International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Development and validation of a questionnaire measuring secondary students' genetic essentialism and teleology (GET) conceptions

Florian Stern, Kostas Kampourakis, Marine Delaval & Andreas Müller

To cite this article: Florian Stern, Kostas Kampourakis, Marine Delaval & Andreas Müller (2020) Development and validation of a questionnaire measuring secondary students' genetic essentialism and teleology (GET) conceptions, International Journal of Science Education, 42:2, 218-252, DOI: 10.1080/09500693.2019.1707905

To link to this article: https://doi.org/10.1080/09500693.2019.1707905

Published online: 12 Jan 2020.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
2020, VOL. 42, NO. 2, 218–252
https://doi.org/10.1080/09500693.2019.1707905

Development and validation of a questionnaire measuring secondary students' genetic essentialism and teleology (GET) conceptions

Florian Stern [a], Kostas Kampourakis[a,b], Marine Delaval[c*] and Andreas Müller[b,c]

[a]Faculty of Science, Section of Biology, University of Geneva, Geneva, Switzerland; [b]University Teacher Training Institute (IUFE), University of Geneva, Geneva, Switzerland; [c]Faculty of Science, Physics Section, University of Geneva, Geneva, Switzerland

ABSTRACT

In this article, we describe the main phases in the development and validation of a questionnaire measuring secondary students' teleology and essentialism conceptions in the context of genetics. The validation process involved 714 Swiss and French secondary school students from 14 different schools. The process included interviews, several pilot studies, experts' feedback, and lasted from January 2017 to May 2018. The final version of the test consists of 20 items based on an ordinal scale, which measure teleology and essentialism conceptions in the context of genetics. Exploratory factor analysis showed a structure with three factors and Cronbach's alpha values indicated a good internal consistency of the items. We also report findings on students' genetic determinism conceptions, even though the respective items were not included in the final questionnaire because their psychometric properties were not as good as those for genetic teleology and genetic essentialism. Implications for science education are discussed.

ARTICLE HISTORY
Received 1 December 2018
Accepted 18 December 2019

KEYWORDS
Assessment; genetics; misconceptions; secondary school; science education; teleology; essentialism; intuitions; genetic essentialism; genetic teleology; genetic determinism

Introduction

Teaching and learning science are based on conceptual understanding. An important distinction is among *concepts* that are mental representations of the world, *conceptions* that are the different meanings associated with particular concepts, and *intuitions* that are spontaneous ways of thinking about phenomena (Kampourakis, 2014). Therefore, concepts, conceptions, and intuitions are not the same, even though they are related and often influence one another, and this is why these terms cannot be used interchangeably. For example, an important concept in evolution is natural selection, which is an unconscious process of differential survival of the individuals of a particular population, which differ from one another in particular inheritable traits, in a particular environment. Different conceptions can be associated with this concept: one may think of natural

CONTACT Florian Stern florian.stern@unige.ch Faculty of Science, University of Geneva, 40 Bd du Pont d'Arve, CH 1205, Geneva, Switzerland
*Present address: INSPÉ Académie de Lille - Hauts de France, Université Lille, EA 4072; PSITEC - Psychologie: Interactions, Temps, Émotions, Cognition, Lille, France

selection as a natural, but goal-directed process of population change towards adaptation (conception 1) or as a process involving an external agent (e.g. Nature or God) that acts intentionally and guides the population to adaptation (conception 2) (for various examples of different conceptions of natural selection, see the chapters in Rosengren, Brem, Evans, & Sinatra, 2012). It is important to note that conceptions may be correct or incorrect, and this is why students' conceptions prior to teaching are better described as *preconceptions*. Those preconceptions that are scientifically incorrect can be described as *misconceptions* (Vosniadou, 2012). Therefore, we consider misconceptions to form a subset of preconceptions, and teaching towards conceptual change as the process of challenging students' preconceptions with the aim of replacing/restructuring them in order for students to acquire a correct understanding of the respective concepts (Kampourakis & Reiss, 2018).

What is most important is that students' conceptions often stem from intuitions that are strongly held and resistant to change. These intuitions can make scientific concepts and explanations seem counterintuitive (see Shtulman, 2017 for an overview). The interplay between intuitive thinking and students' conceptions and explanations about biological entities and processes has been studied in cognitive science research (Carey, 1985; Inagaki & Hatano, 2002, 2006; Keil, 1989; Medin & Atran, 1999). Previous research aims at answering general, important questions such as whether it is knowledge about artefacts that influences knowledge about organisms or the opposite (Keil, Greif, & Kerner, 2007; Kelemen & Carey, 2007); or whether causes are perceived by humans as central in the process of explanation (Keil & Wilson, 2000; Lombrozo & Carey, 2006). Furthermore, it can also provide insights about how teleology (Kelemen, 2012), essentialism (Gelman & Rhodes, 2012), and genetic essentialism intuitions (Heine, Dar-Nimrod, Cheung, & Proulx, 2017) can form obstacles in students' conceptual understanding in particular domains such as evolution and genetics. It is, therefore, necessary to document students' conceptions so that science teachers explicitly address them during teaching.

Eventually, the ultimate aim of documenting students' conceptions and the respective conceptual obstacles is conceptual change. This differs from simple knowledge acquisition or enrichment because it usually occurs due to conceptual conflict. Conceptual change occurs when someone already possesses some prior idea/belief and realises that it is internally inconsistent or contradictory (Carey, 2009; Chi, 2013). It may involve changes in the internal structure of concepts, or changes, central to their meanings, in the relations of concepts to others (see Keil & Newman, 2008). In this sense, conceptual change may involve changes in the meaning of concepts or introduction of new ones in explanations, but also the restructuring of explanations and consequently changes in the relations among concepts. As understanding the role of intuitions adds to the understanding of conceptions, it is of great importance to investigate intuitions and see how they impact students' conceptions and misconceptions.

In our study, we focus on students' understanding of genetics. There are two reasons for this. On the one hand, research in genetics has multiple implications for decisions on important socioscientific issues (Kampourakis, Reydon, Patrinos, & Strasser, 2014). On the other hand, people form an intuitive understanding of heredity from a very young age, and therefore develop preconceptions about how the related phenomena occur (Ergazaki, 2018).

It is well known that genetics is difficult to understand because students hold a variety of misconceptions about the nature of the genetic material, about the role of genes and the nature and potential of genetic technologies. These misconceptions, in turn, stem from a variety of sources such as the mass media, textbooks, teachers themselves and intuitive ways of thinking such as genetic essentialism. To address these misconceptions, teachers should acquire the necessary knowledge both about genetics and about students' misconceptions and prepare teaching materials that are comprehensible; be able to explain problematic metaphors and correct distortions that are so widespread; and be prepared to rearrange the curriculum based on what supports students' understanding (Stern & Kampourakis, 2017).

In this study, we focus on the possible impact of intuitive thinking on students' conceptions about genetics. Conceptual development research has supported the conclusion that persistent human intuitions such as design teleology and psychological essentialism have a significant influence on students' conceptual understanding of evolution (Gelman & Rhodes, 2012; Kelemen, 2012). It may, therefore, be the case that students' misconceptions about genetics are not only due to their misunderstanding of the respective topics, but also due to these intuitions. As design teleology and psychological essentialism are obstacles to understanding evolution, our research hypothesis is that these two intuitions could also be obstacles to understanding genetics. For this reason, it is of great importance to measure the conceptions stemming from these intuitions; we, therefore, investigated whether students express conceptions related to design teleology and psychological essentialism in the context of genetics. As students can exhibit a range of different conceptions related to teleology and essentialism, we considered 'genetic teleology' and 'genetic essentialism' to encompass all these different conceptions and not just 'design teleology' and 'psychological essentialism', respectively. Therefore, hereafter we refer to genetic teleology conceptions and genetic essentialism conceptions. As in many countries, formal genetics education starts in secondary school, it is useful to know whether genetic teleology and genetic essentialism conceptions are already present at this level, in order to address them as soon as possible. We have therefore developed a questionnaire that could be used for this purpose, and in the present article, we report its development in detail.

Theoretical background

Assessment of students' conceptions related to genetics

It must be noted that conceptions, concepts, and intuitions are not directly observable, and as a result cannot be directly measured, unlike a property such as weight or height. Consequently, all we can do is make inferences about students' conceptions and/or intuitions based on their explanations. Various conceptual assessments have been developed for measuring students' understanding of genetics. The aim of the Genetics Concept Assessment is to evaluate science and non-science undergraduate students' understanding of genetics. It consists of 25 multiple-choice items assessing, for example, students' knowledge of the molecular structure of genes through questions such as 'Describe the molecular anatomy of genes and genomes', or students' understanding of the mechanisms of inheritance through questions such as 'Describe the mechanisms by which an organism's genome is passed on to the next generation' (Smith, Wood, & Knight, 2008). In a similar way, the Genetics Literacy Assessment Instrument (GLAI) was developed to measure undergraduate science students' basic understanding of genetics. Through 31 multiple-choice items, students' conceptual knowledge is assessed within six fundamental areas for genetic understanding: the nature of the genetic material, transmission, gene expression, gene regulation, evolution, and genetics and society (Bowling et al., 2008). However, both of these instruments were intended for undergraduates, not secondary students. In order to better investigate the genetics understandings of the latter, a test was developed that specifically focused on secondary students' reasoning in genetics, asking them to provide a written justification for each item. In particular, 13 multiple-choice items were developed, based on six categories that were deemed to be fundamental for reasoning on genetics mechanisms: 'Mapping genotype to phenotype', 'Punnett squares, meiosis and mitosis process', or 'Monohybrid inheritance: Mapping phenotype to genotype' (Tsui & Treagust, 2010).

Among the recent studies, two stand out as directly related to our own. The outcome of one of them is the Public Understanding of Genetics and Genomics ('PUGGS') questionnaire. It consists of 45 multiple-choice items such as Likert-scale or True/False questions, focusing on 11 core ideas and aiming at simultaneously assessing college students' knowledge about modern genetics and genomics, their belief in genetic determinism, and their attitudes towards applications of modern genetics and genomic-based technologies (Carver, Castéra, Gericke, Evangelista, & El-Hani, 2017). Another study investigated the prevalence of teleology, essentialism and anthropocentrism conceptions through 12 Likert-scale items administered to biology major and non-major students. An interesting finding of this study was the absence of correlation between teleology and essentialism conceptions; however, only a few of their items focused on genes (Coley & Tanner, 2015).[1]

As some recent articles on test development within different fields emphasise the importance of the validation process (Adams & Wieman, 2011; Lewthwaite & Fisher, 2005; Summers & Abd-El-Khalick, 2018; Velayutham, Aldridge, & Fraser, 2011), it seems that more attention should be paid to this issue in biology education research. A study of the validity and reliability of 95 assessments in genomics and bioinformatics has concluded that a vast majority of them (90%) failed to mention evidence for reliability or validity. Therefore, researchers should pay special attention to the validation process of a new instrument, because using low-quality instruments can lead to wrong conclusions about students' understandings of genetics (Campbell & Nehm, 2013).

Genetic teleology

An explanation can be described as *teleological* when it refers to a goal or a purpose. For example, if one asked: 'Why do airplanes have wings?', a typical answer would be: 'Airplanes have wings in order to fly'. This is a teleological statement, because airplanes were intentionally designed by humans in order to fly and indeed have wings for the purpose of flying. Teleological statements of this kind have a heuristic value as they help us make sense of common phenomena. But now consider the question: 'Why do birds have wings?'. Does the response 'Birds have wings in order to fly' still make sense? The answer is not as obvious as in the previous case. Many birds like eagles and hawks do in fact use their wings in order to fly; however, other birds like penguins or ostriches that also have wings are not able to fly. The variety of birds' flying ability is the outcome of an unintentional natural process, natural selection; as a result, there is no inherent goal related to the existence of birds' wings. Wings were likely useful for flight to the common ancestor of all birds, but are not necessarily used for this role in all contemporary species.

