Developmental Cognitive Neuroscience 17 (2016) 131–137

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: http://www.elsevier.com/locate/dcn

Commentary

The audacity of specificity: Moving adolescent developmental neuroscience towards more powerful scientific paradigms and translatable models

Jennifer H. Pfeifer[a,b,*], Nicholas B. Allen[a,b,*]

[a] Department of Psychology, University of Oregon, United States
[b] Center for Translational Neuroscience, Prevention Science Institute, University of Oregon, United States

ARTICLE INFO

Article history:
Available online 23 December 2015

Keywords:
Precision
Translational neuroscience
Models
Dual-systems
Social reorientation

In this issue of *Developmental Cognitive Neuroscience*, two articles revisit a pair of seminal models that have permeated developmental neuroscience research focused on adolescence. Shulman and colleagues (2016) "review, reappraise, and reaffirm" research relevant to dual-systems models of adolescent development, while Nelson and colleagues (2016) "expand and update" their proposal regarding the social reorientation model of adolescence and its underlying neural circuitry. The present commentary aims to complement these efforts with a constructive critique that leads to concrete steps we believe can, and should, be taken to improve our models and maximize cumulative scientific progress in the field. We propose here that for adolescent developmental neuroscience to be truly meaningful – and by this we mean precise enough to not only make accurate and testable research predictions, but also be translatable into prevention, intervention, and policy programs that will significantly improve developmental outcomes for adolescents – we need to refocus our priorities and enable our scientific models to evolve.

Nearly two decades ago, Eysenck (1997) described the range of scientific methodologies appropriate to different stages of psychological research:

> "Science begins with a hunch, acquired through observation and induction, which is clearly a preparadigmatic position. If the hunch seems to work, psychologists construct small-scale hypotheses, for which they seek verification. If such verification is forthcoming in sufficient quantity, the level of theory is reached, and one may then consider the demands of falsification. . . The point between hypothesis and theory would seem to mark the advent of a paradigm. . . when the ordinary business of science takes over, that is, the large-scale testing of deductions from the theory, and the attempt to explain anomalies in terms of the theory's apparent failure" (pp. 1225–1226).

We believe that many of our models in adolescent developmental neuroscience, and the resultant research, are persisting in a *verification* stage, where we primarily focus on supportive evidence that is consistent with the model in question. Indeed, the task is so complex that this is no small achievement, and it is not surprising that the field registers some satisfaction at having models that explain a wide range of phenomena. However, greater progress will be achieved if we progress to a more *falsification* oriented approach, where we (i) rigorously examine and account for inconsistent evidence, and (ii) put our models at strong risk of falsification based on more precise predictions.

A precise prediction that is supported by data provides much stronger evidence for a model than does a less precise prediction.

* Corresponding authors at: Department of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227, United States.Acknowledgements: The authors would like to thank members of the fall 2015 graduate seminar on Adolescence at the University of Oregon who provided valuable input on a draft of this manuscript. JHP was funded by P50 DA035763 and R01 MH107418. NBA was funded by R01 MH107418.

E-mail addresses: jpfeifer@uoregon.edu (J.H. Pfeifer), nallen3@uoregon.edu (N.B. Allen).

http://dx.doi.org/10.1016/j.dcn.2015.12.012

In other words, the degree of logical support for a model is greater given the rarity of the observation absent the theory – what Salmon (1984) has called a "damn strange coincidence" and Meehl (1978) has referred to as a "risky prediction." The importance of this kind of precision goes beyond scientific progress and model building (although that is reason enough). One of the great challenges that our field, along with many others, struggles with is finding strong translational applications of our work – ones that can really have an impact at both the population and individual level (Allen and Dahl, 2015). However, for this admittedly lofty goal to ever be achieved we have to have models that make predictions that are sufficiently precise and robust that we can prescribe public policy and clinical innovations that have real impact.

We are well aware that is much easier to sit on the sidelines and encourage others to do better than it is develop models and put them to the test. We have had our own attempts at theorizing and building models, with varying degrees of success, and we know that it is hard and exacting work (e.g., Allen and Badcock, 2003; Davey et al., 2008; Pfeifer and Peake, 2012). In this respect we would like to make it unambiguous that the work represented in the target articles is a brave and necessary part of the scientific process. The authors and their ideas have our respect and admiration. Also, it is fair to note that we are not proposing an alternative model here, but we do believe that the approach we describe herein is important in addition to, and support of, the process of model building and refinement.

1. A precision approach for adolescent developmental neuroscience: PECANS

In recent years, a number of reviews surveying the evidence regarding influential models of adolescent brain and behavioral development (Pfeifer and Allen, 2012; and others, e.g., Bjork et al., 2012; Crone and Dahl, 2012; Telzer, 2016) have noted sets of findings that do not conform to model predictions. It is tempting for supporters of these models to push these inconsistencies to the side, and for both sides to create a qualitative "box-score tally" of studies that do or do not provide support. For example, Shulman and colleagues (2016) list nine articles that show adolescents engage the striatum to a greater extent than both children and adults, four articles that find the opposite pattern, and four more that fail to demonstrate any age differences (p. 20). As is common in qualitative reviews, this list is then summarized as revealing "considerable evidence" in support of dual-systems models, while a "handful of studies" find the opposite pattern or no differences whatsoever. They then go on to explain this inconsistency in terms of separating out reward anticipation from receipt – three of those nine supporting studies are listed to demonstrate that adolescents engage the striatum "consistently" more than adults during reward receipt. Meanwhile, they suggest there is a tendency to see increased striatum during anticipation only when the cue reliably predicts greater likelihood of reward (referencing two studies that observe this, and two that do not).

