

Constructing and Validating a Scale of Inquisitive Curiosity

Kathryn Iurino, University of Oregon
Brian Robinson, Michigan State University
Markus Christen, University of Zurich
Paul Stey, United States Environmental Protection Agency
Mark Alfano, Delft University of Technology

Abstract

We advance the understanding of the philosophy and psychology of curiosity by operationalizing and constructing an empirical measure of Nietzsche's conception of *inquisitive curiosity*, expressed by the German term *Wissbegier*, (“thirst for knowledge” or “need/impetus to know”) and *Neugier* (“curiosity” or “inquisitiveness”). First, we show that existing empirical measures of curiosity do not tap the construct of inquisitive curiosity, though they may tap related constructs such as idle curiosity and phenomenological curiosity. Next, we map the concept of inquisitive curiosity and connect it to related concepts, such as open-mindedness and intellectual humility. The bulk of the paper reports four studies: an Anglophone exploratory factor analysis, an Anglophone confirmatory factor analysis, an informant study, and a Germanophone exploratory and confirmatory factor analysis.

Introduction

Curiosity has been embraced as a virtue in the modern world, where information technologies have facilitated the exchange of and search for information on an unprecedented scale and where the quest for innovation has become a mantra. As Phillips (2015) and others have observed, curiosity is now routinely affirmed in educational institutions and creative industries; and it is promoted by advertising campaigns and corporate communications as a glossy affirmative feature that is expected from learners and leaders alike. Interestingly, the philosophical literature on curiosity is rather sparse: half of the 244 contributions in the Philosopher's Index that pop up when searching for "curiosity" have been published only in the last 10 years (search performed in October 2015); and Inan, who recently (2012) wrote a monograph on the philosophy of curiosity, observed "that philosophers had simply not taken curiosity as a topic that was worthy of philosophical investigation" (Inan 2012, xii-xiii). In psychology, a more sustained interest in curiosity can be observed: Although more than one third of the 636 contributions that are listed in PsychInfo under the major heading "curiosity" have been published in the last 10 years, we find continuous interest in this topic going back to the 1950s starting with the seminal work of Berlyne (1954) and, besides the current years, another peak in interest in the 1970s. However, one has to take into account that within psychology, terms like 'interest', 'intrinsic motivation', 'play', and 'exploration' have often been used interchangeably with 'curiosity', yet these terms do not clearly refer to the same construct (Kreitler & Kreitler, 1994). In the course of this research, several scales have been proposed to measure (facets of) curiosity (for an overview see Reio et al. 2006).

In our contribution, we aim to advance the understanding of the philosophy and psychology of curiosity by following in particular Nietzsche's understanding of curiosity as *inquisitive* curiosity, expressed by the German term *Wissbegier*, ("thirst for knowledge" or "need/impetus to know") and *Neugier* ("curiosity" or "inquisitiveness"). This work emerged out of a larger project that aims to explore the many facets of intellectual humility. In a semantic analysis, we found that one dimension of this construct refers to this Nietzschean understanding of curiosity expressed by terms like 'inquiry', 'exploration' and 'scrutiny' (Figure 1; for details about how this analysis was performed, see Christen et al. 2014). This result motivated us to analyze the curiosity dimension in more detail, because – despite the existence of several curiosity scales (Litman & Spielberger 2003; Kashdan et al. 2004; Kashdan et al. 2009) – existing measures do not adequately tap the construct of inquisitive curiosity. In this paper, we make a contribution to the burgeoning interdisciplinary field of character psychology by building and validating a psychological scale of inquisitive curiosity.

In doing so, we follow the best practices in scale construction which have been laid out by Leonard Simms (2008) as a three-phase process:

- (1) the substantive validity phase,
- (2) the structural validity phase, and
- (3) the external validity phase.

In phase one, experimenters conduct a thorough literature review of the construct to be measured and related constructs. Based on this review, they determine whether a new scale is truly needed,

collect definitions of the construct, develop an initial item pool, and conduct expert-review studies. In phase two, the experimenters develop an item selection strategy, collect responses from appropriate samples, evaluate the items psychometrically, create provisional scales, and modify and add items to address problems. Finally, in phase three, the experimenters conduct studies to evaluate convergent, divergent, discriminant, and criterion-related validity, finalize their scale, and report it.