To make sense of this, it is crucial to distinguish between two kinds of teleology: first, *design teleology* in the case of artefacts, in which an intentional agent created a particular structure for a specific purpose; and *selective teleology* in the case of organisms, in which natural selection has unintentionally brought about a functional structure (Lennox & Kampourakis, 2013; Kampourakis, 2014). Whereas other conceptualisations of teleological thinking exist, distinguishing among 'basic function-based'; 'basic need-based'; 'elaborated need-based' explanations (Kelemen, 2012), we think that the distinction between design teleology and selective teleology is very important as it highlights an important difference between artefacts and organisms: the former have parts designed for a purpose, whereas the latter do not. All artefacts have parts for some role because this is what they were designed for; organisms can be said to have parts for some role only if those parts have become prevalent through natural selection because of what they do. The main issue here is not teleology per se but the underlying consequence aetiology that is based on design (Kampourakis, 2019).

Many studies have found that design-teleological thinking is an obstacle for students' and children's understanding of evolution (Kelemen, 2012). Not only students and children, but also physical scientists seem to intuitively use teleological explanations to make sense of natural phenomena (Kelemen, Rottman, & Seston, 2013). Interestingly, the teleological bias seems to be independent of cultures as non-schooled Romani adults (Casler & Kelemen, 2008) and Chinese people (Rottman et al., 2017) have also been found to provide purpose-based explanations.

In the context of genetics, we thus refer to *genetic teleology* when the existence of genes is explained by a particular goal, which may take different forms, such as design ('genes were designed in order to fulfil a role'), need ('genes exist because of an organism's need to fulfil a role'), or natural selection ('genes exist for a role because they were selected for conferring this role').

Genetic essentialism

Essentialism is the idea that entities have essences, a set of properties that all members of the kind must have, and the combination of which only members of the kind do, in fact, have (Wilkins, 2013). Various forms of essentialism exist; one among them is *psychological essentialism,* the intuition that organisms have underlying essences, which are fixed and thus unchangeable (Gelman, 2003). In particular, psychological essentialism is used to refer to the intuitive beliefs that certain categories are real rather than invented by humans, as well as that these categories possess an underlying causal force, the essence, which is responsible that the members of these categories are the way they are and share many properties (Gelman & Rhodes, 2012).

The essences of artefacts are those among their features that relate to their intended use: wings are part of the essence of airplanes because the intended use of airplanes is flight and wings contribute to that. The essences of artefacts are fixed in the sense that even if an airplane is put in a museum and is no longer used for flying, this will not change the fact that flying is what its wings were designed for. Organisms might also be perceived to have essences, in the sense that they have a developmental potential that produces robust outcomes (a baby will develop to a human, a puppy will develop to a dog, a kitten will develop to a cat, etc.). However, the essences of organisms are not fixed. They are characterised by developmental plasticity that results in variation between members of different species, and also within members of the same species. Therefore, artefact essentialism is significantly different from organism essentialism (Kampourakis, 2014).

Essentialist thinking is popular among children and adults who intuitively not only think in term of essences, but also perceive them as fixed. For instance, children have been found to consider internal, invisible features and properties as more important than external ones (Gelman & Wellman, 1991). Adults have also been found to exhibit essentialist thinking by expecting members of a category to be alike in non-obvious ways (Gelman, 2005). Eventually, studies investigating the influence of essentialism on understanding biological concepts have shown that essentialist thinking is a conceptual obstacle to understanding evolution (Gelman & Rhodes, 2012). Psychological essentialist thinking is correct for artefacts whose essences remain fixed, but not for organisms that undergo major changes during development and evolution (Shtulman & Schulz, 2008).

Genetic essentialism is the idea that genes constitute our essence. It has been conceptualised as comprising four dimensions: the homogeneity of genes within species, which downplays the variation among the members of the same species; the fixity of genes, which entails that they are transmitted immutable across generations; genes as internal, single causes, which means that they directly cause observable traits; and the inference for the presence of gene from a given physical or mental characteristic, which entails that the presence of a particular gene can be inferred from the observation of a related characteristic (Dar-Nimrod & Heine, 2011). In our study, we only investigate two of these dimensions: genes as fixed entities and gene homogeneity within species.

Genetic determinism

Genetic determinism is the idea of the existence of genes '*for*' traits. According to this view, genes determine traits and diseases in a fixed manner, independently of the environment in which an individual lives. This is scientifically inaccurate, as all characters are the outcome of complex interactions between genes and the environment that take place during the development. The relation between genes and traits is a many-to-many one, with many genes being implicated in the development of each trait, and with each gene being implicated in the development of several different traits (Kampourakis, 2017).

Genetic determinism has been shown to be a serious obstacle for understanding genetics and its implications for society (Jiménez-Aleixandre, 2014). A study with British secondary students found that when they were asked the question 'Why are genes important?', 73% of them responded that genes are involved in the determination of characteristics and only 14% that they are involved in the transfer of information (Lewis, Leach, & Wood-Robinson, 2000). In another study with secondary students in the USA, it was concluded that they tended to map the molecular (genotypic) level directly onto higher levels (which may include that of the phenotype), thus overlooking the important role of proteins and of all other phenomena and mechanisms that result in the production of phenotypes (Duncan & Reiser, 2007). In addition, it has been found that even biology teachers may rely on genetic determinism to explain intellectual sameness between twins, gender differences or the superiority of some ethnic groups (Castéra & Clément, 2014).

More recently, two empirical studies investigated genetic determinism conceptions of university students. In one of them, 56 British 1st- and 2nd-year undergraduate students took part in a study comparing two curricula: half of the students were assigned the classical Mendelian curriculum characterised by a strong genetic determinist component; whereas the other half were taught a Weldonian curriculum with a weak genetic determinist component, emphasising the importance of environment and developmental processes. The comparison of students' genetic determinism conceptions before and after teaching showed that the students who had been taught the Mendelian curriculum showed no change in their genetic determinism conceptions, whereas the students who had been taught the Weldonian curricula had less deterministic views about genes (Jamieson & Radick, 2017). In another study, 446 Brazilian 1st-year undergraduate students answered the PUGGS questionnaire about genetic determinism. There were four major results: first, belief in genetic determinism is a two-dimensional construct, as beliefs about social and biological traits are distinct; second, there is little belief in genetic determinism about social traits; third, the association between knowledge of genetics and belief in genetic determinism is weak; fourth, social factors such as age and religiosity show stronger associations with belief in genetic determinism rather than knowledge (Gericke et al., 2017). As mentioned by the authors, the third finding of this study directly raises the question of the usefulness of genetics literacy for decreasing beliefs in genetic determinism, even though this contradicts the results of Jamieson and Radick (2017), who provide evidence for the importance of teaching about this.

Summary of constructs

In sum, several past studies have found misconceptions based on teleology, essentialism and genetic determinism in students' explanations of biological phenomena. For this reason, we decided to investigate further secondary students' genetic determinism, as well as genetic teleology and genetic essentialism conceptions. We were also interested in associations between such conceptions. First, previous research has documented a link between evolution understanding and design teleology, as well as evolution understanding and psychological essentialism. Therefore, as these two intuitions are linked with evolution, our hypothesis is that a similar link may exist between genetic teleology and genetic determinism, and between genetic essentialism and genetic determinism. Second, the potential existence of a genetic teleology/genetic determinism and a genetic essentialism/genetic determinism link is also supported by the theoretical assumption that: (1) if genes are considered as our essence (genetic essentialism), then they are likely to determine phenotypic outcomes (genetic determinism); (2) if genes are considered as having a purpose (genetic teleology), then they are likely to determine phenotypic outcomes (genetic determinism). To the best of our knowledge, no studies have explored the potential links between genetic determinism and genetic teleology, nor between genetic determinism and genetic essentialism. However, in the case of genetic teleology and genetic essentialism, some empirical evidence exists about their relation. Two empirical studies have however found no association between teleological and essentialist thinking (Coley & Tanner, 2015; Stern et al., 2018). Based on this, we also expected to find no link between genetic teleology and genetic essentialism. In the results section, we provide empirical evidence of potential links between these constructs.

In the present study, we aimed at developing an instrument measuring at the same time students' conceptions related to genetic teleology, genetic essentialism and genetic determinism. To the best of our knowledge, no instrument was ever designed to investigate simultaneously these specific conceptions. However, even though here we present the results on students' genetic determinism conceptions, we decided not to include the respective items in the final questionnaire that we present, because their psychometric indicators were not as good as those of the other constructs. We first describe the *structure* of the interviews and the various versions of the questionnaires, as well as the kind of statistical methods used for assessing reliability and validity. Then we present the *content* and the evolution of the items, as well as statistical results from the implementation of the questionnaire.

Overview of the study

Context of study and participants

Our research was funded by the Swiss National Science Foundation. In Geneva, Switzerland, students can attend three different types of upper-secondary schools: 'formation gymnasiale' (hereafter FG), 'formation de culture générale' (hereafter FC), and 'formation professionnelle' (hereafter FP). According to the International Standard Classification of Education of the UNESCO, the first two belong to ISCED level 34 called 'upper secondary general education', whereas the last one belongs to ISCED level 35 named 'upper secondary vocational education' (Unesco, 2012; Office fédéral de la statistique, 2015).

In total, 663 Swiss secondary school students (age 15–19 years old) from 14 different schools took part either in the interviews or in one of the four pilot studies P_1, P_2, P_3, P_4. Additionally, 51 French students from upper-secondary school named Lycée' (hereafter 'L') (denoted as S4, which belongs to ISCED level 34) aged 15–18 years old also took part in the last pilot study P_4 (they were not available to participate in P_1–P_3 due to constraints of their teaching programme). Females were a bit more than males (57% vs. 43%), with the majority of students coming from school S1 (75%), the kind of schools that students typically attend before enrolling in university studies. Overall, we had to work with a convenience sample, based on the availability of students and teachers. Table 1 summarises the characteristics of the sample.

Table 1. Summary of the sample's characteristics. Entries are number of students.

		Qualitative validation				Quantitative validation	
		Interviews (*n* = 52)	P_1 (*n* = 12)	P_2 (*n* = 158)	P_3 (*n* = 96)	P_4 (*n* = 396)	Total (*n* = 714)
Sex	Male	28	5	61	46	164	304
	Female	24	7	97	50	232	410
School Type	FG	33	12	130	80	280	535
	FC	5	–	28	16	65	114
	FP	14	–	–	–	–	14
	L	–	–	–	–	51	51
Age	15–16	16	–	12	–	67	95
	16–17	13	–	64	16	107	200
	17–18	12	–	56	80	180	328
	18–19	11	12	26	–	42	91

Process of development and validation of the instrument

The process of questionnaire development, revision and validation involved qualitative and quantitative methods, through interviews and questionnaires (Cohen, Manion, & Morrison, 2007, pp. 317–383). Generally speaking, questionnaires can be closed-ended or open-ended. In the first case, students have to choose among a certain number of already formulated responses; in the second case, they can provide their responses without being limited in choosing among a predetermined set of responses. Two-tier tests combine the advantages of both types of questionnaires. In the first tier, students have to choose among a set of predetermined responses, whereas in the second-tier they have to justify their choice, or provide an alternative response. Two-tier tests are thus very useful tools for investigating students' conceptions in more details (Tsui & Treagust, 2010).

The development of the GET questionnaire consisted of two distinct phases. In the first one, 52 students were interviewed and two-tier questionnaires were administered to 266 students during three pilot studies (P_1, P_2, P_3). In the second phase, a closed-ended questionnaire was administered to 396 students during the last pilot study (P_4). These studies are described here as pilot studies because the aim of all of them was to develop a reliable and valid questionnaire. The main study, during which we will use the questionnaire presented here for documenting students' conceptions using a large sample is currently underway. Versions of the questionnaires Q_1, Q_2, Q_3, and Q_4 were respectively used in P_1, P_2, P_3, and P_4. Figure 1 provides an overview of the whole project, while Figure 2 provides a more detailed view of the various steps of the project, which are described in detail in the subsequent sections.

Development and validation of the instrument by qualitative methods

As the development of the questionnaire was an iterative process, in this section we present the methods used in each stage and the results of that stage. The reason for this is that the results of each stage were considered for deciding upon the methods of the next stage, and so it is necessary to present together both the methods and the results of each stage.

