



Measuring Virtue: An Aristotelian Perspective on Advancing Positive Psychology

James McManus^{1,2}  · Oleg N. Medvedev² · Liezl van Zyl¹ · Dan Weijers¹

Accepted: 6 May 2024
© The Author(s) 2024

Abstract

In philosophical discourse, flourishing is widely thought to require the development and exercise of virtues. Positive psychology, as the science of happiness and flourishing, should, therefore, include the scientific investigation of the virtues. However, the main classification of virtues and strengths, the Values in Action Inventory of Strengths (VIA-IS), faces major criticisms that, if not addressed, may undermine the credibility of positive psychology as a science. One such criticism is that virtue-based classifications lack a conceptual foundation; another is that these classifications may harm individuals by attaching potentially stigmatising labels to them. In this paper, we propose that positive psychologists can address these critiques by taking an interdisciplinary approach, which involves adopting a philosophically rich Aristotelian theory of virtue, and further developing this approach through empirical research. Adopting a philosophical theory will provide a strong conceptual base for psychologists, guiding the construction of measures and the formation of hypotheses about virtues. However, as many scholars acknowledge, it is difficult to measure all aspects of Aristotelian virtues, such as emotions, reasons, and attitudes. Another aim of this paper is, therefore, to suggest appropriate methods for measuring these aspects of Aristotelian virtues.

Keywords Positive psychology · Philosophy · Interdisciplinary research · Aristotelian virtue theory · Psychometric assessment

Oleg N. Medvedev, Liezl van Zyl and Dan Weijers are jointly contributed 20 percent.

✉ James McManus
jamesmcmanus128@gmail.com

¹ School of Social Sciences, University of Waikato, Philosophy, Hamilton, New Zealand

² School of Psychology, University of Waikato, Hamilton 3216, New Zealand

1 Introduction

Positive Psychology is the scientific study of human happiness, well-being, and flourishing (Gable & Haidt, 2005). In Philosophy, human flourishing is thought to require the exercise and development of the virtues (Kraut, 2009). For this reason, Positive Psychology should include a conception of the virtues that can form the basis of empirical study. However, the most popular model of virtues and strengths in Positive Psychology, the Values-in-Action Inventory of Strengths (VIA-IS), faces major criticisms (Efendic & Van Zyl, 2019; Van Zyl et al., 2023). Among the strongest of these criticisms is the claim that it lacks an adequate conceptual foundation, and the claim that it forms part of a neoliberal ideology that harms individuals by labelling and blaming them for qualities that are not within their control (Van Zyl et al., 2023; Bright et al., 2014; Fowers et al., 2021; Kristjánsson, 2012; Kristjánsson, 2013; Fernández-Ríos & Novo, 2012; Burr & Dick, 2021; Thompson, 2018). These critiques of the study of virtues in Positive Psychology threaten to undermine the credibility of the discipline (Efendic & Van Zyl, 2019).

To respond to these critiques, we propose that positive psychologists should take an interdisciplinary approach to virtue measurement by drawing from Aristotelian virtue theory, which is considered the dominant virtue theory in Philosophy (Swanton, 2021). We argue that positive psychologists should focus on developing measurement tools that capture a complete account of virtue, based on an Aristotelian theory of virtue. This entails that scholars operationalise each main component of the virtues, including appropriate behaviours, emotions, reasoning, motivations, and practical wisdom as well as the corresponding vices. As some scholars have noted, measuring each of these components is challenging (Fowers et al., 2021; Shahab et al., 2020). For this reason, we propose measurement strategies and specific statistical analyses that we believe will be useful for measuring virtues, such as Confirmatory Factor analysis (CFA), Exploratory Structural Equation Modelling (SEM), Latent Profile analysis (LPA), and Network analysis.

In the first section of this paper, “How the VIA-IS Lacks a Conceptual Foundation”, we begin by introducing the VIA-IS and the critique that it lacks a strong conceptual foundation. In the second section we introduce Aristotelian virtue theory by explaining the key components of virtues, such as appropriate behaviours, emotions, reasoning, motivations, and practical wisdom as well as the corresponding vices. Section three, “Measuring Virtue”, explains how Aristotelian virtues should be operationalised. We propose that psychometric virtue measurements should include specially designed facets that measure each component of virtue individually. In section four, “Measurement Analyses for Virtue Scales”, we explain specific statistical techniques, such as Exploratory Structural Equation Modeling, which we believe will be helpful for validating and investigating virtue scales. In the final section we discuss the critique that Positive Psychology is a neo-liberal ideology that causes harm to individuals by blaming them for character traits that are the result of situational influences rather than personal responsibility. We discuss this critique in relation to an Aristotelian approach to virtue, explaining that care needs to be taken to minimise stigma and harm.

1.1 How the VIA-IS Lacks a Conceptual Foundation

Positive psychologists focus on the positive aspects of human beings — they are interested in what makes it possible for people to flourish or thrive (Gable & Haidt, 2005). It is commonly accepted that the development and exercise of virtuous character traits are required for flourishing or happiness (Kraut, 2009). For this reason, virtues should be fundamental objects of concern and importance to positive psychologists generally. Indeed, they underpin the popular VIA-IS — currently the dominant positive psychological tool for assessing and studying virtues (Peterson & Seligman, 2004; Kristjánsson, 2013). Although there are useful approaches to happiness and well-being in psychology, such as hedonic accounts that utilise measures of positive emotions and life satisfaction, flourishing and virtue-based approaches offer an alternative conception of well-being which can complement existing theories and research (Delle Fave, 2020; Diener et al., 1985; Watson et al., 1988). Numerous empirical studies have been conducted with the VIA-IS, and the evidence strongly supports the notion that character strengths are related to various aspects of well-being; in particular, they correlate with the PERMA model, which is designed to measure flourishing (PERMA stands for Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment) (Green, 2022; Wagner et al., 2020; Bruna et al., 2019; Wagner & Ruch, 2015; Boiman-Meshita & Littman-Ovadia, 2021).

The VIA-IS is a classification of 24 character strengths that are arranged under six broad virtue categories: Wisdom, Courage, Humanity, Justice, Temperance and Transcendence (Peterson & Seligman, 2004). Each of the 24 character strengths represents one way of exhibiting a particular virtue. For instance, Temperance is represented by the strengths of forgiveness, gratitude, humility, prudence, and self-regulation. The virtues and strengths included in the VIA-IS were partly selected because they are valued across cultures and historical periods and are thought to contribute to the flourishing of all human beings (Peterson & Seligman, 2004). However, despite promising findings, and as noted in a systematic review conducted by Van Zyl et al. (2023), several critics of positive psychology have argued that although the virtues are a fundamental element of the VIA-IS, they are poorly conceptualised and viewed merely as a set of behaviors. Virtue theorists have similarly argued that the VIA-IS is inconsistent with an Aristotelian account of virtue, despite being inspired by it (Bright et al., 2014; Fowers et al., 2021; Kristjánsson, 2012; Kristjánsson, 2013).

One of the significant drawbacks of the VIA-IS is its focus on observable behaviours as opposed to the internal characteristics of individuals. Aristotelian virtue theory, being the dominant philosophical theory, views virtues as deep traits, reflecting people's inner states, including their values, desires, emotions, reasoning and motivations (Hursthouse, 2001). Moreover, Aristotelians define virtues as excellences of character, constituting appropriate responses to virtue-relevant stimuli. Despite the richness of the Aristotelian virtue theory offered by philosophers, the VIA-IS pays little attention to the non-behavioural aspects of virtues. Several scholars, including Fowers et al. (2021), Snow (2022), and Bright et al. (2014), have criticised Positive Psychology for its focus on measurable behaviours as well as its neglect of the Aristotelian theory in relation to excellence. Bright et al. (2014) observe that the tools employed by positive psychologists (namely, self-report questionnaires) take

a continuous approach to virtue, implying that virtue is something everyone possesses to some degree. This behaviourally driven and continuous approach is, they argue, incompatible with the idea that virtues are excellences. For instance, they cite several studies that suggest that virtue expression can be maladaptive; for example, forgiveness has been found to perpetuate relationship abuse, and generosity to impair workplace functioning (McNulty, 2011; Flynn, 2003). Based on these findings, some scholars claim that virtues are not inherently valuable, but neutral and context-dependent (Bright et al., 2014). This contrasts with the Aristotelian idea that virtues are excellences, consisting of appropriate responses to any given circumstance. They also argue that the continuous approach to virtues comes apart from the traditional understanding of virtues as types of traits that consist of a mean between two vices, one of excess, the other of deficiency (Bright et al., 2014; Aristotle, ca. 350 B.C.E./2020). Taking a purely continuous approach may also restrict the influence of Aristotelian theory in Positive Psychology by neglecting other character dispositions, such as continence and incontinence, which are dispositions that fall short of full virtue due to a lack of one or more necessary component of virtue (these terms will be explained in more detail shortly).

