & Thompson, D. (1996). Democracy and Disagreement. Cambridge, MA: Harvard University Press;
Gutmann, A. & Thompson, D. (1997). Deliberating about Bioethics. The Hastings Center Report, 27(3), 38–41;
Gutmann, A. & Thompson, D. (2004). Why Deliberative Democracy? Princeton: Princeton University Press;
Hendriks, C. M. (2006). Integrated Deliberation: Reconciling Civil Society's Dual Role in

Deliberative Democracy. Political Studies, 54(3), 486–508;
Himmelroos, S. & Christensen, H. S. (2013). Deliberation and Opinion Change: Evidence from a Deliberative Mini-Public in Finland. Scandinavian Political Studies, 37(1), 41–60;
Huckfeldt, R., Mendez, J. M. & Osborn, T. (2004). Disagreement, Ambivalence, and Engagement: The Political Consequences of Heterogeneous Networks. Political Psychology, 25(1), 65–95;
Karpowitz, C. F., Mendelberg, T. & Shaker, L. (2012). Gender Inequality in Deliberative Participation. American Political Science Review, 106(03), 533–547;
Landemore, H. (2012). Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many. Princeton: Princeton University Press;
Landwehr, C. (2014). Facilitating Deliberation: The Role of Impartial Intermediaries in Deliberative Mini-Publics. Deliberative Mini-Publics: Involving Citizens in the Democratic Process. Colchester: ECPR Press. 77-92;
MacLean, S. & Burgess, M. M. (2010). In the Public Interest: Assessing Expert and Stakeholder Influence in Public Deliberation about Biobanks. Public Understanding of Science, 19(4), 486–496;
Manin, B. (2005). Democratic Deliberation: Why We Should Focus on Debate rather than Discussion. Presented at the Paper delivered at the Program in Ethics and Public Affairs Seminar, Princeton University;
Mansbridge, J., Hartz-Karp, J., Amengual, M. & Gastil, J. (2006). Norms of Deliberation: An Inductive Study. Journal of Public Deliberation, 2(1), 7;
McDermott, R. (2002). Experimental Methods in Political Science. Annual Review of Political Science, 5, 31–51;
Meagher, K. M. & Lee, L. M. (2016). Integrating Public Health and Deliberative Public Bioethics: Lessons from the Human Genome Project Ethical, Legal, and Social Implications Program. Public Health Reports, 131, 44–51;
Moore, A. (2010). Public Bioethics and Public Engagement: The Politics of "Proper Talk", Public Understanding of Science, 19(2), 197–211;
Moore, A. (2012). Following from the Front: Theorizing Deliberative Facilitation. Critical Policy Studies, 6(2), 146–162;
Moreno, J. D. (1995). Deciding Together: Bioethics and Moral Consensus (Vol. 61). New York: Oxford University Press;
Mutz, D. C. (2002). The Consequences of Cross-Cutting Networks for Political Participation. American Journal of Political Science, 46(4), 838–855;
Mutz, D. C. (2006). Hearing the Other Side: Deliberative versus Participatory Democracy (1 edition). Cambridge/New York: Cambridge University Press;
Nabatchi, T. (2010). Deliberative Democracy and Citizenship: In Search of the Efficacy Effect. Journal of Public Deliberation, 6(2);
Neblo, M. A. (2015). Deliberative Democracy between Theory and Practice. Cambridge: Cambridge University Press;
Parkinson, J. (2006). Deliberating in the Real World: Problems of Legitimacy in Deliberative Democracy. Oxford: Oxford University Press;
Rasmussen, L. M. (2005). Introduction: In Search of Ethics Expertise. In L. Rasmussen (Ed.), Ethics Expertise (pp. 1–12). Dordrecht: Springer;
Rawls, J. (1993). Political Liberalism. New York: Columbia University Press. Paperback edition, 1996; Second edition, 2005;
Rei, W., & Yeh, J. W. (2009). Promises and Pitfalls of Using National Bioethics Commissions as an Institution to Facilitate Deliberative Democracy-Lessons from the Policy Making of Human Embryonic Stem Cell Research. National Taiwan University Law Review, 4, 69;
Robertson, J. A. (1983). Procreative Liberty and the Control of Conception, Pregnancy, and Childbirth. Virginia Law Review, 405-464;

Robertson, J. A. (1985). Embryos, Families, and Procreative Liberty: The Legal Structure of the New Reproduction. Southern California Law Review, 59(5):939-1041;
Sanchini, V. (2015). Bioethical Expertise: Mapping the Field. Biblioteca della libertà, L-213: 43-61;
Schicktanz, S., Schweda, M. & Wynne, B. (2012). The Ethics of "Public Understanding of Ethics"—Why and How Bioethics Expertise Should Include Public and Patients' Voices. Medicine, Health Care, and Philosophy, 15(2), 129–139;
Smith, G. (2009). Democratic Innovations: Designing Institutions for Citizen Participation. Cambridge: Cambridge University Press;
Steinkamp, N. & Gordijn, B. (2003). Ethical Case Deliberation on the Ward. A Comparison of Four Methods. Medicine, Health Care, and Philosophy, 6(3), 235–246;
Steinkamp, N. L., Gordijn, B. & ten Have, H. A. M. J. (2008). Debating Ethical Expertise. Kennedy Institute of Ethics Journal, 18(2), 173–192;
Stewart, J., Kendall, E. & Coote, A. (1994). Citizens' Juries. London: IPPR;
Sunstein, C. R. (2000). Deliberative Trouble – Why Groups Go to Extremes. 110 Yale Law Journal, 71–119;
Trotter, G. (2006). Bioethics and Deliberative Democracy: Five Warnings from Hobbes. Journal of Medicine and Philosophy, 31(3), 235-250;
Tukey, J. W. (1977). Exploratory Data Analysis, London: Pearson;
Varelius, J. (2008). Is Ethical Expertise Possible? Medicine, Health Care, and Philosophy, 11(2), 127–132;
Verba, S. (1961). Small Groups and Political Behavior: A Study of Leadership. Princeton University Press;
Weinstein, B. D. (1994). The Possibility of Ethical Expertise. Theoretical Medicine, 15(1), 61–75.
Young, I. M. (2002). Inclusion and Democracy (1 edition). Oxford/New York: Oxford University Press.

## FIGURE AND TABLES

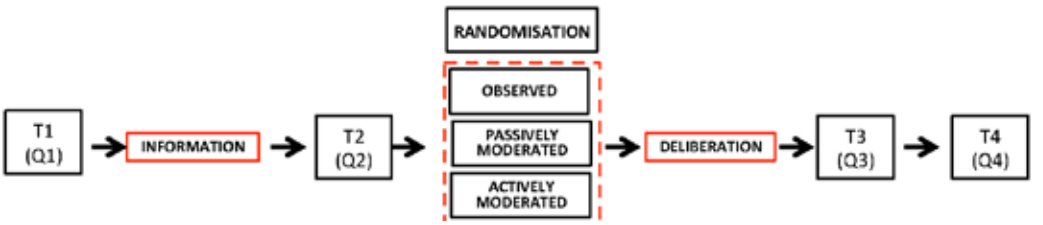



**Figure 1:**
**Design of the Randomized Controlled Trial (RCT)**

The flow-chart of the RCT is shown. Details are in the Main Experiment section of the text. T1, T2, T3 and T4 are the times of intervention at which the questionnaires, Q1, Q2, Q3 and Q4, were administered, respectively.

**Table 1: Scores per group at the various time points**
Questionnaire scores. Observed means and MICs, divided into groups and time (N=243)

| Groups | | Time | | | |
|---|---|---|---|---|---|
| | | T1 | T2 | T3 | T4 |
| **Observed** | Mean (STD) | 35.34 (7.96) | 35.10 (7.98) | 35.53 (7.20) | 34.71 (7.45) |
| | MIC[$] (STD) | --- | -0.24 (2.47) | 0.19 (3.71) | -0.56 (3.92) |
| | P* | --- | 0.53 | 0.81 | 0.20 |
| **Actively Moderated** | Mean (STD) | 36.32 (7.97) | 37.04 (8.12) | 37.02 (7.98) | 37.24 (8.24) |
| | MIC[$] (STD) | --- | 0.71 (2.29) | 0.70 (3.76) | 0.96 (3.78) |
| | P* | --- | 0.13 | 0.14 | 0.04 |
| **Passively Moderated** | Mean (STD) | 36.06 (7.46) | 36.30 (7.61) | 36.09 (7.65) | 35.51 (8.21) |
| | MIC[$] (STD) | --- | 0.24 (2.55) | 0.03 (3.63) | -0.16 (4.01) |
| | P* | --- | 0.56 | 0.86 | 0.84 |

The scores (N=243), expressed as mean values (standard deviation, STD, is in parentheses), are reported together with the differences *vs.* T1, for the three branches of the study at the various time points. There were no drop-outs in T2 and T3. In T4, there were 8 drop-outs (3.3%). There were no statistically significant differences among the three arms of the study in the number of drop-outs at T4 (P=0.94). See Appendix 4J. *P: p-value from t-test linear regression model for repeated measures, considering the correlation between groups of discussion, and correcting for the level of the score at T1, for age, and for degree. [$]: The differences were calculated on the number of available students. Note that in this and all following analyses, we adopted a rather stringent criterion to define significance, setting our significant p-value at 0.01, instead of the customary 0.05.