Interviews

During September and October 2016, the first two authors conducted semi-structured interviews with 52 students from all three types of upper secondary schools. All interviews had an average length of 15 min, were audio-recorded and then transcribed. The interview protocol was based on four basic questions and several others that followed, depending each time on students' responses, knowledge, and interest in the topic. Students' responses

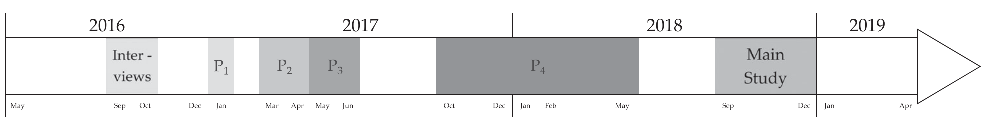

Figure 1. Overview of the whole research project's timetable, from September 2016 to April 2019.

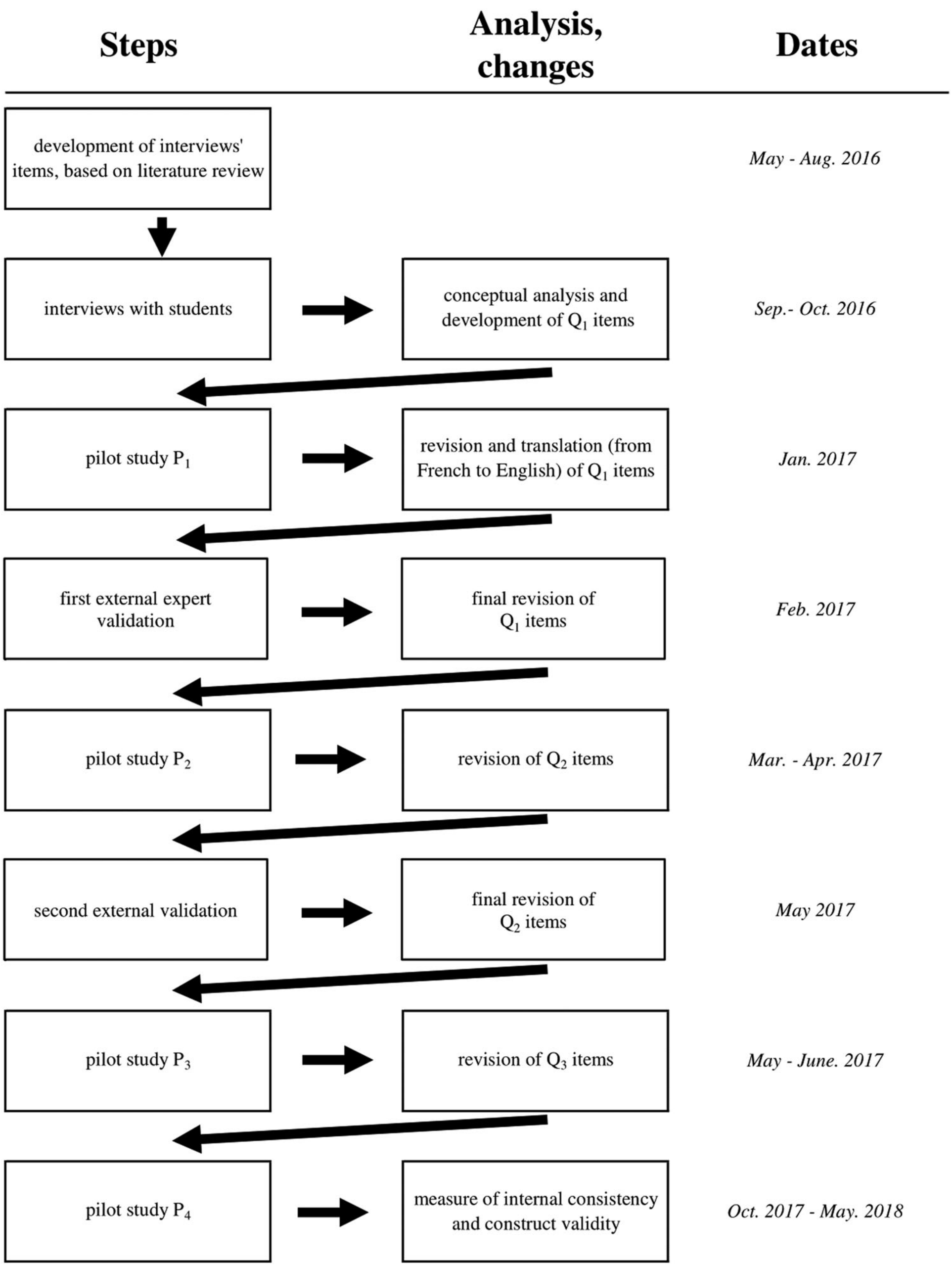

Figure 2. Development and validation of the research instrument.

were analysed using the ATLAS software and we noted conceptions whenever we inferred that they were reflected in the response given. The codes used were not determined in advance, but emerged through students' responses.

The questions we used during the interviews were simple ones on animal and plant traits that students were generally expected to be familiar with. The first question asked students to predict the type of hair that the offspring of particular couples would have, as well as to explain the origin of the type of hair that particular individuals had. On the basis of their predictions and explanations, we managed to explore in detail their understanding of heredity, as well of basic concepts such as 'DNA' and 'genes'. The focus of the second question was on the interaction between genes and the environment, and their impact on the observed phenotypes. Students were asked about the possible differences or similarities between phenotypes of two genetically identical seeds growing up under the same and under different environmental conditions. In the third question, we asked students to predict the adaptation of a population of brown flies to a new, predominantly green environment, in which they would not initially be easily concealed and in which they would thus have a disadvantage. Students were asked to explain whether this population of flies would adapt or die out, the exact mechanism of how this would happen, as well as whether this would happen intentionally or randomly. Finally, in the last question, we asked students to explain how a white rabbit could be born in a family of grey rabbits, and whether this new trait would be due to a change in genes or to environmental factors such as the kind of food that its mother had eaten during gestation. The conceptions were grouped into different categories such as 'genetics basics', 'genetic determinism', 'genetic teleology', and 'genetic essentialism' conceptions.

Figure 3 presents students' most frequent conceptions. These results made us aware that some concepts could easily fit into several categories; for example, we considered *genes as main causes*, held by 75% of the students, to be a genetic essentialism conception; however, it could also be categorised as a genetic determinism conception. Another conception of the same kind is the *inference of genes from traits,* which was held by 67% of the students. This overlap of conceptions is further discussed in the development of items for questionnaires Q_1–Q_4. In contrast, the most popular conception was not a genetic determinism one: *the influence of environment on phenotypes* and *environment,* which was held by 80% of students. One explanation for this is the context of the questions asked (e.g. those environmental factors such as sunlight or water are required for plants to grow). For a similar reason, we also found that the conception *environment influences some traits but not all* was held by 67% of them. Finally, genetic teleology conceptions were also prevalent, for example, *adaptation is purposeful*, and *adaptation will happen if there is time* were respectively expressed by 63% and 51% of the students. Other conceptions about the basic genetics knowledge that were reported were *genes mentioned*, *DNA mentioned* and *siblings inherit different combinations of genes* which were respectively mentioned by 33%, 53% and 53% of students.[2] Additionally, we noticed that some students were not familiar with some words, such as 'allele', whereas all of them were familiar with the term 'gene'. Therefore, to avoid misunderstandings, we decided as a compromise to only use the term 'gene' in the further development of the items, in order to refrain from including the term 'allele' that several students might not know. We should also note also that some other conceptions, such as *change of behaviour to survive* were not explicitly or implicitly related to genetics, although they were expressed by almost a third of the students.

We, thus drew on students' conceptions, explanations and vocabulary related to genetic teleology, genetic essentialism and genetic determinism as documented during the interviews in order to develop items for a questionnaire.

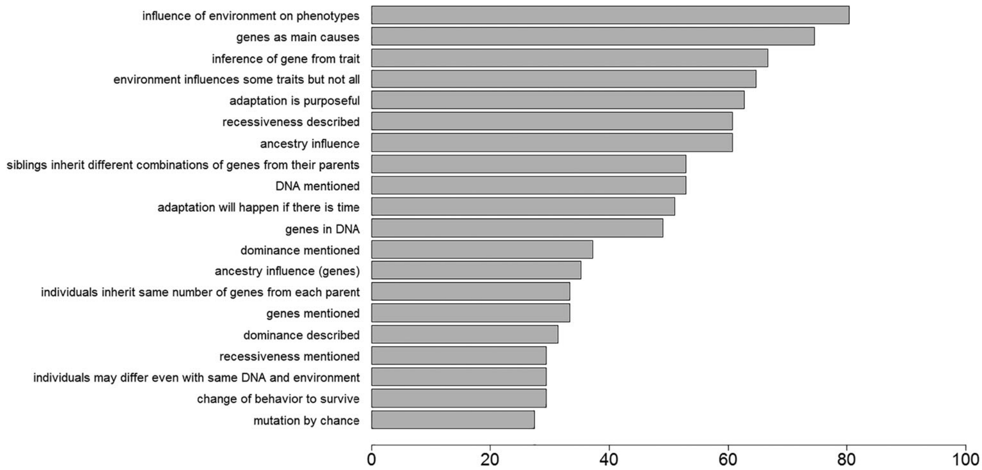

Figure 3. The 20 conceptions most commonly found in 52 students' answers during the interviews, in percent. Note that *Ancestry influence* was used when the influence of ancestors did not relate to genes; while *Ancestry influence (genes)* refers to the same conception whose description did mention genes.

Pilot studies 1–3 (P_1–P_3)

The goal of the first three pilot studies was to develop items for measuring, in a reliable and valid manner, students' genetic determinism conceptions (the degree to which genes determine traits), genetic teleology conceptions (the degree to which genes have a certain purpose) and genetic essentialism conceptions (the degree to which genes are considered as essences). Based on students' most popular conceptions identified through the analysis of interviews and on past research findings, items were developed to test the understanding of relevant conceptions about genetics, teleology and essentialism. These items were the core of the development of 2-tier tests that were qualitatively improved for six months, during the three pilot studies: P_1 (January 2017), P_2 (March–April 2017) and P_3 (May–June 2017) (see Figure 1). Thus, the details in the content of items slightly evolved, and is the outcome of successive revisions based on students', teachers' and experts' feedback about the items.

Questionnaire Q_1 was administered in paper format, whereas questionnaires Q_2 and Q_3 were made available online using the survey software 'LimeSurvey'. Students individually provided their answers in a computer room at their school, under the supervision of their teacher and the first author, and needed on average 30 min to provide their responses. Before the test, students were shortly introduced to the topic of the questionnaire; they were also told that their responses would be treated in an anonymous and confidential manner. In order to be able to distinguish among students, a special code was attributed to each one of them. However, we did not know to which student each code corresponded, but could only use the codes to distinguish between different students. Finally, all items were administered in a random order. A summary of the characteristics of these pilot studies can be found in Table 2.

Pilot Study 1 (P_1)

Pilot Study P_1 was conducted in January 2017 with 12 students. Q_1 consisted of 22 2-tier items that were developed based on the interviews. The genetic determinism items were

Table 2. Evolution of the characteristics of the first three versions of the GET questionnaire.

	Q_1 ($n = 12$)	Q_2 ($n = 158$)	Q_3 ($n = 96$)
Number of items	22	12	12
Number of answers per item	5	4	4
Several answers allowed	No	No	Yes

about various traits such as eye colour, obesity or aggressiveness, for which students were asked to choose an explanation for their origin, among five available ranging from a strong genetic determinism ('genes only') to its total absence ('environment only').[3] The genetic teleology items were about the potential purpose associated with the existence of a particular gene, for example, a gene shared by humans and other organisms, or a gene found in some environments only. Finally, the genetic essentialism items were about the homogeneity of genes within the same species, the fixity of genes, genes as internal, single causes and the inference for the presence of gene from a characteristic. One major confusion that arose was the mix of concepts in some explanations we developed, which was pointed out by students. For example, students were asked to choose between several explanations for the origin of obesity, and one of these simultaneously mentioned the importance of the diet and the regularity of physical activity. The problem with this formulation was that some students partly agreed with the importance of the diet, but not with the regularity of physical activity, or vice-versa. This mixture of concepts did not allow for an accurate description of students' conceptions, and all items were therefore further revised so that each one of them corresponded to a single concept. Another practical issue was the length of the test, which was a bit long as nearly half of the students did not manage to complete it in time. Therefore, the next version of the questionnaire had a reduced number of items.