In response to these criticisms, we propose that rather than merely taking inspiration from some aspects of Aristotelian virtue theory, positive psychologists should develop measures that are based on all major aspects of this theory. As a rich and well-developed position in Philosophy, Aristotelian virtue theory offers a conceptual foundation or basis for empirical models of virtue, which, as noted earlier, is currently lacking in the most popular classification of virtues and character strengths in Positive Psychology, the VIA-IS. A number of interdisciplinary scholars have attempted to measure virtue as informed by Aristotelianism (Wright et al., 2020; Morgan et al., 2017; Darnell et al., 2022). Before discussing these scales and offering further suggestions on measuring virtue, we will briefly outline the important aspects of Aristotelian virtue theory.

1.2 The Aristotelian Theory of Virtue

According to Aristotelian virtue theorists, virtues such as honesty, kindness, and courage, are multifaceted dispositions of character that are necessary for achieving *eudaimonia* (Aristotle, ca. 350 B.C.E./2020). ‘*Eudaimonia*’ originates from ancient Greek and is variously translated as ‘human flourishing’, ‘true happiness’, ‘a good life’ and ‘well-being’ (e.g., Blackburn, 2016). Aristotelians prefer ‘flourishing’ as it captures Aristotle’s naturalistic approach to living well (Hursthouse, 2001). Regardless of the specific translation, *eudaimonia* is seen as an end that is good or desirable in itself (Hursthouse, 2001; Kraut, 2022; Blackburn, 2016; Aristotle, ca. 350 B.C.E./2020). Thus, although the virtues are descriptive of real psychological tendencies, they are also defined in terms of normative standards — virtues are traits that enable people to flourish.

According to Aristotle, flourishing is not merely a matter of experiencing happy or positive affective states (although he accepts that a happy life will include pleasant feelings). Rather, to flourish is to function well as a human being, which he argues involves virtuous activity (Aristotle, ca. 350 B.C.E./2020). For Aristotle, humans are

by nature rational and social beings. He therefore considers the virtues— the traits required for functioning well as human beings— to be rational excellences that allow people to thrive in human societies. Following Aristotle, contemporary virtue theorists claim that being virtuous is good for the individual, given that it makes it possible for them to live a life that is both desirable or worthwhile as well as admirable. But virtuous activity (e.g. being honest in one's dealings with others, courageously pursuing worthwhile goals, etc.) also contribute to the happiness of others as well as the good of society as a whole (Aristotle, ca. 350 B.C.E./2020; Hursthouse, 2001).

Virtues are deep traits, in that they involve a person's behavioural, affective, and cognitive propensities, including desires, reasoning, emotions, and motivations (Hursthouse, 2001). A virtuous person has a tendency to behave in certain ways (e.g. tell the truth, help friends in need, etc.), but what distinguishes them from someone who does these things out of habit or from selfish or deplorable motives, is that their behaviour aligns with appropriate and praiseworthy forms of reasoning, feeling, and desiring (Hursthouse, 2001). For example, giving expensive gifts is not truly generous if it is motivated by a desire to embarrass the recipients or bolster their own social status. Instead, someone with the virtue of generosity will be motivated by desire to contribute to the welfare or happiness of others, because they view this as a desirable or worthwhile goal (Aristotle, ca. 350 B.C.E./2020).

Further, and somewhat less intuitively, Aristotle argues that someone who aims to act virtuously but fails to do so due to ignorance or incompetence, does not possess the relevant virtue either (Aristotle, ca. 350 B.C.E./2020). For example, the person who is motivated by compassion for starving children, but donates large sums of money to ineffective charities, does not have generosity as a virtue. In the Aristotelian view, virtues require *phronesis* (practical wisdom), which can be defined as the ability to make good judgments about which goals are worth pursuing, as well as possessing the necessary practical skills, knowledge, and experience to successfully pursue these goals (Russell, 2009). Indeed, this nuanced understanding of practical wisdom is important, yet inappropriately captured by most extant measures of virtue, such as the VIA-IS.

As practical wisdom enables one to see which acts are good or virtuous in any given situation, people with the virtues behave, reason, and feel in ways that are appropriate in each context (Russell, 2009). For Aristotle, virtue is defined in terms of appropriateness, as an excellence. He sees each virtue as a mean corresponding to vices of excess and deficiency, that is, as inappropriate extremes in behaviour, reasoning, and emotion (Aristotle, ca. 350 B.C.E./2020). For example, the virtue of courage is the mean between the vices of cowardice and recklessness. Cowards experience too much fear relative to the situation, thereby not performing behaviours they should, whereas reckless people don't experience enough fear and perform risky behaviours that are unlikely to serve a desirable end. By contrast, the virtuous mean (courage) involves experiencing an appropriate amount of fear relative to the situation. In the Aristotelian view, expressing emotions such as fear and anger is not always consistent with vice. For example, expressing a certain amount of anger in response to an act of injustice can be appropriate in a given situation and be conducive to the achievement of worthwhile ends. It is important to note that extant measures of virtue currently lack a sufficiently complex understanding of appropriateness in

relation to behaviour, emotion, and reasoning), and although scholars acknowledge the difficulty in measuring it, we think it is useful to incorporate the idea of virtue as a mean in assessments of virtue (Fowers et al., 2021).

The vices are extremes (through either excess or defect), but there are ways of failing to be virtuous that do not amount to vice. Aristotle describes various other character types that people may exhibit, namely continence, incontinence, natural virtue and habit (Aristotle, ca. 350 B.C.E./2020). Continence refers to the disposition to behave appropriately based on a correct understanding of why the behaviour is called for. However, continent people do not experience appropriate affect— they typically don't want to do what is right, and so have to resort to the use self-control or will power to behave appropriately. Continence differs from full virtue because a fully virtuous person has a desire to do what is right and good, and therefore does so effortlessly and without having to resist conflicting feelings or desires (Aristotle, ca. 350 B.C.E./2020). Although they are not yet virtuous, continent people have made significant progress in learning how to be virtuous.

Incontinence is described as weakness of will. Like both the continent and fully virtuous person, an incontinent person reasons appropriately and knows what things are important. However, they do not experience the appropriate emotional impetus to perform right actions, and neither do they have the self-control to do so. As a result, they often act inappropriately, despite knowing better (Aristotle, ca. 350 B.C.E./2020).

Finally, Aristotle uses the term 'natural virtue' to refer to a state that is characterised by appropriate emotions and desires but accompanied by a lack of practical wisdom (Aristotle, ca. 350 B.C.E./2020). This disposition is often found in well-natured children who have praiseworthy motivations who lack the practical experience to successfully promote desirable ends. People with this disposition cannot be relied upon to act appropriately. For instance, a well-motivated child might remove a goldfish from its tank to warm it up. Although this act is motivated by compassion, it is not a virtuous action because it is not informed by practical wisdom— an understanding of what is required to promote the welfare of goldfish. A child who possesses natural virtue are in the beginning states of developing full virtue. It may therefore be useful to study this disposition, perhaps offering insights into improving character development in children and adolescents.

These character types are presented in Table 1. Note that there are more possible combinations involving the different components of character; however, Table 1 presents the main dispositions Aristotle discusses.

2 Measuring Virtue

In Psychology, to date, there are no comprehensive accounts of Aristotle's virtue theory. This is unfortunate, because it is the dominant theory in Philosophy and offers a detailed conceptual framework for the study of virtue (Swanton, 2021). However, in recent years there has been some very promising interdisciplinary work, which we will discuss here. We propose that such work should continue but that it should pay closer attention to all important aspects of Aristotelian virtue theory.

Table 1 Aristotelian Character Dispositions

Note Appropriate cognition subsumes practical wisdom, including the broad deliberative ability to determine what ends are valuable and worth promoting and preserving. When someone has both appropriate cognition and affect, they have appropriate motivations

	Appropriate Cognition (Practical Wisdom)	Appropriate Motivations	
		Appropriate Affect	Appropriate Behaviours
Full Virtue	Yes	Yes	Yes
Contenance	Yes	No	Yes
Incontenance	Yes	No	No
Natural virtue	No	Yes	Sometimes
Habit	No	No	Yes
Vice (excess or deficiency)	No	No	No

The Multi-Component Gratitude Scale, developed by Morgan et al. (2017), is one example of a scale created by interdisciplinary scholars to measure Aristotelian virtues. Morgan et al. (2017) contend that any measure of virtue should focus on several psychological components. Consequently, the Multi-Component Gratitude Scale is designed to assess four components of gratitude: a behavioural component, an affective component, an attitudinal component, and a conceptual component. This attempt to measure virtue is commendable and, in our view, one of the best examples of a psychometric scale designed to measure virtue. Morgan et al.'s (2017) scale measures nearly all of the important aspects of virtue emphasised in Aristotelian virtue theory; it has various items relating to emotions, behaviours, and beliefs. Moreover, this scale doesn't just capture important elements of Aristotle's theory, it also predicts well-being scores on the Satisfaction with Life Scale (Diener et al., 1985), the Subjective Happiness Scale (Lyubomirsky & Lepper, 1999), and the Positive and Negative Affect Schedule (Watson et al., 1988). The developers found that according to these scales, the more components of gratitude that participants scored highly on, the higher they reported their well-being to be. This serves to demonstrate the possibility and importance of developing and testing fuller accounts of virtue, as well as the usefulness of Aristotle's virtue theory.