Regardless of whether such lists generated by qualitative reviews (including both the Shulman and Nelson papers in this issue, but definitely not limited to them) are intended to be comprehensive or illustrative, we propose that it is well past time for us all to move beyond qualitative box-score tallies, and engage in more precise assessment of how robustly the evidence supports or contradicts the models. A recently published quantitative meta-analysis (Silverman et al., 2015) observed that across 26 studies, adolescents activated a number of regions more than adults during reward processing, including ventral and dorsal striatum, insula, amygdala, anterior and posterior cingulate cortex, orbitofrontal cortex, and frontal poles. This is an essential step away from box-score tallies. We applaud this work and will be the first to note that this kind of an assessment is much stronger and more satisfying, even in terms of *verifying* dual-systems or imbalance models. However, Meehl (1997) suggested that ideally, model assessment will be influenced by *both* the precision of prediction and extent to which observed data are close to those predictions.

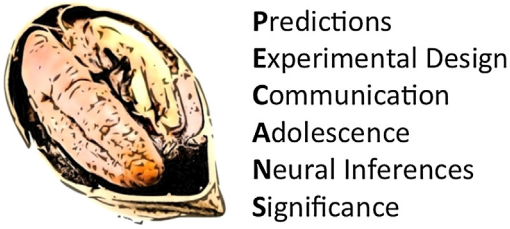

Fig. 1. The PECANS checklist, a mnemonic to enhance precision in adolescent developmental neuroscience.

As we have written about previously (Pfeifer and Allen, 2012), dual-systems and imbalance models in particular are applied so widely that they can be (and are) used to explain the vast majority of adverse outcomes in adolescence (and beyond; e.g., risk taking, anxiety, depression, violence, substance abuse, borderline personality, and indeed adolescence itself). Theories that include description of specific social brain systems, such as those of Nelson and colleagues (2016), have been less susceptible to this type of indiscriminant application – not because they are inherently less vulnerable to it, but mainly because to date there has been significantly less specific research informed by the predictions of these models. We suspect that as research interest in this subfield continues to intensify, the social reorientation model and other similar theories will rapidly face similar challenges in specificity. This is foreshadowed by the efforts of Nelson and colleagues (2016) to expand the social reorientation model beyond adolescence, and illustrate potential tension between deep precision and broad application just as experienced in previous years by dual-systems and related models.

In addition to a lack of precision in the applications of these models, there is a corresponding lack of audacity in the ways we (ourselves included) have tested these models. For a field so concerned about adolescent risk, we all have played it remarkably safe! There are many contributing factors to our collective risk aversion, not least of which include the need in new areas of study to build foundational knowledge bases from which to turn our hunches into hypotheses (as described by Eysenck, 1997), as well as the expense and difficulty of conducting adolescent developmental neuroscience research, particularly longitudinal and ecologically valid assessments. As such, we suggest that the field will be best served by taking the following concrete steps towards "audacious specificity" in the following domains (see Fig. 1): **P**rediction, **E**xperimental Design, **C**ommunication, **A**dolescence (Developmentally Meaningful Indicators), **N**eural Inferences, and **S**ignificance (Ecological and/or Translational Outcomes).

Precision in the first three areas (prediction, experimental design, and communication) is simply good scientific practice across disciplines. Additionally, accomplishing precision in these areas also requires precision in the last three areas (adolescent development, neural inferences, and significant outcomes) that are more particular to the field. In many cases, issues that are presented below with respect to one area actually contribute to others as well. Together, these six keywords provide a guiding mnemonic, following in the tradition of utilizing the PICO checklist to guide evidence-based medicine, which has been credited with improving both research practice and research synthesis through systematic reviews and meta-analyses. (The PICO checklist is a method used to frame and answer a clinical question – the mnemonic stands for

Patient, problem or population, *Intervention*, *Comparison, control or comparator*, and *Outcomes;* Straus et al., 2004). In the following sections, we briefly outline each facet of the PECANS checklist. We then discuss concrete examples from two recent studies intended to test dual-systems models, as well as one of our own previous studies, highlighting ways in which these studies were more or less precise and illustrating how being mindful of these suggestions may improve future research.