First external Expert Validation

The items of Q_1 were translated from French to English, and eventually submitted to a panel of seven biology researchers and biology educators for a first revision. These were asked whether they agreed or disagreed that the conceptions we intended to test were really present in our items. The interrater reliability of agreement was computed with Fleiss' Kappa coefficient (Fleiss, 1971), and showed strong interrater reliability of agreement for genetic determinism conceptions (0.84), genetic teleology conceptions (0.81), and genetic essentialism conceptions (0.88). Additionally, some relevant comments helped us to clarify certain statements, as for example some terms were judged by experts to be too complicated for students.

Pilot Study 2 (P_2)

Pilot Study P_2 took place from March to April 2017, involving 158 students. Following the revision of Q_1 items, 12 2-tier items related to genetic determinism, genetic teleology, and genetic essentialism were retained. After the administration of Q_2 and the analysis of students' justifications in the second-tier, we concluded that the items were understood by the majority of students. However, some words were still confusing, especially for the youngest students, such as 'mutations', and they were therefore removed. Furthermore, many students asked for the possibility to choose two answers instead of one for some items (and in a few cases more than one answers could indeed be considered as somehow correct).

Second external Validation

The questionnaire Q_2 was then administered to 16 second-year biology undergraduates from the University of Geneva who were asked to also comment on the comprehensibility of the items. These students were chosen because they are quite close to the age of our participants, but still had a significantly deeper understanding of biology concepts. As expected, they provided relevant and useful comments. For example, they noted that some genetic determinism items answers provided additive explanations (e.g. the influence of genes *and* environment) instead of including the concept of interaction (e.g. the influence of genes *interacting* with the environment), which is more accurate. As a result, the concept of interaction was included in the genetic determinism items in Q_3. The biology undergraduates also suggested that the term 'gene(s)' that appeared in genetic determinism items could be confusing and was replaced by 'genes' (without brackets) for clarity purposes. Finally, some inconsistencies in the use of terms were reported across items, and were accordingly corrected.

Pilot Study 3 (P_3)

Between May and June 2017, a third pilot study P_3 was conducted with 96 secondary school students. Questionnaire Q_3 was a 2-tier questionnaire and included the 12 items based on the revision of the items of Q_2. The major change was the introduction of the concept of interaction in genetic determinism items, and the possibility for students to select two answers instead of only one. In some students' justifications, we found the contrast between genes and environment not only in the genetic determinism items, but also in the genetic essentialism items; this indicated that students might not conceptualise genetic determinism and genetic essentialism as different constructs. As a result, the genetic essentialism items that overlapped with the genetic determinism items were not included in the final questionnaire (Q_4). In addition, the concept of interaction that was included in Q_3 was found to be rather unknown to students, except to the older ones. For this reason, we decided not to mention *interaction* explicitly in Q_4, even though we implicitly referred to this concept when we wrote that both genes and environment contribute to the development of traits. Finally, and most importantly, several responses were irrelevant to the conceptions we intended to measure. This was a major problem, and we decided that the scope of items should be narrowed down. Examples of students' conceptions found in Q_1, Q_2, and Q_3 that were considered relevant to test further are presented in Table 3.

In sum, the three pilot studies P_1 to P_3 provided the grounds for arriving at a new set of items, which were included in Q_4. During P_4 we aimed at arriving at a final set through the selection of specific items, which is described in the next section.

Validation of the Instrument by Quantitative Methods

Pilot study P_4

The goal of the fourth pilot study P_4 (October 2017–May 2018) was to investigate the psychometric indicators of the Q_4 questionnaire such as item difficulties, item discrimination,

Table 3. Example of genetic determinism, genetic teleology and genetic essentialism conceptions found in students' questionnaires Q_1–Q_3.

	Students' Statements
Genetic Determinism	'it is genes that determine the height of a person'
	'some persons gain weight more easily than others, this is written in genes'
Genetic Teleology	'because genes are selected according to our needs'
	'the same genes appear in different organisms because they fulfil functions that organisms require to work correctly. These genes have thus been designed to fulfil some roles that organisms need'
Genetic Essentialism	'there is nothing which changes in genes'
	'each ethnic group has specific genes: it is in fact what allows us to classify people in their ethnic group'

item-to-test correlations, Cronbach Alpha coefficients (Ding & Beichner, 2009; Cronbach, 1951). Furthermore, the structure of the questionnaire and its underlying dimensions were investigated through exploratory factor analysis.

Pilot Study P_4 was conducted with 396 secondary school students, under the same conditions of administration previously described. Questionnaire Q_4 consisted of 30 items: 10 related to genetic determinism, 10 related to genetic teleology and 10 related to genetic essentialism. All previous versions of the questionnaires (Q_1–Q_3) had a lower number of items because there was also a second-tier in all of them. In Q_4, the number of items was substantially higher in order to be able to select items improving the psychometric properties of the instrument. The new items had the same features with the older items, and only referred to different traits. Due to time constraints regarding the administration of the questionnaire in schools, we could not have a second-tier in Q_4. In other words, a compromise was made between the number of tested items and the presence of the second-tier, in order to match a specific time constraint, and so Q_4 was a closed-ended questionnaire. A three-level ordinal scale ('2' for a *strong* intuition underlying a conception, '1' for a *moderate* intuition underlying a conception, '0' for a *weak* intuition underlying a conception) was used to categorise students' answers. A summary of the core conceptions tested in Q_4 is presented in Table 4.

Description of the questionnaire

There were eight genetic determinism items about different human traits, which were divided into two categories. On the one hand, there were four items about *mental* traits: aggressiveness (GD1), intelligence (GD2), high painting skills (GD3) and high sport skills (GD4); on the other hand, there were four items about *biological* traits: obesity (GD5), myopia (GD6), height (GD7) and infarction (GD8). Additionally, the items about French language skills (GD9) and eye colour (GD10) were added and considered as *control* items, in order to see if the influence of the environment for the

Table 4. Summary of the 9 core conceptions tested in the fourth version Q_4 of the GET questionnaire.

Core conceptions		
Genetic determinism (GD)	Genetic teleology (GT)	Genetic essentialism (GE)
Genetic determinism	Design-based teleology	Psychological essentialism
Contribution of both genes and environment	Need-based teleology	Moderate essentialism
Environmental determinism	Natural teleology	Weak essentialism

former and the influence of the genes for the latter would be assumed to be very high by a large majority of students. Students were asked to explain the origins of these traits by selecting one of three kinds of explanations expressing the following conceptions: (1) genetic determinism; (2) interaction of both genes and environment; (3) environmental determinism.[4] These explanations were respectively associated with a *strong*, *moderate* and *weak* genetic determinism intuition. We considered the second explanation as the correct answer, because all traits are the outcome of complex interactions between genes and environment that take place during development.

There were 10 genetic teleology items about various human traits, which were subdivided into two sub-categories. The first one had five items about genes associated with traits that are outcomes of *past* events: opposable thumbs (GT1), big brain (GT2), bipedalism (GT3), communication (GT4), and sociability (GT5). The second one had five more items about genes associated with traits that could be outcomes of hypothetical *future* events: cellulose digestion (GT6), antibody production (GT7), protection from ultraviolet radiation (GT8), protection from high heat (GT9), and protection from radiation (GT10). Students were asked to explain the origins of these traits or abilities by selecting one out of three kinds of explanations expressing the following conceptions: (1) design-based teleology; (2) need-based teleology; (3) natural teleology.[5] These explanations were associated with a *strong*, *moderate* and *weak* design intuition respectively, and we considered the third explanation as the correct answer, as natural selection is what best explains adaptations.[6]

Finally, there were 10 genetic essentialism items about various human traits, which were also subdivided into two sub-categories. The first one had five items about the *homogeneity* of genes between western humans and the following groups: Neanderthals (GE1), Chinese people (GE2), chimpanzees (GE3), Eskimos (GE4), and baboons (GE5).[7] The second one had five items about the *fixity* of genes in the following conditions: daltonism (GE6), breast cancer (GE7), dwarfism (GE8), diabetes (GE9), and Alzheimer's disease (GE10). In the former category, students were asked to explain the genetic differences between themselves and the other groups by selecting one from three kinds of explanations expressing the following conceptions: (1) existence of group-specific genes; (2) large genetic differences between the compared groups; (3) small genetic differences between the compared groups. In the latter category, students were asked to explain how likely the changes in genes were by selecting one from three kinds of explanations expressing the following conceptions: (1) all genes are fixed entities; (2) some genes are fixed entities; some others are changeable entities; (3) all genes are changeable entities. Note that in both cases, explanations (1), (2), (3) were respectively associated with a *strong*, *moderate* and *weak* genetic essentialism intuition, and we considered the third explanation as the correct answer. Once we had a final list of items in English, we had the new ones back-translated in French and we asked three bilingual experts to examine and confirm the wording of the final list of terms in French. Examples of such items are shown in Table 5.

It should be noted that the format of the items is the same within each dimension. We opted for this solution because there is already a contextual variation between items, and we did not want to add construct irrelevant variance in our data (Haladyna & Downing, 2004; Messick, 1984).

Table 5. Theoretical structure and examples of items from the GET questionnaire.

Genetic Determinism Items		
Tested conception	**Example 1 (mental trait)** *Whether a person has outstanding intellectual skills depends on:*	**Example 2 (biological trait)** *Whether a person is tall depends on:*
(1) Genetic determinism	her/his genes only.	her/his genes only.
(2) Contribution of both genes and environment*	her/his genes and external factors (e.g. the education provided by her/his family).	her/his genes and external factors (e.g. the quality of food received during her/his childhood).
(3) Environmental determinism	external factors only (e.g. the education provided by her/his family).	external factors only (e.g. the quality of food received during her/his childhood).
Genetic Teleology Items		
Tested conception	**Example 1 (past event)** *Our brain is much bigger than the one of cave men. Thus, genes associated with a big brain:*	**Example 2 (future event)** *Humans cannot digest cellulose, which is the most abundant sugar on earth. Thus, in the future genes associated with cellulose digestion:*
(1) design-based teleology	have been designed for several roles such as solving complex problems.	will be designed for digesting cellulose.
(2) need-based teleology	have appeared for satisfying several needs such as solving complex problems.	will appear for satisfying the need for digesting cellulose.
(3) natural teleology*	have appeared by chance and were selected for several effects such as solving complex problems.	if they appear by chance, may be selected for digesting cellulose.
Genetic Essentialism Items		
Tested conception	**Example 1 (within-group homogeneity of genes)** *If we analyze the genes of Neanderthals (a prehistoric human group), we will identify:*	**Example 2 (fixity of genes)** *A person with Alzheimer's disease has memory difficulties because of a dysfunctional brain. We assume that in a given family everyone has a good memory. Therefore there are only genes associated with a well-functioning brain. A descendant in this family:*
(1) psychological essentialism	genes specific to them	will have a good memory, because the genes associated with a well-functioning brain always remain fixed.
(2) moderate essentialism	many genes different from ours	will have a good memory, because the genes associated with a well-functioning brain are fixed, even though other may change.
(3) weak essentialism*	few genes different from ours	may have Alzheimer's disease, if the genes associated with a well-functioning brain change into genes associated with Alzheimer.

Conceptions with a star (*) were considered as correct. Note that each item had an additional 'I have no opinion', not shown here for clarity.

Statistical analyses

In this section, we present statistical analyses providing further evidence of the quantitative validation of the items of the GET questionnaire. All calculations were done with the R software, in which the *psych* (Revelle, 2019) package was used.

Items were first administered in two sets with two different random orders on a subsample of 89 students, in order to test for potential order effects. The analysis was carried out through Chi-square tests with significance level p <.05, adjusted for multiple comparisons with the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). No evidence for an order effect was found and we thus kept only one version of the questionnaire for further use (for more details, see Appendix 1). After the administration of Q_4, the quality of items was further investigated through a map (Figure 4) of item difficulty versus item discrimination (Jorion et al., 2015).