However, despite adhering to a full Aristotelian account of virtue more closely than the VIA-IS classification, the Multi-Component Gratitude Scale still suffers from certain limitations. For one, although it includes an account of appropriateness, the scale gives a very limited account of the vices of excess, with no items measuring the tendencies for excessiveness in emotions and behaviours. As such, it doesn't successfully measure gratitude as a mean between a deficiency (ingratitude) and an excess ('overgratefulness') and is therefore incomplete (Manela, 2019).

Moreover, whereas Aristotelian virtue theory holds that virtuous people have virtuous motivations, the Multi-Component Gratitude Scale doesn't include items about the general and fundamental motivations underlying participants' actions, such as concern for improving their own and others' welfare. Another limitation concerns practical wisdom. The Multi-Component Gratitude Scale assesses people's beliefs and conceptions, tapping into their practical wisdom, but there are no items to assess whether people know (or believe they know) how to express their gratitude successfully. Overall, although this scale is useful, these limitations mean that it is not a suitable model for measuring the Aristotelian virtue of gratitude in all its dimensions.

Another good example of interdisciplinary research on virtue measurement is the Aristotelian Phronesis Model (APM) (Darnell et al., 2022). The APM is an empirical model that conceptualises practical wisdom as involving four different functions. These functions include the constitutive, blueprint, emotional regulation, and integrative functions. Briefly, the constitutive function is about being able to perceive the morally relevant stimuli in any given situation. The blueprint function is about comprehending which behaviours are likely to result in flourishing. The integrative function relates to the ability to weigh and balance different virtues in situations where the demands of these virtues conflict. Finally, the emotional regulation function relates to experiencing emotions that are appropriate and in line with reason, which can be achieved through cognitive appraisals of the situation. In our view, this model of practical wisdom is comprehensive and aligns well with Aristotelian theory.

Recently, Darnell et al. (2022) performed a ‘proof of concept study’, taking items from various scales to assess the APM. This scale draws from different assessment types, such as self-report and ability test questions, including multi-choice, short answer, and vignette style questions, to assess the four functions of practical wisdom. The researchers found that their four-function model was supported by confirmatory factor analysis and that higher scores on the scale are related to performing more pro-social behaviours. Despite the acknowledgement of the researchers that this scale, as a proof of concept, only approximates the real construct, this research is another good example of how interdisciplinary work by psychologists and philosophers can achieve conceptually rich scales. However, as the APM scale assesses practical wisdom, which is an intellectual virtue, it cannot measure every aspect of character virtues, as understood by Aristotelian virtue theorists, such as appropriate behaviours and emotions (Kraut, 2022).

We propose that interdisciplinary approaches, such as the ones taken by Morgan et al. (2017) and Darnell et al. (2022), offer the best method for developing conceptually rich and theory-driven models of virtue. However, current measures of virtues, including those that take an interdisciplinary approach to Aristotelian virtue theory, have shortcomings. For this reason, our aim is to offer suggestions for measuring all the important aspects of virtue highlighted by Aristotelian virtue theorists.

We suggest that psychometric measures of virtue, based on Aristotelian virtue theory, should include all the character dispositions discussed by Aristotle (virtue, vice, continence, incontinence, habit, and natural virtue) and should include the different dimensions or elements of each of these dispositions (behaviour, emotions, motivation, and reasoning). (Aristotle, ca. 350 B.C.E./2020). This will allow us to investigate the accuracy and consistency of an Aristotelian theory of character. It requires designing separate facets for measuring each dimension of virtue independently. For example, an Aristotelian virtue scale should have separate facets intended to measure appropriate behaviours, emotions and reasons separately. However, to properly measure the coalescence of the different components of virtue, we also propose an additional facet that measures motivations (appropriate motivations involve appropriate reasoning and emotions). Motivations should be measured in a way that ties specific behaviours to praiseworthy and appropriate actions. For example, if we had a behavioural item about donating to charity, we should have a corresponding motivational

item assessing whether these behaviours are associated with praiseworthy motives, such as concern for the welfare of others.

One limitation of measuring praiseworthy motives is that there are numerous appropriate motives that could lead someone to perform any particular virtuous act. However, for the sake of a general scale designed to measure virtue, motives can be reduced to the fundamental motivations at the core of Aristotle's secular theory of virtue. Fundamentally, that is, virtues are about living an excellent human life and promoting the flourishing of oneself and others (Aristotle, ca. 350 B.C.E./2020). Therefore, we suggest that any account of virtuous motivations should at least focus on motivations related to promoting and preserving one's own flourishing and the flourishing of others. Moreover, as the virtues are excellences, and people with virtue strive towards excellence, additional motivations that tie behaviours to a general aim to achieve excellence can be included. Motivations of concern for others, self-concern, and appreciation of excellence, measured as separate and distinguishable facets of virtue, should be adequate to measure the general appropriateness of one's motivations, especially in combination with the independent behavioural, emotional, and reasoning items. Of course, more specific motivations can be added to address more nuanced research questions and to study virtuous motivations in greater depth.

Another concern amongst researchers is how to measure the vices, particularly the vice of excess. Typically, parametric psychology scales measure linear constructs — the higher you score on a measure, the more of the trait you have. However, this poses challenges for Aristotelian virtue theory because exceeding what is appropriate in terms of motivations and behaviours results in the vice of excess, which is no longer a virtue (Aristotle, ca. 350 B.C.E./2020). In the case of honesty, for instance, an individual would not have the virtue of honesty if they were committed to *always* telling the truth, regardless of the situation. A well-known example in this regard is the person who, when asked by a Nazi officer whether there are Jewish people in the house, tells the truth about their Jewish friend hiding in the attic. In this case, telling the truth is not an expression of virtue, especially if the motivation to be honest (just for the sake of honesty) was stronger than their motivation to protect their friend. Rather, telling the truth to the Nazis, in this case, manifests vice because it expresses an inappropriate and excessive concern for the truth. Vice is a way of failing to be virtuous, and, as a disposition, forms a significant part of someone's character. Determining whether someone tells the truth in the wrong ways, in inappropriate circumstances, and so on, is important for determining how virtuous the person is.

Our proposal for measuring the vice of excess is to design an independent facet to measure it (rather than incorporating it into other facets that cover deficient to appropriately high levels). The reason for this is that when measuring vice as an independent facet, it will be possible to investigate vice scores effectively and independently of the other facets, such as appropriate behaviours. This will make it easier to assess how the vice of excess correlates with other constructs, such as life satisfaction and depression, anxiety and stress. Moreover, although it may sound counterintuitive, we would expect vice items to correlate with the other virtue items, such as behavioural tendencies and motivations. This is because people who are virtuously honest (or courageous) and people who are excessive in honesty (or reckless) are likely to report similar behaviours, such as telling the truth (or standing up to adversities) on scales

designed for the general population and in ordinary contexts. The difference between these groups can be found in their sensitivity to context. Whereas people who have the virtue of courage will know when the risks of a situation outweigh the benefits, the person with the vice of recklessness will not.

In terms of the vice of excess and practical wisdom, people higher in practical wisdom will theoretically be less likely to feel and behave in excessive ways (Aristotle, ca. 350 B.C.E./2020). However, when it comes to self-report, people with vices of excess may rate themselves highly, as they might believe (wrongly) that they know which ends are valuable and how to preserve and promote these. In this regard, distinguishing between virtue and the vice of excess may be tricky — what we propose is to investigate and identify the group of participants who score highly on the virtue facets, namely, appropriate behaviours, emotions, practical wisdom and motivations, without scoring highly on a specially designed vice of excess facet. This group of people may be considered high in virtue, whereas people scoring high on the virtue items as well as the excess items (or extremely high on the total scale) can be thought to have the vice of excess. (We will suggest statistical analyses for doing this). For a detailed investigation of the relationship between the vices and practical wisdom, another promising avenue of research would be to assess how people who score highly on the vices respond to a comprehensive practical wisdom scale, ideally one that includes ability test items such as the one employed by Darnell et al. (2022). For an example of how to organise facets and items to measure a virtue, see Supplementary File 1, in which we present a model for measuring the virtue of conscientiousness.

When it comes to scale construction, we recommend that items be written carefully and in accordance with the relevant philosophical literature. This means that the authors should have a working understanding of Aristotelian virtue theory and the philosophical research on the virtue targeted for measurement. This can be a daunting task, especially for someone with little training in Philosophy; for this reason, scale construction can benefit from collaboration with at least one expert in Aristotelian virtue theory.

In some cases, however, establishing what types of items constitute expressions of virtue or what types of situations and stimuli are the most relevant to one virtue rather than others may be difficult. For this reason, some scholars working on virtue measurement recommend using prototype analysis (Wright et al., 2020). Prototype analysis can be performed by creating an index of items denoting situational stimuli relevant to the virtue being measured. These indexes of items can then be given to experts in virtue theory to rate these items in order of how important and prototypical they are of the target virtue. This strategy may be particularly useful for establishing the conceptual boundaries between virtues that appear very similar, such as generosity and compassion. Using this method could also lessen the chances of researchers committing the jingle jangle fallacy, which occurs when the same construct is measured with different names (Wright et al., 2020).