2. Prediction

In many ways, one core goal of this commentary is to assert that, in recognition of their broad accomplishments thus far, the models presented in this issue – particularly the dual-systems model of Shulman and colleagues (2016), as well as related models such as imbalance (Casey, 2015) or triadic models (Ernst et al., 2006), are ready to be challenged to make more precise predictions. In other words, we believe that there is sufficient *verification* in the published literature to warrant a reorientation towards *falsification* in this subfield; with respect to the social reorientation model of Nelson and colleagues (2016), we suspect that the rapidly growing research interest in this subfield means a similar level of verification is not far behind. But in order for a theory to be falsifiable, it needs to make precise predictions. We note that terms like imbalance and mismatch are unfortunately quite imprecise in this regard, requiring the null hypothesis to be that they are either precisely matched or in balance during adolescence (which as Meehl (1978, 1990a, 1990b, 1997) was fond of pointing out, is almost certainly never true), or that there is no difference in the degree of imbalance or mismatch among children, adolescents and adults. Additionally, the precision with which we define "adolescents" and the "variable" in question is also critical. In other words, even if individual models make directional predictions (as many do), this precision frequently gets lost in translation via imprecise definitions of adolescent development and/or neural mechanisms (as is discussed in more detail below).

Shulman and colleagues (2016) identify several refinements of the dual-systems model that we endorse, especially the prediction that social contexts (especially peer contexts) may be significant moderators of many developmental effects. There are other changes to the models that we are concerned about, however. One trend that has caught our attention is the gradual expansion of constructs and processes in one of the systems, without a similar inclusivity with respect to neural circuitry. Early conceptualizations of one system in these models focused primarily on motivational processes like reward sensitivity, subserved by ventral striatum (Steinberg, 2008; Casey et al., 2008), despite frequent mention of broader affective and emotional factors (but see the triadic model for an approach that has always strongly differentiated these two; Ernst et al., 2006). As acknowledged by Shulman and colleagues (2016), there is now enhanced emphasis on the contributions of social context and social cognition. But instead of considering how this expands the relevant neural circuitry, which is significantly (Blakemore, 2008) but not entirely (Nelson et al., 2016) distinct from the key regions and networks of interest in these models at present, these moderating and mediating processes are generally subsumed into a "socioemotional," "affective," or "motivational" system. If we are being precise about networks and processes, and want to be more precise in our predictions, we believe these constructs should be unmerged. The original description of the social information processing network (Nelson et al., 2005) attempted to distinguish these constructs in some ways, but the updated review in this issue largely did not revisit what we now know a decade later about the differential functions and relevant regions or networks of interest. In our view, compelling models must differentiate among social, affective, and motivational factors (and their sub-components) and associated neural networks, which each are likely to have specific reciprocal relationships with the lateral frontoparietal circuitry supporting cognitive functions such as attention and control.

3. Experimental design

A second goal is to encourage us all to design stronger tests of these more precise predictions. For example, with respect to dual-systems and related models, if the literature indeed is converging to suggest "a mid-adolescent peak in reward sensitivity, particularly during reward receipt" (Shulman et al., p. 25), we should design studies with the best chance of falsifying this refined and more precise prediction. In that case, we may want to shift towards more targeted samples that, instead of spanning the entire decade of adolescence more or less evenly, emphasize and more carefully sample around mid-adolescence (which highlights how carving up adolescence into more precise and developmentally meaningful phases is critical, to be discussed below). Ideally, we should also include multiple assessments of neural and behavioral responses to reward receipt and reward anticipation in the same study with the same participants. Without careful planning, a given fMRI paradigm may be ill-suited to assess both anticipation and receipt of reward, and directly compare them, particularly in some forms of event-related designs. Researchers may plan studies that assess both constructs equally well, or optimize different tasks and counterbalance the order of administration in the scan protocol. Regardless of the approach chosen, pitting the two directly against each other in the same sample and not across studies would be a more "risky" test of the refined hypothesis put forth by Shulman and colleagues (2016). A similar suggestion could be made with respect to social reorientation model hypotheses. One strong prediction that can be inferred from Nelson and colleagues (2016) is that neural responses will be particularly attuned to social targets according to developmental stage, and this could be tested with careful experimental design both within stages (e.g., directly comparing adolescent reactivity to mothers, peers, romantic partners) and across stages (e.g., assessing whether attunement to each stage-matched social target is evident). However, we suspect the former approach will be much more tractable than the latter, at least in the near future.

An additional benefit of the emphasis on precise experimental design is the attention it draws to construct validity in our tasks. What precise constructs or processes (e.g., neurocognitive, affective, interpersonal) do our tasks assess? For example, in a recent systematic review of fMRI paradigms used to study reward in adolescent versus adults, the authors concluded that the various tasks produced mixed results and that it is "difficult to clearly map the role of specific neural mechanisms onto. . . developmental changes" (p. 988; Richards et al., 2013). Furthermore, they conclude that, "we lack knowledge in the multiple ways that different variables, including subject characteristics, experimental factors, and environmental contexts, can influence the neural systems underlying reward-related behavior" (p. 988). This is just one example of an area where a variety of tasks that putatively tap into a coherent underlying construct (i.e., functioning of the reward system) can show strongly dissociated effects across different studies of that construct. Indeed, rarely (if ever) are more than one of these tasks used in a single neuroimaging study, so we actually have little information on the degree to which performance on these tasks covaries. Nevertheless, we often interpret these tasks as if they do all tap one construct. This highlights the importance of very careful attention to construct validity in our experimental tasks. In fact, it would serve us well to return to the rich and venerable literature on construct validity in the area of psychometrics. Cronbach and Meehl

(1955) originally presented the concept of a "nomological network" as a method of evaluating the construct validity of psychological tests, and proposed that in order to demonstrate that your measure has construct validity researchers must develop a nomological network that consists of both a theoretical framework and an empirical framework regarding methods of measurement, and specification of the links among and between these two frameworks. Perhaps because of the expense and complexity of neuroimaging compared to psychometric measures, very few imaging tasks have been rigorously subjected to this kind of evaluation, and yet the absence of such an analysis can be a critical barrier to progress if studies that are putatively investigating a particular construct are not in fact doing so.

