



THE
JUBILEE CENTRE
FOR CHARACTER & VALUES

UNIVERSITY OF
BIRMINGHAM

Can Virtue Be Measured?

Jennifer Cole Wright

These are unpublished conference papers for the 'Can Virtue Be Measured?', held by the Jubilee Centre for Character and Values at Oriol College, Thursday 9th – Saturday 11th January 2014. These papers are works in progress and should not be cited without author's prior permission.



**School of Education
University of Birmingham
Edgbaston
Birmingham**

Jennifer Cole Wright: Can Virtue Be Measured?

(Subtitle: Sure – But, it Ain't Gonna Be Easy!)

1. The Question: Can Virtue Be Measured?

The question before us – indeed, the framing question of this conference – is “can virtue be measured”? It strikes me that this question could be a question about (at least) two things – first, whether any of the existing empirical research claiming to measure the presence (or absence) of virtue should be taken seriously (and if so, which parts); second, whether the first question is a reasonable question to ask not just because there is worry about the empirical methods employed thus far, but because the very idea of measuring something like *virtue* is implausible on the face of it. In other words, the question framing this conference is a question about whether virtue is the sort of thing that is scientifically “untouchable”, either due to the existence of *practical* barriers (not enough time, funding, resources; inadequate technology) or *in principle*.

I take it that most (if not all) of us present are disinclined towards the latter view. I myself see no reason to think that virtue is sort of thing that science cannot, at least in principle, touch – but I also think that the work done thus far has only begun to scratch the surface (and some of it, the wrong surface). That is to say, while I am inclined to think that virtue is the sort of thing empirical science can measure, any endeavor to do so – if done properly – will require a tremendous amount of time, energy, resources, and creativity and will only bear fruit through the convergence of different methodological strategies employed in different (but overlapping) ways.

With this in mind, I'd like to discuss some ideas – born out of a tremendously productive and enjoyable collaboration with Nancy Snow – about how a robust empirical program studying virtue could be built. As such, I will be focusing here on two things: 1) offering for your consideration a working “operational definition” of virtue(s), and 2) providing ideas about different empirical strategies for measuring it, borrowing from the approaches being developed to study other related (at least insofar as they share a common set of challenges) constructs, such as personality.

2. A Working Definition Of Virtue(s)

Step 1: Distinguishing between *virtue intelligence* and *full-blown virtue*

In her keynote address, Nancy Snow introduced the notion of “virtue intelligence.” Virtue intelligence, as we’ve come to think about it, is more or less shorthand for identifying the inner mental states of the virtuous agent, as understood in Aristotelian terms. Similar to constructs such as social intelligence and emotional intelligence, it identifies the perceptual processes, motivational states, cognitions, affective responses and other mechanisms that interact to lead to virtuous action.

Virtue intelligence involves, at a minimum, the presence of (some set of) virtues, such as honesty, bravery, kindness, and generosity. (It remains an open question how many of – and which – virtues would have to be present in order for a person to possess virtue intelligence.) It is important to note that while crucial, by itself virtue intelligence is not (yet) full-blown virtue. After all, it seems right to hold that people can have virtue intelligence without having the consciously acknowledged goal to develop and maintain their virtues – they can, for example, be courageous or generous without consciously wanting to be so or deliberately seeking to cultivate and sustain these virtues.

We would not want to deny that such people are virtuous *in some sense* – insofar as they possess certain virtues, they possess “virtue intelligence” (of some degree). However, at the same time, we would not want to say that they possess full-blown virtue. Full-blown virtue, we would argue, is a form of *properly motivated* virtue intelligence – i.e., virtue intelligence that has become linked with the right sort of motivational structure, such as a long-standing and consciously acknowledged commitment to the development and maintenance of virtue(s). People possessing full-blown virtue act in virtue-relevant ways not only when, but *because* such actions are called for – they are motivated to do the right things for the right reasons.

An initial categorization of the distinction between virtue intelligence and full-blown virtue, then, is based on the conscious adoption, or lack thereof, of the goal of developing and maintaining virtue(s).

Step 2: Operational definition of *virtue(s)*

Walter Mischel and Yuichi Shoda (two social-cognitive psychologists) proposed a theory of personality called the “cognitive-affective personality system” (CAPS), which includes a conception of a “trait” that meshes well with the Aristotelian conception of virtue (see e.g.,

Mischel 1973; Mischel and Shoda 1995). Snow has argued elsewhere (2010) that virtues, including Aristotelian virtues, function like (and, indeed, are a subset of) CAPS traits.