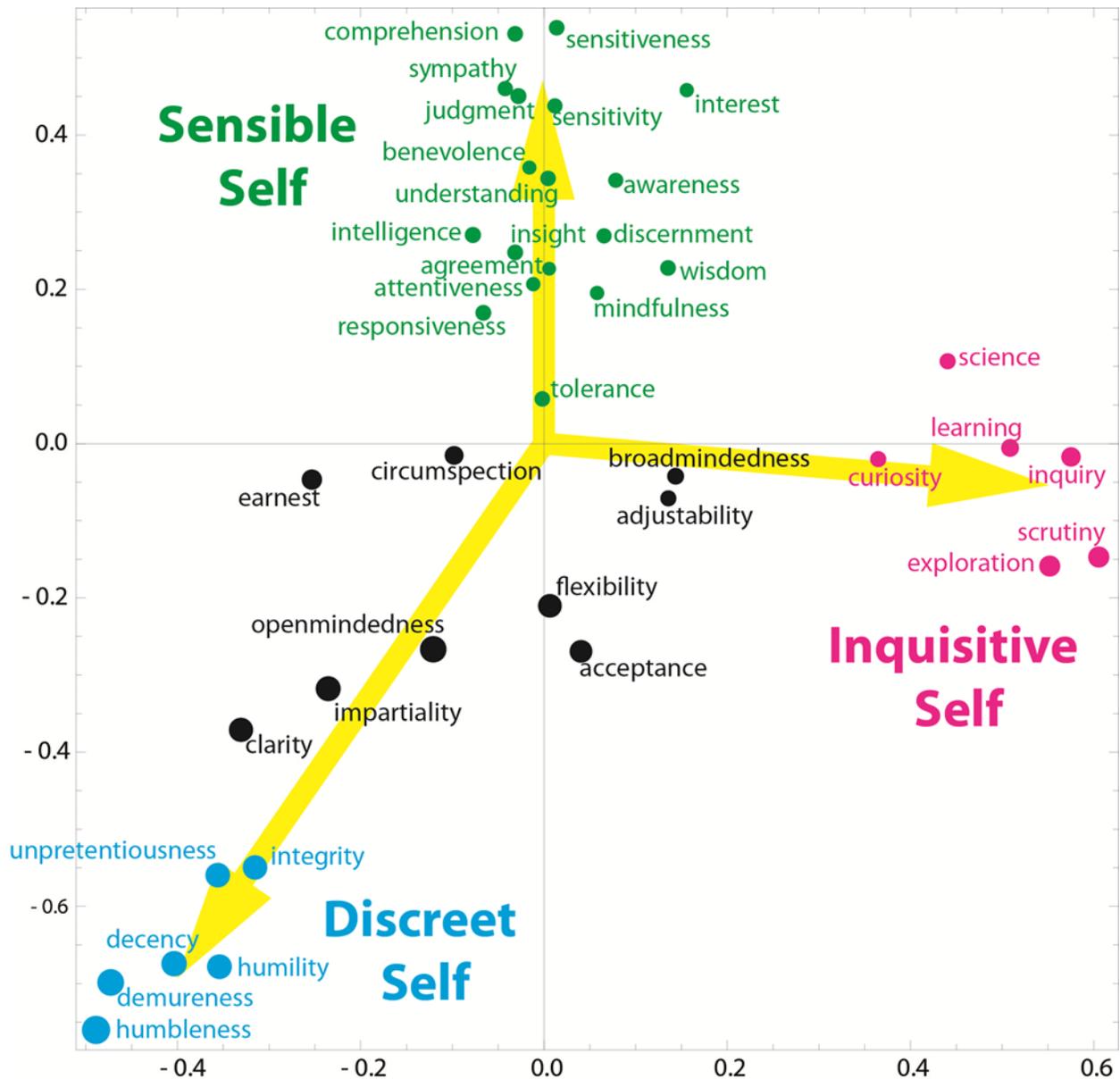


Figure 1: Semantic dimensions of intellectual humility, including the “inquisitive self”, which served as motivation for developing a scale of inquisitive curiosity (from Christen et al. 2014).

Here is the plan for this paper: first we review some existing scales of curiosity and explain our concerns with them – especially with their face validity. In section 2, we draw on

recent work on Nietzschean curiosity to explore the concept of inquisitive curiosity that our new scale is meant to measure. The bulk of the paper is section 3, where we review four empirical studies of our new scale. We conclude with a brief general discussion.

1 Extant curiosity scales

Several curiosity scales have already appeared in the empirical literature. In this section, we review a few of the most recent and most representative. The first is Kashdan et al. (2004) – a 7-item self-report scale with two factors: *exploration* and *absorption*.

Exploration items include:

- I would describe myself as someone who actively seeks as much information as I can in a new situation.
- I frequently find myself looking for new opportunities to grow as a person (e.g., information, people, resources).
- I am *not* the type of person who probes deeply into new situations or things. [reverse-scored]
- Everywhere I go, I am out looking for new things or experiences.

Absorption items include:

- When I am participating in an activity, I tend to get so involved that I lose track of time.
- When I am actively interested in something, it takes a great deal to interrupt me.
- My friends would describe me as someone who is “extremely intense” when in the middle of doing something.

We have no doubt that these factors measure interesting and important psychological constructs. However, we question the face validity of the absorption factors as a measure of *inquisitive curiosity*. Face validity, as the concept is used in psychology, refers to the degree to which the items in a scale or other measure “on their face” seem to match the target construct, so our point here is that we think these items may tap an interesting construct, but not inquisitive curiosity. In particular, the items in the absorption factor seem to measure the disposition to enter the state that Mihaly Csikszentmihalyi (2008) calls “flow.” It’s possible for someone to have a flow experience even when they are not learning anything new or encountering novel experiences, things, or properties. Furthermore, one item in the exploration factor (about growing as a person) does not seem to involve curiosity at all, but rather some kind of perfectionist eudaimonism. And the remaining three exploration items seem to measure only a specific aspect of curiosity: openness to new experiences. They primarily characterize how someone responds when encountering novel situations, such as while traveling. One can easily imagine a person who scores high on exploration because she enjoys sight-seeing but who has no patience for the intensive study of physics, poetry, or cooking. Thus, we think that the Kashdan et al. (2004) scale does not quite hit the mark.

The next extant measure of curiosity, Kashdan et al. (2009), is meant to be an improvement on Kashdan (2004). This updated scale also has two factors: *stretching* and *embracing*.

Stretching items include:

- I actively seek as much information as I can in new situations.
- I am at my best when doing something that is complex or challenging.
- I view challenging situations as an opportunity to grow and learn.
- I am always looking for experiences that challenge how I think about myself and the world.
- I frequently seek out opportunities to challenge myself and grow as a person.

Embracing items include:

- I am the type of person who really enjoys the uncertainty of everyday life.
- Everywhere I go, I am out looking for new things or experiences.
- I like to do things that are a little frightening.
- I prefer jobs that are excitingly unpredictable.
- I am the kind of person who embraces unfamiliar people, events, and places.

Once again, we have no doubt that this scale measures an interesting and important psychological construct. However, we continue to have concerns about face validity. The embracing factor, in particular, does not seem primarily to measure inquisitive curiosity. It includes multiple items that have to do with comfort with uncertainty or even risk. Perhaps it is easier to acquire new knowledge or understanding when one is comfortable in this way, but it is also possible to experience such comfort without any motivation to learn. This is exemplified by what the poet Keats (1899, p. 277) called “negative capability,” a disposition to be “in uncertainties, mysteries, doubts, without any irritable reaching after fact and reason,” which is opposed to “being incapable of remaining content with half-knowledge.”

The items in the stretching factor fare somewhat better when assessed for face validity, but they too seem to miss the motivated inquisitiveness of curiosity, emphasizing instead enjoyment of growing through facing challenges. Not all growth of this sort involves the acquisition of knowledge or understanding, however – or the destruction of ignorance or error, for that matter.

Finally, consider the epistemic curiosity scale developed by Litman & Spielberger (2003). In this paper, the authors first distinguish perceptual curiosity, which has to do with a thirst for new phenomenal experiences, from epistemic curiosity, which has to do with a thirst for knowledge. They further divide epistemic curiosity into diversive and specific facets. Diversive epistemic curiosity drives its bearer to acquire broad but relatively shallow knowledge, while specific curiosity drives its bearer to acquire deep knowledge of a narrow domain.

Items measuring the diversive facet include:

- I enjoy discussing abstract concepts.
- I find it fascinating to learn new information.
- I enjoy learning about subjects which are unfamiliar to me.
- I enjoy exploring new ideas
- When I learn something new, I like to find out more about it.

Items in the specific facet include:

- When I see an incomplete puzzle, I like to try and imagine the final solution.
- I am interested in discovering how things work.
- When I encounter a new kind of arithmetic problem, I enjoy imagining solutions.
- When I encounter complicated machinery, I like to ask how it works.
- When I hear a riddle, I am interested in trying to solve it.

These items, especially those associated with the specific facet, fare better on a test of face validity for inquisitive curiosity. However, Litman & Spielberg (2003) report significant gender differences for this factor, with men scoring significantly higher than women. This is presumably because several items in it primarily focus on stereotyped STEM topics (machinery, math). We worry that the items measuring this facet are too narrowly focused and may even trigger stereotype threat. Additionally, from a psychometric point of view, it is unfortunate that all of the items in this scale are positively scored. Acquiescence is a well-known problem in personality psychology: basically, people have a tendency to agree with whatever you put in front of them. This is one reason why scales with reverse-scored items can be useful.

2 Inquisitive curiosity

In order to construct a scale of inquisitive curiosity that isn't plagued by concerns, it is helpful to first consider previous philosophical discussions of curiosity. While traditionally not a popular topic of philosophical analysis, two major philosophers have discussed it: Augustine and Nietzsche.¹ In this section, we explore the concept of inquisitive curiosity with an eye to its operationalization based on what we find to be absent in Augustine's conception of curiosity and celebrated in Nietzsche's.