Writing virtue-based items can also be challenging because some parts of Aristotelian theory, such as motivation-based facets, may be more complex and abstract than others. For this reason, specific and clear instructions may be required to guide participants and provide context. Pilot testing on a small sample is also highly recom-

mended before full administration. This is to ensure that the items are comprehensible by non-experts in virtue theory and that they contain accessible wording.

Perhaps the most important thing to consider when creating items to assess an Aristotelian account of virtue is whether the items cover a complete account of the virtue. For this reason, it is important to develop a large enough item pool to capture these various aspects of virtues. Typically, eight to ten items per domain are written during scale development (e.g., Wood et al., 2008; Kun et al., 2017; Pratscher et al., 2019). Additionally, to fulfil key psychometric requirements, scales should contain three items per measurement domain after poor items have been removed (Robinson, 2018). For this reason, virtue scales designed to measure the seven distinctive aspects of virtues discussed in this paper (e.g., appropriate behaviours, emotions, concern for others, self-concern, appreciation of excellence, practical wisdom, and the vice of excess) should contain anywhere between 21 and 70 items, as three items are required for each measurement domain to establish reliability (Robinson, 2018). However, items in the mid-range of this bracket (i.e., around 30–40) may be more ideal, as 21 items might be insufficient for capturing the entirety of each aspect of virtue. On the other hand, longer scales may become impractical to administer due to response fatigue, making the scale burdensome, especially if administered with other measures of interest. Nevertheless, the exact number of items required to appropriately measure any particular virtue depends on multiple factors, such as the number of facets included and the exact purpose of the scale.

When developing virtue scales, it is also worth considering the particular virtue being targeted for measurement. This is because virtues can be more or less complex; as such, some virtues may require more items to operationalise comprehensively. For instance, Miller (2021) proposes that honesty consists of truthfulness, forthrightness, respect for property, proper compliance, and fidelity to promises. Such an account is very broad and would require a large pool of items. In cases like these, it is also acceptable to measure one domain of the virtue, such as truth-telling, as its own sub-virtue. This may help create deeper measures that provide sufficient insight into each component of the virtues while still being practical for various purposes.

Response options should also be clear; we suggest four to five response options measuring agreement with each item. This is because fewer than four response categories can impede the sensitivity and the ability of the scale to differentiate between people with different characteristics properly, and more than five items can lead to confusion, making it difficult for participants to determine the differences between each response category (Medvedev et al., 2016; Sprague et al., 2018; Robinson, 2018; Simms et al., 2019). Response options that measure how much people agree with each item (e.g., disagree to strongly agree) work better for assessing virtues compared to response options that ask about the frequency of manifesting virtue (e.g., never to very often). This is because some aspects of virtue necessitate the measurement of core beliefs and motivations that people hold. Attitudes, beliefs, and values are more appropriately assessed by asking the participants whether they agree with such values and beliefs and so on. This contrasts with emotional experiences and behaviours that are performed and experienced with varying frequencies; these items can either be measured with frequency or agreement-styled response options.

Two of the main challenges associated with measuring virtues involve culture and social desirability. So far, research in positive psychology has indeed been criticised for relying on data from Western, Educated, Industrialised, Rich, and Democratic (W.E.I.R.D) populations (Van Zyl et al., 2023). This is concerning, given that virtues are constructs that are, in part, culturally relative. This means that what counts as an expression of virtue in one culture may not constitute an expression of virtue in another, given the its social norms, and so on. In this sense, virtue scales that are designed to be used in Western societies should not be presented as universal assessment tools — researchers should acknowledge the context and purpose for which the scale has been created.

The lack of tools designed to measure non-W.E.I.R.D populations is also concerning in its own right, as studying how virtue manifests differently across cultures and contexts (and learning how to measure these virtues) will benefit more people across the world. More balanced cultural research will also advance our global understanding of the similarities and differences between the virtues of different countries and the potential impacts these have on the well-being of different people. For these reasons, investigating culturally specific manifestations or expressions of virtues is a commendable research aim. However, it is recommended that culturally specific virtue research be conducted in collaboration with experts from the culture in question.

Social desirability and limiting biased responses are two further concerns in virtue research, as there may be substantial differences between reported virtue and actual virtue (Fower et al., 2021; Miller, 2017; Grimm, 2010). Social desirability can be a problem for any scale that measures behaviour or traits that are socially desirable, and that people may therefore feel pressure to conform to. In the case of virtue measurement, this is particularly problematic, as being virtuous is considered to be both desirable and admirable or praiseworthy, whereas vice is strongly disapproved of. For this reason, some people may be more inclined to rate themselves as more virtuous than they really are, which could possibly reduce the accuracy of scales designed to measure virtues.

Although social desirability is concerning, it is important to note that this type of bias is likely to introduce a consistent type of error variance, as people will tend to be biased in the same direction. Provided that the data that is gathered is normally distributed, scholars will be able to see differences in people's self-assessment of their own traits, despite social desirability effects. In these cases, useful and effective comparisons between people are still possible for research. Additionally, one way to reduce the error variance introduced to datasets via social desirability is to employ the right sort of analysis, such as applying Rasch analysis and converting ordinal scores into an interval level of measurement, thereby improving the precision of the assessment (Medvedev & Krägeloh, 2022).

Another step that can be taken to reduce the effect of social desirability is to emphasise the anonymity of the research participants in consent forms, advertisements, and invitations. Additionally, social desirability tests can be included in surveys, such as the Balanced Inventory of Desirable Responding Short Form (Hart et al., 2015). Correlations can then be conducted between these scales and the virtue assessment to determine whether the virtue measure correlates to a concerning degree with social desirability or not (e.g., Fowers et al., 2022). So far, previous research has found

weak correlations between virtue scales and social desirability measures (Fowers et al., 2021). Another potential way to reduce social desirability is to include some ability test questions, perhaps ones designed to assess the practical wisdom of respondents. Ideally, if the self-report questions are working properly, then they should be able to predict scores on the ability questions and vice versa. Likewise, self-report measures can be administered along with behavioural tests to determine whether particular self-report scales can actually predict behaviour. For instance, Fowers et al. (2022) found that participants who scored highly on their Interpersonal Fairness Scale were less likely to be influenced by situational stimuli that can influence people to not act fairly.

2.1 Measurement Analyses for Virtue Scales

To investigate this model of Aristotelian virtues, we propose that scholars use particular statistical analyses, including Exploratory and Confirmatory Factor analysis (EFA & CFA), Exploratory Structural Equation Modelling (ESEM), Rasch analysis, Latent Profile analysis (LPA), and Network analysis. EFA, CFA, and ESEM are essential for validating new scales, especially those with multiple facets, such as scales designed to measure virtues. For this reason, these should be the first major types of analyses applied after developing a virtue measure. CFA and ESEM are useful techniques for confirming whether a proposed theoretical factor structure, such as an Aristotelian model of virtue, appropriately fits and explains the variance of the observed data (Brown, 2015; Van Zyl & ten Klooster, 2022). In particular, these analyses can be used to investigate whether distinguishable but related components of virtues, such as motivational, behavioural, affective, and reasoning components, can be discovered. CFA and ESEM are also important in ensuring that the data fits an idealized and unbiased measurement model.

Assessing whether the variance in the observed data fits the measurement model is important, as it helps to ensure that the data is not biased and that there are no problematic items in the scale. In terms of virtue, it is also important to determine whether theoretical accounts of the factor structure align well with the patterns of variance observed in the data. If the theoretical account does not adequately align with the observed data, it may be inappropriate to use the proposed factor solution to inform subsequent analyses. This could be problematic, especially if one theoretical account of the structure of virtue is necessary to test a particular claim made by Aristotelian theorists, such as the claim that frequently performing virtuous actions allows us to acquire practical wisdom, and whether the vices are related to undesirable outcome variables. In this case, domain validity needs to be established to determine whether a proposed factor structure is appropriate for further study.

Several analyses can be used to establish the domain validity of virtue scales. In most cases, exploratory methods such as EFA should initially be employed. EFA is a good preliminary evaluation of the factor structure of a scale, and it can be used to determine whether items are loading onto theoretically proposed domains (Yong & Pearce, 2013). After this, confirmatory techniques can be employed with an independent sample to confirm the exploratory observations. Confirmatory techniques, such as CFA, more effectively allow researchers to test a preconceived model based on

theory, whereas EFA is primarily an exploratory method for generating hypotheses about possible models (Chumney, 2012). However, because virtues are theory-driven concepts, it is important for researchers to not merely take a data-driven approach by basing their proposed factor structure model only on EFA results. Rather, they should use EFA to provide hints about which potential models are theoretically plausible, striking a good balance between psychometric acceptability and theoretical coherence.