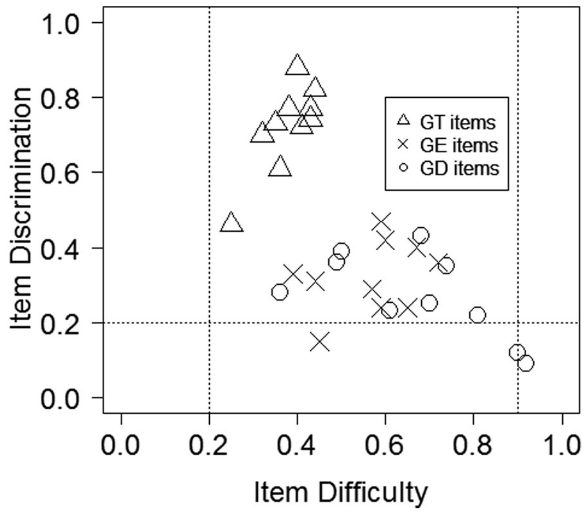

Figure 4. Scatterplot of item difficulty and discrimination value of the GET questionnaire. Symbol represent genetic teleology items (Δ), genetic essentialism items (X) and genetic determinism items (O). Recommendations for minimum and maximum values are in dotted in lines.

Figure 4 provides evidence for an overall good quality of items with respect to difficulties and discrimination recommendations (Ding & Beichner, 2009), except for the two genetic determinism items about French language skills (GD9) and eye colour (GD10) at the bottom right of the map (their difficulty and discrimination values were both out of recommended values and were therefore excluded from further analysis). The genetic essentialism item daltonism (GE6) had a rather low discrimination value (0.13), but was nevertheless kept as it had an acceptable difficulty. Note that most genetic teleology items were highly discriminant (range: 0.41–0.86) and of rather high difficulty (range: 0.25–0.44). In contrast, the discrimination of genetic determinism items (range: 0.22–0.41) and genetic essentialism items (range: 0.13–0.46) were overall lower, while there was a great variation of item difficulty for the genetic determinism items (range: 0.36–0.81) and the genetic essentialism items (range: 0.39–0.72). Finally, construct validity was further explored through exploratory factor analysis using weighted least square method on the polychoric matrix of correlation, which is relevant for ordinal data (Maerten-Rivera, Huggins-Manley, Adamson, Lee, & Llosa, 2015). The Kaiser-Meyer-Olkin (KMO) index checked the adequacy for the analysis, and middling to meritorious KMO = 0.78 was computed (Kaiser & Rice, 1974). Bartlett sphericity test ($\chi^2(378, N = 396) = 5265.74$, $p = 0.00$) supported that the correlation matrix of items was useable for factor analysis (Bartlett, 1951).

In the following sections, we present and discuss the results of two analyses of the structural properties of the questionnaire; the first one including genetic determinism items, and the second one leaving them out.

Analysis 1 (genetic determinism, genetic teleology,& genetic essentialism)

We consider the first 8 genetic determinism, 10 genetic teleology and 10 genetic essentialism items here. As shown in the scree plot (see Figure 5), four eigenvalues are greater

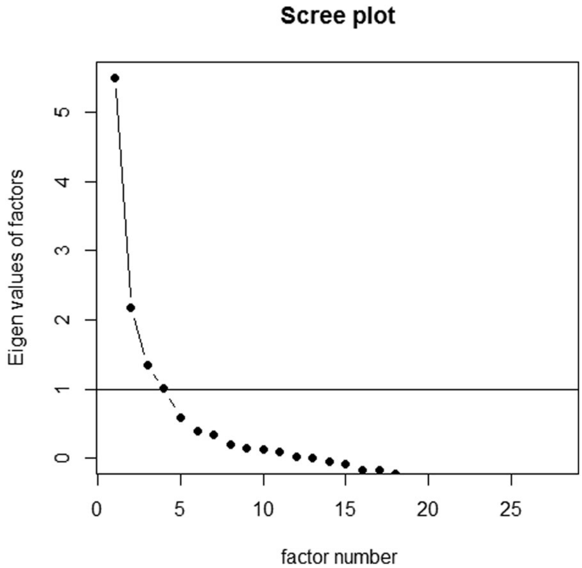

Figure 5. Scree plot considering genetic determinism items The horizontal line emphasises the separation between eigenvalues greater or smaller than 1.

than 1; therefore according to Kaiser rule a structure with four factors is retained (Field, Miles, & Field, 2012).

Table 6 presents the results of factor analysis, which shows evidence for an interpretable structure consisting in 19 items splitting into four factors. This is obtained after the elimination of items that do not load on any factor such as items about myopia (GD6), height (GD7) and infarction (GD8). This four-dimensional structure is also resulting from the removal of items simultaneously loading on several factors and therefore hardly interpretable, such as items about Chinese people (GE2) and Eskimoos (GE4). Besides, items about big brain (GT2), communication (GT4), cellulose digestion (GT6) and high heat protection (GT9) were removed as they had smaller loadings than other items on the same factor.

Additionally, quality fit indexes support an overall acceptable fit (RMSR = 0.05, RMSEA = 0.096 and TLI = 0.76, see Jorion et al., 2015), and Cronbach alpha values suggest a variable internal consistency of items depending on the factor (genetic teleology items (0.82), genes fixity items (0.65); group homogeneity of genes items (0.65), genetic determinism items (0.53), see Table 6).

Overall, the results of the factor analysis showed a clear separation between the different dimensions: genetic teleology; genetic essentialism-homogeneity; genetic essentialism-fixity; genetic determinism. However, there were some concerns and observations about the genetic determinism items.

- First, three genetic determinism items – myopia (GD6), height (GD7) and infarction (GD8), which were about *biological traits* – did not load at all on any factor. In contrast, all genetic determinism items about *mental* traits – aggressiveness (GD1), intelligence (GD2), high painting skills (GD3) and (GD4) high sport skills – loaded on the same

Table 6. Factor loadings of GET. Those smaller than 0.30 are not shown (according to Stevens' recommendations, 2002). Cronbach alpha and associated interval of confidence are also reported in the same table.

	Factors			
Items	Genetic Teleology	Genetic Essentialism (fixity of genes)	Genetic Essentialism (within-group homogeneity of genes)	Genetic Determinism (mental traits)
GD1 Agressiveness				0.58
GD2 Intelligence				0.49
GD3 High Painting Skills				0.80
GD4 High Sport Skills				0.43
GD5 Obesity				0.33
GT1 Opposable Thumbs	0.83			
GT3 Bipedalism	0.81			
GT5 Sociability	0.66			
GT7 Antibodies Production	0.63			
GT8 Ultraviolet Protection	0.79			
GT10 Radiation Protection	0.72			
GE1 Neanderthals			0.48	
GE3 Chimpanzees			0.86	
GE5 Baboons			0.87	
GE6 Daltonism		0.56		
GE7 Breast Cancer		0.62		
GE8 Dwarfism		0.56		
GE9 Diabetes		0.65		
GE10 Alzheimer's disease		0.74		
Cronbach Alpha	0.82	0.65	0.65	0.53
Cronbach Alpha Interval of Confidence (at 5% level)	[0.79, 0.84]	[0.60, 0.71]	[0.59, 0.72]	[0.46, 0.60]

factor. One hypothesis that might explain this would be that the readily observable variation among biological traits may be substantially greater than the variation of mental traits we considered; as a result, the latter are more likely to load together than the former. Such a consideration would require further empirical research in order to be confirmed. Additionally, the remaining genetic determinism item about obesity (GD5) loaded (though with a small correlation) on the same factor together with mental traits, which could be explained if obesity is considered as both a biological trait and a mental trait (in the sense that one cannot control one's appetite). As a consequence, the five remaining items all refer to mental traits; and the removal of the other genetic determinism items leads to the absence of information about biological traits, as the conceptions about those are not measured any more.

- Second, even though the fit indexes of the factor analysis were overall good, the level of internal consistency strongly varied from one dimension to another. The alpha value of genetic determinism items was rather low (0.53), contrary to the ones of genetic teleology and genetic essentialism items, which both were in the range between good and acceptable for a test for conceptual understanding (Jorion et al., 2015).
- Third, the correlation between genetic determinism and each one of the other constructs (genetic teleology, genetic essentialism-fixity, and genetic essentialism-homogeneity) were found to be low (respectively −0.14 (±0.05), 0.20 (±0.05) and −0.08 (±0.05)); This result was not expected according to our initial hypothesis; as we could have expected an association between the genetic determinism and genetic teleology constructs, or an association between the genetic determinism

and genetic essentialism constructs, which was not the case. A reason for the absence of an association between genetic determinism and genetic essentialism may stem from our choice to keep the two dimensions of genetic essentialism that are least overlapping with genetic determinism, according to students' answers in P_1–P_3. In the case of genetic determinism and genetic teleology, even though the concept of 'end' seems to underlie both constructs, this conceptual closeness was not reflected by students' answers. The absence of correlation between the two constructs requires further research.

- Fourth, the five genetic determinism items had rather low loadings compared to those of other dimensions. Among these, three items had loadings below 0.5, whereas in the other dimensions all items but one had loadings greater than 0.5. These low loadings indicate that the genetic determinism dimension could not be empirically established as well as the other ones were.

For all these reasons, we considered genetic determinism items to be less informative than the others, and ran a new analysis (analysis 2, see below) to validate a version of the questionnaire without these items.

Even though the genetic determinism items were eventually removed from the GET questionnaire and are therefore not considered in the rest of the article, it is of interest to note at least two similar results about genetic determinism found in the study using the PUGGS questionnaire (Gericke et al., 2017) and in the present study. First, through a principal component analysis, Gericke and al. exhibited the distinction between biological and social traits when considering beliefs in genetic determinism. With a related analysis, we obtained similar results, as biological traits were separated from mental traits. Note that our items and the ones of Gericke's and al. are sometimes about the same trait, e.g. height that was considered as a biological trait; and aggressiveness (violent behaviour in their study) which was considered as a mental (social in their study) trait. At the same time, some other biological items were only used in our study (e.g. the item about myopia) or in theirs (e.g. the item about schizophrenia); similarly, some other mental/social items were only used in our study (e.g. the item about high sport skills) or in theirs (e.g. the item about addiction to gambling). Second, similarly to Gericke et al.'s study (2017), who found low levels of belief in genetic determination of social traits, in our study the analysis of students' answers to genetic determinism items about mental traits has also shown a weak endorsement of genetic determinism explanations (on average, genetic determinism explanations for such items were endorsed by only 5% of the students). In sum, their study and ours provide evidence that the level of genetic determinism strongly depends on the nature of the traits considered, and that it is low for mental/social traits.

Analysis 2 (genetic teleology & genetic essentialism)

In this analysis, we only considered the 10 genetic teleology and 10 genetic essentialism items. The factor analysis and its associated scree plot show evidence for an interpretable structure with 20 items splitting into three factors: genetic teleology, genetic essentialism-fixity, and genetic essentialism-homogeneity (see Table 7 and Figure 6). The final version of the questionnaire consists of all 10 genetic teleology items, 5 genetic essentialism items about fixity and 5 genetic essentialism items about homogeneity.[8] Cronbach alpha values

Table 7. Factor loadings of GET. Those smaller than 0.30 are not shown (according to Stevens' recommendations, 2002). Cronbach alpha and associated interval of confidence are also reported in the same table.