Though Augustine has nothing but condemnation for curiosity, he discusses two kinds. He primarily understands curiosity as a desire "to acquire new experiences," typically sinful ones, such as sex and seeing a "mangled corpse" (*Confessions* X.35). This sort of phenomenal curiosity is an "appetite for knowledge" of what it is like to have certain experiences; it mainly aligns with what Berlyne (1954) has called "perceptual curiosity" as the kind of curiosity that "leads to increased perception of stimuli" (p. 180). Some of the items on the curiosity scale of Kashdan et al. (2009) do seem to be tapping this phenomenal curiosity. However, it appears in both factors of stretching and embracing, and several items appear unrelated to curiosity in this way. Therefore, the scale provided Kashdan et al. (2009) appears to be only related to, but not directly measuring, phenomenal curiosity.

Augustine then briefly notes that phenomenal curiosity can easily lead to another variety. "Because of this morbid curiosity, monstrous sights are exhibited in the show places. Because of it, men proceed to search out the secrets of nature, things beyond our end, to know which profits us nothing, and of which men desire nothing but the knowledge" (*Confessions* X.35). We can call this sort of curiosity *idle curiosity*. Whether it stems from phenomenal curiosity is immaterial, at least for present purposes. They are distinct kinds of curiosity in that have

¹ A third is Hume, on whom see Gelfert (2013).

different motivations and aim at different kinds of knowledge. Idle curiosity can have a variety of forms. One might be curious about the latest celebrity gossip, not because one cares who in Hollywood is sleeping with whom, but just to know; it achieves nothing independent to know this. A desire to conduct scientific research into a natural phenomenon can also be a form of idle curiosity, if the scientist is only driven by a desire to know for no other reason than to find the answer. She doesn't care what the answer is. Whatever the answer is, it will neither upset any of her other beliefs, nor will it be of any use to her or anyone else. It is merely a puzzle to be solved.

Nietzsche viewed curiosity in a different light. He is arguably the most important philosopher of curiosity in the Western tradition. We draw on historical and conceptual work on Nietzsche's conception of curiosity by Mark Alfano (2013) and Bernard Reginster (2013; 2015) to explore his notion of inquisitive curiosity. Nietzsche references curiosity (i.e., *Neugier*, *Wissbegier*, and their cognates) seventy-five times in his published and authorized works (Alfano 2013). Based on Alfano's and Reginster's complementary reconstructions of Nietzsche's conception of curiosity, several points can be made that distinguish it from Augustine's two kinds of curiosity. First, Alfano stresses that Nietzschean curiosity is *not* merely the desire to acquire more true beliefs, contrary to idle curiosity. Both Nietzschean curiosity and idle curiosity have the same end goal of acquiring *true beliefs*. Nietzsche, however, recognized that there is value in curiosity even if one fails to attain knowledge. There is something worthwhile in being a "seeker after knowledge," (HH I.630) as Reginster emphasizes. If we seek knowledge but fail, we are left in uncertainty. Such a state, however, is not lamentable; curiosity entails a delight in the "danger of uncertainty" (GS Preface). Idle curiosity is entirely disappointed when knowledge is not obtained. Nietzschean curiosity, on the other hand, still finds value in confronting uncertainty through inquiry, which leads us to our second point.

The reason that Nietzschean curiosity "involves an attraction for uncertainty and ignorance" (Reginster 2013:456-7) is that it is fundamentally opposed to "faith" understood as "the disposition to stop investigating" (Alfano 2013:12). Even if curiosity only produces uncertainty, that uncertainty still undermines the dogmatic acceptance of a previously held belief. Beliefs that we dogmatically hold are beliefs that we very much want to be true. It matters to us whether these beliefs are true. Nietzschean curiosity requires inquiry into these matters. This is different from idle curiosity, for which it does not matter to the inquirer one way or the other.

Finally, in contrast with other forms of curiosity, such as those measured by Kashdan et al. (2004; 2009), Nietzschean curiosity is highly inquisitive, protracted, and even unpleasant. As he puts it, "let [knowledge] be something else for others; for example, a bed to rest on, or the way to such a bed, or a diversion, or a form of leisure – for me it is a world of dangers and victories in which heroic feelings, too, find places to dance and play. 'Life as a means to knowledge' – with this principle in one's heart one can live not only boldly but even gaily" (GS 324). As Alfano (2013) elaborates, Nietzschean curiosity is "characterized by an insatiable desire to solve novel, difficult problems and puzzles, and to discover and invent them when none are ready to hand." It matters both how hard the inquiry is and how important it is. Again the contrast with idle curiosity is clear. Nothing is riding on such inquiry and it may or may not be easy to undertake.

We find Nietzsche a useful precedent for inquisitive curiosity. His conception of the construct involves the forceful asking of questions that are difficult to answer (either for purely epistemic reasons or because the answers run counter to our squeamishness or faith in humanity). In the next section, we describe four studies that we conducted to help build a new scale of inquisitive curiosity. We take inquisitive curiosity at least to be consistent with Nietzschean curiosity, even if not identical to it. Furthermore, inquisitive curiosity as we have described it fits well with contemporary philosophical work on the concept, such as Whitcomb (2010), who argues that curiosity is a desire, the contents of which are questions and the satisfier of which is knowledge. It also accords with popular cultural insistence on the importance of questioning well, in addition to answering well – a distinction hilariously drawn by Douglas Adams in *The Hitchhiker's Guide to the Galaxy* (1979) when the supercomputer Deep Thought cryptically announces, after 7.5 million years of computation, that the answer to the ultimate question (which was never made explicit) is 42.

3 Four empirical studies of inquisitive curiosity

We conducted four studies to establish the validity and reliability of a new measure of inquisitive curiosity:

- 1) an exploratory factor analysis,
- 2) a confirmatory factor analysis,
- 3) an “informant” analysis, and
- 4) a test of the validity and reliability of a translation of the scale in German.

In the following subsections, we explain these four studies, which were part of a larger project that aimed to explore factors of intellectual humility. As the semantic analysis suggested (see Introduction), we expected inquisitive curiosity to be one of the dimensions of the construct.

3.1 Curiosity from the inside: An exploratory study

In the first study, we conducted what is known as an exploratory factor analysis. For those who are unfamiliar with psychometrics, what this essentially means is that we wrote several dozen statements that express 1) the attitude of someone who embodies the target construct, 2) the attitude of someone who embodies dispositions *contrary* to the target construct, and 3) the attitudes of someone who embodies related but distinct dispositions. We then invited several hundred people to read each of these statements and agree or disagree with them. Next, we examined participants' responses for clustering patterns. A “factor” in the resulting analysis is a set of statements that participants tend to respond to in the same way (either mostly agreeing or mostly disagreeing).