**Table 2: Baseline distribution (T1) of the questionnaire scores, by group**

| Group | N | Mean (STD) | Median (q1;q3) | Coefficient (95% CI) | P |
|---|---|---|---|---|---|
| All | 243 | 35.9 (7.78) | 37.0 (30.0;42.0) | | |
| Observed | 79 | 35.3 (7.96) | 37.0 (29.0;42.0) | Ref. | |
| Passively Moderated | 84 | 36.3 (7.97) | 38.0 (30.5;42.5) | 0.98 (-1.40;3.36) | 0.42 |
| Actively Moderated | 80 | 36.1 (7.46) | 37.0 (30.5;42.0) | 0.72 (-1.69;3.13) | 0.56 |

N: number of students, STD: standard deviation, q1: first quartile, q3: third quartile, Ref: reference group, CI: Confidence interval, P: p-value. Coefficients and p-values were calculated using a univariate linear regression model. The coefficient of the linear regression model is the average difference in Q1 scores between the comparison group (Actively Moderated or Passively Moderated group) and the reference group (ref. Observed group).

**Table 3: Differences in MIC transformation among the groups**

| Time | Comparison | Coefficient (SE) | P-Value |
|---|---|---|---|
| T2 | Actively Moderated/Observed | 0.61 (0.53) | 0.25 |
| | Passively Moderated/Observed | 1.07 (0.52) | 0.04 |
| | Actively Moderated / Passively Moderated | -0.46 (0.52) | 0.38 |
| T3 | Actively Moderated /Observed | -0.03 (0.53) | 0.96 |
| | Passively Moderated/Observed | 0.63 (0.52) | 0.23 |
| | Actively Moderated / Passively Moderated | -0.66 (0.52) | 0.21 |
| T4 | Actively Moderated /Observed | 0.55 (0.54) | 0.31 |
| | Passively Moderated/Observed | 1.65 (0.53) | **0.0019** |
| | Actively Moderated / Passively Moderated | -1.10 (0.53) | 0.04 |

| Time | Comparison | Coefficient (SE) | P-Value |
|---|---|---|---|
| T2 | Actively Moderated / Observed | 0.53 (0.54) | 0.33 |
| | Passively Moderated/ Observed | 0.98 (0.53) | 0.07 |
| | Actively Moderated / Passively Moderated | -0.45 (0.53) | 0.40 |
| T3 | Actively Moderated / Observed | -0.11 (0.54) | 0.83 |
| | Passively Moderated/ Observed | 0.53 (0.53) | 0.32 |
| | Actively Moderated / Passively Moderated | -0.65 (0.53) | 0.22 |
| T4 | Actively Moderated / Observed | 0.47 (0.55) | 0.39 |
| | Passively Moderated/ Observed | 1.55 (0.54) | **0.004** |
| | Actively Moderated / Passively Moderated | -1.08 (0.54) | 0.05 |

Top: Coefficient is the average difference in MIC between the groups of intervention at each time point and was estimated using a linear regression model for repeated measures, considering the correlation between groups of intervention, and correcting for the level of the score at T1, for age and for degree. P: p-value, SE: Standard error.

Bottom: Coefficient is the average difference in MIC between the groups of intervention at each time point and was estimated using a linear regression model for repeated measures, considering the correlation between groups of discussion, without adjusting for covariates. P: p-value, SE: Standard error.

## APPENDICES MAIN STUDY

### APPENDIX 1

In Appendix 1, we provide additional information concerning the main experiment. In particular:

1A. A description of the modalities of training and of the instructions received by Observers, Passive Moderators and Active Moderators for the performance of their duties during the deliberative sessions.
1B. The Questionnaire (Q) administered at T1, T2, T3 and T4.
1C. The Informative Material provided after T1.
1D. The Comprehension Questionnaire administered together with Q2.
1E. The Evaluation Questionnaire administered together with Q3.
1F. Results of correct answers to the Comprehension Questionnaire
1G. The results of the Evaluation Questionnaire.
1H. The scoring matrix adopted to quantitatively evaluate the Questionnaires at the various time points.
1I. The comparison between the analysis group and the outlier group.
1J. Number of study participants for each time point/group.
1K. Results of the Questionnaires (Q1-Q4) at T1-T4.

**1A. Training and performance of Observers, Passive Moderators and Active Moderators before and during the deliberative sessions**

To ensure uniformity of treatment in the various sub-groups of each experimental arm, supervisors were trained and given specific rules of behavior (see Table below) and a statement to deliver to the participants at the beginning of each session. Statements were as follows:

*Observer:* "My name is X and I will be observing your group as you discuss the ethical issues regarding genetic testing in the context of reproduction. I will not intervene in any way. I cannot provide you with any additional information. You can either start a discussion based on the informative material or based on the questions you found in the questionnaires. I will only tell you when the time for your discussion is over".

*Passive Moderator:* "My name is X and I will be your moderator today as you discuss the ethical issues regarding genetic testing in the context of reproduction. I will only intervene so that everyone gets a chance to express his or her opinion. I will keep time of your interventions so that everyone can speak for roughly the same amount of time. I will not provide you with additional scientific or ethical information. You can either start a discussion based on the informative material or on the questions you found in the questionnaires. I suggest that you start by presenting yourselves, your background, and by expressing your preferences on the topic at hand".

*Active Moderator:* "My name is X and I will facilitate this group today as you discuss the ethical issues regarding genetic testing in the context of reproduction. I will keep time of your interventions, making sure that everyone gets the chance to express his or her opinion. Moreover, I will help to promote an open and respectful discussion on different perspectives on the issue at hand. My role in facilitating this group is that of helping you to elaborate your own position. You are just asked to justify your preferences – that is, provide reasons for them that can be considered acceptable by reasonable people even though they may not share your perspective -, and I will help you do that. Any reasonable position you defend will be considered equally valid. I will not judge your position, I will only help you understand and consider various possible implications and consequences of it, nor will I provide you with any scientific additional information. If necessary, I will just refer you back to the material that you have read. I suggest that you start by presenting yourselves, your background, and by expressing your preferences on the topic at hand".

The rules of behavior for the supervisors are summarized in the following Table:

| TASKS | PASSIVE MODERATOR | ACTIVE MODERATOR |
|---|---|---|
| Ensure that all the participants have the chance to speak. | | |
| Curb talkative participants. | | |
| Keep time. | | |
| Give the participants time to think and reflect. * | | |
| Prevent episodes of domination amongst participants. | | |
| Ensure that all the participants express a *preference.* | | |
| Ensure that the preference is grounded on validated scientific information. | | |
| Encourage participants to provide arguments to justify their preferences. | | |
| Ensure that presented arguments are logically consistent and do not show logical fallacies. | | |
| Establish a dialogic relationship with the participants so as to understand their viewpoint/preference, with the final aim of making them aware of it. | | |
| Ensure that the participant is aware of the implications of having adopted one position over another, both at an individual level and a societal level. | | |
| Ensure that all viewpoints have been pointed out in the discussion and, if not, do this, so as to allow the participants to be aware of all the possible scenarios. | | |
| Encourage participants to interact with one another, promoting a cooperative attitude. | | |
| Encourage participants to pay attention to what other participants are saying. | | |
| Maintain a neutral position. | [Since he/she does not intervene in the discussion] | [Since, despite intervening, he/she does not reveal his/her own viewpoint] |
| Provide participants with additional scientific information with respect to that already present in the supplied material. | **NO** | **NO** |
| Refer back to the supplied material in order to provide context for discussion (if necessary). | | |

* In the passively moderated setting, this translated essentially in an attitude of the moderator towards shy participants exemplified by the dynamics: "If you do not want to say something now, why don't you take the time to think and reflect and I will make sure that we come back to you when you are ready". In the actively moderated settings, this function was executed in a more proactive way by helping shy participants to articulate their thoughts in a maieutic fashion.

In addition to providing the above guidelines and rules, we were concerned that during the deliberative sessions, some degree of unconscious manipulation by the supervisors might occur: a situation that might apply especially in the actively moderated groups. Thus, training of the supervisors (Observers, Passive Moderators, Active Moderators) was implemented with great care.

In particular:

1. As pointed out by Karpowitz and Mendelberg, it is important to find out how experimenters are trained (Karpowitz and Mendelberg 2012). In our setting, all supervisors met three times before the experiment to receive instructions and, raise and discuss possible questions and, importantly, to simulate the actual modalities of the intervention. In particular:

- In the first meeting, one of the study designers (V.S.) met with all supervisors, to allow them to introduce themselves to the others. Then, the experimental design was presented and discussed, including all the propaedeutic work derived from Pilot Study 1 and Pilot Study 2. The study population was described and discussed. Finally, the schedule of the experiment was presented and discussed.
- In the second meeting, the "rules of engagement" in the three arms of the intervention were presented and discussed, with particular attention to procedure standardization emphasizing the DOs and DON'Ts pertaining to the various roles, and "what to say and how to say it".
- In the final meeting, a role-play was set up in order to simulate the real experimental setting and test all the details previously discussed.

2. In preparation for the actual deliberative sessions, supervisors were asked to dress similarly and avoid revealing their academic background.
3. Before the actual experiment, all supervisors were asked to fill in the Questionnaire (Q), so that their preferences would be disclosed and recorded in advance.
4. Finally, and most importantly, participants were asked to evaluate the figure supervising their group and to declare whether they thought they had been manipulated (see comments on the "evaluation questionnaire" and in the "Additional analyses" section in the main text).

**1B. The Questionnaire (Q)**

The questionnaire shown below was used at T1, T2, T3 and T4.