Following this work, we have operationalized virtues as “traits” – which, as such, are composed of a set of:

- (a) trait-appropriate cognitive/affective/behavioral responses that are
- (b) consistently triggered by
- (c) trait-relevant stimuli in the person’s
- (d) environment.

So, let’s unpack this.

Regarding (a): We consider traits to be a set of linked, trait-appropriate “cognitive/affective/ behavioral” responses because trait activation typically involves a variety of trait-appropriate *cognitive* states (e.g., beliefs, desires, goals, values, etc.), and *affective* states (e.g., emotions, moods, etc.) and corresponding *behavioral* responses. In other words, trait activation produces a cascade of corresponding trait-appropriate responses – “inner” mental states, emotional reactions, and “outer” behaviors. (This is not to say that the cognitive/ affective and the behavioral responses don’t come apart – they can and do. In order for someone to be said to fully possess a trait, however, they need typically to be linked so that they co-occur.)

Regarding (b): When we say that a trait’s cognitive/affective/behavioral responses are “consistently” triggered by trait-relevant stimuli, we mean at the very least that they are *typically* triggered – that is, activated more often than not. Ultimately, however, the cultivation (and/or socialization) of virtue traits will result in these responses becoming habitually activated. For any trait-relevant response to become habitual, it must be triggered repeatedly, becoming chronically accessible; it must be automatic, that is, not require conscious deliberation or processing for activation; and, at least in the case of full-blown virtue, it must be aligned in the right sort of way with the person’s values/goals – something a person does and endorses.

Regarding (c): When we say that “trait-relevant stimuli” activate cognitive/affective/behavioral responses, we are referring to stimuli in two senses – that is, we are referring to both the *objective* features of the situation to which a person is responding (features that “call for” a trait-appropriate response, such as, in the case of virtue, seeing a person slumped over and moaning – a person in apparent need – as requiring a compassionate, helpful response), but also, and just as importantly, the *subjective* interpretation that a person gives to those objective

features. Thus, the person herself must “see the situation as” being trait-relevant and calling for a trait-appropriate response. Ideally – perhaps in some cases of full-blown virtue – these two will overlap completely, but most of the time they will diverge.

This means that the development of virtue traits will require both that the person be able to accurately assess and interpret the objective features of a situation (i.e., to know when virtue-relevant stimuli are present and when they are not) and be able to determine what the virtue-appropriate response would be.

(Obviously, this raises an important problem for anyone measuring virtue traits, insofar as it assumes that there is a *right answer* about which stimuli are/are not virtue-relevant, objectively speaking, and that we can *know* what the right answers are, in order to measure the degree of overlap. Importantly, though, this is not just a problem for the measurement of virtue(s) – or for virtue ethics more generally. I will get to some thoughts about how to address this empirically in a bit).

Regarding (d): When we refer to a person’s “environment” we are referring to both the *external* environment (the physical/social world by which the person is surrounded and in which the person is immersed), and also the *internal* environment of the person him/herself (e.g., their ideas, imaginings, etc.).

I will now use this account of virtue(s) to discuss possible methodological strategies of the measurement of virtue intelligence and full-blown virtue.

3. Measuring Virtue(s) – Part I: Virtue Intelligence

According to the account given above, the measurement of “virtue intelligence” involves the measurement of the presence of (a set of) virtue traits, which we’ve operationalized as being a set of (a) virtue-appropriate cognitive/affective/behavioral responses that are (b) consistently triggered by (c) virtue-relevant stimuli in the person’s (d) environment. This means that to empirically measure the level of people’s virtue intelligence, we need to be able to measure three variables:

- 1) *Sensitivity to the presence of virtue-relevant (external/internal) stimuli, which includes the accurate attribution of virtue-relevant significance to stimuli.* For example, does a person notice that another is in need, is upset or struggling? If so, then they have the

opportunity to express compassion, which seems called for. If not, the opportunity for compassion will be missed.

- 2) *Recognition and generation of virtue-appropriate (cognitive/affective/behavioral) responses.* Suppose an instructor notices that a student is upset, perhaps on the verge of tears. Does the teacher react with appropriate gentleness and compassion? Simply ignore the student? Acknowledge the student's distress, but react brusquely or with annoyance? And would the teacher recognize another's compassionate response to the student's distress as appropriate?
- 3) *Consistency and habituality of the connection between the sensitivity to virtue-relevant stimuli and the recognition and generation of virtue-relevant responses – what we call “dispositionality”.* For example, does a person typically have a compassionate response to someone else's perceived need, or is it relatively rare? And how many different expressions of need does the person respond compassionately towards?