3.1.1 Participants and Procedure

Participants were college students at a large Midwestern public university. We sent an invitation email to a random sample of 5000 students, of whom 442 responded. The average age of participants was 20.9; about 90% of participants were between the ages 18–22. 30.5% of the

sample was male. 85.5% of the sample was White/Caucasian, 4.5% was African-American/Black, 2.9% was Asian, and 1% was Hispanic. The majority of participants reported taking fewer than 2 philosophy classes (83.7%) and fewer than 2 psychology classes (58.8%). Only .5% of the participants were psychology majors and there were no philosophy majors in the sample. The two most popular majors of participants were Business Administration (9.6%) and Religious Studies (8.2%). Average GPA of the sample was 3.29, the average SAT score was 1346.67, and the average ACT score was 25.44.

3.1.2 Measure

The original item pool consisted of 52 items, of which about a dozen could be associated with curiosity (see appendix A for a complete list of items). Items were informed by a variety of sources. Some items were adapted from the items in the HEXACO personality inventory, including most subscales of the honesty/humility dimension, as well as the dependence subscale of emotionality, the diligence subscale of conscientiousness, and the inquisitiveness and unconventionality subscales of openness. Other items were informed by a thorough consideration of the defining aspects of Nietzschean curiosity, as well as some of its more penumbral elements. These items were developed after conducting a thorough literature review of philosophical and psychological research on intellectual virtues.

A balance of reverse and forward-keyed items were included (27 of the 52 items indicated the *absence* of the intended construct). Participants were asked to rate on a 7-point scale their level of agreement with the items, with 1 anchored as “strongly disagree” and 7 anchored as “strongly agree.”

3.1.3 Results

To determine the factor structure, we used *Mplus* Version 7 (Muthén & Muthén, 2012) to conduct an exploratory factor analysis using promax rotation. A parallel analysis suggested 7 factors should be extracted by strict cut-off criteria, but the 8-factor solution was more interpretable. Items that had a loading greater than .30 in magnitude on their primary factor were retained (see Appendix A for a list of items that met this criterion for each factor). Though ideally each item would also have low cross-loadings ($< .10$) on all factors other than the primary factor, only four items satisfied this criterion. However, compared to solutions with fewer factors, the average proportion of cross-loadings $> .10$ for each item was lowest in the 8-factor solution.

Upon inspection of the content of each factor, we found that only the first four factors were relevant to *intellectual humility*, the construct we initially set out to measure. A factor with content consistent with *inquisitive curiosity* also clearly emerged as a separate factor from these four humility factors. Four items loaded above .3 on the *inquisitive curiosity* factor, and the content of these items aligned with a central aspect of *inquisitive curiosity*: motivated inquiry that *perseveres* in order to achieve the highest possible level of mastery. The items loading on this factor include “I try to learn as much as I can” and “I often push myself very hard when trying to master a new idea.” This kind of curiosity is quite different from possessing a disposition to be comfortable with new ideas, as the Kashdan scales define curiosity. The present definition of *curiosity* assumes no such comfort, and in fact, the motivation to achieve mastery implies quite the opposite, as truly mastering an idea characteristically brings with it uncertainty and difficulty. We think the *inquisitive curiosity* measure better reflects the virtue of curiosity in

the Nietzschean sense than previous measures, as in order to endorse these items, a person has to approach learning with a great degree of tenacity.

Though *inquisitive curiosity* clearly emerged as a distinct factor from the other four intellectual humility factors, it was related to the intellectual humility factors. Below are the four intellectual humility factors listed in order of the strength of their scale scores' correlation with the scale score for *inquisitive curiosity*, with brief descriptions of the content of each.

Open-mindedness ($r=.44$ with *curiosity*): The tendency to acknowledge the limitations of one's knowledge irrespective of one's status, and relatedly, to be open to learning from those of lower status.

Engagement ($r=.38$ with *curiosity*): The motivation to understand ideas one disagrees with or ideas that come from diverse sources.

Corrigibility ($r=.2$ with *curiosity*): A lack of defensiveness in reaction to being told one has made an intellectual mistake.

Modesty ($r=.11$ with *curiosity*): The tendency not to be highly preoccupied with one's intellectual reputation.

Three other factors also emerged which were conceptually peripheral to both intellectual humility and curiosity (these were *uniqueness*, which was composed of three items about feeling special when one has knowledge; *intellectual machiavellianism*, which was composed of four items about manipulating others in order to get more information, and *intellectual kleptomania*, which was composed of three items about stealing others' ideas).

3.2 Curiosity from the inside: A confirmatory study

In the second study, we conducted what is known as a confirmatory factor analysis (CFA). The basic idea behind a confirmatory factor analysis is to double-check the structure or pattern that emerged in an exploratory analysis (something, we note, which was not done by Litman & Spielberger 2003). As philosophers of science have long been aware, it is remarkably easy to over-fit a model to a data-set. Confirmatory factor analysis checks the degree to which the best-fitting model from the exploratory analysis fits the data generated by a fresh set of participants.

3.2.1 Participants and Procedure

Participants ($N = 465$; $M_{\text{age}} = 33.29$, 240 female) were recruited and compensated using Amazon.com's Mechanical Turk platform. Ages ranged from 18–82, and median education completed was an Associates degree; 41.9% had a Bachelors or higher level of education. 53.2% had taken at least one philosophy class; 69.4% had taken at least one psychology class. The median number of philosophy classes taken was 1, and the median number of psychology classes was also 1. Seventy-five percent of participants were White/Caucasian, 8.6% were African-American or Black, 7.7% were Asian, 5.8% were Hispanic, 0.4% were Pacific Islander, and 1.9% were Multiracial.

3.2.2 Measure

The measure was identical to the one used in study 1. Participants responded to 52 items presented in random order. Participants rated their agreement on a 7-point Likert-type scale.

3.2.3 Results: Confirmatory factor analysis

To determine the extent to which the 8-factor structure replicated in a second sample, we conducted a CFA using Maximum Likelihood estimation with robust standard errors.² In the CFA we included all items that loaded above .3 on their factors in study 1. The CFA had marginal fit: $\chi^2(1052) = 2400.97$, CFI = .826, RMSEA = .053, SRMR = .075. Inspection of the modification indices in *Mplus* showed that a few items had a tendency to cross-load on multiple factors, and were clear candidates for removal. This led us to drop items 14 and 16 from the modesty factor, and 39 and 41 from the open-mindedness factor. Dropping these items resulted in a model with better fit, $\chi^2(874) = 1742.62$, CFI = .874, RMSEA = .046, SRMR = .061.

Factor loadings on the curiosity factor were fairly consistent in this model, ranging from a magnitude of .6 to .72, suggesting that each item related to the curiosity factor to a similar degree (no item was notably more characteristic of the curiosity factor than another). The curiosity factor showed a similar pattern of relationships to the intellectual humility factors as in the first study: the factor correlation between engagement and curiosity was the highest ($r = .733$), followed by the correlation between open-mindedness and curiosity ($r = .691$), followed by the correlation between curiosity and corrigibility ($r = .517$). Unlike in the first study, modesty and curiosity showed no relation at all ($r = .004$). Of the additional three factors that emerged in the eight-factor solution that were more peripheral to intellectual humility (machiavellianism, kleptomania, and uniqueness), only machiavellianism and kleptomania showed a relation with curiosity ($r = -.214$, $r = -.498$, respectively).