4. Communication

Assuming we have been precise in our predictions and experimental design, we might still stumble at the brink of success if we are imprecise in communicating our findings. Below we identify needs for precision in two distinct communication streams.

The first area of precise communication is internal to the field. Authors, reviewers, editors, and journals should provide and demand much more precise reporting in papers, and a suite of associated practices which will facilitate cumulative progress. One step in particular we now feel strongly about is the need to report main effects, by group, at the whole-brain level. It is particularly common to focus on interactions (e.g., task versus control for adolescents versus adults), but that hampers our ability to combine information across studies (e.g., in meta-analyses). For example, in the Silverman et al. (2015) meta-analysis, 62 studies were identified in a literature search addressing adolescent reward sensitivity, but according to the methods only 26 could be included in the meta-analysis, and 21 were excluded specifically because the manuscript did not provide sufficient detail or relied solely upon ROI analysis.

We also feel that more precision in the labeling of regions is critical. For example, there are meaningful differences between components of the basal ganglia – caudate, nucleus accumbens, putamen – and the way in which we often confuse and/or equate these certainly undermines our precision. Other regions that are less well-defined structurally, such as the temporal–parietal junction (TPJ), are sometimes labeled as such with too large a degree of latitude, especially given that different subregions exhibit different anatomical and functional connectivity (for example with respect to the TPJ, see Carter et al., 2012). Still other often-used regional labels – chief among them "prefrontal cortex" – are so imprecise as to be largely uninformative. Even simple lateral and medial distinctions in PFC provide only a fraction more precision. In dual-systems and related models, we propose that a great deal more precision must be achieved especially with respect to subregions of the PFC. An advance in the dual-systems review by Shulman and colleagues (2016) is its assignment of medial PFC and orbitofrontal cortex with the striatum to the "socioemotional system," distinct from lateral PFC, anterior cingulate cortex, and lateral parietal cortex in the "control system." However, this quickly becomes complicated since ventromedial PFC has consistently been implicated in regulation networks (Etkin et al., 2011), and dorsomedial PFC is strongly associated with social cognition circuitry (Eickhoff et al., 2014; Bzdok et al., 2013). We believe a key challenge for the development of these models is not only to differentiate social cognitive processes but also to meaningfully integrate them as a key feature of adolescent development, especially in terms of their specific neurobiological substrates. This is an issue that Shulman and colleagues do not address explicitly, although social cognitive processes and networks are clearly a central concern in the social reorientation model by Nelson and colleagues.

Interrogating ROIs as the sole reported analytical approach can be problematic even if defined with a high degree of precision. A targeted ROI analysis may be seen to some extent as a more risky test of one's theory, presuming that one has specified directional effects. However, this neglects the exponentially increasing emphasis in the field on networks and circuits (Pfeifer and Allen, 2012; Casey, 2015). Additionally, reporting only ROI analyses limits the contribution that rich whole-brain datasets may provide, particularly when combined with many other studies.

Some conservative data thresholding procedures bias us towards detecting more circumscribed regions with high magnitudes. This is an important practice that supports making more precise regional inferences, unlike the use of lower magnitude thresholds that produce low spatial sensitivity (Woo et al., 2014), but there may be reliable peaks at lower thresholds that will only be identified using whole-brain big data approaches. Furthermore, inclusion of whole-brain data allows for direct comparison of effects observed in predicted and non-predicted regions, once again improving the testing of the specificity of the predicted effects.

Finally, we believe that we all could do more to ensure that conclusions are communicated accurately and with the right degree of circumspection. This is true within our peer-reviewed papers, in part to prevent misunderstandings for readers who do not have as much "neuroimaging literacy". It is perhaps even more true of, and important to, our communications with the public through various means (such as media interviews, popular press books, or blog posts). Simply admitting that our models are "oversimplified" and "heuristic devices" prior to describing an oversimplified heuristic model is a great disservice, if that model is still likely to be taken as a scientific fact by lay audiences despite such nuanced caveats. It also undermines the refinements that have evolved as further studies reveal the need to modify theory.