*Note that at T2 and T3 additional questionnaires were administered: the Comprehension Questionnaire (T2) and the Evaluation Questionnaire (T3).*

**IDENTIFICATION NUMBER:**

Before completing this questionnaire please read the following points carefully:

1. The questionnaire is completely anonymous and the answers will be used only for statistical analyses.
2. When filling out the questionnaire, please note that there are no right or wrong answers.
3. The purpose of the questionnaire is solely to assess how the participants' preferences are distributed with respect to the statements in the questionnaire.
4. The questionnaire focuses on the following subject: genetic testing in the context of reproductive choices.
5. Choosing the response "I neither agree nor disagree" may mean that you do not have sufficient information to answer the question, or that you are not yet certain of your preference despite having sufficient information, or that there are other reasons for not giving or being able to give a definitive answer to the question.
6. Please **mark with an "X" only one answer**, and provide an answer for **each question**.
7. Remember to enter your identification number.

**Before starting the questionnaire, please fill in the demographic information**

**GENDER: M / F**

**AGE: _____ in years**

**DEGREE:**

- ☐ Medicine
- ☐ Nursing
- ☐ Physiotherapy
- ☐ Cognitive Sciences
- ☐ Philosophy
- ☐ Radiology

**Abbreviations Used**

PGD = Preimplantation genetic diagnosis

PD = Prenatal diagnosis

**QUESTIONNAIRE**

**For each of the following statements, please mark with an X the answer that most accurately reflects your opinion.**

*Note: each question was followed by the following options*

- ☐ Strongly disagree
- ☐ Disagree
- ☐ Neither agree nor disagree
- ☐ Agree
- ☐ Strongly agree

**Question 1**

*A person who wants to have a child and suspects to be at risk of giving birth to an individual with a genetic disease can freely choose whether or not to verify this risk through genetic testing i.e., he/she is not obliged to undergo genetic testing.*

**Question 2**

*A person who wants to have a child after being informed, following genetic testing, to be at risk of giving birth to an individual with a genetic disease, should be forced towards a specific set of reproductive choices: reproductive abstinence; adoption; heterologous fertilization; PGD and implantation in the uterus of unaffected embryos; conception, PD and therapeutic abortion.*

**Question 3**

*It is ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases, because frequently it is the parents who will bear the greater burden of the child's genetic disease.*

**Question 4**

*It is ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases, as this is consistent with the aims of medicine: to prevent and to cure disabilities.*

**Question 5**

*It is NOT ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases, as it is not the fault of the embryo/fetus if it is suffering from a genetic disease. Not implanting or aborting an affected embryo/fetus will harm it unjustly.*

**Question 6**

*It is NOT ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases because, by doing so, one assumes to have the right to choose whom to allow or to deny the possibility of life.*

**Question 7**

*It is NOT ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases in the case of low-penetrance diseases, as by doing so they may eliminate a future individual who will not develop the disease.*

**Question 8**
*It is NOT ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases because the affected embryo/fetus has only two alternatives: to be born with the disease or not to be born at all. In fact, PGD/PD is not a therapy: the affected embryo/fetus is not treated; on the contrary, a healthy one is chosen in its place.*

**Question 9**
*It is NOT ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases because; by doing so, there will be fewer and fewer sick people in the world and therefore their voices and their rights will be less and less heard or considered to be politically relevant.*

**Question 10**
*It is NOT ethically acceptable for parents to use PGD or PD with the aim of having a child free of genetic diseases because, in the long-term, this practice is likely to promote social rejection of people suffering from those diseases.*

**DID YOU ANSWER TO ALL OF THE QUESTIONS?**
**DID YOU FILL IN YOUR IDENTIFICATION NUMBER?**
**PLEASE CHECK ONE LAST TIME!**

*Note that the original Q administered to participants contained 14 questions. We noticed, however, that in 4 cases some ambiguities in the formulation of the questions (or in their possible interpretation) prevented the assignment of an unambiguous quantitative score (see 3G) to the answers. For this reason, these questions were excluded from further consideration and are not shown here.*

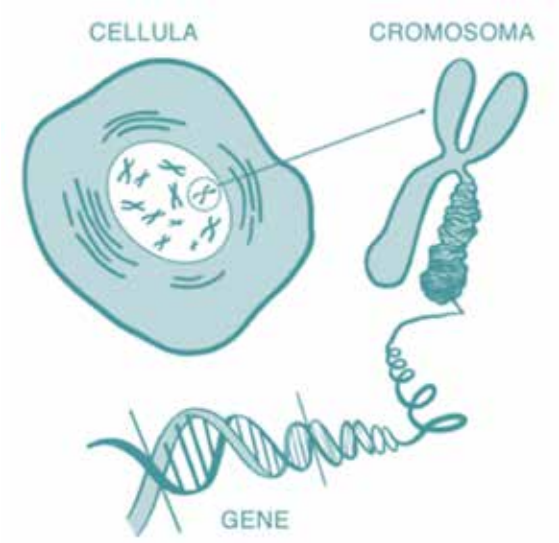
## 1C. The Informative Material

### GENETIC TESTS AND REPRODUCTIVE CHOICES
### INFORMATIVE MATERIAL

WORDS HIGHLIGHTED IN RED ARE DEFINED IN THE GLOSSARY AT THE END OF THE INFORMATIVE MATERIAL

**Introduction: genes and chromosomes**

The human body is made up of approximately 100,000 billion cells. Almost all cells contain a set of chromosomes, which carry genetic information. A gene is a heritable region on the DNA, from which an RNA molecule, associated with a particular function, is synthesized.
The human genome is made up of thousands of genes: 20,000 – 25,000 depending on the particular calculation. Genes control all cellular functions and have a fundamental role in the determination of many characteristics, such as eye color, blood group and height.
Genes are contained on long, linearly condensed filaments, called chromosomes. *Homo Sapiens* have 46 chromosomes: 22 pairs of autosomal or non-sex chromosomes, and one pair of sex chromosomes, X and Y. A person's chromosomes are inherited from his/her parents, 23 from the mother and 23 from the father. Thus, there are usually two copies or versions of each gene, termed alleles.
Chromosomes and genes are made up of a chemical substance called deoxyribonucleic acid or DNA.




A genetic disorder is a disease caused by an alteration in the genetic material present in the cells, involving one or more genes. A genetic disorder can be *inherited*, if passed from parent to child. In this case, the mutation is present in the DNA in the oocyte or sperm. Alternatively, a genetic disorder can emerge after conception or during pregnancy, in which case the disease is referred to as a *congenital*, rather than inherited, genetic disorder.
A separate discussion applies to cancer, where, in general, cells accumulate genetic mutations during a person's life that lead to their uncontrolled proliferation.

**Genetic disorders are usually classified as:**

a. *Chromosomal disorders.* Chromosomal disorders derive from variations in the set of human chromosomes. Since each chromosome contains thousands of genes, chromosomal alterations usually result in very serious clinical syndromes, i.e., a set of medical signs and symptoms that are associated with one or more somatic abnormalities, growth retardation, mental delay, etc. There are two types of chromosomal variations that can determine the onset of a disorder: numerical alterations in the number of whole chromosomes, referred to as aneuploidy or polyploidy, and structural alterations in the integrity, copy number and sequence direction within the chromosomes, due to translocations, insertions, deletions, duplications, etc.
*An example of a chromosomal disorder is Down's syndrome. This disorder is a numerical chromosomal disorder, specifically an aneuploidy disorder. It is also known as trisomy 21 because all the body's cells contain 3 copies of chromosome 21. The life expectancy of individuals with Down's syndrome is about 60 years. This syndrome is the most common chromosomal abnormality in humans: it appears in 1 out of 700/1000 live births. The only other viable trisomies are Edward's syndrome (abnormality in chromosome 18) and Patau's syndrome (abnormality in chromosome 13) and Kleinfelter's syndrome. All other trisomies are non-viable. The only viable monosomy is Turner's syndrome.*

b. *Monogenic or single-gene disorders.* Monogenic or single-gene disorders are caused by mutations in a single gene (point mutations or genetic mutations). Monogenic disorders are classified as autosomal if the mutation occurs in a gene on a non-sex chromosome or X/Y-linked if the mutation occurs in a gene on a sex chromosome. Autosomal disorders can also be classified as dominant or recessive. An autosomal disorder is dominant if the mutation of a single allele is sufficient for the disease to manifest itself, and recessive if both alleles need to be mutated.
*An example of a monogenic disorder is Huntington's disease, which is a dominant autosomal disorder. This disease is caused by the mutation of one of the two alleles of the Huntingtin gene. Disease onset usually occurs in individuals between 30 to 50 years of age, after which the disease progresses slowly, but is fatal after 16-20 years. The incidence of this syndrome is 5-10 cases per 100,000 people.*

c. *Multifactorial inheritance disorders.* Multifactorial inheritance disorders are caused by a combination of multiple factors, including genetic and environmental factors and their reciprocal interactions.
*An example of a multifactorial inheritance disorder is diabetes mellitus. Diabetes is a chronic disease that is characterized by the presence of elevated levels of glucose in the blood due to alterations in the amount or function of insulin. Insulin is a hormone produced by the pancreas that allows the absorption of blood glucose into intestinal mucosal cells, where it is used as an energy source. When this mechanism is impaired, glucose builds up in the bloodstream. There are different types of diabetes (type 1, type 2 and gestational diabetes), all of which are considered as multifactorial disorders. The incidence of this disease is about 1 in every 20 people*[1].