The fact that curiosity shows strong relationships with three other components of our intellectual humility scale suggest that it may be worth considering how open-mindedness, engagement, and corrigibility relate to inquisitive curiosity.

3.2.4 Results: Item response theory analysis

In order to further investigate the psychometrics of this new measure, we supplemented the confirmatory factor analysis with analyses from the perspective of item response theory (IRT). IRT is a model-based framework used for investigating item and test properties; it assumes a latent trait or ability that is a function of both the participant's responses, and the properties of the items (Emberton & Reise, 2000). Thus, IRT allows us to simultaneously estimate an individual's trait level and also the relevant item parameters. This is distinctly different from the classical approach.

The goal of this further examination was two-fold. First, we aimed to identify the characteristics of individual items; and second, we wanted to estimate the overall reliability of

² Exploratory data analysis revealed the assumption of multivariate normality had been violated (Mardia's coefficient = 69.82).

the measure in a manner distinct from the classical testing theory approach. In order to investigate the individual properties of items, we used a graded response model (Samejima, 1969, 1996) implemented in the `ltm` package (Rizopoulos, 2006) in the R statistical language (R Core Team, 2013).

Of particular interest were item slopes (a) and threshold parameters (b_j). Item slopes describe an item's ability to differentiate between participants having levels of the latent trait above or below the item's location (Baker, 2001). Item slopes are frequently referred to as discrimination parameters. Threshold parameters can be considered cut points on the latent trait's continuum where a participant with that level of the latent trait is equally likely to select the response category j rather than category $j + 1$. In the testing literature, threshold parameters are referred to as difficulty parameters. This is because, historically, the first area in which this sort of psychometric analysis was used was intelligence testing. Harder questions on an intelligence test have a higher difficulty parameter. Similarly, items that are more difficult to endorse tend to have higher threshold parameters.

Table 1 below illustrates the parameter estimates from the fitted response model of the curiosity items. Several features are important to note. First, the threshold parameter estimates are distributed more towards the negative end of the latent continuum. This relates to the amount of measurement precision these items provide on the negative end of the latent continuum; we revisit this below. Second, the standard error estimates are all rather small even at the extremes of the response scale. This suggests participants made use of the entirety of the response categories, which is encouraging. Finally, we note that the slope parameters are quite similar across the various items. This is also encouraging in that it supports the use of simple un-weighted sum scoring to compute the scale (Cheng, Yuan, & Lui, 2012).

Table 1. *IH-Curiosity item parameter estimates.*

Item	b_1	b_2	b_3	b_4	b_5	b_6	a
19	-3.18 (0.33)	-2.69 (0.45)	-2.33 (0.43)	-1.71 (0.39)	-0.93 (0.34)	0.64 (0.29)	2.33 (0.26)
20	-3.76 (0.52)	-2.51 (0.56)	-1.83 (0.53)	-1.32 (0.51)	-0.32 (0.49)	1.2 (0.64)	2.23 (0.23)
21	-3.91 (0.48)	-2.21 (0.46)	-1.2 (0.41)	-0.66 (0.39)	0.11 (0.38)	1.52 (0.66)	1.56 (0.16)
22	-4.05 (0.61)	-2.41 (0.61)	-1.71 (0.58)	-1.25 (0.57)	-0.55 (0.55)	0.68 (0.54)	1.92 (0.2)

Note: b indicates a threshold parameter, a indicates slope, SE estimates are in parantheses.

We next examined the item information curves for each of the curiosity items (see Figure 2 below). This gives a visual display of the level of measurement precision (known as *information*) each item provides, across all levels of the latent trait (θ). Similarly, the test information curve (see Figure 3 below) sums the individual item information curves to give an indication of the measurement precision of the entire scale across all levels of θ . As can be seen in Figure 3, the curiosity scale provides the greatest amount of information toward the negative end of the latent continuum, and as Figure 2 suggests, this pattern seems to hold for each of the individual items as well. Item 21 appears to provide the least amount of information overall, but provides slightly more information than the other items on the upper end of the latent continuum. Overall, though, these results suggest the items in the scale behave similarly. The curiosity scale is most adept at discriminating between people with low to moderate levels of curiosity; it may not be as adept at discriminating between people with moderate to high levels of curiosity.

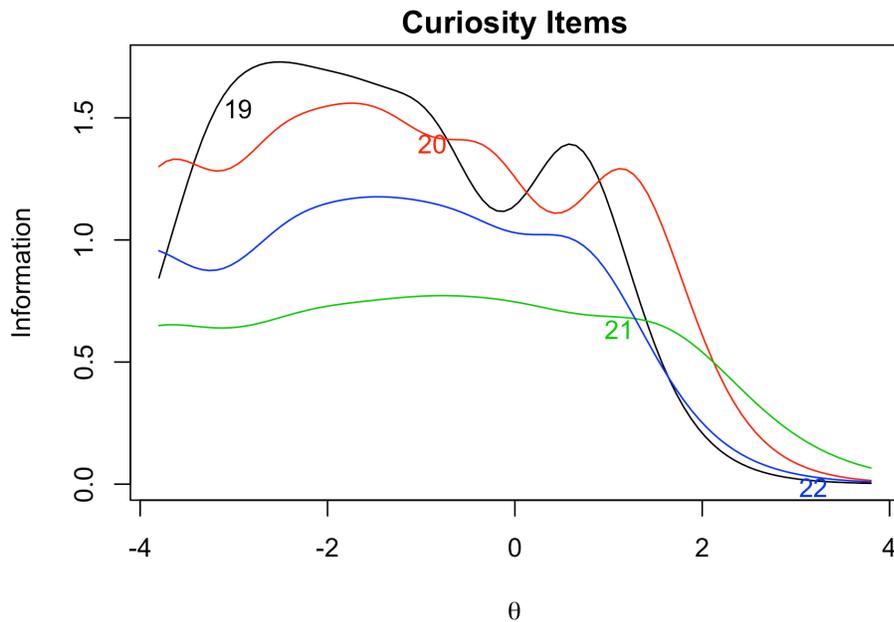


Figure 2. Item information curves for Curiosity items.

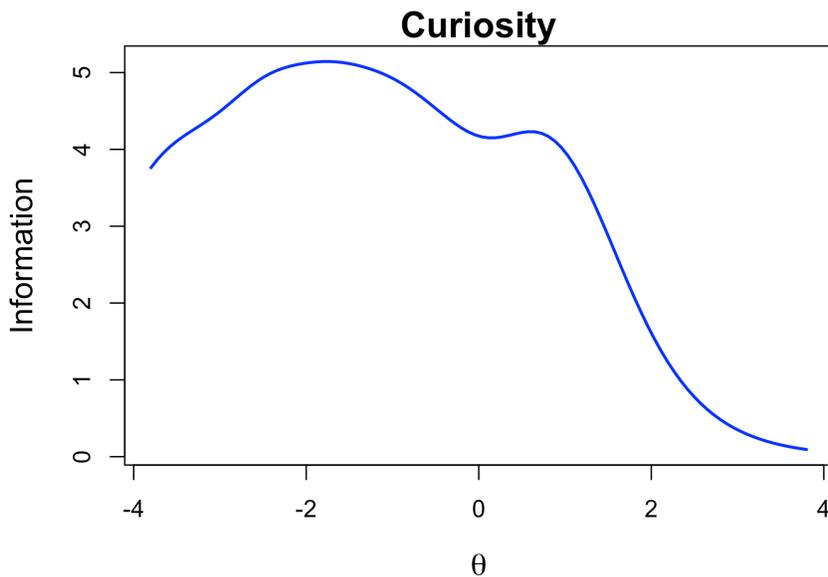


Figure 3. Test information function for the Curiosity scale.