	Factors		
Items	Genetic Teleology	Genetic Essentialism (within-group homogeneity of genes)	Genetic Essentialism (fixity of genes)
GT1 Opposable Thumbs	0.83		
GT2 Big Brain	0.78		
GT3 Bipedalism	0.82		
GT4 Communication	0.45		
GT5 Sociability	0.66		
GT6 Cellulose Digestion	0.68		
GT7 Antibodies Production	0.67		
GT8 Ultraviolet Protection	0.78		
GT9 High Heat Protection	0.75		
GT10 Radiation Protection	0.72		
GE1 Neanderthals		0.57	
GE2 Chinese People		0.55	
GE3 Chimpanzees		0.67	
GE4 Eskimoos		0.53	
GE5 Baboons		0.78	
GE6 Daltonism			0.57
GE7 Breast Cancer			0.64
GE8 Dwarfism			0.58
GE9 Diabetes			0.64
GE10 Alzheimer's disease			0.71
Cronbach Alpha	0.87	0.64	0.65
Cronbach Alpha Interval of Confidence (at 5% level)	[0.85, 0.89]	[0.59, 0.68]	[0.60, 0.71]

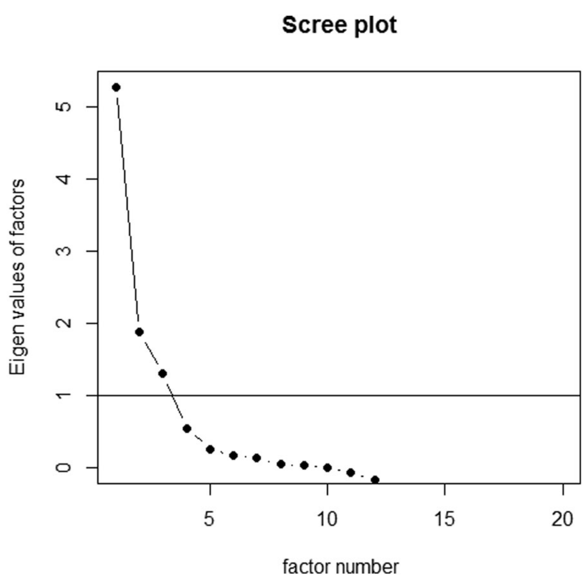

Figure 6. Scree plot without genetic determinism items. The horizontal line emphasises the separation between eigenvalues greater or smaller than 1.

suggest that the internal consistency of items (genetic teleology items, (0.87); genetic essentialism items about fixity, (0.65); genetic essentialism items about homogeneity, (0.64), see Table 7) is in the range between good and acceptable for a test for conceptual understanding (Jorion et al., 2015).

The quality of fit indexes of the factor analysis with 20 items is questionable (RMSR = 0.07, RMSEA = 0.149 and TLI = 0.59, see Jorion et al., 2015), but improves and reaches acceptable values when removing the items about Chinese people (GE2) and Eskimos (GE4) (RMSR = 0.06, RMSEA = 0.117 and TLI = 0.76, see Jorion et al., 2015). We thus worked with the improved questionnaire version, but as the two items also load on the genetic essentialism dimension about homogeneity and are of educational interest, we have retained them in the questionnaire (Appendix 2). Additionally, and similarly to Analysis 1, the correlation between the genetic teleology and genetic essentialism-fixity constructs, and the correlation between the genetic teleology and genetic essentialism-homogeneity, were also found to be very weak (respectively 0.00 (±0.05) and −0.06 (±0.05)).

In sum, the final version of the GET questionnaire has 20 items: 10 about genetic teleology, and 10 about genetic essentialism (see Appendix 2; additionally, and for information purposes only, the list of removed genetic determinism items can be found in Appendix 3). This selection of items took in account different psychometric indicators, such as the difficulty, the discrimination and the internal consistency of items, as well as the factor analysis, and educational interests.

Discussion

Overall, it seems that our conceptualisation of genetic teleology and genetic essentialism is mostly consistent with empirical data, according to the results of the factor analysis above. The two dimensions of essentialism were clearly separated, whereas the two dimensions of teleology were not. This is not problematic, however, because the only difference between the two dimensions of teleology is the temporal scale (past/future). The genetic teleology items loaded together on the same factor, without any distinction between past or future events. This difference between past (associated with an explanation) and future events (associated with a prediction) may have not been considered by students as significant. In contrast, the two essentialism dimensions (homogeneity/fixity) loaded on two different factors in the factor analysis and so we can infer that students perceived them as conceptually distinct.[9]

The psychometric characteristics of the proposed questionnaire also appear to be satisfactory. We should note that there are only a few tests of this kind where an attempt on structural analysis (factor analysis) has been made. The review by Liu (Liu, 2012) reported only 4 out of 15 instruments where such an attempt was made. Moreover, for most of the few examples where factor structures are actually known, no reliabilities for the subscales found were reported (evolution/CINS: Anderson, Fisher, & Norman, 2002; electrical circuits/DIRECT: Engelhardt & Beichner, 2004; motion, force/FMCE: Ramlo, 2008; matter, energy/MIET: Ding & Beichner, 2009; motion, force/FCI: Scott, Schumayer, & Gray, 2012). For the few exceptions, where subscale reliabilities are known (actually the only ones, to the best of our knowledge), they turned out to be quite low (statistics/SCI: 0.3–0.5, Allen, 2006; dynamics/DCI: 0.2–06, Jorion et al., 2015), the best-known examples being one in physics (static forces/CATS: 0.5–0.7), and, in the domain of the present study, PUGGS (0.67–0.85; Carver et al., 2017). Three of the subscales of the GET have reliabilities comparable to those of PUGGS (genetic teleology and the two genetic essentialism scales: 0.65–0.82). In sum, the psychometric properties of GET appear as being acceptable to good within the area of conceptual assessment, and as a sufficient basis for an instrument useful for teaching and research purposes.

In addition, no correlations were found between the genetic teleology and genetic essentialism constructs, in agreement with our hypothesis and with previous research (Coley and Tanner, 2015; Stern et al., 2018). A reason for this may be that these constructs are not overlapping, even though causality seems to underlie both of them. Finally, and even though it was not the main aim of the present study, the development of the GET questionnaire seems to challenge Dar-Nimrod and Heine's conceptualisation of genetic essentialism. Based on the four-dimensional structure of genetic essentialism suggested by these authors, we initially investigated each of these four sub-dimensions, which were: genes as fixed entities, gene homogeneity within species; genes as internal, single causes; inference for the presence of gene from a given biological or mental characteristic. While the two first sub-dimensions were clearly found to be separated from students' genetic determinism conceptions, this was not the case for the last two sub-dimensions, which showed some overlap with genetic determinism. Our study suggests that the concepts of genes as internal, single causes and the inference of the presence of genes from a given characteristic are related to genetic determinism. Therefore, it is unclear whether these concepts should be considered as relating to genetic determinism or to genetic essentialism in future studies.

We must also note a few limitations related to the questionnaire and our results. First, the questionnaire was administered to students with various backgrounds and knowledge, and so we had to develop a version that uses simplified terms and that can be understood by the youngest students (aged 15), while at the same time being of interest to the oldest students (aged 19). Second, another limitation stems from the sample itself. While we overall had students from a range of school types, most students came from type FG. We have no indications of strong differences of the GET properties, but a more detailed investigation of possible differences among the school types in this study (and beyond) is of interest. Third, the three-level scale used for measuring genetic teleology and genetic essentialism could be improved with a more fine-grained distinction between the strength of the intuition underlying the conceptions (which would include, for example, four or five levels instead of three in our final version). Last but not least, it is interesting to test convergent validity with other questionnaires, such as the PUGGS (Gericke et al., 2017) or Coley and Tanner instrument (Coley & Tanner, 2015). Cross-validation has been partially established between the PUGGS and the GET with respect to genetic determinism, as similar results about the separation of biological/mental traits (or biological/social) were found using two independent and different approaches. However, convergent validity analysis about genetic teleology and genetic essentialism could not be established with other tests. The PUGGS instrument does not consider such constructs, and the Coley and Tanner (2015) instrument was not about genetic teleology (except one of the items) and genetic essentialism but teleology and essentialism in general. Even though so far there are no overlapping instruments that could be used to compare with our results about the genetic teleology and genetic essentialism constructs, it would be interesting to make such a comparison in future research.

Being aware of this, the validated scales of the GET questionnaire could now be further used with a larger sample for investigating whether genetic teleology and genetic essentialism conceptions vary between ages, schools, different kinds of trainings. This might be of great interest for understanding better which conceptions are, or not, addressed at schools. Additionally, this questionnaire could be used by teachers as a valuable tool to detect the presence of intuitive thinking among their students. In the next section, we provide some guidelines for using the GET questionnaire and for interpreting its results.

Potential use of the GET questionnaire

The GET questionnaire is intended for secondary students, and its goal is to detect the strength of their conceptions related to genetic teleology and genetic essentialism. This can provide teachers with evidence for their students' preconceptions before and after the lesson. For example, if a student has a high score for genetic essentialism or genetic teleology, that student is more likely to have misconceptions based on such intuitions (e.g. that genes have a purpose, or that genes are our essences) than another one with a lower score. This would be a valuable information for teachers, in order for them to design teaching interventions that would explicitly address these preconceptions.

In practice, each of the three categories of items (teleology, essentialism-homogeneity, essentialism-fixity) are considered separately. For each of them, the average range of a student's scores lies between 0 and 2, based on the scores of the three choices of each item: weak intuition (0), moderate intuition (1) and strong intuition (2). From this, we can arbitrarily define a 4-categories framework for interpretation purposes, as described in Table 8.

This interpretation framework can be used to compare different students' scores. For example, let us consider a student with a mean score of 1.12 for answering essentialism items about fixity; as this student's score lies between 1.00 and 1.50, he or she would, therefore, be considered as exhibiting 'moderate to strong' genetic essentialism about fixity, according to Table 8. Another student with a mean score of 0.23 for the same set of essentialism items about fixity would be considered as exhibiting a 'weak' genetic essentialism about fixity. This framework can also be used to compare different populations, such as different age/grade groups. As an illustration, Table 9 shows the mean scores for each dimension and age group in our sample. For example, students aged 15–16 years old had an average score of 1.25 for genetic teleology items; we, therefore, considered the 15–16 age group to have a 'moderate to strong' genetic teleology intuition. At the same time, older students aged 18–19 had an average score of 0.32, which was considered as exhibiting a 'weak' genetic teleology intuition.

Table 8. Interpretation framework of the GET questionnaire.

Mean score within:	Interpretation of the strength of the teleology or essentialism intuition
[0,00, 0.50]	weak
[0.50, 1,00]	weak to moderate
[1,00 1.50]	moderate to strong
[1.50, 2.00]	strong

Table 9. Entries are overall mean scores. Standard deviations are also reported in brackets.

	Intuition Scores 0: weak, 1: moderate, 2: strong		
Age	GT items (10 items)	GE items (homogeneity, 5 items)	GE items (fixity, 5 items)
15–16 ($n = 67$)	1.25 (0.37)	0.64 (0.57)	0.48 (0.46)
16–17 ($n = 107$)	0.84 (0.55)	0.54 (0.47)	0.55 (0.49)
17–18 ($n = 180$)	0.84 (0.55)	0.54 (0.47)	0.55 (0.50)
18–19 ($n = 42$)	0.32 (0.31)	0.60 (0.56)	0.74 (0.49)
all ($n = 396$)	0.82 (0.53)	0.56 (0.52)	0.58 (0.48)

We are aware that these cut-off values are arbitrarily chosen from the way we designed items and their associated hypothesised degree of intuition. There is no acceptable threshold *per se*, but we recommend that users of the test to investigate further mean scores greater than 1, which indicate the presence of a moderate or even stronger intuition (which is, for example, the case for the teleology intuition of students aged 15–16 years old in our sample, whose mean score is 1.25). Additionally, mean scores between 0.50 and 1.00 are of interest, as they suggest that an underlying intuition seems to be present in a mild form.

Teachers or researchers could develop more personalised interpretations and revise such cut-offs, e.g. based on the intrinsic characteristics of their own sample, such as median or mean scores. More generally, further research using this questionnaire among populations of various characteristics would be of high interest to empirically set reference values for interpreting the results. These may be of great help for secondary education teachers or education researchers to better evaluate the prevalence of intuitive thinking in schools, and the need to address it in biology education.

Notes

1. We have conducted a quasi-replication study, based on the Coley & Tanner (2015), and we arrived at quite similar results (Stern, Kampourakis, Huneault, Silveira, & Müller, 2018).
2. Note that the code "genes mentioned" or "DNA mentioned" refers to students who only mentioned the word "genes" or "DNA", and did not exhibit any more complex conceptions when asked about genetics.
3. These responses including absolute words such as "only", were inspired by students' answers given in the interviews.
4. Even though answers (1) and (3) may seem extreme cases, in P_4 they overall accounted for 37% of students' answers to the genetic determinism items - and 46% in the youngest ages. Therefore, these high percentages of answers using absolute language seem crucial for understanding better students' conceptions.
5. Both design-based teleology and need-based teleology are types of design teleology, with the difference that the intentional agent related to design is external in the first case and internal in the second case (see Kampourakis, 2019).
6. It can be argued that students might simply guess the correct answer (3) because of the use of scientific language. In our opinion, it was necessary to include an answer using scientific language, as this is exactly what some students did in the interviews. Indeed, during those, some students used elaborated scientific words, describing *natural selection* ("there could be a fly that will be born green by chance, which will survive, because birds cannot eat it. And then, it will reproduce more often than the others flies, which will be eaten. Next, this will begin a new lineage of green flies"). At the same time, some others used an intuitive

vocabulary, such as expressing a *need* ("flies will realise they have to change a bit. Actually, they will need other things that they did not need before."). Therefore, due to this variety of scientific accurateness observed in students' answers, we developed item choices with a variety of scientific precision, as a guarantee of the representativeness of the scope of students' conceptions.