**Genetic analysis**

A genetic test or analysis aims to detect or exclude the presence of DNA modifications associated with genetic disorders through the analysis of specific genes or chromosomes. Genetic analyses are usually performed on blood or tissue samples.

**What are genetic tests used for?**

A genetic test is a tool used to determine:

i) If a person has a genetic disorder – *diagnostic purpose*.
ii) A person's predisposition to develop a specific genetic disorder, particularly, in cases where there is a family history of the disease – *predictive purpose*.

1 This estimation is based on a study according to which there are 347 million people with diabetes mellitus worldwide today (for further information: http://www.who.int/mediacentre/factsheets/fs312/en/).

iii) Individual genetic variations, knowledge of which permits the selection of the most appropriate treatment for a specific person – *pharmacogenomics purpose*.

**What can genetic tests tell us?**

To understand what a genetic test can tell us about a given genetic disorder, it is important to understand the concepts of penetrance and genetic risk.

***Penetrance***

Penetrance is the frequency (expressed as a percentage) with which a characteristic linked to a particular gene, and thus to a corresponding genetic disease, is displayed in individuals carrying a given mutation. The concept of penetrance is of primary importance in the debate on genetic testing because it indicates the frequency with which a particular genotype determines, *at the population level,* the appearance of a corresponding genetic disorder.
There are two types of disease penetrance: complete and incomplete. Penetrance is *complete* when 100% of carriers of a certain genotype display the typical phenotype associated with that genotype (e.g., Down's syndrome is a genetic disorder with complete penetrance because everyone who has a trisomy of chromosome 21 is affected by the syndrome). Penetrance is *incomplete* when less than 100% of carriers display the typical phenotype (e.g., Huntington's disease is a genetic disorder with incomplete penetrance because not all individuals carrying a mutation in the disease-causing gene develop the disease).
For diseases with complete penetrance, the individual will know that, at the population level, the presence of the genotype determines the presence of the disease in all cases. For diseases with incomplete penetrance, the individual is less facilitated in the choice he/she has to make because he/she does not know whether the observed genotype will give rise to the corresponding genetic disorder.

***Genetic risk***

"Genetic risk" is the probability that an individual carrying one or more mutations associated with a genetic disorder will actually suffer from the disease. Penetrance is linked to single mutations, while genetic risk takes into account all of the mutations present in an individual. Thus, there may be individuals carrying several low penetrance mutations, which when considered together, increase the genetic risk of that individual.

**Genetic tests and reproductive choices**

By "reproductive choices" we mean the decisions that one has to make as a prospective parent regarding whether to procreate, with whom, under what conditions, when, etc.
To help a person make these decisions, genetic testing can be carried out on the prospective parents and on the embryo, either before implantation in the uterus or during pregnancy. Genetic tests on prospective parents are performed using small blood samples and/or saliva and are used to determine whether the parent is a healthy carrier, suffers from a certain disease, or neither of these alternatives.
For the embryo/fetus, two types of genetic tests can be performed: prenatal diagnosis and preimplantation genetic diagnosis.

**Prenatal Diagnosis (PD)**

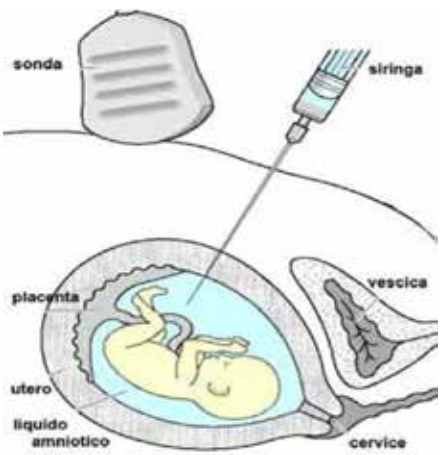
PD refers to all techniques that reveal the presence of disease (genetic and non-genetic) in the fetus. These techniques are performed during pregnancy and may be invasive or non-invasive. Invasive techniques (e.g., amniocentesis and chorionic villus sampling) are reimbursed by the National Health Service for pregnant women over 35 years old at the time of delivery. In contrast, non-invasive techniques, such as maternal blood tests, are paid for by the pregnant woman.

***Non-invasive techniques include:***

- *Ultrasound.* Ultrasound is a radiological investigation that does not use ionizing radiation but ultrasound. It is therefore risk-free, and is used routinely during pregnancy to assess gestational age, to monitor fetal growth, to identify twin pregnancies, and to determine the sex of the unborn child. Ultrasound tests are able to diagnose anatomical malformations that are often transmitted as a multifactorial disorder, but cannot identify specific biochemical or molecular defects.
- *Screening of maternal blood in particular, triple and quadruple tests on maternal blood.* Triple and quadruple screening tests are carried out between the 15$^{th}$ and 18$^{th}$ gestational week and are performed using a simple blood test. These tests assess the concentrations of specific substances present in the maternal blood that are produced by the fetus and the placenta. The triple test measures the amounts of three substances: alpha-fetoprotein AFP, beta-human chorionic gonadotropin bHCG and unconjugated estriol E3 FREE. The quadruple test measures the amounts of inhibin A in addition to the substances in the triple test. These analyses evaluate the fetus' genetic risk for developing a particular disease, but cannot diagnose with certainty the actual presence of the genetic disease.
- *Non-invasive tests to detect fetal DNA in maternal blood.* These tests are early diagnostic tests that are performed from the 9$^{th}$ week of gestation. They are precise and reliable tests, as well as safe as they require a normal sample of maternal blood. This technique assesses the risk of having some fetal chromosomal abnormalities, such as Down's syndrome or other syndromes that are derived from alterations of the sex chromosomes. The reliability of these tests in detecting these abnormalities is 99%.

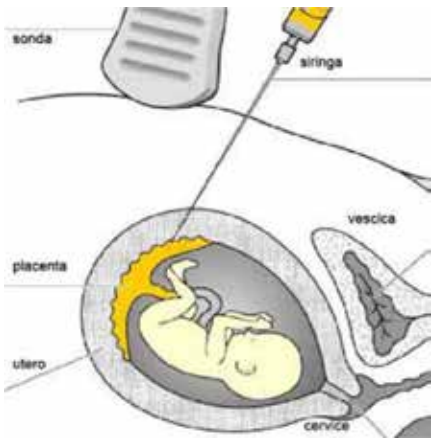
**Invasive techniques**[2]**:**

- *Amniocentesis.* Amniocentesis is performed through trans-abdominal sampling of the amniotic liquid after the 15$^{th}$ week of gestation under ultrasound guidance. The risk of miscarriage is low but not negligible (less than 1%).




2 Invasive diagnosis can be performed in the following cases: a) in women older than 35 years at time of delivery; b) in parents carrying chromosomal translocations or aneuploidy of sex chromosomes; c) in women who previously gave birth to a child with chromosomal abnormalities; d) following detection of fetal malformations by ultrasound scan; e) following a positive nuchal translucency ultrasound scan assessing the quantity of the fluid in the nape of the fetal neck, or a positive triple test biochemical analysis performed on a blood sample, which quantifies the risk of chromosomal abnormalities in the fetus; f) for the detection of infective agents in the amniotic fluid; g) for studies on fetal DNA; h) for the determination of metabolites in the amniotic fluid.

- *Chorionic villus sampling.* Chorionic villus sampling involves trans-abdominal sampling of placental villi under ultrasound guidance after the 10th gestational week. The risk of miscarriage is the same as or slightly higher than that of amniocentesis[3].




- *Cordocentesis.* Cordocentesis involves sampling of fetal blood after the 18th gestational week. The risk of miscarriage is 2-3%.

**How to choose between the different invasive and non-invasive techniques?**
Both amniocentesis and chorionic villus sampling allow the detection of chromosomal abnormalities (karyotype and microscopic rearrangements). Genetic testing is not carried out unless there is some indication that a specific genetic disease might be present, such as a family history. This is because it is not possible to test for all genetic disorders since they are numerous and not all are known. It is therefore possible for a child to be born with a genetic disorder despite having a karyotype result that appears negative for chromosomal mutations.

3 There are some reports indicating a higher risk of miscarriage for chorionic villus sampling with respect to amniocentesis. In reality, the higher rate of miscarriage reflects the higher risk of a spontaneous miscarriage in the first trimester, when chorionic villus sampling is performed. Thus, the two methods carry equivalent risks of miscarriage.

The main differences between amniocentesis and chorionic villus sampling are the time at which the tests are performed (chorionic villus sampling is usually performed between the 11th-12th gestational week and amniocentesis between the 16th-18th gestational week) and the length of time required to obtain results (a few days for chorionic villus sampling and 2-3 weeks for amniocentesis).
The choice of technique depends on the following factors: gestational week, the likelihood that a chromosomal abnormality is present, and the desired level of confidence in the results, which is influenced by the efficacy and sensitivity of the test.
The reliability of PD varies depending on the technique. The reliability of non-invasive techniques, such as ultrasound, is between 59-80%, while that of invasive techniques, such as amniocentesis and chorionic villus sampling, is close to, although not quite, 100% (99%).
The reliability of the non-invasive technique, maternal blood screening, is 99% but, unlike amniocentesis and chorionic villus sampling, this technique is limited to just a few specific chromosomal abnormalities.