3.3 Curiosity from the outside

Inquisitive curiosity isn't just a matter of what you think about yourself. Like any character trait, inquisitive curiosity is a disposition to think, feel, and act. Since the characteristic manifestation of this disposition is *protracted inquiry*, it would be surprising if someone who embodied inquisitive curiosity were not seen as such by those who know her well. Of course, it is possible to conduct one's inquiries in secret or in such a way that other people are simply

baffled by one's behavior. Nevertheless, even Nietzsche would say that the people who know you best are typically in a decent epistemic position to evaluate your character. Indeed, if Alfano (2015a, 2015b, 2016) is right, Nietzsche held that the character one publically projects and the character that is attributed to one may causally or even constitutively determine the character one actually has.

For these reasons, we conducted an “informant” study of inquisitive curiosity. Informant studies have become popular in personality psychology (John & Robins 1993; Kenny & West 1994; Klonsky et al. 2002; Connolly et al. 2007; Vazire & Carlson 2010) in part because they lend external verification to self-report measures of dispositions. I might think that I'm the life of the party, but if everyone I know sees me as a wallflower, I'm probably wrong. Of course, intersubjective agreement isn't infallible. I and my friends might all think that I'm uproariously funny despite my being a charmless bore. That said, informant reports shed some evidential light on people's dispositions.

In this study, we replicated the confirmatory analysis of study 2, then asked participants to nominate several “informants” who knew them well. We then invited the nominees to rate the original participants on the same statements, transformed into third-personal attributions rather than first-personal expressions of attitudes. For instance, “I try to learn as much as I can,” became, “(S)he tries to learn as much as (s)he can.” We then compared the pattern of responses by participants with the patterns of responses by their informants.

3.3.1 Participants and Procedure

Two groups of participants were recruited for this study. The first, self-report group of participants ($N = 1185$; $M_{\text{age}} = 33.6$, 56% female) were recruited and compensated using Amazon.com's Mechanical Turk platform. Ages ranged from 18-79, and median education completed was an associate's degree (2-year); 43.5% had a Bachelors or higher level of education. 56.4% had taken at least one philosophy class; 71.2% had taken at least one psychology class. The median number of philosophy classes taken was 1, and the median number of psychology classes was also 1. 71.2% percent of participants were White/Caucasian, 13% were African-American or Black, 5.4% were Asian, 5.1% were Hispanic, .4% were Pacific Islander, and 1.8% were Multiracial.

The second, informant group of ($N = 107$; $M_{\text{age}} = 36.12$ years, 58% female) were recruited by emailing up to five informants per self-report participant. Informants' given names and email addresses were provided by participants.³ We then emailed all potential informants inviting them to tell us about the participant and offering as compensation a \$5 online gift card to Target. Of the 1402 informants contacted, $N=107$ completed the survey, giving us informant ratings on 89 of our main participants (74 of our main participants had one informant, 14 had two informants, and one had five informants). Ages of the informants ranged from 16-65, 70.1%

³ We should note that some participants objected to being asked for such potentially-identifying information. Future informant-report studies on Mechanical Turk should take this privacy concern into account.

percent of informant participants were White/Caucasian, 6.5% were African-American or Black, 1.9% were Asian, 8.4% were Hispanic, .9% were Pacific Islander, and 3.7% were Multiracial.

3.3.2 Measure

Two distinct measures were used. For the self-report group, the measure was identical to the one used in studies 1 and 2 with six new items added as potential replacements for items that performed less well in the intellectual humility scales. Participants rated their agreement with these 58 items on a 7-point Likert-type scale.

For the informant group, the measure was identical to the one just described for the self-report group, with one obvious but important exception: instead of asking participants to agree or disagree with statements about themselves (“I am...”), we asked them to agree with the same statements about the person who referred them to us (“s/he is...”). We also included items meant to gauge how well the informants knew the participants:

How do you know XX?
How long have you known XX?
How well do you know XX?
How close are you to XX?

Responses to these questions might, we presumed, help us to subdivide our informants into the more and the less informed.

3.3.3 Results: Agreement between Informant and Self-Ratings

Our informants appeared to be fairly well-informed: on average, informants reported knowing the main participant 14.18 years, and 90.7% of informants reported that they knew the participant “quite well” or “about as well as anybody does.” The highest proportion of informants were friends of the main participant (44.9%) followed by family members (19.6%) and romantic partners (17.8%), and the lowest proportion were colleagues (10.3%).

To check for agreement between informant-ratings and self-ratings of inquisitive curiosity, we averaged the informants’ responses on each item for main participants who had more than one informant, and then checked how highly the self-report scale scores correlated with those of the scales composed by the aggregated informants’ responses. The Curiosity scale showed a modest level of agreement ($r=.359$) that was comparable to the other intellectual humility scales. The Modesty scale showed the highest level of agreement between self- and informant-reports ($r=.472$), which is perhaps not surprising, given that the Modesty items capture a preoccupation with how one is perceived by others, e.g. “I like to be the smartest person in the room.” The Open-mindedness scale had a similar level of self-informant agreement ($r=.338$) as Curiosity. Corrigibility and Engagement had the lowest levels of agreement ($r=.285$ and $r=.283$, respectively).

Since there have been no other informant studies of curiosity scales in the literature, we cannot evaluate whether this level of agreement is typical for a construct like curiosity. However,

according to three meta-analyses, the average agreement for personality traits is somewhere in the range .40-.60 (Connolly et al., 2007; Kenny, 1994; Klonsky et al., 2002; as described in Vazire & Carlson, 2010). Self-other agreement tends to be lower for more evaluative personality traits, such as intellect and conscientiousness, and higher for traits that are constituted by more observable behaviors, such as extraversion (John & Robins, 1993). For a construct like Curiosity, we might expect that only informants who know the main participant enough to comment on the participant's most important motivations and goals would be knowledgeable enough to generate a high degree of consensus.

Among the informant ratings, the correlations between the curiosity scale and other intellectual humility scales showed a similar pattern as in the second study; curiosity was most strongly related to open-mindedness ($r=.643$), followed by engagement ($r=.520$) and corrigibility ($r=.461$). Also similar to the previous studies, the correlation between modesty and curiosity was not significantly different from zero. The correlations between scales composed of the same people rating themselves were very similar to the correlations between informant-ratings. It is worth noting that similar relationships between curiosity, open-mindedness, and engagement have been observed across different samples and with the addition of this study, across different methods (in both self-ratings and informant-ratings). This suggests that the relationship between curiosity and the other intellectual humility scales is more than just a method artifact, and deserves further attention.