7. Note that Chinese people were included because we wanted to investigate participants' perception of the genetic differences between European and non-European populations.
8. Note that comparing these results with the ones in Analysis 1, it appears that the questionnaire is structurally stable, in the sense that genetic teleology items load together in both analyses, as well as all genetic essentialism items about homogeneity, and all genetic essentialism items about fixity.
9. Note that fit indexes improved when Chinese (GE2) and Eskimo (GE4) items were removed. This change can be explained because these two items are about living human groups and as such may contrast with other genetic essentialism items about homogeneity which are about extinct human species (Neanderthals (GE1) or non-human species (Chimpanzees (GE3); Baboons (GE5)).

Acknowledgements

We are indebted to all teachers and students who participated in this study. We also thank the following scholars for their useful feedback: Rebecca Carver (Department of Communication, Norwegian Institute of Public Health, Oslo, Norway), Jérémy Castéra (Aix-Marseille Universite, EA4671 ADEF, France), Ravit Duncan (Graduate School of Education and the School of Environmental and Biological Sciences, Rutgers University, United States of America), Charbel El-Hani (History, Philosophy, and Biology Teaching Lab, Institute of Biology, Federal University of Bahia, Salvador, Brazil), Nicklas Gericke (Department of Environmental and Life Sciences, Karlstad University, Sweden), Marcus Hammann (Center for Didactics of Biology, University of Münster, Germany), Rémy Kopp (Section of Biology, University of Geneva, Switzerland), François Lombard (Section of Biology, University of Geneva, Switzerland), Michael Reiss (Institute of Education, University College London, United Kingdom), Bruno J. Strasser (Section of Biology, University of Geneva, Switzerland).

Disclosure statement

No conflict of interest was reported by the authors.

Funding

This work was supported by the Swiss National Science Foundation [project number 100019_162679].

ORCID

Florian Stern http://orcid.org/0000-0001-5713-5545

References

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education*, *33*(9), 1289–1312. doi:10.1080/09500693.2010.512369

Allen, K. (2006). *The Statistics Concept Inventory: Development and analysis of a Cognitive Assessment Instrument in Statistics* (SSRN Scholarly Paper No. ID 2130143). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2130143

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, *39*(10), 952–978. doi.org/10.1002/tea.10053

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Statistical Psychology*, *4*(1), 1–2. doi:10.1111/j.2044-8317.1951.tb00299.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57* (1), 289–300.

Bowling, B. V., Acra, E. E., Wang, L., Myers, M. F., Dean, G. E., Markle, G. C., … Huether, C. A. (2008). Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Genetics*, *178*(1), 15–22. doi:10.1534/genetics.107.079533

Campbell, C. E., & Nehm, R. H. (2013). A critical analysis of assessment quality in genomics and bioinformatics education research. *CBE Life Sciences Education*, *12*(3), 530–541. doi:10.1187/cbe.12-06-0073

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.

Carver, R. B., Castéra, J., Gericke, N., Evangelista, N. A. M., & El-Hani, C. N. (2017). Young adults' belief in genetic determinism, and knowledge and attitudes towards modern genetics and genomics: The PUGGS Questionnaire. *PLoS ONE*, *12*(1), doi:10.1371/journal.pone.0169808

Casler, K., & Kelemen, D. (2008). Developmental continuity in Teleo-functional explanation: Reasoning about nature among Romanian Romani adults. *Journal of Cognition and Development*, *9*(3), 340–362. doi:10.1080/15248370802248556

Castéra, J., & Clément, P. (2014). Teachers' conceptions about the genetic determinism of human behaviour: A survey in 23 countries. *Science & Education*, *23*(2), 417–443. doi:10.1007/s11191-012-9494-0

Chi, M. T. H. (2013). Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (2nd ed., pp. 49–70). New York, NY: Routledge.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). London: Routledge.

Coley, J. D., & Tanner, K. D. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. *CBE Life Sciences Education*, *14*(1), doi:10.1187/cbe.14-06-0094

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin*, *137*(5), 800–818. doi:10.1037/a0021860

Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics – Physics Education Research*, *5*(2), 020103, doi:10.1103/PhysRevSTPER.5.020103

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching*, *44*(7), 938–959. doi:10.1002/tea.20186

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, *72*(1), 98–115. doi:10.1119/1.1614813

Ergazaki, M. (2018). Biology and young children. In K. Kampourakis & M. J. Reiss (Eds.), *Teaching biology in schools: Global research, issues and trends* (pp. 22–34). New York, NY: Routledge.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage Publications Ltd.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378–382. doi:10.1037/h0031619

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford: Oxford University Press.
Gelman, S. A. (2005). Essentialism in everyday thought. *Psychological Science Agenda*, *19*(5), 1–6.
Gelman, S. A., & Rhodes, M. (2012). "Two-thousand years of stasis": How psychological essentialism impedes evolutionary understanding. In K. S. Rosengren, S. Brem, E. M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. Oxford: Oxford University Press.
Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, *38*(3), 213–244.
Gericke, N., Carver, R., Castéra, J., Evangelista, N. A. M., Marre, C. C., & El-Hani, C. N. (2017). Exploring relationships among belief in genetic determinism, genetics knowledge, and social factors. *Science & Education*, *26*(10), 1223–1259. doi:10.1007/s11191-017-9950-y
Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, *23*(1), 17–27. doi:10.1111/j.1745-3992.2004.tb00149.x
Heine, S. J., Dar-Nimrod, I., Cheung, B. Y., & Proulx, T. (2017). Essentially biased: Why people are fatalistic about genes. In J. M. Olson (Ed.), *Advances in experimental social psychology: Vol. 55. Advances in experimental social psychology* (pp. 137–192). San Diego, CA: Elsevier Academic Press.
Inagaki, K., & Hatano, G. (2002). *Young children's naive thinking about the biological world*. New York, NY: Psychology Press.
Inagaki, K., & Hatano, G. (2006). Young children's conception of the biological world. *Current Directions in Psychological Science*, *15*(4), 177–181. doi:10.1111/j.1467-8721.2006.00431.x
Jamieson, A., & Radick, G. (2017). Genetic determinism in the genetics curriculum. *Science & Education*, *26*(10), 1261–1290. doi:10.1007/s11191-017-9900-8
Jiménez-Aleixandre, M. P. (2014). Determinism and underdetermination in genetics: Implications for students' engagement in argumentation and epistemic practices. *Science & Education*, *23*(2), 465–484. doi:10.1007/s11191-012-9561-6
Jorion, N., Gane, B. D., James, K., Schroeder, L., DiBello, L. V., & Pellegrino, J. W. (2015). An analytic framework for evaluating the validity of concept inventory claims. *Journal of Engineering Education*, *104*(4), 454–496. doi:10.1002/jee.20104
Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, *34*(1), 111–117. doi:10.1177/001316447403400115
Lennox, J. G., & Kampourakis, K. (2013). Biological teleology: the need for history. In K. Kampourakis (Ed.). *The philosophy of biology: A companion for educators* (pp. 421–454). Dordrecht: Springer.
Kampourakis, K. (2014). *Understanding evolution*. Cambridge: Cambridge University Press.
Kampourakis, K. (2017). *Making Sense of Genes*. Cambridge: Cambridge University Press.
Kampourakis, K. (2019). Students' "teleological misconceptions" in evolution education: Why the underlying design stance, not teleology per se, is the problem. *Evolution Education and Outreach*. doi:10.1186/s12052-019-0116-z
Kampourakis, K., & Reiss, M. (2018). Introduction: Biology didactics. In K. Kampourakis, M. Reiss (Eds.). *Teaching biology in schools: Global research, issues and trends* (pp. 1–8). New York, NY: Routledge.
Kampourakis, K., Reydon, T. A. C., Patrinos, G. P., & Strasser, B. J. (2014). Genetics and society—educating scientifically literate citizens: Introduction to the thematic issue. *Science & Education*, *23*(2), 251–258.
Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
Keil, F. C., Greif, M. L., & Kerner, R. S. (2007). *A world apart: How concepts of the constructed world are different in representation and in development. Creations of the mind: Theories of artifacts and their representation*. Oxford: Oxford University Press, 231–245.
Keil, F. C., & Newman, G. E. (2008). Two tales of conceptual change: What changes and what remains the same. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 83–101). New York, NY: Routledge.

Keil, F. C., & Wilson, R. A. (eds.). (2000). *Explanation and cognition*. Cambridge, MA: MIT Press.
Kelemen, D. (2012). Teleological minds: How natural intuitions about agency and purpose influence learning about evolution. In K. S. Rosengren, S. Brem, E. M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 66–92). Oxford: Oxford University Press.
Kelemen, D., & Carey, S. (2007). The essence of artifacts: Developing the design stance. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 212–230). New York, NY: Oxford University Press.
Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology. General*, *142*(4), 1074–1083. doi:10.1037/a0030399
Lewis, J., Leach, J., & Wood-Robinson, C. (2000). All in the genes? — Young people's understanding of the nature of genes. *Journal of Biological Education*, *34*(2), 74–79. doi:10.1080/00219266.2000.9655689
Lewthwaite, B., & Fisher, D. (2005). The development and validation of a primary science curriculum delivery evaluation questionnaire. *International Journal of Science Education*, *27*(5), 593–606. doi:10.1080/0950069042000230758
Liu, X. (2012). Developing measurement instruments for science education research. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 651–665). Dordrecht: Springer.
Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*(2), 167–204. doi:10.1016/j.cognition.2004.12.009
Maerten-Rivera, J. L., Huggins-Manley, A. C., Adamson, K., Lee, O., & Llosa, L. (2015). Development and validation of a measure of elementary teachers' science content knowledge in two multiyear teacher professional development intervention projects. *Journal of Research in Science Teaching*, *52*(3), 371–396. doi:10.1002/tea.21198
Medin, D. L., & Atran, S. (1999). *Folkbiology*. Cambridge, MA: MIT Press.
Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement*, *21*(3), 215–237. doi:10.1111/j.1745-3984.1984.tb01030.x
Office fédéral de la statistique. (2015). Classification internationale type d'éducation (CITE 2011). https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees/publications.assetdetail.214600.html
Ramlo, S. (2008). Validity and reliability of the force and motion conceptual evaluation. *American Journal of Physics*, *76*(9), 882–886. doi.org/10.1119/1.2952440
Revelle, W. (2019). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois. R package version 1.9.12.
Rosengren, K. S., Brem, S. K., Evans, E. M., & Sinatra, G. M. (2012). *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. Oxford: Oxford University Press.
Rottman, J., Zhu, L., Wang, W., Schillaci, R. S., Clark, K. J., & Kelemen, D. (2017). Cultural influences on the teleological stance: Evidence from China. *Religion, Brain & Behavior*, *7*(1), 17–26. doi:10.1080/2153599X.2015.1118402
Scott, T. F., Schumayer, D., & Gray, A. R. (2012). Exploratory factor analysis of a Force Concept Inventory data set. *Physical Review Special Topics – Physics Education Research*, *8*(2), 020105. doi:10.1103/PhysRevSTPER.8.020105
Shtulman, A. (2017). *Scienceblind: Why our intuitive theories about the world are so often wrong*. New York, NY: Basic Books.
Shtulman, A., & Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. *Cognitive Science*, *32*(6), 1049–1062. doi:10.1080/03640210801897864
Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The genetics concept assessment: A new concept inventory for gauging student understanding of genetics. *CBE Life Sciences Education*, *7*(4), 422–430. doi:10.1187/cbe.08-08-0045
Stern F., & Kampourakis, K. (2017). Teaching for genetics literacy in the post-genomic era. *Studies in Science Education*, *53*(2), 193–225.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences*, 4th ed. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
Stern, F., Kampourakis, K., Huneault, C., Silveira, P., & Müller, A. (2018). Undergraduate biology students' teleological and essentialist misconceptions. *Education Sciences*, *8*(3), 135.
Summers, R., & Abd-El-Khalick, F. (2018). Development and validation of an instrument to assess student attitudes toward science across grades 5 through 10. *Journal of Research in Science Teaching*, *55*(2), 172–205. doi:10.1002/tea.21416
Tsui, C.-Y., & Treagust, D. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, *32*(8), 1073–1098. doi:10.1080/09500690902951429
Unesco. (2012). *International Standard Classification of Education (ISCED) 2011*.
Velayutham, S., Aldridge, J., & Fraser, B. (2011). Development and validation of an instrument to measure students' motivation and self-regulation in science learning. *International Journal of Science Education*, *33*(15), 2159–2179. doi:10.1080/09500693.2010.541529
Vosniadou, S. (2012). Reframing the classical approach to conceptual change: Preconceptions, misconceptions and synthetic models. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 119–130). Dordrecht: Springer.
Wilkins, J. (2013). Essentialism in biology. In K. Kampourakis (Ed.), *The philosophy of biology: A companion for educators* (pp. 395–420). Dordrecht: Springer.