**Preimplantation genetic diagnosis (PGD)**


PGD is a complementary procedure to PD that detects genetic disorders in embryos generated through medically assisted reproduction. PGD is used by couples with a high reproductive risk for a given genetic disorder and is carried out at very early stages of embryonic development, before implantation of the embryo in the uterus. Thus, in contrast to PD tests, PGD tests are not performed during pregnancy, but earlier before the embryo is implanted in the uterus. This allows a choice to be made as to whether or not to implant an embryo presenting a genetic disorder.

PGD is performed through the following steps:

a. Induction of ovulation. Ovulation is artificially induced by ovarian stimulation. The purpose of this stimulation is to induce the maturation of multiple follicles in the patient in order to obtain more oocytes and, thus, increase the probability of obtaining embryos to transfer.
b. Oocyte retrieval. This is performed via transvaginal ultrasound. The aspirated fluid is sent to the laboratory for collection of mature oocytes.

c. Medically assisted reproduction. This is the artificial fertilization of the oocyte by male sperm. The technique typically used for artificial fertilization is *Intracytoplasmic Sperm Injection* (ICSI). This technique ensures a greater precision of the fertilization process by injecting sperm directly into the cytoplasm of a single oocyte.
d. Harvesting of embryonic cells. On the third day after fertilization, the embryo usually consists of 6-8 cells. One/two of these cells are collected by introducing a glass micropipette in an opening in the 'zona pellucida' (the wall that surrounds the embryo until the blastocyst stage) and gently aspirating. This procedure does not interfere with the subsequent development of the embryo.
e. Analysis of harvested cells to test for the presence of genetic mutations associated with the genetic disorder under investigation.
f. Implantation in the uterus of embryos displaying no genetic defects, unless otherwise indicated by the parents.

PGD is able to detect the genetic disorder under investigation in 95% of cases, but fails to detect in 5% of cases[4]. This means that, in the case of a disease with a rate of onset of 1%, the probability that the child who was positive in the PGD test will be born with the disease is 1 in 20 x 1 in 100, i.e., 1 in 2000[5].

**GLOSSARY**

**Allele.** One of a pair of genes that appear at a particular location on a particular chromosome and control the same characteristic.
**Amniotic liquid.** A liquid composed mainly of water, mineral salts, lipids and proteins produced by the placenta and by the membranes that surround the uterine wall in early pregnancy.
**Blastocyst**. The embryo during the early stages of its development. This phase corresponds to the 5 – 7$^{th}$ day of fertilization.
**Chromosome.** Elongated filaments present in the nucleus of animal and plant cells, and comprised of a single DNA molecule that holds the genetic information. Members of each species typically have the same number of chromosomes in their cells.
**Chronic disease.** A stationary or slowly progressive disease.
**DNA.** Deoxyribonucleic acid, which carries hereditary information and is found almost exclusively in the nucleus of the cell.
**Follicle.** Spheroidal cellular aggregation present in the ovary that contains the oocyte.
**Genome.** The set of DNA sequences in the nucleus, including all genes and other sequences.
**Genotype.** The genetic and hereditary characters of an individual or population that result in a phenotype.
**Gestation.** The period between conception and birth during which the development of the fetus takes place.
**Implantation.** Implantation of the fertilized oocyte in the wall of the uterus.
**Karyotype.** The profile of chromosomes in a cell defined by their number, size, shape and dimension. The karyotype is specific for each species, organism and cell type.

4 This is due to various factors: i) possible contamination of the sample with foreign material; ii) inability to amplify one of the two alleles for technical reasons, and consequently the mutation is not detected (phenomenon known as Allele Drop Out); iii) mosaicism, when cells derived from the same embryo present different karyotypes. Thus, some cells within an embryo could be normal, while others are mutated. Depending on the precise cells that are sampled, the cytogenetic analysis will give varying results.
5 Diagnostic error: less than 1%.

**Medically assisted reproduction.** All procedures involving the processing of human oocytes, sperm or embryos with the aim of resulting in a pregnancy.
**Miscarriage.** Miscarriage is the premature termination of a pregnancy. This may be due to natural causes (spontaneous) or induced.
**Mutation**. A random variation in the genetic makeup of an individual animal or plant that causes a change in protein synthesis and in the transmission of characteristics.
**Oocyte.** The female gamete.
**Ovarian stimulation.** Application of a stimulus to the ovaries to stimulate the production of oocytes.
**Phenotype.** The set of morphological characteristics of an individual, resulting from the interaction between their genetic material and environmental factors.
**RNA.** Ribonucleic acid is a molecule similar to DNA that is contained in the nucleus and cytoplasm of cells and is required for protein synthesis.
**Translocation.** The physical movement of genome sequences inside the nucleus that change their position on chromosomes.

**1D. The Comprehension Questionnaire (administered at T2 together with Q2)**
**Question 1**
Genetic diseases include:
- All chromosomal disorders
- All chromosomal disorders, monogenic/single-gene disorders and multifactorial inheritance disorders
- Only monogenic disorders

**Question 2**
Genetic tests/analyses are able to:
- Determine only whether a person has a genetic disorder at the time of testing
- Determine only a person's predisposition to developing a specific genetic disorder
- Determine both of the above points, as well as individual genetic variations thereby allowing the selection of the most appropriate treatment for a specific individual

**Question 3**
Penetrance tells us:
- The relationship between genotype and phenotype for a specific genetic disease in a given population
- The relationship between genotype and phenotype for a specific genetic disease in a specific individual
- How severe a given disease will be in a specific individual

**Question 4**
Prenatal tests:
- Are performed on the embryo to determine whether it is affected by a specific genetic disorder
- Are performed on the fetus, already implanted in the uterus, during different stages of pregnancy to determine whether it is affected or not by a specific genetic disorder
- Are performed on the fetus, already implanted in the uterus, during different stages of pregnancy to determine whether it is affected or not by any of the known genetic disorders

**Question 5**
Preimplantation genetic diagnosis:
- Is performed on the fetus during the second month pregnancy to check for chromosomal abnormalities
- Is performed on embryos, created through various assisted reproduction techniques, before their implantation in the uterus, to test for a given genetic disorder
- Is performed on embryos, created through various assisted reproduction techniques, before their implantation in the uterus to test for multifactorial inheritance disorders.

**1E. The Evaluation Questionnaire (administered at T3 together with the Q3)**

**Please fill in the following table expressing your opinion on the experience. Please tick one box for each question.**

| QUESTIONS | Not at all | Small degree | Moderate degree | High degree | Very high degree |
|---|---|---|---|---|---|
| 1. Did the discussion promote an attitude of **higher respect towards the preferences of the other participants**? | | | | | |
| 2. Did the discussion **prompt your group to reach a consensus**? | | | | | |
| 3. Did the discussion have **an impact on the transformation of your preferences concerning the issue at hand**? | | | | | |
| 4. How much has the discussion **allowed you to express your preferences in an unconstrained way**? | | | | | |
| 5. Do you think you have been somehow manipulated towards a specific position **by the person who supervised the discussion**? | | | | | |
| 6. How **clear** were the questions of the questionnaire? | | | | | |

**Do you have any additional comments and/or suggestions?**

write here

**1F. Results of correct answers to the Comprehension Questionnaire**

| N of correct answers | N of students | % | p-value* |
|---|---|---|---|
| **0** | **0** | **0** | **0.0001** |
| **1** | **4** | **1.5** | |
| **2** | **6** | **2.2** | |
| **3** | **34** | **12.4** | |
| **4** | **117** | **42.7** | |
| **5** | **113** | **41.2** | |

*: p-value of multinomial distribution test. The multinomial distribution test was used to test the hypothesis that the distribution of answers could originate from random answers.

**1G. Results of the Evaluation Questionnaire**

| Question | Answers | All % | Observed (%) | Passively Moderated % | Actively Moderated % | *P** |
|---|---|---|---|---|---|---|
| **1** | Not at all | 0.8 | 0 | 1.2 | 1.3 | 0.15 |
| | Small degree | 2.9 | 2.5 | 1.2 | 5.1 | |
| | Moderate degree | 8.7 | 11.4 | 4.8 | 10.3 | |
| | High degree | 33.3 | 27.8 | 31.3 | 41.0 | |
| | Very high degree | 54.2 | 58.2 | 61.4 | 42.3 | |
| **2** | Not at all | 2.9 | 1.3 | 3.6 | 3.8 | 0.15 |
| | Small degree | 6.3 | 5.1 | 6 | 7.7 | |
| | Moderate degree | 21.7 | 12.7 | 22.9 | 30.8 | |
| | High degree | 42.9 | 46.8 | 44.6 | 37.2 | |
| | Very high degree | 25.8 | 34.2 | 22.9 | 20.5 | |
| **3** | Not at all | 18.8 | 15.2 | 26.5 | 14.1 | 0.31 |
| | Small degree | 43.8 | 43.0 | 45.8 | 42.3 | |
| | Moderate degree | 27.5 | 31.6 | 18.1 | 33.3 | |
| | High degree | 8.3 | 8.9 | 8.4 | 7.7 | |
| | Very high degree | 1.7 | 1.3 | 1.2 | 2.6 | |
| **4** | Not at all | 0.4 | 0 | 1.2 | 0 | 0.30 |
| | Small degree | 1.3 | 2.5 | 0 | 1.3 | |
| | Moderate degree | 2.9 | 3.8 | 1.2 | 3.8 | |
| | High degree | 16.3 | 12.7 | 22.9 | 12.8 | |
| | Very high degree | 79.1 | 81.0 | 74.7 | 82.1 | |
| **5** | Not at all | 90.4 | 93.7 | 95.2 | 82.1 | 0.011 |
| | Small degree | 7.5 | 3.8 | 3.6 | 15.2 | |
| | Moderate degree | 0.8 | 1.2 | 1.2 | 0 | |
| | High degree | 0.8 | 1.3 | 0 | 1.3 | |
| | Very high degree | 0.4 | 0 | 0 | 1.3 | |
| **6** | Not at all | 0.4 | 0 | 1.2 | 0 | 0.15 |
| | Small degree | 0.8 | 1.3 | 1.2 | 0 | |
| | Moderate degree | 16.2 | 24.1 | 13.3 | 11.5 | |
| | High degree | 51.3 | 51.9 | 51.8 | 50.0 | |
| | Very high degree | 31.1 | 22.8 | 32.5 | 38.5 | |