5. Adolescence (developmentally meaningful indicators)

As noted several times above, general aspects of precision can be especially complex in the field due to our target population. Adolescence is a particularly long phase of development, spanning nearly a decade or more depending on whether one includes emerging adulthood and how one defines pubertal onset. This is illustrated by the fact that we sometimes carve up adolescence into phases (early, middle, late, and even pre; as noted by Shulman et al. on pp. 7–8), but also often treat adolescence as if it is developmentally homogeneous. For example, the recently released meta-analysis on adolescent reward processing (Silverman et al., 2015) was only able to report that adolescents between 10 and 19 years of age (mean age does not appear to be provided) exhibit greater activation during reward processing (collapsing across anticipation versus receipt, and positive versus negative outcomes) than do adults. Unfortunately, there is a lack of agreement in the field about the relevant demarcations across adolescence. In addition, even when these phases are investigated more directly, we often slip back into the shorthand of referring to "adolescents" in our conclusions, abstracts, and other forms of communication.

Another issue is that chronological age is typically used to define an adolescent's developmental stage or maturation. However age is always, at best, an imperfect measure of maturation, a fact that has been recognized within developmental research for some time (Wohlwill, 1970). This is especially true during adolescence, when developmental processes such as puberty and physical maturation can vary markedly between individuals of the same chronological age – even more so than is typical at other stages of life (Mirwald et al., 2002; Ellis, 2004). As the influence of pubertal development wanes in middle to late adolescence, measures of biological processes may need to be supplemented with description of social

role transitions and other contextually meaningful sociocultural indicators of maturation (e.g., entering the workforce or higher education, living independently, acquiring a driver's license, forming stable long-term romantic relationships). Such individual variation in the timing and tempo of maturation is also likely to apply to brain development itself. As such, in order to aid precision we need to develop ways of quantifying maturation that directly assess neurodevelopment (e.g., Dosenbach et al., 2010, 2013).

There is an astounding lack of empirical detail about when imbalances or mismatches are detected within this extended developmental stage. As Shulman and colleagues note, extensive behavioral research provides a fairly strong expectation that the population average peak response in ventral striatum – if indeed it does represent a neural basis for sensation-seeking – should fall specifically around age 15 (Cauffman et al., 2010; Steinberg et al., 2009). However, they conclude "the neuroimaging literature does not allow for a precise estimation of age of peak striatal response" (p. 25). In fact, the field likely does possess the data to answer this question, it just needs to be examined more precisely and collaboratively. For example, in parallel to the aforementioned meta-analysis by Silverman et al. (2015), we have been working on a different kind of quantitative meta-analysis of these studies. A primary goal of ours has been to assemble the studies in a manner that allows us to actually ask *when* (chronologically, as we do not have enough published studies assessing pubertal development to ask the comparable question) across the decade of adolescence this peak appears to occur (for more information, see http://dsn.uoregon.edu/research/arcs). This goal will only be achievable with widespread participation across the field in reporting and sharing whole-brain group-level summary statistics (Poldrack and Yarkoni, in press).

Another important issue is to contrast putative adolescent specific phenomena with a wider range of alternative non-adolescent developmental phases. In particular, studies that show differences between adolescent and adult groups often conclude that the differences must represent phenomena that are specific to adolescence, and then seek to explain these phenomena in terms of adolescent-emergent or adolescent-specific neurodevelopmental processes. But the comparison with pre-adolescent child groups is less often made, and when it is made, does not always confirm the conclusion of adolescent specificity. For example, an important recent meta-analysis of risky decision-making tasks across development found evidence for greater risk-taking in adolescents versus adults, but not adolescents versus children (Defoe et al., 2015). Data like these need to be addressed by models that presume increased risky decision-making is an adolescent-specific phenomenon.

6. Neural inferences

One area where precision is of concern across the cognitive neurosciences is in the specificity of our neural inferences. As we previously suggested (Pfeifer and Allen, 2012), and as nicely summarized by Casey (2015) when considering several models related to dual-systems: "new findings have moved the field away from simplistic one-to-one mappings of the ventral striatum and amygdala to reward and avoidant behaviors, toward the recognition of distinct computational roles they each play in learning that influence adaptive action in response to both positive and negative outcomes" (p. 299). Here Casey (2015) is identifying the problem of one-to-many and many-to-one (or many-to-many) relations in neuropsychological inference. This issue has been extensively explored by Cacioppo and Tassinary (1990) with respect to an allied problem – psychophysiological inference. They argue that the strongest form of inference is associated with being able to achieve a one-to-one mapping between the psychological construct of interest and the physiological process being measured – however, such strong inferences are the exception rather than the rule. For example, many psychological states can be associated with increased amygdala activity (fear, pleasure, surprise, uncertainty), and moreover, a given psychological state, such as depression, can be associated with changes in multiple brain regions (e.g., amygdala, hippocampus, subgenual anterior cingulate cortex (ACC)). However, if we want to use a neurobiological marker (e.g., activity in the nucleus accumbens) to infer a psychological state (e.g., reward anticipation) then we need to work towards as close to a one-to-one mapping between these domains as we can achieve. (This of course is not to suggest that a single psychological function is ever likely to be purely associated with the activity of a single brain region in terms of a comprehensive model of the neurobiological substrate of that function. Rather, the strongest regional neuropsychological inferences will be possible when activity in the region is a highly sensitive and specific *marker* of the psychological processes of interest.) Developmental questions also dictate that we map patterns of change in neurobiological systems to patterns of change in behavior across time, further specifying the nature of the one-to-one mapping needed for strong neuropsychological inference.