I will go through some suggestions for measuring each of these in turn. The inspiration for many of the ideas presented below comes from Mehl and Conner's (2012) *Handbook of Research Methods for Studying Daily Life* (especially Chs. 4-5, 7-15, and 29). Though the focus of this volume is not the study of virtue, many of the conceptual and methodological tools presented there apply.

Step 1: Measuring sensitivity to the presence of *virtue-relevant stimuli*

First, we can operationalize “sensitivity” in a number of ways: here, I will do so in terms of people's *perception* (visual/auditory), *identification*, and/or *generation* of virtue-relevant stimuli.

There are a variety of techniques that could be used to measure people's *perception* of virtue-relevant stimuli. First, we could measure people's perception of virtue-relevant stimuli in actual or fictional social interactions. For example, people could be asked to “hit a button” while watching a video or listening to an audio clip of other people's social interactions whenever they observe that a virtue-relevant stimulus is present (that is, whenever a virtue-appropriate response would be *called for*). We could then have them go back through their “hits” and verbally describe/explain what they take the virtue-relevant stimuli to be and why (as well as what they think the virtue-appropriate responses would be – also helpful in addressing step 2).

Second, we could measure people's *identification* of virtue-relevant stimuli in actual or fictional narratives (e.g., stories, short vignette-style cases, etc.) or in the abstract (decontextualized lists of stimuli). Two measurement methods will be useful here. First, we could ask people to read the narratives/vignettes, highlight the presence of virtue-relevant stimuli, and explain why they take them to be virtue-relevant stimuli (see Narvaez, 2001). They could also be asked what they think the virtue-appropriate responses would be (helpful in addressing step 2). Second, we could use a method of abstraction by giving people a virtue to consider (for example, honesty) and then presenting them with different sets of circumstances, asking them to identify for which of them virtue-relevant stimuli are present – that is, for which circumstances virtue-appropriate response would be called for.

Finally, we could use both narratives and abstraction to measure people's capacity to spontaneously *generate* virtue-relevant stimuli. For example, we can have people write narratives that contain virtue-relevant stimuli, and explain why they take them to be virtue-relevant stimuli, as well as what they think the virtue-appropriate responses would be (which would also provide information that addresses step 2). We could also have people consider a virtue, and then spontaneously generate a list of as many circumstances as they can think of in which a display of that virtue would be called for.

At some point we will need to take into account the degree of overlap between people's sensitivity to the *objective* features of the situation to which the person is responding and the *subjective* interpretation that the person gives to those objective features – what we might call their degree of “accuracy” (raising the problem mentioned earlier). While, this is a worry that simply cannot be fully deflated, I think there are several strategies we could employ to at least partially address it. One would be to develop a “paradigmaticity” index that would allow us to measure how paradigmatic particular virtue-relevant stimuli are, in terms of the speed at which they are perceived and/or how frequently they are generated or identified as examples. Participants' responses could then be measured according to how well they track paradigmatic virtue-relevant stimuli. (Of course, in this regard, we might expect a curvilinear trajectory, with people both low and high in virtue intelligence being lower on paradigmaticity than those mid-range – people high in virtue intelligence are likely to be sensitive to virtue-relevant stimuli not commonly recognized as such.) Another important consideration (distinct from paradigmaticity)

is the extent to which people provide good reasons (even if post-hoc) for why they consider things to be virtue-relevant stimuli, or not.

Step 2: Measuring recognition and generation of *virtue-appropriate responses*

People's "recognition" of virtue-appropriate responses can be operationally defined as their ability to identify cognitive/affective/behavior responses as being virtue-appropriate or not, whether we take virtue-appropriateness to be a yes/no matter, or something that admits of degree. There are several ways this could be measured:

First, we could use a third-person response method, in which we give people narratives, short-vignettes, or even shorter, decontextualized, examples of *other people's* cognitive/affective/behavior responses to virtue-relevant stimuli and ask them to identify which are virtue-appropriate and which not (and why).