*P: Fischer's exact test

N=240. Observed, N=79; Passively Moderated, N=83; Actively Moderated N=78

**1H. Scoring matrix for the Questionnaire (Q1 – Q4)**

| Question | Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree |
|---|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 4 | 5 |
| 2 | 5 | 4 | 3 | 2 | 1 |
| 3 | 1 | 2 | 3 | 4 | 5 |
| 4 | 1 | 2 | 3 | 4 | 5 |
| 5 | 5 | 4 | 3 | 2 | 1 |
| 6 | 5 | 4 | 3 | 2 | 1 |
| 7 | 5 | 4 | 3 | 2 | 1 |
| 8 | 5 | 4 | 3 | 2 | 1 |
| 9 | 5 | 4 | 3 | 2 | 1 |
| 10 | 5 | 4 | 3 | 2 | 1 |

The scoring matrix, used to evaluate quantitatively the questionnaire is reported. The column "Question" displays the question number. Scores were assigned on a 5-point scale: a score of 5 was given for answers closest to a perspective in favor of freedom in reproduction, while a score of 1 was given for answers most distant from that perspective.

**1I. Comparison between the analysis group and the outlier group at T1.**

| Variable | All<br>N (% col) | Analysis group<br>N (% row) | Outlier group<br>N (% row) | p |
|---|---|---|---|---|
| **All** | 274 (100) | 243 (88.7) | 31 (11.3) | |
| **Degree** | | | | 0.02 |
| Philosophy | 64 (23.4) | 60 (93.7) | 4 (6.2) | |
| Medicine | 104 (38.0) | 96 (92.3) | 8 (7.7) | |
| Professional | 106 (38.7) | 87 (82.1) | 19 (17.9) | |
| **Age** | | | | 0.99 |
| <21 | 124 (45.3) | 110 (88.7) | 14 (11.3) | |
| >=21 | 150 (54.7) | 133 (88.7) | 17 (11.3) | |
| **Gender** | | | | 0.40 |
| F | 152 (55.5) | 137 (90.1) | 15 (9.9) | |
| M | 122 (44.5) | 106 (86.9) | 16 (13.1) | |

By analyzing the mean individual change (MIC) between T1 and T2, we identified 31 (11%) students as outliers defined as external to the median range ± (1.5*interquartile range), i.e., score ≤ -6 or score ≥ 6.
Professional degrees: Physiotherapy, Cognitive Sciences, Nursing, and Radiology.
p: p-value calculated using the Chi-square test.

**5J. Number of study participants for each time point/group**

| Group | Time | | | |
|---|---|---|---|---|
| | **T1<br>N** | **T2<br>N** | **T3<br>N** | **T4<br>N (% T1)** |
| **Observed** | **79** | **79** | **79** | **78 (98.7)** |
| **Passively Moderated** | **84** | **84** | **84** | **82 (97.6)** |
| **Actively Moderated** | **80** | **80** | **80** | **75 (93.8)** |
| **Total** | **243** | **243** | **243** | **235 (96.7)** |

**1K. Results of the Questionnaires (Q1-Q4) at T1-T4**

| Answers T1 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 1 | | | | | 0.447 |
| No answer | 1 (0.4) | 0 (0) | 0 (0) | 1 (100) | |
| Strongly agree | 6 (2.5) | 1 (16.7) | 4 (66.7) | 1 (16.7) | |
| Agree | 21 (8.6) | 6 (28.6) | 8 (38.1) | 7 (33.3) | |
| Neither agree nor disagree | 5 (2.1) | 0 (0) | 3 (60) | 2 (40) | |
| Disagree | 66 (27.2) | 27 (40.9) | 20 (30.3) | 19 (28.8) | |
| Strongly disagree | 144 (59.3) | 45 (31.2) | 49 (34) | 50 (34.7) | |
| QUESTION 2 | | | | | 0.764 |
| Strongly agree | 124 (51) | 45 (36.3) | 41 (33.1) | 38 (30.6) | |
| Agree | 77 (31.7) | 24 (31.2) | 25 (32.5) | 28 (36.4) | |
| Neither agree nor disagree | 10 (4.1) | 2 (20) | 5 (50) | 3 (30) | |
| Disagree | 30 (12.3) | 7 (23.3) | 13 (43.3) | 10 (33.3) | |
| Strongly disagree | 2 (0.8) | 1 (50) | 0 (0) | 1 (50) | |
| QUESTION 3 | | | | | 0.682 |
| Strongly agree | 17 (7) | 4 (23.5) | 8 (47.1) | 5 (29.4) | |
| Agree | 40 (16.5) | 17 (42.5) | 10 (25) | 13 (32.5) | |
| Neither agree nor disagree | 30 (12.3) | 8 (26.7) | 12 (40) | 10 (33.3) | |
| Disagree | 77 (31.7) | 27 (35.1) | 23 (29.9) | 27 (35.1) | |
| Strongly disagree | 79 (32.5) | 23 (29.1) | 31 (39.2) | 25 (31.6) | |
| QUESTION 4 | | | | | 0.633 |
| Strongly agree | 41 (16.9) | 12 (29.3) | 16 (39) | 13 (31.7) | |
| Agree | 41 (16.9) | 17 (41.5) | 15 (36.6) | 9 (22) | |
| Neither agree nor disagree | 29 (11.9) | 10 (34.5) | 10 (34.5) | 9 (31) | |
| Disagree | 85 (35) | 27 (31.8) | 24 (28.2) | 34 (40) | |
| Strongly disagree | 47 (19.3) | 13 (27.7) | 19 (40.4) | 15 (31.9) | |
| QUESTION 5 | | | | | 0.945 |
| Strongly agree | 79 (32.5) | 25 (31.6) | 30 (38) | 24 (30.4) | |
| Agree | 79 (32.5) | 27 (34.2) | 25 (31.6) | 27 (34.2) | |
| Neither agree nor disagree | 32 (13.2) | 9 (28.1) | 11 (34.4) | 12 (37.5) | |
| Disagree | 42 (17.3) | 13 (31) | 16 (38.1) | 13 (31) | |

| Answers T1 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| Strongly disagree | 11 (4.5) | 5 (45.5) | 2 (18.2) | 4 (36.4) | |
| QUESTION 6 | | | | | 0.184 |
| Strongly agree | 64 (26.3) | 23 (35.9) | 22 (34.4) | 19 (29.7) | |
| Agree | 78 (32.1) | 18 (23.1) | 32 (41) | 28 (35.9) | |
| Neither agree nor disagree | 22 (9.1) | 12 (54.5) | 7 (31.8) | 3 (13.6) | |
| Disagree | 52 (21.4) | 16 (30.8) | 15 (28.8) | 21 (40.4) | |
| Strongly disagree | 27 (11.1) | 10 (37) | 8 (29.6) | 9 (33.3) | |
| QUESTION 7 | | | | | 0.59 |
| Strongly agree | 15 (6.2) | 6 (40) | 6 (40) | 3 (20) | |
| Agree | 55 (22.6) | 15 (27.3) | 18 (32.7) | 22 (40) | |
| Neither agree nor disagree | 51 (21) | 15 (29.4) | 23 (45.1) | 13 (25.5) | |
| Disagree | 77 (31.7) | 28 (36.4) | 24 (31.2) | 25 (32.5) | |
| Strongly disagree | 45 (18.5) | 15 (33.3) | 13 (28.9) | 17 (37.8) | |
| QUESTION 8 | | | | | 0.793 |
| Strongly agree | 61 (25.1) | 21 (34.4) | 18 (29.5) | 22 (36.1) | |
| Agree | 74 (30.5) | 19 (25.7) | 30 (40.5) | 25 (33.8) | |
| Neither agree nor disagree | 45 (18.5) | 15 (33.3) | 17 (37.8) | 13 (28.9) | |
| Disagree | 32 (13.2) | 13 (40.6) | 8 (25) | 11 (34.4) | |
| Strongly disagree | 31 (12.8) | 11 (35.5) | 11 (35.5) | 9 (29) | |
| QUESTION 9 | | | | | 0.859 |
| Strongly agree | 108 (44.4) | 30 (27.8) | 42 (38.9) | 36 (33.3) | |
| Agree | 70 (28.8) | 23 (32.9) | 23 (32.9) | 24 (34.3) | |
| Neither agree nor disagree | 21 (8.6) | 8 (38.1) | 6 (28.6) | 7 (33.3) | |
| Disagree | 31 (12.8) | 13 (41.9) | 8 (25.8) | 10 (32.3) | |
| Strongly disagree | 13 (5.3) | 5 (38.5) | 5 (38.5) | 3 (23.1) | |
| QUESTION 10 | | | | | 0.103 |
| Strongly agree | 69 (28.4) | 16 (23.2) | 30 (43.5) | 23 (33.3) | |
| Agree | 64 (26.3) | 19 (29.7) | 23 (35.9) | 22 (34.4) | |
| Neither agree nor disagree | 24 (9.9) | 12 (50) | 6 (25) | 6 (25) | |
| Disagree | 60 (24.7) | 24 (40) | 13 (21.7) | 23 (38.3) | |
| Strongly disagree | 26 (10.7) | 8 (30.8) | 12 (46.2) | 6 (23.1) | |