Cacioppo and Tassinary (1990) do propose a number of approaches to the issue, many of which are applicable to developmental neuroscience. For example they suggest that researchers redefine what constitutes an element – i.e., the psychological or neurodevelopmental variable of interest. In developmental neuroscience this will probably entail moving from regional to network based analyses and/or incorporating "form" (i.e., configural or temporal information) into our definition of critical dependent and independent neurobiological variables. In other words, we need to enhance the precision of our variables such that there is stronger hypothesized specificity (e.g., refrain from talking about PFC as if it is functionally homogeneous), and move away from many-to-one and many-to-many relationships that only support weaker inferences, by precisely characterizing circuits and networks (spatially, temporally, developmentally). In particular we should relinquish our reliance on single ROIs (both structurally and functionally) given the low likelihood that the activity or volume of a given structure will map onto a behavioral phenotype in a one-to-one way. Furthermore, we still only poorly understand the relationship between brain structure and brain function, and should be very careful to avoid assuming that the patterns observed will be equivalent.

7. Significance (ecological and/or translational outcomes)

Perhaps the most important issue with respect to the translation of our science is to ensure that we are also measuring the ultimate ecological outcome of interest (i.e., that thing that can answer the "Who cares?" or "Why should we spend public money on this?" questions). If you think the processes you are studying are implicated in mental health, then measure mental health. If you think they are relevant to risk taking, measure *real life* risk taking. If you think they can affect academic performance, measure academic performance. This is a question of the ultimate construct validity of our work. For example, in a recent review of the neuroscience of adolescent decision making Hartley and Somerville (2015) note that "many tasks employed in neuroeconomic studies fail to capture key qualitative features of naturalistic choice contexts, which may diminish their validity for understanding real-world decision-making" (p. 109). In other words, in some cases we may be spending a lot of time, energy, and money studying the neurodevelopmental correlates of laboratory tasks that may have little or no relationship to phenomena of actual interest if we don't actually take that

critical step and measure the relationship between our laboratory measures and ecologically important, real world behaviors. In developmental science, these processes also need to be measured across development, in order to ensure that when we think *developmental change* in a neural or laboratory measure of interest maps on to *developmental change* in an ecological variable of interest, it does so.

8. Using the PECANS checklist

In this section, we briefly consider three studies using the PECANS checklist – one of our own, and two additional recent studies from other labs, spanning three different methodologies (fMRI, sMRI, and rs-fcMRI). It should be noted that we hold these latter two studies in very high esteem, and in no way should our discussion of them be construed as a negative assessment of their contribution to the literature. In fact, we think these latter two studies provide some exceptional examples of precision in prediction, experimental design, communication, adolescent development, neural inferences, and significant outcomes, and as such reveal the strengths of research that ticks many boxes on the PECANS checklist. We also subject our own work to scrutiny, as well as explore theirs, for ways in which the studies could have been more precise.

Several years ago (Pfeifer et al., 2011), one of us (JHP) examined longitudinal change in reactivity to emotional facial expressions during the transition from late childhood to early adolescence. The study took a relatively precise look at a specific period of adolescent development, facilitated by the narrow range of ages tested (9.5–10.6 and 12.5–13.6 years at waves one and two, respectively). Yet even more developmental precision could have been possible in that domain by examining puberty instead of chronological age (see Moore et al., 2012, for a complementary analysis of these data taking just such an approach). The study also aimed to link neural changes over three years within subjects to significant outcomes, namely changes in self-reported resistance to peer influence and risky behavior, although real-world measurement of these constructs would have provided even more precision. Some aspects of experimental design were a "mixed bag", as the paradigm selected (passive viewing of emotions) was arguably better optimized to assess functional changes in the amygdala than the ventral striatum. The precision of neural inferences was bolstered by functional connectivity analyses to assess whether responses in ventral striatum were down-regulating the amygdala, but these results were not especially robust. Finally, although the study conducted and reported whole-brain analyses, more precision could have been achieved with respect to communication, as whole-brain results from each wave of the longitudinal study were not reported independently.

An exciting recent structural MRI paper (Mills et al., 2014) tested for the mismatch proposed by dual-systems models in another longitudinal dataset where each participant contributed a minimum of 3 datapoints over 10–20 years. The predictions were quite precise and risky, namely that subcortical structures involved in processing affect or reward would develop earlier than cortical structures involved in cognitive control, and the degree of mismatch would relate to risky behavior. However, because this was a secondary data analysis in which significant outcomes were assessed retrospectively, the experimental design did not allow for strong tests of the relationship between structural mismatch and risky behavior, as acknowledged by the authors. Communication of results was extremely thorough, including multiple visualizations of both group and individual level data. But the structural definition of PFC was less precise, limiting neural inferences and communication (as acknowledged by the authors). More precise communication may also have been achieved in the abstract of this study, in that there was considerably more evidence of structural mismatch with respect to amygdala and PFC (26/33 individuals) than nucleus accumbens and PFC (17/33 individuals), and the authors state in their discussion that the group level data did not provide evidence of a clear structural mismatch between the nucleus accumbens and PFC during adolescence, but the abstract stated "the majority of individuals in our sample showed relatively earlier maturation in the amygdala and/or nucleus accumbens compared to the PFC, providing evidence for a mismatch in the timing of structural maturation between these structures" (p. 147).