Second, we could use a first-person/past response method, according to which people are asked to recall (e.g., through daily diary entries) virtue-relevant stimuli from their past and what *their own* cognitive/affective/behavior responses were, and then asked to evaluate their responses as virtue-appropriate (or not) and why (for more on diary methods, see *Handbook*, Chs, 4, 5, and 8). People's ability to identify which of their own past cognitive/affective/behavior responses were virtue-appropriate and which were not (and why) represents a useful measure of the development of virtue intelligence.

This approach could also be expanded upon to see if people can identify how to make non-virtue-appropriate cognitive/affective/behavior responses virtue-appropriate, as well as how to encourage future cognitive/affective/behavior responses to be virtue-appropriate (this would be getting into "generation"). To illustrate: we could measure 1) the extent to which a person both recognized that her treatment of one of her co-workers in the past was unfair, 2) was able to articulate how she could have responded (what alternative response would have been fair), and 3) was able to articulate steps that she could take to insure a fair response in the future.

Finally, we could use the first-person/present method, according to which people are asked to record their "live" cognitive/affective/behavior responses to virtue-relevant stimuli and then asked to evaluate them as virtue-appropriate/not-virtue-appropriate (as well as why). This could be done using daily diaries and/or more timely "event-contingent response" (ECR) technologies. These are typically portable electronic devices that allow people to record their responses at the

moment that an event occurs – in this case, a response to virtue-relevant stimuli (for more on ECR methods, see *Handbook*, Chs. 4, 5, 7, and 9).

People’s “generation” of virtue-appropriate responses can be operationally defined as their ability to articulate what the best virtue-appropriate responses would be, given the presence of particular virtue-relevant stimuli. Again, several methods of measurement are available:

First, we could use a hypothetical method in which we give people narratives, vignettes, and/or decontextualized examples of virtue-relevant stimuli (visual images also being an option), and ask them to spontaneously generate either what *their* cognitive/ affective/behavioral responses would/should be or what *others’* cognitive/affective/ behavior responses would/should be. (There is an important distinction between what people think theirs or others’ responses *would* be (actual) and what they *should* be (ideal). The difference between the actual and the ideal responses can serve as an important measure of the degree to which people both recognize the appropriate response and believe they or others would be able to achieve it.)

Second, we could use an observational/experimental approach to observe people in controlled circumstances in which certain virtue-relevant stimuli are present and then monitor their cognitive/affective/behavioral responses to those stimuli. We could then have people evaluate their responses as virtue-appropriate (or not). We could also have others evaluate them as virtue-appropriate (or not).

Monitoring can involve a video/audio set up at location, as well as recordings, of physiological responses, such as heart rate, galvanic skin response, and pupil dilation. (This is not to suggest that we take physiological responses – e.g., increased heart rate at the sign of another’s distress – to, by themselves, be indicators of virtue. Nonetheless, they are an important part of trait-appropriate cognitive/affective/behavioral responses. For example, in a study of the helping behavior of young children, Hepach, Vaish, & Tomasello (2012) found that when children witness a person in need of help, they experience sympathetic distress, revealed by increased pupil dilation. This pupil dilation decreases when that person receives help, whether it is the child who helps the individual or a third-party, suggesting that is *the person receiving help* that eases the child’s sympathetic distress, not a more egocentric desire to be the helper.)

Perhaps the most interesting method, though, would be to use an observational/naturalistic approach, observing people in natural circumstances in which certain virtue-relevant stimuli are present and then monitor their cognitive/affective/behavior responses, and ask them

(later) to evaluate their responses as virtue-appropriate (or not). We could also have others evaluate them as virtue-appropriate (or not). Monitoring can involve video/audio worn by participants – for example, by using the Electronically Activated Recorder (EAR), which is a portable audio device that records for specified periods of time (see *Handbook* Ch. 10) – or wired at a particular location, for example, in a dorm room or apartment (see *Handbook* Chs. 14-15). Mobile physiological response recording is also possible, though more expensive and difficult (see *Handbook* Chs. 11-13).

Once again, we run into the issue of “accuracy” – this time in the form of “appropriateness.” How do we determine when a response is virtue-appropriate and when it is not? We will have to develop a standard for this, perhaps “paradigmaticity,” as suggested earlier, while at the same time being sensitive to the fact we should expect people high in virtue intelligence to recognize and responses that may not be widely recognized as virtue-appropriate.

Step 3: Measuring *dispositionality*

When it comes to measuring virtue intelligence, we are interested in determining its “dispositionality” – i.e., the degree to which a person has virtue-appropriate responses in the presence of virtue-relevant stimuli across a wide variety of types of situations. The greater the degree, the more claim we have to them possessing a global, as opposed to merely local trait. Therefore, dispositionality is a critical feature of virtue intelligence to be able to measure.