| Answers T2 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 1 | | | | | 0.108 |
| Strongly agree | 8 (3.3) | 2 (25) | 6 (75) | 0 (0) | |
| Agree | 22 (9.1) | 6 (27.3) | 9 (40.9) | 7 (31.8) | |
| Neither agree nor disagree | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| Disagree | 54 (22.2) | 23 (42.6) | 14 (25.9) | 17 (31.5) | |
| Strongly disagree | 159 (65.4) | 48 (30.2) | 55 (34.6) | 56 (35.2) | |
| QUESTION 2 | | | | | 0.517 |
| Strongly agree | 121 (49.8) | 38 (31.4) | 47 (38.8) | 36 (29.8) | |
| Agree | 80 (32.9) | 29 (36.2) | 20 (25) | 31 (38.7) | |
| Neither agree nor disagree | 8 (3.3) | 2 (25) | 4 (50) | 2 (25) | |
| Disagree | 25 (10.3) | 6 (24) | 11 (44) | 8 (32) | |
| Strongly disagree | 9 (3.7) | 4 (44.4) | 2 (22.2) | 3 (33.3) | |
| QUESTION 3 | | | | | 0.171 |
| Strongly agree | 23 (9.5) | 6 (26.1) | 12 (52.2) | 5 (21.7) | |
| Agree | 42 (17.3) | 15 (35.7) | 8 (19) | 19 (45.2) | |
| Neither agree nor disagree | 15 (6.2) | 6 (40) | 6 (40) | 3 (20) | |
| Disagree | 93 (38.3) | 34 (36.6) | 31 (33.3) | 28 (30.1) | |
| Strongly disagree | 70 (28.8) | 18 (25.7) | 27 (38.6) | 25 (35.7) | |
| QUESTION 4 | | | | | 0.607 |
| Strongly agree | 38 (15.6) | 12 (31.6) | 14 (36.8) | 12 (31.6) | |
| Agree | 43 (17.7) | 16 (37.2) | 15 (34.9) | 12 (27.9) | |
| Neither agree nor disagree | 18 (7.4) | 7 (38.9) | 7 (38.9) | 4 (22.2) | |
| Disagree | 93 (38.3) | 29 (31.2) | 26 (28) | 38 (40.9) | |
| Strongly disagree | 51 (21) | 15 (29.4) | 22 (43.1) | 14 (27.5) | |
| QUESTION 5 | | | | | 0.794 |
| Strongly agree | 80 (32.9) | 28 (35) | 29 (36.2) | 23 (28.7) | |
| Agree | 92 (37.9) | 26 (28.3) | 31 (33.7) | 35 (38) | |
| Neither agree nor disagree | 11 (4.5) | 2 (18.2) | 4 (36.4) | 5 (45.5) | |
| Disagree | 47 (19.3) | 17 (36.2) | 16 (34) | 14 (29.8) | |
| Strongly disagree | 13 (5.3) | 6 (46.2) | 4 (30.8) | 3 (23.1) | |

| Answers T2 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 6 | | | | | 0.753 |
| Strongly agree | 66 (27.2) | 19 (28.8) | 26 (39.4) | 21 (31.8) | |
| Agree | 81 (33.3) | 27 (33.3) | 28 (34.6) | 26 (32.1) | |
| Neither agree nor disagree | 15 (6.2) | 4 (26.7) | 7 (46.7) | 4 (26.7) | |
| Disagree | 52 (21.4) | 19 (36.5) | 17 (32.7) | 16 (30.8) | |
| Strongly disagree | 29 (11.9) | 10 (34.5) | 6 (20.7) | 13 (44.8) | |
| QUESTION 7 | | | | | 0.09 |
| Strongly agree | 23 (9.5) | 5 (21.7) | 13 (56.5) | 5 (21.7) | |
| Agree | 51 (21) | 14 (27.5) | 24 (47.1) | 13 (25.5) | |
| Neither agree nor disagree | 37 (15.2) | 14 (37.8) | 8 (21.6) | 15 (40.5) | |
| Disagree | 87 (35.8) | 28 (32.2) | 26 (29.9) | 33 (37.9) | |
| Strongly disagree | 45 (18.5) | 18 (40) | 13 (28.9) | 14 (31.1) | |
| QUESTION 8 | | | | | 0.526 |
| Strongly agree | 72 (29.6) | 24 (33.3) | 28 (38.9) | 20 (27.8) | |
| Agree | 83 (34.2) | 21 (25.3) | 29 (34.9) | 33 (39.8) | |
| Neither agree nor disagree | 23 (9.5) | 7 (30.4) | 7 (30.4) | 9 (39.1) | |
| Disagree | 47 (19.3) | 20 (42.6) | 13 (27.7) | 14 (29.8) | |
| Strongly disagree | 18 (7.4) | 7 (38.9) | 7 (38.9) | 4 (22.2) | |
| QUESTION 9 | | | | | 0.524 |
| Strongly agree | 102 (42) | 30 (29.4) | 41 (40.2) | 31 (30.4) | |
| Agree | 68 (28) | 19 (27.9) | 21 (30.9) | 28 (41.2) | |
| Neither agree nor disagree | 19 (7.8) | 8 (42.1) | 6 (31.6) | 5 (26.3) | |
| Disagree | 39 (16) | 17 (43.6) | 10 (25.6) | 12 (30.8) | |
| Strongly disagree | 15 (6.2) | 5 (33.3) | 6 (40) | 4 (26.7) | |
| QUESTION 10 | | | | | 0.247 |
| Strongly agree | 79 (32.5) | 24 (30.4) | 29 (36.7) | 26 (32.9) | |
| Agree | 65 (26.7) | 13 (20) | 28 (43.1) | 24 (36.9) | |
| Neither agree nor disagree | 18 (7.4) | 8 (44.4) | 5 (27.8) | 5 (27.8) | |
| Disagree | 54 (22.2) | 22 (40.7) | 14 (25.9) | 18 (33.3) | |
| Strongly disagree | 27 (11.1) | 12 (44.4) | 8 (29.6) | 7 (25.9) | |

| Answers T3 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 1 | | | | | 0.355 |
| Strongly agree | 12 (4.9) | 0 (0) | 7 (58.3) | 5 (41.7) | |
| Agree | 16 (6.6) | 8 (50) | 4 (25) | 4 (25) | |
| Neither agree nor disagree | 3 (1.2) | 1 (33.3) | 1 (33.3) | 1 (33.3) | |
| Disagree | 46 (18.9) | 17 (37) | 15 (32.6) | 14 (30.4) | |
| Strongly disagree | 166 (68.3) | 53 (31.9) | 57 (34.3) | 56 (33.7) | |
| QUESTION 2 | | | | | 0.268 |
| Strongly agree | 136 (56) | 46 (33.8) | 49 (36) | 41 (30.1) | |
| Agree | 64 (26.3) | 22 (34.4) | 17 (26.6) | 25 (39.1) | |
| Neither agree nor disagree | 7 (2.9) | 1 (14.3) | 4 (57.1) | 2 (28.6) | |
| Disagree | 26 (10.7) | 10 (38.5) | 8 (30.8) | 8 (30.8) | |
| Strongly disagree | 10 (4.1) | 0 (0) | 6 (60) | 4 (40) | |
| QUESTION 3 | | | | | 0.568 |
| Strongly agree | 21 (8.6) | 10 (47.6) | 6 (28.6) | 5 (23.8) | |
| Agree | 55 (22.6) | 21 (38.2) | 16 (29.1) | 18 (32.7) | |
| Neither agree nor disagree | 11 (4.5) | 2 (18.2) | 4 (36.4) | 5 (45.5) | |
| Disagree | 90 (37) | 29 (32.2) | 30 (33.3) | 31 (34.4) | |
| Strongly disagree | 66 (27.2) | 17 (25.8) | 28 (42.4) | 21 (31.8) | |
| QUESTION 4 | | | | | 0.072 |
| Strongly agree | 35 (14.4) | 13 (37.1) | 12 (34.3) | 10 (28.6) | |
| Agree | 58 (23.9) | 22 (37.9) | 18 (31) | 18 (31) | |
| Neither agree nor disagree | 13 (5.3) | 7 (53.8) | 4 (30.8) | 2 (15.4) | |
| Disagree | 81 (33.3) | 24 (29.6) | 22 (27.2) | 35 (43.2) | |
| Strongly disagree | 56 (23) | 13 (23.2) | 28 (50) | 15 (26.8) | |
| QUESTION 5 | | | | | 0.822 |
| Strongly agree | 75 (30.9) | 22 (29.3) | 29 (38.7) | 24 (32) | |
| Agree | 83 (34.2) | 27 (32.5) | 26 (31.3) | 30 (36.1) | |
| Neither agree nor disagree | 20 (8.2) | 7 (35) | 9 (45) | 4 (20) | |
| Disagree | 48 (19.8) | 18 (37.5) | 13 (27.1) | 17 (35.4) | |
| Strongly disagree | 17 (7) | 5 (29.4) | 7 (41.2) | 5 (29.4) | |