Finally, an impressive recent cross-sectional study of 269 8–25 year-olds (van Duijvenvoorde et al., in press) explored resting-state functional connectivity patterns from the perspective of dual-systems models. The authors selected a more precise dorsolateral PFC seed ROI based on an fMRI task of cognitive feedback-learning collected concurrently (although the affective reward-processing task presumably intended to define nucleus accumbens resulted in too widespread activation, so it was defined structurally). They also complemented this seed-based approach with data-driven independent components analysis to identify whole-brain networks. Their predictions were directional, which provided some precision, but in the case of whole-brain analysis of functional connectivity this was relatively underspecified (as any brain region whatsoever could have provided the increasing intrinsic connectivity with dorsolateral PFC, or adolescent-specific pattern of connectivity with the nucleus accumbens). An even more precise and risky prediction might specify age (or puberty) related changes in connectivity directly between dorsolateral PFC and nucleus accumbens, but no such relationship was hypothesized or observed. Although the reporting was generally exceptionally comprehensive, even more precision could be achieved in some communicative respects, such as the labeling of a subgenual ACC/posterior medial OFC region (acknowledged as such in the table) as ventromedial PFC throughout the text; in our opinion, a more precise structural label and definition is always preferable to a less precise one. The authors found that linear increases between dorsolateral PFC and thalamus mediated age-related changes in a meaningful adaptive outcome (learning rate), but perhaps more conservative assessment of these results with respect to their consistency with dual-systems models is warranted, given the historical lack of attention to the thalamus in these approaches. We also found the adolescent-specific peak in intrinsic functional connectivity within a cognitive-control network from the data-driven ICA approach, as well as the lack of an adolescent-specific peak in intrinsic functional connectivity with nucleus accumbens (largely monotonic increases were observed, except with hippocampus), both intriguing and unexpected if one took a traditional dual-systems model perspective. Indeed, the authors note that their hypotheses were deductions from dual-systems models, which have not previously made specific predictions about resting-state functional connectivity.

We consider these latter two studies to be outstanding exemplars of research that is raising the standards in terms of operationalizing and testing models and adolescent neurodevelopment. The PECANS checklist helps to clarify the specific ways in which studies like these are moving the field forward. However, the PECANS checklist can also be used to identify those areas where even high quality studies can do a better job of providing the precision required for strong inferences and impactful translational outcomes.

9. Conclusion

The origin of the word *pecan* traces back to an Algonquin word referring to "nuts that require a stone to crack." But hammering away at a pecan shell with a blunt, round, heavy stone mainly

results in squashing the nut. We need thin, sharp stones – precise tools for extracting the meat of the nut. To further advance our science of adolescent development, which to this point has significantly benefited from the seminal dual-systems and social reorientation models (Shulman et al., 2016; Nelson et al., 2016), we propose that "audacious specificity" as concretized by the PECANS checklist will help us all work together to conduct even more cumulatively meaningful work. Ideally, the PECANS checklist would be consulted prior to initiating a study, allowing researchers to enhance precision from the outset, as well as revisited during data analysis and communication phases. We may be nuts (or just naïve), but it is our sincere hope that providing the PECANS mnemonic as a concrete, guiding rubric will facilitate more precise and "risky" research that can refine our models and enable transformative translational work to improve and protect adolescents as they develop.

Acknowledgements

The authors would like to thank members of the fall 2015 graduate seminar on Adolescence at the University of Oregon who provided valuable input on a draft of this manuscript. J.H.P. was funded by P50 DA035763 and R01 MH107418. N.B.A. was funded by R01 MH107418.