We can operationally define “dispositionality” in terms of both *consistency* and *habituality*. “Consistency” is a matter of the frequency with which a person has virtue-appropriate responses in the presence of virtue-relevant stimuli. It could be measured along two distinct dimensions: 1) *depth*, that is, how frequently people have virtue-appropriate responses to the same virtue-relevant stimuli, and 2) *breadth*, that is, how many different virtue-relevant stimuli to which people have virtue-appropriate responses. Together, these measurements indicate how “global” the trait can be considered to be.

“Habituality” is the extent to which the response has become a *dynamically automatic* response to virtue-relevant stimuli. By this, we mean the skill-based automaticity that develops in various forms of expertise. While rapid and largely unconscious, automaticity is also appropriately sensitive to the changing virtue-relevant stimuli of the environment. It allows for virtue-appropriate responses to “flow” from the individual without the need for extended reflection and cognitive evaluation. For example, a parent who sees her child fall from a swing

does not need to stop and reflect about whether and how to help. She “automatically,” that is, without the need for conscious reflection, goes to the child and offers the aid and comfort she deems necessary, perhaps assuaging the child’s fears while tending to a scraped knee. As people develop virtue intelligence, this habituality expands to include such automatic, yet appropriate, responses to many different people in a wide range of circumstances. (It is important to note that just because something is *habitual* does not mean that it is *robust* – indeed, many habitual behavior is actually quite fragile, vulnerable to interruption/extinction with only slight shifts in situational cues. Nonetheless, the habitual system plays a critical part in the development of virtue intelligence).

4. Measuring Virtue(s) – Part II: Full-Blown Virtue

When virtue intelligence become linked with the right sort of motivational structure – that is, when people become motivated to do the right things for the *right reasons* – we would argue that then virtue intelligence becomes full-blown virtue. We take the “right sort of motivational structure” to be the desire to possess, maintain, and develop sensitivity to virtue-relevant stimuli and the ability to recognize and generate virtue-appropriate responses for the sake of developing and possessing virtue *because* one believes:

- 1) that the virtue itself is valuable and important, and
- 2) that it is important to be a person who possesses and has the ability to appropriately exercise that virtue.

This involves the measurement of two additional components: first, the *chronic accessibility of virtue-oriented values/goals* – that is, the goal to be, become, or remain a person who possesses that virtue; and second, the *chronic accessibility of virtue-relevant identity attributes* – that is, attributes that identify one as a person who actually or ideally possesses and can appropriately exercise that virtue.

The chronic accessibility of virtue-oriented value/goals can be operationally defined as the extent to which people *explicitly* and *implicitly* identify virtue-oriented values/goals as being important values/goals in their lives. For example, as an explicit measurement, people could be asked to spontaneously generate a list of their most important values/goals (e.g., to be good parent; to volunteer their time at a homeless shelter) and/or to rank a list of values/goals in terms of their importance. To get at the same thing implicitly, people could be asked to document their

daily activities and how much time they spend on them, and then these activities can be coded for the values/goals reflected by them – the importance of each value/goal being implied by the amount of time spent embodying, pursuing, and reflecting upon it (see Frimer and Oaks, forthcoming).

The chronic accessibility of virtue-relevant identity attributes can be operationally defined as the extent to which people *explicitly* and *implicitly* identify virtue-relevant attributes as being central to their actual and/or ideal identities. As an explicit measurement of this, people could be asked to spontaneously generate a list of the attributes that are most central to who they take themselves to be and/or who they *ideally want* to be, and/or to rank a list of attributes given to them in terms of their centrality to their actual or ideal self-conceptions. So, for example, when asked to describe themselves, we would be interested in whether attributes like “honest,” “loyal,” “kind,” etc. show up on their attribute lists.

This could also be ascertained implicitly by asking people to write about themselves (at either a one-time sitting, or on a daily basis for a period of time) and then these descriptions can be coded for the attributes reflected in their writing – the centrality of each attribute is implied by the frequency with which it shows up in self-descriptions. People could also be asked to write about how they would ideally like to be, in order to assess not only those attributes central to their current self-conception, but those central to the conception of themselves they are striving to achieve.