Although CFA is a traditional and common validation technique in Positive Psychology, positive psychologists have been critiqued for using CFA because it can result in poorer measurement quality, with biased factor loadings (Van Zyl & ten Klooster, 2022; Ng et al., 2017). Indeed, critics also point out that some assessment tools in Positive Psychology produce inconsistent factor structures, which is likely due to methodological issues, including the use of standard CFA instead of CFA adaptations and ESEM (Wong & Rory, 2018; Van Zyl & ten Klooster, 2022; Van Zyl et al., 2023). For this reason, it is recommended that future validation studies either apply recent CFA adaptations that mitigate these problems or use ESEM, which may be a useful tool for validating virtue models, as it provides more flexibility than standard CFA models.

One of the limitations of CFA is that cross-loadings are restrained to zero, limiting the dynamic interaction between factors (Van Zyl & ten Klooster, 2022). In CFA, factors are often referred to as 'pure', only loading onto their latent factor. This is problematic, as it can lead to stronger relationships between the items and factors and better fit statistics than what would otherwise be the case, causing unrealistic indicators of measurement quality (Van Zyl & ten Klooster, 2022). Moreover, researchers also note that CFA may be inappropriate for virtue and personality measurements, as these tests contain items that can be interpreted in various ways at once, causing cross-loadings between facets (Ng et al., 2017; Hopwood & Donnellan, 2010; Van Zyl & ten Klooster, 2022).

One strategy for mitigating concerns about cross-loadings is to use CFA adaptations, such as bifactor and hierarchical modelling and the correlated uniqueness model, which can reduce item-specific and method effects (Morin, 2020). Bifactor and hierarchical modelling can be used to separate the effects of general and specific factors and account for both shared and unique variance among items. Essentially, these models assume an underlying latent trait that accounts for the shared variance amongst the items (Reise, 2012). The Correlated uniqueness model is also useful as it can be used to specify the correlations between residuals of items that are conceptually related and cross-load in a way that deviates from the latent factors.

Using ESEM is another potential strategy to mitigate the concerns associated with standard CFA. ESEM allows for more flexibility by enabling a limited number of cross-loadings (close to zero) between items on different factors. This has been said to result in a more realistic model because psychological traits are often complex and correlate with multiple variables in various ways (Marsh et al., 2010, 2013; Van Zyl & ten Klooster, 2022). For example, when it comes to virtues, it is likely the case, and theoretically assumed, that one factor, such as virtuous motives, will correlate with virtuous behaviours and appropriate emotions and so on.

ESEM offers several advantages over standard CFA. For one, ESEM is considered to reduce measurement bias and generate models that are more consistent with theoretical conceptions (Van Zyl & ten Klooster, 2022). Moreover, while ESEM has its limitations, such as the inability to model hierarchical structures and other complex relationships, the recently developed ESEM-within-CFA and SET-ESEM can resolve these and related limitations. Overall, ESEM is considered a more rigorous and robust technique that can provide models more suitable for psychometric virtue and well-being measures (Van Zyl & ten Klooster, 2022).

Rasch analysis is another more advanced statistical technique that can be used for validation and further enhancing a scale reliability after employing EFA, CFA or ESEM. Rasch analysis is mainly used to examine whether an assessment conforms to the fundamental principles of measurement proposed by Thurstone (1931). These principles include unidimensionality (the idea that assessments should measure only one construct), measurement invariance (the measure should work equally well for everyone), and equal distance between measurement units (e.g., the scale should measure at least at an interval level of measurement). Rasch analysis is a generally underused technique in positive psychology, with a few exceptions (e.g., McManus et al., 2024; Medvedev, 2017); however, it is a powerful tool for eliminating error variance, as it can convert ordinal data into interval data, resulting in more reliable measurement instruments (Medvedev & Krägeloh, 2022). It can also assess differential items functioning (DIF), which can inform researchers of whether the scale works differently for different groups, such as gender or age groups. Another notable way in which Rasch is useful is that it can produce person-item threshold distribution plots that can be used to determine whether the scale can effectively measure the range of a person's abilities (e.g., how virtuous people are) in the sample and that there are no floor or ceiling effects.

These various features of Rasch analysis are very relevant to virtue measurement. As previously mentioned, the ordinal to interval conversions can eliminate error variance caused by social desirability. However, the person-item threshold distribution plot can also help mitigate concerns about social desirability. This is because the person-item threshold distribution plot can show whether a virtue scale can measure the full distribution of different amounts of virtue in the sample. If a scale can do so, and the Rasch model fit is acceptable, it is a good indication that the scale can differentiate between different people, regardless of whether social desirability bias influences them. Additionally, Rasch can also be used for cultural research, a topic that's highly relevant to virtues, by revealing whether certain cultural groups respond differently to particular items versus other cultural groups (i.e., whether there is DIF). This can be highly informative about whether meaningful comparisons can be made across cultures. Lastly, the stringent test of unidimensionality in Rasch analysis can provide stronger evidence than other types of analyses about the unidimensionality of virtue scales and whether the components of virtue really relate to each other in a way that constitutes a unified construct. For a recent example of how Rasch analysis can be used to validate a virtue assessment see McManus et al. (2024).

After a virtue measure based on Aristotelian theory has been validated, the next step in the research process is to use the scale to test claims made by Aristotelian theorists. This is vital for establishing an empirical basis for Aristotelianism and determining

whether it offers a realistic account of virtue and character. In this regard, LPA will be a particularly important tool. LPA is a statistical technique for investigating personal profiles and patterns of responses. This analysis is often compared to factor analysis; however, instead of identifying groups of related items, LPA identifies groups of individuals that respond similarly to particular groups of items (Spurk et al., 2020). These groups of individuals are typically referred to as latent populations, and people can be clustered according to a broad range of responses, such as attitudes, behaviours, emotions, and so on. In this sense, LPA is a type of categorical analysis that divides people into groups. LPA will be useful for investigating Aristotelian virtue theory, in particular, whether it is true that people roughly fall into seven character categories, that is, being fully virtuous, continent, incontinent, vicious (as an excess or defect), naturally virtuous or merely having (good or bad) habits (see Table 1). Thus, LPA could be useful for establishing empirical support for Aristotle's overall theory of character and how different components of virtue interact in different types of people. For this reason, LPA analysis, alongside other statistical analyses, will be useful for assessing Aristotelian virtue theorists' complex ideas about character dispositions, including the idea that traits are continuous but also roughly categorical.

LPA can be used to test and identify whether the particular character categories proposed by Aristotle can be discovered in the data. It will also allow for the possible identification of those individuals who score highly on the vice items compared to individuals who score highly on the virtue items. In this sense, we can investigate the character dispositions proposed by Aristotle and assess how these different dispositions relate to valuable outcome variables such as life satisfaction, educational success and health. This analysis is a useful tool for evaluating the empirical viability of Aristotelian virtue theory, as, unlike other analyses, it can identify important subcategories of people in a data set. For this reason, it can potentially advance theory, development and practice, especially if new subcategories are identified.

Another method that can be used to test claims made by virtue theorists and advance theory about the virtues is Network analysis, which is an advanced statistical method used to establish a nuanced picture of the nomological net of associations between sets of variables (Epskamp et al., 2012). It is particularly useful for post-validation analysis as it can display unique interactive links between facets of virtue that can be used to find clues and generate hypotheses about which components of virtues are the most important. Network analysis has a decided advantage over standard correlational analyses because it can be used to identify direct and indirect relationships between variables. It can also provide visual representations of complex correlational networks, thereby aiding interpretability (Åkerblom et al., 2021; Costantini et al., 2015).

Network analysis will be useful for investigating virtues, given that the virtues have multiple components that are expected to relate to each other (and perhaps to other variables) in different ways (Aristotle, ca. 350 B.C.E./2020). Thus, it can be used to identify which variables directly correlate with each other within the network of variables, thereby revealing the internal structure of the virtues, showing how each component is connected to the others. This will allow us to test an important claim made by virtue theorists, namely that people can learn to become more virtuous by habitually performing virtuous actions, which, in due time, leads them to acquire

practical wisdom and have appropriate feelings and motivations. Directional network plots display probability estimates about the expected direction of casualty between these variables (Heeren et al., 2021), thereby helping us to determine whether performing virtuous acts causes increases in virtuous motivation. Network analysis is also useful because it can display indirect relationships, such as between appropriate behaviours and emotions. It could be the case that appropriate behaviours correlate with practical wisdom, and that practical wisdom correlates with appropriate emotions. Thus, seeing these indirect relationships, as well as the direct relationships, helps form a deeper understanding of the correlational network.

Overall, Network analysis will be a useful tool for potentially advancing virtue theory. It may, for instance, identify one aspect of virtue as the most vital for attaining full virtue or experiencing life satisfaction or health. It may also be the case that each separate aspect of virtue offers unique benefits. This will be useful for discovering the components of virtue an individual is lacking in, and which aspects they should focus on in order to develop full virtue. In this way, Network analysis can be used to advance future theory. In short, we believe that Network analysis offers unique advantages, as it helps in comprehending constructs as networks of interrelated and self-sustaining variables. This understanding might be more realistic than the traditional view that well-being constructs exist as underlying latent constructs (Borsboom, 2013).