3.4 Curiosity and Neugier

Nietzsche wrote in German. Concerns about the indeterminacy of translation notwithstanding (Quine, 1960, chapter 2; Davidson 1984, p. 62), it stands to reason that a measure of inquisitive curiosity in English should translate well to a measure of inquisitive *Neugier* in German. Furthermore, we are interested in potential cultural differences regarding the understanding of inquisitive curiosity. The first successful psycholexical replication of the Big Five model outside of English was in German (Ostendorf 1990). This suggests that our methodology is especially likely to replicate in German, though we hope eventually to study other languages, including non-Germanic and even non-Indo-European languages. To establish this, we translated our entire set of statements into German, then conducted both an exploratory and a confirmatory factor analysis, thereby replicating both study 1 and study 2 in a different language.

3.4.1 Participants and Procedure

We used the participant recruitment system of the University of Zurich that allows approaching students and staff of all faculties. In total, 579 participants provided valid answers. The average age of participants was 34.5; 38.7% of the sample was male. The distribution across faculties was as follows: philosophy 4.5%, psychology 12.1%, other humanities discipline 24.4%, social sciences 13.0%, sciences 18.1%, law 9.5%, medicine 12.8%, other 5.7%. 39.0% of the sample had a master's degree, 31.1% had a PhD. We used a random sub-sample of 279 participants for exploratory and the remaining 300 participants for confirmatory factor analysis.

3.4.2 Measure

The goal of the German-language study was to precisely replicate the procedure used in the English study. This approach should allow better identification of cultural differences regarding the understanding of intellectual humility, such as differences in the composition of the factors, in particular inquisitive curiosity. Therefore, we refrained from simply translating the final items of the English scale. Instead, all 52 items of the original English scale were translated by one author (M.C.) and the translations were checked independently by two experts who were not involved in the study (one social psychologist and one English teacher in a technical university). When comparing the German with the English result, we checked for “configural invariance,” which obtains when the translation has the same factor structure as the original (though not necessarily the same means, variance, etc.).

3.4.3 Results

We used the same methodology for determining the factor structure as outlined in studies 1 and 2. In the exploratory factor analysis, parallel analysis suggests that the number of factors equals 10 and the number of components equals 7. Given this situation and to enable a comparison with the result of the English study, we used the 8-factor solution, as this was also more interpretable. As in the English study, we retained items with a factor load greater than .30 in magnitude on their primary factor. Then, items that cross-loaded on more than one factor with weight $>.30$ were deleted. The result is outlined and compared with the result of the English EFA in Table 2; the item numbers match with the translations. We find that three factors show a complete match (machivellianism, kleptomania, corrigibility) and two factors show an item-overlap of 50% and more (engagement and curiosity). Interestingly, the factor *Neugier* was broader in German compared to the English “Curiosity”. Three other factors are semantically rather distinct. Considerably more items have been excluded in the German study compared to the English study using the same exclusion criteria. We then performed the CFA analogous to the English study, leading to the final scale of inquisitive curiosity.

When comparing both scales (Table 2), we find that the German scale includes a facet not present in the English scale, namely some degree of openness and self-understanding, expressed by the items “I enjoy reading about the ideas of different cultures” and “I have a good understanding of what I know and what I don’t know.”

English Curiosity Scale Items	German Curiosity Scale Items
I try to learn as much as I can.	Ich versuche so viel wie möglich zu lernen.
I often push myself very hard when trying to master a new idea.	Ich strenge mich oft sehr an, wenn ich eine neue Idee verstehen will.
Often when I’m in the process of learning something, I end up quitting without having really mastered it [reverse scored].	Wenn ich daran bin, etwas zu lernen, ist es häufig so, dass ich aufgabe, bevor ich es wirklich verstanden habe [reverse scored].

<p>I learn only the minimum amount needed to get by [reverse scored].</p>	<p>Ich lerne nur das Minimum, um gerade noch durchzukommen [reverse scored].</p> <p>Ich lese gerne über Ideen anderer Kulturen.</p> <p>Ich kann gut einschätzen, was ich weiss und was ich nicht weiss.</p>
---	---

Table 2: Comparison of English and German items for curiosity and *Neugier* scales.

The latter item indicates a meta-cognitive awareness of one's own strengths and limitations – a feature which, in recent philosophical work, has been connected with open-mindedness (Adler 2004) and intellectual humility (Hazlett 2012; Whitcomb et al. forthcoming). This once gain suggests that inquisitive curiosity – at least as it is understood in German-speaking Switzerland – is deeply related to other intellectual virtues.

4 General discussion

In *The Gay Science* 2, Nietzsche remarks:

But to stand in the midst of this *rerum concordia discors* and the whole marvelous uncertainty and ambiguity of existence *without questioning*, without trembling with the craving and rapture of questioning, without at least hating the person who questions, perhaps even being faintly amused by him – that is what I feel to be *contemptible*, and this is the feeling I look for first in anyone. Some folly keeps persuading me that every person has this feeling, simply as human.

In this chapter, we have shown how to measure the extent to which every person does in fact have this feeling, this drive to ask questions in order to achieve knowledge.

References

- Adams, D. (1979). *The Hitchhiker's Guide to the Galaxy*. Pocket Books.
- Adler, J. (2004). Reconciling open-mindedness and belief. *Theory and Research in Education*, 2: 127-41.
- Alfano, M. (2016). How one becomes what one is: The case for a Nietzschean conception of character development. In I. Fileva (ed.), *Perspectives on Character*. Oxford University Press.
- Alfano, M. (2015a). An enchanting abundance of types: Nietzsche's modest unity of virtue thesis. *Journal of Value Inquiry*, 49: 417-35.
- Alfano, M. (2015b). How one becomes what one is called: On the relation between traits and trait-terms in Nietzsche. *Journal of Nietzsche Studies*, 46(1): 261-9.
- Alfano, M. (2013). The most agreeable of all vices: Nietzsche as virtue epistemologist. *British Journal for the History of Philosophy*, 21(4): 767-90.
- Augustine. (trans. 1960). *The Confessions of Saint Augustine*. J. K. Ryan (Trans.). New York: Doubleday.
- Baker, F. (2001). *The basics of item-response theory* (ERIC Clearinghouse on Assessment and Evaluation). College Park, MD: University of Maryland.
- Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, 45: 180-191.
- Cheng, Y., Yuan, K., & Lui, C. (2012). Comparison of reliability measures under factor analysis and item response theory. *Educational and Psychological Measurement*, 72, 52-67.
- Christen M, Alfano M, Robinson B: The Semantic Space of Intellectual Humility. Proceedings of the European Conference on Social Intelligence, ECIS 2014, Barcelona, November 3-5 2014.
- Connolly, J., Kavanagh, E., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, 15, 110-117.
- Csikszentmihalyi, M. (2008). *Flow: The Psychology of Optimal Experience*. New York: Harper.
- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Emberton, S. & Reise, S. (2000). *Item Response Theory for Psychologists*. Mahway, NJ: Lawrence Erlbaum.
- Gelfert, A. (2013). Hume on curiosity. *British Journal for the History of Philosophy*, 21(4): 711-32.
- Hazlett, A. (2012). Higher-order epistemic attitudes and intellectual humility. *Episteme*, 9(3): 205-23.
- Inan, I. (2012). *The Philosophy of Curiosity*. New York: Routledge.
- John, O. & Robins, R. (1993). Determinants of interjudge agreement on personality traits: Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61, 521-551.
- Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment*, 82: 291-305.
- Kashdan, T., Gallagher, M., Silvia, P., Winterstein, B., Breen, W., Terhar, D., & Steger, M. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research on Personality* 43(6): 987-98.