Appendices

Appendix 1

Genetic determinism		Genetic teleology		Genetic essentialism	
Items	Critical value, *p*-value	Items	Critical value, *p*-value	Items	Critical value, *p*-value
aggressiveness (GD1)	$\chi^2(1) = 1.06$, $p = 0.74$	opposable thumbs (GT1)	$\chi^2(1) = 2.91$, $p = 0.51$	Neanderthals (GE1)	$\chi^2(1) = 0.32$, $p = 0.95$
intelligence (GD2)	$\chi^2(1) = 4.88$, $p = 0.49$	big brain (GT2)	$\chi^2(1) = 4.04$, $p = 0.49$	Chinese people (GE2)	$\chi^2(1) = 0.58$, $p = 0.90$
high painting skills (GD3)	$\chi^2(1) = 0.42$, $p = 0.94$	bipedalism (GT3)	$\chi^2(1) = 7.65$, $p = 0.22$	chimpanzees (GE3)	$\chi^2(1) = 9.96$, $p = 0.21$
high sport skills (GD4)	$\chi^2(1) = 0.00$, $p = 1$	communication (GT4)	$\chi^2(1) = 3.84$, $p = 0.49$	Eskimos (GE4)	$\chi^2(1) = 2.85$, $p = 0.51$
obesity (GD5)	$\chi^2(1) = 1.84$, $p = 0.57$	sociability (GT5)	$\chi^2(1) = 2.59$, $p = 0.51$	baboons (GE5)	$\chi^2(1) = 2.04$, $p = 0.57$
myopia (GD6)	$\chi^2(1) = 2.41$, $p = 0.53$	cellulose digestion (GT6)	$\chi^2(1) = 4.58$, $p = 0.49$	daltonism (GE6)	$\chi^2(1) = 3.83$, $p = 0.49$
height (GD7)	$\chi^2(1) = 0.10$, $p = 1$	antibody protection (GT7)	$\chi^2(1) = 2.89$, $p = 0.51$	breast cancer (GE7)	$\chi^2(1) = 1.43$, $p = 0.67$
infarction (GD8)	$\chi^2(1) = 3.15$, $p = 0.51$	protection from ultraviolet radiation (GT8)	$\chi^2(1) = 2.01$, $p = 0.57$	dwarfism (GE8)	$\chi^2(1) = 1.90$, $p = 0.57$
French language skills (GD9)	$\chi^2(1) = 1.29$, $p = 0.51$	protection from high heat (GT9)	$\chi^2(1) = 4.28$, $p = 0.49$	diabetes (GE9)	$\chi^2(1) = 7.76$, $p = 0.22$
eye colour (GD10)	$\chi^2(1) = 0.00$, $p = 1$	protection from radiation (GT10)	$\chi^2(1) = 1.13$, $p = 0.74$	Alzheimer's disease (GE10)	$\chi^2(1) = 3.41$, $p = 0.51$

Tests of the difference of answers between two random orders of administration of questionnaire Q_4 in P_4. Critical values and corresponding *p*-values of multiple comparisons are reported.

Appendix 2

List of items of the final version of the GET.

GT1 *Humans have opposable thumbs. Thus, genes associated with opposable thumbs:*
have been designed for several roles such as holding objects.
have appeared for satisfying several needs such as holding objects.
have appeared by chance and were selected for several effects such as holding objects.
I have no opinion.

GT2 *Our brain is much bigger than the one of cave men. Thus, genes associated with a big brain:*
have been designed for several roles such as solving complex problems.
have appeared for satisfying several needs such as solving complex problems.
have appeared by chance and were selected for several effects such as solving complex problems.
I have no opinion.

GT3 *Humans are `bipeds`, which means they walk on legs. Thus, genes associated with bipedalism:*
have been designed for several roles such as having hands available for other tasks.
have appeared for satisfying several needs such as having hands available for other tasks.
have appeared by chance and were selected for several effects such as having hands available for other tasks.
I have no opinion.

GT4 *Humans have the ability to communicate with one another. Thus, genes associated with the ability to communicate:*
have been designed for several roles such as communicating between humans.
have appeared for satisfying several needs such as communicating between humans.
have appeared by chance and were selected for several effects such as communicating between humans.
I have no opinion.

GT5 *Humans developed the ability to help one another in bigger and bigger groups. Thus, genes associated with the ability to socialize:*
have been designed for several roles such as helping one another.
have appeared for satisfying several needs such as helping one another.
have appeared by chance and were selected for several effects such as helping one another.
I have no opinion.

GT6 *Humans cannot digest cellulose, which is the most abundant sugar on earth. Thus, in the future genes associated with cellulose digestion:*
will be designed for digesting cellulose.
will appear for satisfying the need for digesting cellulose.
if they appear by chance, may be selected for digesting cellulose.
I have no opinion.

GT7 *A new microbe appears in a human population. No vaccine exists against it. Thus, in the future genes associated with antibody production:*
will be designed for protecting from this microbe.
will appear for satisfying the humans need to protect from this microbe.
if they appear by chance, they may be selected for protecting from this microbe.
I have no opinion.

GT8 *A white-skin human population moves into an extremely sunny place where ultraviolet radiation is high.Thus, in the future genes associated with ultraviolet protection:*
will be designed for protecting from ultraviolet radiation.
will appear for satisfying the need to protect from ultraviolet radiation.
if they appear by chance, may be selected for protecting from the ultraviolet radiation.
I have no opinion.

GT9 *The average temperature of Earth increases every year. But humans cannot survive too high temperature.Thus in the future genes associated with high temperatures:*
will be designed for adapting to high temperatures.
will appear for satisfying the need to adapt to high temperatures.
if they appear by chance, may be selected for adapting to high temperatures.
I have no opinion.

GT10 *Electronic devices such as antenna emit electromagnetic radiations. Thus in the future genes associated with protection from radiation:*
will be designed for protecting from electromagnetic radiation.
will appear for satisfying the need to protect from electromagnetic radiation.
if they appear by chance, may be selected for protecting from electromagnetic radiation.
I have no opinion.

GE1 *If we analyze the genes of Neanderthals (a prehistoric human group), we will identify:*
genes specific to them.
many genes different from ours.
few genes different from ours.
I have no opinion.

GE2* *If we analyze the genes of Chinese people, we will identify:*
genes specific to them.
many genes different from ours.
few genes different from ours.
I have no opinion.

GE3 *If we analyze the genes of chimpanzees, we will identify:*
genes specific to them.
many genes different from ours.
few genes different from ours.
I have no opinion.

GE4* *If we analyze the genes of Eskimos, we will identify:*
genes specific to them.
many genes different from ours.
few genes different from ours.
I have no opinion.

GE5 *If we analyze the genes of baboons, we will identify:*
genes specific to them.
many genes different from ours.
few genes different from ours.
I have no opinion.

GE6 *A person with daltonism can hardly see particular colors. We assume that in a given family everyone sees colors well. Therefore there are only genes associated with a good color-vision. A descendant in this family:*
will have a good color-vision, because the genes associated with a good color-vision always remain fixed.
will have a good color-vision, because the genes associated with a good color-vision are fixed, even though others may change.
may have daltonism, if the genes associated with a good color-vision change into genes associated with daltonism.
I have no opinion.

GE7 *A woman with breast cancer has a tumor in this organ. We assume that in a given family no woman has a breast cancer. Therefore there are only genes associated with a 'normal breast'. A descendant in this family:*
will have unaffected breasts, because the genes associated with a 'normal breast' always remain fixed.
will have unaffected breasts, because the genes associated with a 'normal breast' are fixed, even though others may change.
may have breast cancer, if the genes associated with a 'normal breast' change into genes associated with breast cancer.
I have no opinion.

GE8 *A dwarf is a person who has an extremely small size. We assume that in a given family everyone has a normal size. Therefore there are only genes associated with a normal size. A descendant in this family:*
will have normal size, because the genes associated with a normal size always remain fixed.
will have a normal size, because the genes associated with a normal size are fixed, even though others may change.
may be a dwarf, if the genes associated with a normal size change into genes associated with dwarfism.
I have no opinion.

GE9 *A diabetic person has high blood sugar levels. We assume that in a given family everyone has a normal blood sugar levels. Therefore there are only genes associated with a normal blood sugar levels. A descendant in this family:*
will have normal blood sugar levels, because the genes associated with normal blood sugar levels always remain fixed.

will have normal blood sugar levels, because the genes associated with normal blood sugar levels are fixed, even though others may change.
may be diabetic, if the genes associated with normal blood sugar levels change into genes associated with diabetes.
I have no opinion.

GE10 *A person with Alzheimer's disease has memory difficulties because of a dysfunctional brain. We assume that in a given family everyone has a good memory. Therefore there are only genes associated with a well-functioning brain. A descendant in this family:*
will have a good memory, because the genes associated with a well-functioning brain always remain fixed.
will have a good memory, because the genes associated with a well-functioning brain are fixed, even though other may change.
may have Alzheimer's disease, if the genes associated with a well-functioning brain change into genes associated with Alzheimer.
I have no opinion.

* these items were included because of their educational interest.

Appendix 3

Additionally, for information purposes only, we present the list of genetic determinism items which were eventually not included in the GET questionnaire.

GD1	*Whether a person is aggressive depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he was maltreated during her/his childhood). external factors only (e.g. if she/he was maltreated during her/his childhood). I have no opinion.
GD2	*Whether a person has outstanding intellectual skills depends on:* her/his genes only. her/his genes and external factors (e.g. the education provided by her/his family). external factors only (e.g. the education provided by her/his family). I have no opinion.
GD3	*Whether a person is a talented painter depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he had favourable conditions for practicing painting). external factors only (e.g. if she/he had favourable conditions for practicing painting). I have no opinion.
GD4	*Whether a person is a high-level athlete depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he had favourable conditions for training in sports). external factors only (e.g. if she/he had favourable conditions for training in sports). I have no opinion.
GD5	*Whether a person is obese depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he has the habit to eat a lot). external factors only (e.g. if she/he has the habit to eat a lot). I have no opinion.
GD6	*Whether a person is myopic depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he has been looking a lot at screens). external factors only (e.g. if she/he has been looking a lot at screens). I have no opinion.

GD7	*Whether a person is tall depends on:* her/his genes only. her/his genes and external factors (e.g. the quality of food received during her/his childhood). external factors only (e.g. the quality of food received during her/his childhood). I have no opinion.
GD8	*Whether a person has an infarction depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he had smoked a lot). external factors only (e.g. if she/he had smoked a lot). I have no opinion.
GD9	*Whether a foreign person speaks French depends on:* her/his genes only. her/his genes and external factors (e.g. if she/he had French classes). external factors only (e.g. if she/he had French classes). I have no opinion.
GD10	*Whether a person has eyes of a certain colour depends on:* her/his genes only. her/his genes and external factors (e.g. her/his country of origin). external factors only (e.g. her/his country of origin). I have no opinion.