2.2 A Potentially Harmful Neo-Liberal Ideology

Even though Aristotelian virtue theory offers a rich conceptual foundation for empirical work on the virtues, it is worth noting that this theory may still be susceptible to a different challenge that is sometimes directed at Positive Psychology generally. New Aristotelian virtue scales could be another example of Positive Psychology being a ‘decontextualised neo-liberal ideology that causes harm’ (Van Zyl et al., 2023). A neo-liberal ideology, in this context, refers to the idea that individuals are responsible for their own choices, misfortunes, successes and flourishing, and so on (Van Zyl et al., 2023; Fernández-Ríos & Novo, 2012). Some of the strongest proponents of this kind of view, such as Burr and Dick (2021), engage in a social constructionist critique of Positive Psychology, in particular, they critique the individualistic approach taken by psychologists. It is important to note that this critique is based on the highly controversial assumption that there is no objective reality and that people’s perceptions are mere products of how language shapes reality through discourse and power dynamics. Despite its controversial nature, this view is still worth considering because it highlights potential harms associated with Positive Psychology in general, as well as an Aristotelian approach to flourishing and well-being. This is especially the case, as flourishing, in the Aristotelian view, partly depends on the character of individuals.

Burr and Dick (2021) observe a tension between the positive psychological and social constructionist’s understandings of people. According to Burr and Dick, positive psychologists take an individualistic approach to psychology; that is, they subscribe to “the idea that people are self-contained psychological units that exist prior to society and social relationships” (Burr & Dick, 2021, p.156). They contrast this

with their social constructionist view, which emphasises social relations and the construction of knowledge through discourse and power dynamics. From a social constructionist perspective, concepts such as traits and virtues are merely social constructions resulting from power dynamics in modern Western societies (Burr & Dick, 2021). They compare the discourse of Positive Psychology to a supposedly current conflation between large body sizes and poor health in the U.K. They claim that healthy weight standards are a socio-political construction stemming from discourses intended to encourage citizens to take personal accountability for their weight and to lower the cost of public health spending. Similarly, they suggest that personality constructs are designed to get people to take accountability for their own flourishing. These powerful discourses, they argue, often favour certain groups of people and disadvantage others. Burr and Dick (2021) view Positive Psychology as a highly political science that promotes a neoliberal ideology that positions individuals as consumers who are responsible for their own flourishing and character; the quality of their lives is primarily a consequence of their own personal successes or failures. This approach, they argue, neglects the broad external factors that determine how individuals view themselves, placing blame on individuals rather than their cultural and structural situations.

Other theorists also critique this individualistic approach taken by Positive Psychology, arguing that it causes harm. Thompson (2018) notes that classification systems, like diagnoses, necessarily make distinctions between different groups of people. Although making distinctions between things may be fundamental to human cognition, Thompson notes that the labels given to these things are not neutral, carrying meanings that can have varying repercussions on people's lives. For instance, Thompson notes that the label of being Jewish had detrimental implications for people living in Germany during the 1930s. As such, these labels cause stigma and can determine how people are treated and what resources are allocated to them (Thompson, 2018). Thompson (2018) argues that psychological disorders are types of classifications that can cause stigma by labelling individuals as dysfunctional and in need of intervention. With the creation of positive psychological classifications, the number of people thought of as dysfunctional may get even larger, as labelling some people as optimal and flourishing implies that others are languishing or dysfunctional.

This critique about labels and stigma is particularly relevant to the Aristotelian account of flourishing and well-being, as Aristotelians believe people can be classed as virtuous, vicious, continent, and so on, and that these are fundamental evaluative terms to describe people's character (Aristotle, ca. 350 B.C.E./2020). Classifying a particular individual as having a vice, say the vice of laziness, seems to position the problem within the individual, and such a label could cause harm to the individual and affect how other people view and treat them, leading to essentialism about the person's identity.

Concerns about harm and stigma are valid and need to be taken seriously. However, despite the potential risks of Positive Psychology and evaluative individual measurement, happiness, well-being, and flourishing are still important topics. Many people strive for well-being and want to flourish. Moreover, although it is true that social structures and society play vital roles in this domain, it is also true that individual differences, including the behaviours and attitudes of individuals, are important

factors for well-being (Anglim, 2020). This is one of the reasons psychologists are interested in investigating this set of factors.

It's also important to note that although external circumstances are important and can shape the ways in which people view and think about themselves and the world, interventions targeted at individuals and studies investigating individual differences are still helpful because they can reveal which characteristics are the most conducive to desirable outcomes like well-being, educational success and relationship quality (e.g., see Green, 2022; Wagner et al., 2020; Bruna et al., 2019; Wagner & Ruch, 2015; Boiman-Meshita & Littman-Ovadia, 2021). Furthermore, the study of Positive Psychology does not have to impede the work of anthropologists, political scientists, social psychologists and sociologists taking a broader sociological approach to well-being. After all, having facts about individuals, even if these are contingent on culture, can help inform scholars about how people can flourish in the cultural context they find themselves in. Furthermore, this information can complement broader movements to improve society and the political conditions for individual welfare. The evidence for this point is readily available; positive psychological interventions have been found to improve well-being in both non-clinical and clinical populations where individuals face more internal and external barriers (van Agteren et al., 2021).

The importance of external resources regarding flourishing is acknowledged by Aristotle himself, who observed that virtue alone is not necessary for flourishing and that people also need some degree of luck and access to external resources in order to live well (Aristotle, ca. 350 B.C.E./2020). This is an important part of Aristotelian virtue theory, and the conditions that help cultivate and sustain virtue is a very important area of research — one that can be complementary to developing assessment tools to measure virtue. Indeed, the proper measurement of Aristotelian virtues allows scholars to empirically investigate the link between external circumstances and virtue.

Power dynamics and oppression are, in fact, areas of interest for some virtue theorists, such as Tessman's (2005) discussion of the phenomenon of 'burdened virtue', which occurs when someone faces oppression to the extent that behaving virtuously doesn't contribute to their flourishing, or when someone who has virtuous motivations is unable to perform virtuous acts because of social or physical restraints. Burdened virtue is, therefore, an important (and testable) aspect of Aristotelian virtue theory. Overall, although virtues are seen as individual characteristics, there have been theoretical analyses of how virtues might interact, what they require, and how they develop in particular conditions. This theoretical work offers a good bridge and theoretical basis for investigating virtues within a social context and accounting for these external conditions. Moreover, burdened virtue may be a promising and valuable area of potential research that can help account for oppressive circumstances, mitigating blame and discouraging the placement of full responsibility on individuals.

When it comes to classifications, although they can cause stigma, it is still true that in order to learn about virtues and flourishing, we need ways of classifying and measuring them. Indeed, because of the potential well-being benefits of psychological measurements for assessing desirable traits, we do not suggest abandoning assessment tools for virtue measurement. Instead, we propose that caution should be taken in their investigation and understanding. Methods should be used to reduce the harm

and stigmatisation of vices through anonymising research and emphasising the flexible nature of the virtues, including how they relate to social contexts and conditions. Of course, it is important to note that virtue is just one variable that influences behaviour, and individual psychology is influenced by many different factors, often beyond the individual's control (Węziak-Białowolska et al., 2019; Steel et al., 2018; Yu et al., 2018; Doris, 2002). This is why it is essential for researchers to be clear about the context in which these instruments are constructed and tested, without making sweeping generalisations based on limited evidence and inappropriate methodologies. Of course, investigating how virtue manifests differently depending on culture and context is an interesting avenue for future research — one that will be aided by more effective assessment tools.

Whether and how character categorisation should be assigned based on individual assessment and intervention is a controversial topic that warrants much discussion. In our view, applications of character categories may be useful in various fields, including employment screening and applied organisation psychology, education, and clinical and forensic psychology. For instance, in education and clinical settings, virtue assessments can be used to test the effectiveness of particular interventions designed to aid virtue cultivation. In education, this is important, as developing virtues may help children flourish. Moreover, integrating effective character development programs into schools may result in a better society where people are more equipped to deal with challenges and work cooperatively. Making categorical classifications in education can also help identify which children may benefit from particular interventions designed to enhance certain aspects of their character. The same is true in clinical and forensic psychology, where patients' lack of virtue may be contributing to maladaptive behaviour that causes harm to themselves or others. Thus, having ways of identifying these problematic character traits and assessing them for improvement may, in some cases, be warranted.

In terms of employment screening, certain jobs require or might benefit from combinations of different character dispositions — to be effective, police officers clearly need courage, but also compassion, conscientiousness, and patience. These virtues are important for police officers as they take on social responsibility to protect members of society from harm. Employing people with virtuous traits for roles such as policing may have a direct influence on the well-being and safety of citizens. In cases like this, the potential harms associated with assessing applicants' character are likely outweighed by the value of a more effective recruitment process, especially if assessment is handled with professionalism and confidentiality. The same is true in other high-stakes jobs, such as in health care. These are fields where virtue assessments will be useful, as the values and characteristics of these professionals need to align with the goals and purpose of the institutions. Moreover, virtue assessments cannot only be used to screen potential employees but also for professional development in these fields to enhance employees' well-being and effectiveness. Virtue assessments may even benefit employees by steering them away from unsuitable careers and steering them towards a career that they find more fulfilling.