| Answers T3 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 6 | | | | | 0.802 |
| Strongly agree | 64 (26.3) | 20 (31.2) | 25 (39.1) | 19 (29.7) | |
| Agree | 81 (33.3) | 25 (30.9) | 26 (32.1) | 30 (37) | |
| Neither agree nor disagree | 19 (7.8) | 8 (42.1) | 5 (26.3) | 6 (31.6) | |
| Disagree | 46 (18.9) | 18 (39.1) | 16 (34.8) | 12 (26.1) | |
| Strongly disagree | 33 (13.6) | 8 (24.2) | 12 (36.4) | 13 (39.4) | |
| QUESTION 7 | | | | | 0.391 |
| Strongly agree | 22 (9.1) | 5 (22.7) | 12 (54.5) | 5 (22.7) | |
| Agree | 54 (22.2) | 17 (31.5) | 18 (33.3) | 19 (35.2) | |
| Neither agree nor disagree | 26 (10.7) | 10 (38.5) | 10 (38.5) | 6 (23.1) | |
| Disagree | 105 (43.2) | 36 (34.3) | 29 (27.6) | 40 (38.1) | |
| Strongly disagree | 36 (14.8) | 11 (30.6) | 15 (41.7) | 10 (27.8) | |
| QUESTION 8 | | | | | 0.488 |
| Strongly agree | 69 (28.4) | 20 (29) | 28 (40.6) | 21 (30.4) | |
| Agree | 86 (35.4) | 28 (32.6) | 25 (29.1) | 33 (38.4) | |
| Neither agree nor disagree | 22 (9.1) | 9 (40.9) | 10 (45.5) | 3 (13.6) | |
| Disagree | 44 (18.1) | 14 (31.8) | 13 (29.5) | 17 (38.6) | |
| Strongly disagree | 22 (9.1) | 8 (36.4) | 8 (36.4) | 6 (27.3) | |
| QUESTION 9 | | | | | 0.576 |
| Strongly agree | 109 (44.9) | 37 (33.9) | 38 (34.9) | 34 (31.2) | |
| Agree | 64 (26.3) | 15 (23.4) | 27 (42.2) | 22 (34.4) | |
| Neither agree nor disagree | 21 (8.6) | 10 (47.6) | 5 (23.8) | 6 (28.6) | |
| Disagree | 34 (14) | 13 (38.2) | 9 (26.5) | 12 (35.3) | |
| Strongly disagree | 15 (6.2) | 4 (26.7) | 5 (33.3) | 6 (40) | |
| QUESTION 10 | | | | | 0.382 |
| Strongly agree | 99 (40.7) | 29 (29.3) | 40 (40.4) | 30 (30.3) | |
| Agree | 63 (25.9) | 19 (30.2) | 20 (31.7) | 24 (38.1) | |
| Neither agree nor disagree | 15 (6.2) | 6 (40) | 6 (40) | 3 (20) | |
| Disagree | 44 (18.1) | 18 (40.9) | 14 (31.8) | 12 (27.3) | |
| Strongly disagree | 22 (9.1) | 7 (31.8) | 4 (18.2) | 11 (50) | |

| Answers T4 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 1 | | | | | 0.77 |
| Strongly agree | 7 (3) | 2 (28.6) | 4 (57.1) | 1 (14.3) | |
| Agree | 21 (8.9) | 6 (28.6) | 6 (28.6) | 9 (42.9) | |
| Neither agree nor disagree | 2 (0.9) | 1 (50) | 1 (50) | 0 (0) | |
| Disagree | 49 (20.9) | 18 (36.7) | 14 (28.6) | 17 (34.7) | |
| Strongly disagree | 156 (66.4) | 51 (32.7) | 57 (36.5) | 48 (30.8) | |
| QUESTION 2 | | | | | 0.525 |
| Strongly agree | 123 (52.3) | 43 (35) | 47 (38.2) | 33 (26.8) | |
| Agree | 62 (26.4) | 17 (27.4) | 18 (29) | 27 (43.5) | |
| Neither agree nor disagree | 10 (4.3) | 3 (30) | 3 (30) | 4 (40) | |
| Disagree | 26 (11.1) | 11 (42.3) | 9 (34.6) | 6 (23.1) | |
| Strongly disagree | 14 (6) | 4 (28.6) | 5 (35.7) | 5 (35.7) | |
| QUESTION 3 | | | | | 0.118 |
| Strongly agree | 17 (7.2) | 4 (23.5) | 9 (52.9) | 4 (23.5) | |
| Agree | 50 (21.3) | 21 (42) | 11 (22) | 18 (36) | |
| Neither agree nor disagree | 14 (6) | 7 (50) | 3 (21.4) | 4 (28.6) | |
| Disagree | 100 (42.6) | 33 (33) | 33 (33) | 34 (34) | |
| Strongly disagree | 54 (23) | 13 (24.1) | 26 (48.1) | 15 (27.8) | |
| QUESTION 4 | | | | | 0.563 |
| Strongly agree | 44 (18.7) | 17 (38.6) | 15 (34.1) | 12 (27.3) | |
| Agree | 57 (24.3) | 21 (36.8) | 16 (28.1) | 20 (35.1) | |
| Neither agree nor disagree | 13 (5.5) | 5 (38.5) | 5 (38.5) | 3 (23.1) | |
| Disagree | 75 (31.9) | 25 (33.3) | 24 (32) | 26 (34.7) | |
| Strongly disagree | 46 (19.6) | 10 (21.7) | 22 (47.8) | 14 (30.4) | |
| QUESTION 5 | | | | | 0.389 |
| Strongly agree | 76 (32.3) | 19 (25) | 31 (40.8) | 26 (34.2) | |
| Agree | 78 (33.2) | 25 (32.1) | 28 (35.9) | 25 (32.1) | |
| Neither agree nor disagree | 24 (10.2) | 12 (50) | 7 (29.2) | 5 (20.8) | |
| Disagree | 36 (15.3) | 16 (44.4) | 9 (25) | 11 (30.6) | |
| Strongly disagree | 21 (8.9) | 6 (28.6) | 7 (33.3) | 8 (38.1) | |

| Answers T4 | All (%) | Observed (%) | Passively Moderated (%) | Actively Moderated (%) | P |
|---|---|---|---|---|---|
| QUESTION 6 | | | | | 0.915 |
| Strongly agree | 66 (28.1) | 19 (28.8) | 26 (39.4) | 21 (31.8) | |
| Agree | 74 (31.5) | 25 (33.8) | 24 (32.4) | 25 (33.8) | |
| Neither agree nor disagree | 19 (8.1) | 9 (47.4) | 6 (31.6) | 4 (21.1) | |
| Disagree | 46 (19.6) | 16 (34.8) | 16 (34.8) | 14 (30.4) | |
| Strongly disagree | 30 (12.8) | 9 (30) | 10 (33.3) | 11 (36.7) | |
| QUESTION 7 | | | | | 0.021 |
| Strongly agree | 16 (6.8) | 3 (18.7) | 9 (56.2) | 4 (25) | |
| Agree | 58 (24.7) | 11 (19) | 27 (46.6) | 20 (34.5) | |
| Neither agree nor disagree | 28 (11.9) | 11 (39.3) | 12 (42.9) | 5 (17.9) | |
| Disagree | 92 (39.1) | 36 (39.1) | 22 (23.9) | 34 (37) | |
| Strongly disagree | 41 (17.4) | 17 (41.5) | 12 (29.3) | 12 (29.3) | |
| QUESTION 8 | | | | | 0.128 |
| Strongly agree | 58 (24.7) | 16 (27.6) | 26 (44.8) | 16 (27.6) | |
| Agree | 84 (35.7) | 25 (29.8) | 31 (36.9) | 28 (33.3) | |
| Neither agree nor disagree | 20 (8.5) | 10 (50) | 7 (35) | 3 (15) | |
| Disagree | 49 (20.9) | 20 (40.8) | 9 (18.4) | 20 (40.8) | |
| Strongly disagree | 24 (10.2) | 7 (29.2) | 9 (37.5) | 8 (33.3) | |
| QUESTION 9 | | | | | 0.892 |
| Strongly agree | 98 (41.7) | 34 (34.7) | 37 (37.8) | 27 (27.6) | |
| Agree | 74 (31.5) | 21 (28.4) | 26 (35.1) | 27 (36.5) | |
| Neither agree nor disagree | 17 (7.2) | 6 (35.3) | 5 (29.4) | 6 (35.3) | |
| Disagree | 36 (15.3) | 13 (36.1) | 10 (27.8) | 13 (36.1) | |
| Strongly disagree | 10 (4.3) | 4 (40) | 4 (40) | 2 (20) | |
| QUESTION 10 | | | | | 0.851 |
| Strongly agree | 91 (38.7) | 30 (33) | 32 (35.2) | 29 (31.9) | |
| Agree | 62 (26.4) | 20 (32.3) | 25 (40.3) | 17 (27.4) | |
| Neither agree nor disagree | 14 (6) | 3 (21.4) | 6 (42.9) | 5 (35.7) | |
| Disagree | 55 (23.4) | 21 (38.2) | 14 (25.5) | 20 (36.4) | |
| Strongly disagree | 13 (5.5) | 4 (30.8) | 5 (38.5) | 4 (30.8) | |