- Keats, J. (1899). *The Complete Poetical Works and Letters of John Keats, Cambridge Edition*. Houghton-Mifflin.
- Kenny, D. & West, T. (1994). Similarity and agreement in self- and other-perception: A meta-analysis. *Personality and Social Psychology Review*, 14, 196-213.
- Klonsky, D., Oltmanns, T., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and Practice*, 9, 300-311.
- Kreitler, S., & Kreitler, H. (1994). Motivational and cognitive determinants of exploration. In H. Keller, K. Schneider, & B. Henderson (Eds.), *Curiosity and Exploration* (pp. 259–284). Berlin: Springer–Verlag.
- Litman, J. & Spielberger, C. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment*, 80(1): 75-86.
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.
- Nietzsche, F. (1882 / 2001). *The Gay Science*. B. Williams (ed.); J. Nauckhoff & A. Del Caro (trans.). Cambridge University Press.
- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Factoren-Modells der Persönlichkeit*. Regensburg, Germany: S. Roderer Verlag.
- Phillips R (2015): Curiosity: Care, Virtue and Pleasure in Uncovering the New. *Theory, Culture & Society* 32(3) 149–161.
- Quine W. V. O. (1960) *Word and Object*. MIT Press.
- Reginster, B. (2015). Nietzsche, proficiency, and the (new) spirit of capitalism. *Journal of Value Inquiry*, 49:453-77.
- Reginster, B. (2013). Honesty and curiosity in Nietzsche's free spirits. *Journal of Value Inquiry*, 51(3): 2013.
- Thomas G. Reio Jr., Joseph M. Petrosko , Albert K. Wiswell & Juthamas Thongsukmag (2006). The Measurement and Conceptualization of Curiosity. *The Journal of Genetic Psychology*, 167:2, 117-135
- Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, 17(5): 1–25.
- Samejima, L. (1969) Estimation of latent ability using a pattern of graded scores. *Psychometrika Monograph Supplement, No. 17*. Greensboro, NC: Psychometric Society.
- Samejima, L. (1996). The graded response model. In W. J. van der Linden & Hambleton, R. K. (Eds.), *Handbook of modern item response theory*. New York: Springer.
- Simms, L. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass*, 2:1, 414-33.
- Vazire, S. & Carlson, E. (2010) Self-knowledge of personality: Do people know themselves? *Social and Personality Psychology Compass*, 4, 605-620.
- Whitcomb, D., Battaly, H., Baehr, J., & Howard-Snyder, D. (forthcoming). Intellectual humility: Owing our limitations. *Philosophy and Phenomenological Research*.
- Whitcomb, D. (2010). Curiosity was framed. *Philosophy and Phenomenological Research*, 81(3): 664-87.

Appendix A. Complete List of Items in English, organized by scale

Items marked with (-) are reverse-scored.

Curiosity scale
19 I try to learn as much as I can. 20 I often push myself very hard when trying to master a new idea. 21 (-) Often when I'm in the process of learning something, I end up quitting without having really mastered it. 22 (-) I learn only the minimum amount needed to get by.
Intellectual humility scale
<i>Open-mindedness items</i> 17 Whenever I don't understand something, I want to ask another person for guidance. 23 When someone disagrees with me, I try to find out why. 27 (-) I think that paying attention to people who disagree with me is a waste of time. 28 I like talking to people with different viewpoints than mine. 33 (-) When a person disagrees with me, I usually assume there is something wrong with that person. 34 I feel no shame learning from someone who knows more than me. 35 If I do not know much about some topic, I don't mind being taught about it, even if I know a lot about other topics. 36 Teachers can learn a lot from their students. 41 I would never brag about how much I know. 45 Even when I have high status, I don't mind learning from others who have lower status. 50 (-) Only wimps admit that they've made mistakes. 51 (-) I don't take people seriously if they're very different from me 52 (-) People from other countries have weird ideas.
<i>Intellectual Modesty items</i> 8 Being smarter than other people is not especially important to me. 9 (-) I would like to have more access to information than everyone else. 10 (-) I would like to be seen explaining ideas that no one else understands. 11 (-) I would get a lot of pleasure from knowing more than other people. 12 I don't feel that I know or understand more than most other people. 13 I wouldn't want people to treat me as though I were intellectually superior to them. 14 (-) I think I am entitled to more intellectual respect than the average person is. 15 (-) I want people to know that I am an unusually intelligent person. 16 (-) I can solve difficult puzzles without needing intellectual support from anyone else. 32 (-) I like to be the smartest person in the room.
<i>Corrigibility items</i> 37 (-) I find it annoying to be told that I've made an intellectual mistake. 38 (-) If someone points out an intellectual mistake that I've made, I tend to get angry. 39 I appreciate being corrected when I make a mistake. 40 When someone corrects a mistake that I've made, I do not feel embarrassed.
<i>Engagement items</i>

- | |
|---|
| 18 (-) I rarely discuss things that I wish I understood better with other people.
24 I enjoy reading about the ideas of different cultures.
25 (-) I would be very bored by a book about ideas I disagreed with.
26 (-) I've never really enjoyed figuring out why people disagree with me.
29 (-) I find it boring to discuss things I don't already understand.
31 (-) A disagreement is like a war. |
|---|

Uniqueness items

- | |
|--|
| 42 (-) I don't feel special when I realize that I know a lot.
44 (-) If I know a lot about some topic, I don't feel special about it.
46 (-) I don't feel special when I realize that I know more than other people. |
|--|

Machiavellianism items

- | |
|--|
| 1 If I want someone I dislike to tell me what I want to know, I will act very nicely toward that person in order to get the information.
2 (-) I wouldn't use flattery to get information from someone, even if I thought it would succeed.
3 If I want information from someone, I will laugh at that person's worst jokes.
4 (-) I wouldn't pretend to like someone just to get that person to tell me what I want to know. |
|--|

Kleptomania items

- | |
|---|
| 5 If I knew I could never get caught, I would be willing to steal someone else's ideas.
6 I would be tempted to steal someone else's ideas if I didn't have the time or interest to come up with my own ideas.
7 I'd be tempted to copy someone else's work, if I were sure I could get away with it. |
|---|

Items not loading above .3 on any factor:

- | |
|--|
| 30 I would rather be convincing but wrong than unconvincing but right.
47 I have a good understanding of what I know and what I don't know.
48 Being prejudiced against someone in a discussion is not a big problem.
49 When arguing with someone, tend to exaggerate how much I know. |
|--|