Overall, although there seem to be some fields where virtue assessment and categorisation are appropriate, it's important to be aware of the potential of categorisation to cause stigma. For this reason, virtue assessments and character labels should be

applied with care. If causation is taken, it seems possible to apply Aristotelian virtue theory without having to blame particular people as vicious or weak-willed. A sensitive application of virtue theory in psychology is also consistent with advice from virtue theorists regarding moral evaluation. For instance, we should note that most people fall short of full virtue, and that virtue is an ideal we should all strive towards as a community (Aristotle, ca. 350 B.C.E./2020). Further, Van Zyl (2019) cautions against using a virtue- and vice-framework to judge the actions and characters of other people, as this can amount to being judgemental and hypocritical. Instead, we should focus on improving our own character.

3 Conclusion

In conclusion, Positive Psychology is a relatively new science that faces challenges and criticism. In particular, the most popular classification of strengths and virtues, the VIA-IS, has been criticised for lacking a strong conceptual foundation, thereby undermining the status of Positive Psychology as a respected science (Efendic & Van Zyl, 2019; Van Zyl et al., 2023). In this paper, we have proposed that positive psychologists should respond to these criticisms by engaging in interdisciplinary work on virtue measurement. They can do this by working alongside philosophers (and perhaps scholars from other disciplines) to create measurements based on Aristotelian virtue theory. An interdisciplinary approach may be particularly helpful in terms of developing empirical models that are theoretically accurate. This will also help in designing appropriate methodology, as the important aspects of Aristotelian virtue theory will be more apparent. Given the complexity and nuance of Aristotelian virtue theory, a careful approach to virtue measurement must be taken. For this reason, we have suggested analyses to help guide researchers in validating and investigating new virtue assessments. These analyses include innovative techniques such as Rasch analysis ESEM, LPA and Network analysis.

Despite the promise of the Aristotelian approach to virtue, care must be taken to avoid causing stigmatisation and harm (Burr & Dick, 2021; Thompson, 2018). For this reason, future research should follow appropriate ethical guidelines, while also interpreting their findings in respect to cultural contexts. Future research can also focus on investigating ‘burdened virtues,’ that is, virtues that are restrained by oppression and unfortunate social circumstances (Tessman, 2005). Such work, we argue, will help to maintain a high standard of science, reduce harm to individuals, and promote well-being. Overall, taking an Aristotelian approach to virtue is promising, as it will lead to new and beneficial instruments that will help scholars understand virtue, flourishing and well-being. The creation of new and rich virtue measurements will also be useful for designing interventions aimed at helping people develop virtue. As such, additional work, involving the careful reflection of critiques of Positive Psychology, can help the discipline advance as a science, promoting a reputation of rigor and effectiveness.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s41042-024-00167-0>.

Acknowledgements The authors would like to acknowledge and express their gratitude to the University of Waikato and the University of Waikato Doctoral Scholarship for their invaluable support in facilitating this research.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Open Access funding enabled and organized by CAUL and its Member Institutions

Declarations

Ethical Approval Not applicable.

Informed Consent Not applicable.

Conflict of Interest The Authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Åkerblom, S., Cervin, M., Perrin, S., Rivano Fischer, M., Gerdle, B., & McCracken, L. M. (2021). A network analysis of clinical variables in chronic pain: A study from the Swedish Quality Registry for Pain Rehabilitation (SQRP). *Pain Medicine*, 22(7), 1591–1602. <https://doi.org/10.1093/pm/pnaa473>.
- Anglim, J., et al. (2020). Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin*, 146(4), 279. <https://doi.org/10.1037/bul0000226>.
- Aristotle (2020). *The Nicomachean Ethics* (A. Beresford, Trans.). Penguin Classics. (Original work published ca. 350 B.C.E).
- Blackburn, S. (2016). *The Oxford dictionary of philosophy* (3rd edition.). Oxford University Press. <https://doi.org/10.1093/acref/9780198735304.001.0001>.
- Boiman-Meshita, M., & Littman-Ovadia, H. (2021). Is it me or you? An actor-partner examination of the relationship between partners' character strengths and marital quality. *Journal of Happiness Studies*, 23(1), 195–210. <https://doi.org/10.1007/s10902-021-00394-1>.
- Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91–121. <https://doi.org/10.1146/annurev-clinpsy-050212-185608>.
- Bright, D. S., Winn, B. A., & Kanov, J. (2014). Reconsidering virtue: Differences of perspective in virtue ethics and the positive social sciences. *Journal of Business Ethics*, 119, 445–460. <https://doi.org/10.1007/s10551-013-1832-x>.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford.
- Bruna, M. O., Brabete, A. C., & Izquierdo, J. M. A. (2019). Reliability generalisation as a seal of quality of substantive meta-analyses: The case of the VIA Inventory of strengths (VIA-IS) and their relationships to life satisfaction. *Psychological Reports*, 122(3), 1167–1188. <https://doi.org/10.1177/0033294118779198>.
- Burr, V., & Dick (2021). A social constructionist critique of positive psychology. *Routledge International Handbook of Theoretical and Philosophical Psychology: Critiques Problems and Alternatives to Psychological Ideas*, 151–169. <https://doi.org/10.4324/9781003036517-11>.

- Chumney, F. (2012). Principal components analysis, exploratory factor analysis, and confirmatory factor analysis. *Reading and understanding multivariate statistics* (pp. 99–136). American Psychological Association.
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRT personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality, 54*, 13–29. <https://doi.org/10.1016/j.jrp.2014.07.003>.
- Darnell, C., Fowers, B. J., & Kristjánsson, K. (2022). A multifunction approach to assessing aristotelian phronesis (practical wisdom). *Personality and Individual Differences, 196*, 111684. <https://doi.org/10.1016/j.paid.2022.111684>.
- Delle Fave, A. (2020). Eudaimonic and hedonic happiness. *Encyclopedia of quality of life and well-being research*, 1–7. https://doi.org/10.1007/978-3-319-69909-7_3778-2.
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13.
- Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge University Press.
- Efendic, E., & Van Zyl, L. E. (2019). On reproducibility and replicability: Arguing for open science practices and methodological improvements at the South African journal of industrial psychology. *SA Journal of Industrial Psychology, 45*(1), 1–10. <https://doi.org/10.4102/sajip.v45i0.1607>.
- Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualisations of relationships in psychometric data. *Journal of Statistical Software, 48*, 1–18. <https://doi.org/10.18637/jss.v048.i04>.
- Fernández-Ríos, L., & Novo, M. (2012). Positive psychology: Zeigeist (or spirit of the times) or ignorance (or disinformation) of history? *International Journal of Clinical and Health Psychology, 12*(2), 333–344.
- Flynn, F. J. (2003). How much should I give and how often? The effects of generosity and frequency of favor exchange on social status and productivity. *Academy of Management Journal, 46*, 539–553. <https://doi.org/10.2307/30040648>.
- Fowers, B. J., Carroll, J. S., Leonhardt, N. D., & Cokelet, B. (2021). The emerging science of virtue. *Perspectives on Psychological Science, 16*(1), 118–147. <https://doi.org/10.1177/1745691620924473>.
- Fowers, B. J., Lane, A. A., Abbamonte, J. M., Lang, S. F., Cokelet, B., Anderson, A. R., & Cioffi, K. (2022). Does trait interpersonal fairness moderate situational influence on fairness behavior? *Personality and Individual Differences, 193*, 111615. <https://doi.org/10.1016/j.paid.2022.111615>.
- Gable, S. L., & Haidt, J. (2005). What (and why) is positive psychology? *Review of General Psychology, 9*(2), 103–110. <https://doi.org/10.1037/1089-2680.9.2.103>.
- Green, Z. A. (2022). Character strengths intervention for nurturing well-being among Pakistan's university students: A mixed-method study. *Applied Psychology: Health and Well-Being, 14*(1), 252–277. <https://doi.org/10.1111/aphw.12301>.
- Grimm, P. (2010). Social desirability bias. In J. Sheth and N. Malhotra (Eds.), *Wiley international encyclopedia of marketing*. <https://doi.org/10.1002/9781444316568.wiem02057>.
- Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The balanced inventory of desirable responding short form (BIDR-16). *Sage Open, 5*(4), 2158244015621113. <https://doi.org/10.1177/2158244015621113>.
- Heeren, A., Lannoy, S., Coussment, C., Hoebeke, Y., Verschuren, A., Blanchard, M. A., & Gierski, F. (2021). A network approach to the five-facet model of mindfulness. *Scientific Reports, 11*(1), 15094. <https://doi.org/10.1038/s41598-021-94151-2>.
- Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*(3), 332–346. <https://doi.org/10.1177/1088868310361240>.
- Hursthouse, R. (2001). *On virtue ethics*. Oxford University Press.
- Kraut, R. (2009). *What is good and why: The ethics of well-being*. Harvard University Press.
- Kraut, R. (2022). Aristotle's ethics. *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/fall2022/entries/aristotle-ethics/>.
- Kristjánsson, K. (2012). Positive psychology and positive education: Old wine in new bottles? *Educational Psychologist, 47*(2), 86–105. <https://doi.org/10.1080/00461520.2011.610678>.
- Kristjánsson, K. (2013). *Virtues and vices in positive psychology: A philosophical critique* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139177818>.
- Kun, Á., Balogh, P., & Krasz, K. G. (2017). Development of the work-related well-being questionnaire based on Seligman's PERMA model. *Periodica Polytechnica Social and Management Sciences, 25*(1), 56–63. <https://doi.org/10.3311/PPso.9326>.

- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, 46, 137–155. <https://doi.org/10.1023/A:1006824100041>.
- Manela, T. (2019). The virtue of gratitude and its associated vices. *The Moral Psychology of Gratitude*, 296–316.
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 22(3), 471. <https://doi.org/10.1037/a0019227>.
- Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, 49(6), 1194. <https://doi.org/10.1037/a0026913>.
- McManus, J., Medvedev, O. N., Zyl, L. V., & Weijers, D. (2024). Development and validation of the aristotelian virtue of conscientiousness scale using Rasch methodology. *Current Psychology*, 1–16. <https://doi.org/10.1007/s12144-024-05887-9>.
- McNulty, J. K. (2011). The dark side of forgiveness: The tendency to forgive predicts continued psychological and physical aggression in marriage. *Personality and Social Psychology Bulletin*, 37(6), 770–783. <https://doi.org/10.1177/0146167211407077>.
- Medvedev, O. N., Siegert, R. J., Feng, X. J., Billington, D. R., Jang, J. Y., & Krägeloh, C. U. (2016). Measuring trait mindfulness: How to improve the precision of the mindful attention awareness scale using a Rasch model. *Mindfulness*, 7, 384–395. <https://doi.org/10.1007/s12671-015-0454-z>.
- Medvedev, O. N., Siegert, R. J., Mohamed, A. D., Shepherd, D., Landhuis, E., & Krägeloh, C. U. (2017). The Oxford Happiness Questionnaire: Transformation from an ordinal to an interval measure using Rasch analysis. *Journal of Happiness Studies*, 18, 1425–1443. <https://doi.org/10.1007/s10902-016-9784-3>.
- Medvedev, O. N., Krägeloh, C. U., Titkova, E. A., & Siegert, R. J. (2020). Rasch analysis and ordinal-to-interval conversion tables for the Depression, anxiety and stress scale. *Journal of Health Psychology*, 25(10–11), 1374–1383. <https://doi.org/10.1177/1359105318755261>.
- Miller, C. B. (2021). *Honesty: The philosophy and psychology of a neglected virtue*. Oxford University Press. <https://doi.org/10.1093/oso/9780197567494.001.0001>.
- Morgan, B., Gulliford, L., & Kristjánsson, K. (2017). A new approach to measuring moral virtues: The multi-component Gratitude measure. *Personality and Individual Differences*, 107, 179–189. <https://doi.org/10.1016/j.paid.2016.11.044>.
- Morin, A. J., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM), and bifactor-ESEM. *Handbook of Sport Psychology*, 1044–1073. <https://doi.org/10.1002/9781119568124.ch51>.
- Ng, V., Cao, M., Marsh, H. W., Tay, L., & Seligman, M. E. (2017). The factor structure of the values in Action Inventory of strengths (VIA-IS): An item-level exploratory structural equation modeling (ESEM) bifactor analysis. *Psychological Assessment*, 29(8), 1053. <https://doi.org/10.1037/pas0000396>.
- Peterson, C., & Seligman, M. (Eds.). (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press.
- Pratscher, S. D., et al. (2019). Interpersonal mindfulness: Scale development and initial construct validation. *Mindfulness*, 10, 1044–1061. <https://doi.org/10.1007/s12671-018-1057-2>.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>.
- Robinson, M. A. (2018). Using multi-item psychometric scales for research and practice in human resource management. *Human Resource Management*, 57(3), 739–750. <https://doi.org/10.1002/hrm.21852>.
- Russell, D. C. (2009). *Practical intelligence and the virtues*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199565795.001.0001>.
- Shahab, S. O., & Adil, A. (2020). Development and validation of Temperance Scale in Pakistan. *PsyCh Journal*, 9(6), 911–923. <https://doi.org/10.1002/pchj.394>.
- Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychological Assessment*, 31(4), 557–566. <https://doi.org/10.1037/pas0000648>.
- Snow, N. E. (2022). What is a science of virtue? *Journal of Moral Education*, 51(1), 9–23. <https://doi.org/10.1080/03057240.2020.1773777>.

- Sprague, E., Siegert, R. J., Medvedev, O., & Roberts, M. H. (2018). Rasch analysis of the Edmonton symptom assessment system. *Journal of Pain and Symptom Management*, 55(5), 1356–1363. <https://doi.org/10.1016/j.jpainsymman.2018.01.016>.
- Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and how to guide of its application within vocational behavior research. *Journal of Vocational Behavior*, 120, 103445. <https://doi.org/10.1016/j.jvb.2020.103445>.
- Steel, P., Taras, V., Uggerslev, K., & Bosco, F. (2018). The happy culture: A theoretical, meta-analytic, and empirical review of the relationship between culture and wealth and subjective well-being. *Personality and Social Psychology Review*, 22(2), 128–169. <https://doi.org/10.1177/1088868317721372>.
- Swanton, C. (2021). *Target Centred Virtue Ethics*. Oxford University Press. <https://doi.org/10.1093/oso/9780198861676.001.0001>.
- Tessman, L. (2005). *Burdened virtues: Virtue ethics for liberatory struggles*. Oxford University Press. <https://doi.org/10.1093/0195179145.001.0001>.
- Thompson, S. (2018). Positive psychology, mental health, and the false promise of the medical model. In N. J. Brown, T. Lomas, & F. J. Eiroa-Orosa (Eds.), *The Routledge international handbook of critical positive psychology* (1st ed., pp. 71–83). Routledge.
- Thurstone, L. L. (1931). The measurement of social attitudes. *Abnormal and Social Psychology*, 27, 249–269. <https://doi.org/10.1037/h0070363>.
- van Agteren, J., Iasiello, M., Lo, L., Bartholomaeus, J., Kopsaftis, Z., Carey, M., & Kyrios, M. (2021). A systematic review and meta-analysis of psychological interventions to improve mental well-being. *Nature Human Behaviour*, 1–22. <https://doi.org/10.1038/s41562-021-01093-w>.
- Van Zyl, L. (2019). Does virtue ethics allow us to make better judgments of the actions of others? *Virtue Ethics: Retrospect and Prospect*, 99–110. https://doi.org/10.1007/978-3-030-15860-6_8.
- Van Zyl, L. E., & Ten Klooster, P. M. (2022). Exploratory structural equation modeling: Practical guidelines and Tutorial with a Convenient Online Tool for Mplus. *Frontiers in Psychiatry*, 12, 795672. <https://doi.org/10.3389/fpsy.2021.795672>.
- Van Zyl, L. E., Gaffaney, J., van der Vaart, L., Dik, B. J., & Donaldson, S. I. (2023). The critiques and criticisms of positive psychology: A systematic review. *The Journal of Positive Psychology*, 1–30. <https://doi.org/10.1080/17439760.2023.2178956>.
- Wagner, L., & Ruch, W. (2015). Good character at school: Positive classroom behavior mediates the link between character strengths and school achievement. *Frontiers in Psychology*, 6, 610. <https://doi.org/10.3389/fpsyg.2015.00610>.
- Wagner, L., Gander, F., Proyer, R. T., & Ruch, W. (2020). Character strengths and PERMA: Investigating the relationships of character strengths with a multidimensional framework of well-being. *Applied Research in Quality of Life*, 15, 307–328. <https://doi.org/10.1007/s11482-018-9695-z>.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063. <https://doi.org/10.1037/0022-3514.54.6.1063>.
- Węziak-Białowolska, D., McNeely, E., & VanderWeele, T. J. (2019). Human flourishing in cross cultural settings. Evidence from the United States, China, Sri Lanka, Cambodia, and Mexico. *Frontiers in Psychology*, 10, 1269. <https://doi.org/10.3389/fpsyg.2019.01269>.
- Wong, P. T. P., & Roy, S. (2018). Critique of positive psychology and positive interventions 1. In *The Routledge International Handbook of Critical Positive Psychology* (1st ed., pp. 142–160). Routledge. <https://doi.org/10.4324/9781315659794-12>.
- Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the authenticity scale. *Journal of Counseling Psychology*, 55(3), 385. <https://doi.org/10.1037/0022-0167.55.3.385>.
- Wright, J. C., Warren, M. T., & Snow, N. E. (2020). *Understanding virtue: Theory and measurement*. Oxford University Press. <https://doi.org/10.1093/oso/9780190655136.001.0001>.
- Yu, S., Levesque-Bristol, C., & Maeda, Y. (2018). General need for autonomy and subjective well-being: A meta-analysis of studies in the US and East Asia. *Journal of Happiness Studies*, 19, 1863–1882. <https://doi.org/10.1007/s10902-017-9898-2>.