In all the above examples, the “chronic accessibility” can be determined by the order (those reported early in the list being more chronically accessible than those reported later) and/or the speed (those reported the fastest being more chronically accessible than those reported more slowly) with which virtue-oriented goals/values and virtue-relevant attributes are identified and/or rated. For example, while two individuals might both list “honest” or “generous” as attributes of their identity, if they show up at the top of the list for one person (coming to mind quickly when asked to reflect on she would describe herself) and close to the bottom of the list for the other (following a host of non-virtue attributes like “funny” or “intelligent”), then we would conclude that these virtues are more chronically accessible for the former person than the latter (see *Handbook* Ch. 29).

5. Final thoughts

I hope that the above discussion has provided some food for thought about the challenges associated with the empirical measure of virtue(s) and some strategies for overcoming them. The first step is to provide an operational definition of virtue(s) that gives us empirical “traction”. The definition given here is an attempt to do just that. The second step is to then identify/generate the empirical tools needed to measure the variables identified by your operational definition.

As I hope has been made clear, the measurement of virtue (however you define it) is not going to be a simple project. To adequately assess the development and maintenance of virtue (whether virtue intelligence or full-blown virtue), a multi-pronged, interdisciplinary approach – one that employs a variety of methodological techniques to collect different types of data from different sources – will be necessary. Luckily, many such techniques have already been developed to study other related psychological constructs (such as personality) and new ones are always being created. Thus, a robust empirical program on virtue is possible – it’s just a matter of taking advantage of what is already available, putting our heads together, and getting creative!

I look forward to hearing your thoughts about the efficacy and viability of this approach.

References

Frimer, J. and Oakes, H. Forthcoming. “Spoken words reveal selfish motives: An individual difference approach to moral motivation.” In *Advances in Experimental Moral Psychology: Moral Persons, Moral Groundings, and Measuring Morality*. Eds H. Sarkissian and J. C. Wright (Bloomsbury Publishing).

Hepach, R., A. Vaish, and M. Tomasello 2012 “Young children are intrinsically motivated to see others helped,” *Psychological Science*, 23(9), 967-972. doi:10.1177/0956797612440571.

Mehl, M., and Conner, T.S. 2012 *Handbook of Research Methods for Studying Daily Life*. (The Guilford Press).

- Ch 4: Wilhelm, P., Perrez, M., and Pawlik, K. “Conducting research in daily life: A historical review,” 62-86.
- Ch 5: Conner, T.S. and Lehman, B.J. “Getting started: Launching a study in daily life,” 89-107.
- Ch 7: Kubiak, T. and Krog, K. “Computerized sampling of experiences and behavior,” 124-143.
- Ch 8: Gunthert, K.C. and Wenzel, S.J. “Daily diary methods,” 144-159.

- Ch 9: Moskowitz, D.S. and Sadikaj, G., “Event-contingent recording,” 160-175.
- Ch 10: Mehl, M. and Robbins, M.L. “Naturalistic observation sampling: The electronically activated recorder (EAR),” 176-192.
- Ch 11: Schlotz, W. “Ambulatory psychoneuroendocrinology: Assessing salivary cortisol and other hormones in daily life,” 193-209.
- Ch 12: Wilhelm, F.H., Grossman, P., and Muller, M.I. “Bridging the gap between the laboratory and the real world: Integrative ambulatory psychophysiology,” 210-234.
- Ch 13: Bussmann, J.B.J. and Ebner-Priemer, U.W. “Ambulatory assessment of movement behavior: Methodology, measurement, and application,” 235-250.
- Ch 14: Goodwin, M.S. “Passive telemetric monitoring: Novel methods for real-world behavioral assessment,” 251-266.
- Ch 15: Intille, S.S. “Emerging technology for studying daily life,” 267-282.
- Ch 29: Fleeson, W. and Nofhle, E.E. “Personality research,” 525-538.

Mischel, W. 1973 “Toward a cognitive social learning reconceptualization of personality,” *Psychological Review* 80(4): 252-283.

Mischel, W. and Shoda, Y. 1995 “A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure,” *Psychological Review* 102(2): 246-268.

Narvaez, D. 2001 “Moral text comprehension: Implications for education and research. *Journal of Moral Education*, 3(1), 43-52.

Narvaez, D. 2006 “Integrative Ethical Education.” In Killen, M. and Smetana, J. (eds.) *Handbook of Moral Development* (Erlbaum Press)

Snow, N. 2010 *Virtue as Social Intelligence: An Empirically Grounded Theory* (Routledge Press).