References

Allen, N.B., Badcock, P.B., 2003. The social risk hypothesis of depressed mood: evolutionary, psychosocial, and neurobiological perspectives. Psychol. Bull. 129 (6), 887–913.
Allen, N.B., Dahl, R.E., 2015. Multi-level models of internalizing disorders and translational developmental science: seeking etiological insights that can inform early intervention strategies. J. Abnorm. Child Psychol. 43 (5), 875–883.
Bjork, J.M., Lynne-Landsman, S.D., Sirocco, K., Boyce, C.A., 2012. Brain maturation and risky behavior: the promise and the challenges of neuroimaging-based accounts. Child Dev. Perspect. 6 (4), 385–391.
Blakemore, S.J., 2008. The social brain in adolescence. Nat. Rev. Neurosci. 9 (4), 267–277.
Bzdok, D., Langner, R., Schilbach, L., Engemann, D.A., Laird, A.R., Fox, P.T., Eickhoff, S.B., 2013. Segregation of the human medial prefrontal cortex in social cognition. Front. Hum. Neurosci. 7.
Cacioppo, J.T., Tassinary, L.G., 1990. Inferring psychological significance from physiological signals. Am. Psychol. 45 (1), 16.
Carter, R.M., Bowling, D.L., Reeck, C., Huettel, S.A., 2012. A distinct role of the temporal–parietal junction in predicting socially guided decisions. Science 37 (6090), 109–111.
Casey, B.J., 2015. Beyond simple models of self-control to circuit-based accounts of adolescent behavior. Annu. Rev. Psychol. 66, 295–319.
Casey, B.J., Jones, R.M., Hare, T.A., 2008. The adolescent brain. Ann. N. Y. Acad. Sci. 1124 (1), 111–126.
Cauffman, E., Shulman, E.P., Steinberg, L., Claus, E., Banich, M.T., Graham, S., Woolard, J., 2010. Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. Dev. Psychol. 46 (1), 193.
Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. Psychol. Bull. 52 (4), 281.
Crone, E.A., Dahl, R.E., 2012. Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat. Rev. Neurosci. 13 (9), 636–650.
Davey, C.G., Yücel, M., Allen, N.B., 2008. The emergence of depression in adolescence: development of the prefrontal cortex and the representation of reward. Neurosci. Biobehav. Rev. 32 (1), 1–19.
Defoe, I.N., Dubas, J.S., Figner, B., van Aken, M.A., 2015. A meta-analysis on age differences in risky decision making: adolescents versus children and adults. Psychol. Bull. 141 (1), 48.
Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Schlaggar, B.L., 2010. Prediction of individual brain maturity using fMRI. Science 329 (5997), 1358–1361.
Dosenbach, N.U., Petersen, S.E., Schlaggar, B.L., 2013. The teenage brain: functional connectivity. Curr. Dir. Psychol. Sci. 22 (2), 101–107.
Eickhoff, S.B., Laird, A.R., Fox, P.T., Bzdok, D., Hensel, L., 2014. Functional segregation of the human dorsomedial prefrontal cortex. Cereb. Cortex, bhu250.
Ellis, B.J., 2004. Timing of pubertal maturation in girls: an integrated life history approach. Psychol. Bull. 130 (6), 920–958.
Ernst, M., Pine, D.S., Hardin, M., 2006. Triadic model of the neurobiology of motivated behavior in adolescence. Psychol. Med. 36 (3), 299–312.
Etkin, A., Egner, T., Kalisch, R., 2011. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn. Sci. 15 (2), 85–93.
Eysenck, H.J., 1997. Personality and experimental psychology: the unification of psychology and the possibility of a paradigm. J. Personal. Soc. Psychol. 73 (6), 1224.
Hartley, C.A., Somerville, L.H., 2015. The neuroscience of adolescent decision-making. Curr. Opin. Behav. Sci. 5, 108–115.
Meehl, P.E., 1990a. Appraising and amending theories: the strategy of Lakatosian defense and two principles that warrant it. Psychol. Inq. 1 (2), 108–141.
Meehl, P.E., 1990b. Why summaries of research on psychological theories are often uninterpretable. Psychol. Rep. 66 (1), 195–244.
Meehl, P.E., 1997. The problem is epistemology, not statistics: replace significance tests by confidence intervals and quantify accuracy of risky numerical predictions. In: Harlow, L.L., Mulaik, S.A., Steiger, J.H. (Eds.), What If There Were No Significance Tests? Erlbaum, Mahwah, NJ, pp. 393–425.
Meehl, P.E., 1978. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. J. Consult. Clin. Psychol. 46 (4), 806.
Mills, K.L., Goddings, A.L., Clasen, L.S., Giedd, J.N., Blakemore, S.J., 2014. The developmental mismatch in structural brain maturation during adolescence. Dev. Neurosci. 36 (3–4), 147–160.
Mirwald, R.L., Baxter-Jones, A.D.G., Bailey, D.A., Beunen, G.P., 2002. An assessment of maturity from anthropometric measurements. Med. Sci. Sports Exerc. 34 (4), 689–694.
Moore, W.E., Pfeifer, J.H., Masten, C.L., Mazziotta, J.C., Iacoboni, M., Dapretto, M., 2012. Facing puberty: associations between pubertal development and neural responses to affective facial displays. Soc. Cogn. Affect. Neurosci. 7 (1), 35–43.
Nelson, E.E., Leibenluft, E., McClure, E., Pine, D.S., 2005. The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. Psychol. Med. 35 (02), 163–174.
Pfeifer, J.H., Allen, N.B., 2012. Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. Trends Cogn. Sci. 16 (6), 322–329.
Pfeifer, J.H., Peake, S.J., 2012. Self-development: integrating cognitive, socioemotional, and neuroimaging perspectives. Dev. Cogn. Neurosci. 2 (1), 55–69.
Pfeifer, J.H., Masten, C.L., Moore, W.E., Oswald, T.M., Mazziotta, J.C., Iacoboni, M., Dapretto, M., 2011. Entering adolescence: resistance to peer influence, risky behavior, and neural changes in emotion reactivity. Neuron 69 (5), 1029–1036.
Poldrack, R.A., Yarkoni, T., 2015. From brain maps to cognitive ontologies: informatics and the search for mental structure. Annu. Rev. Psychol., http://dx.doi.org/10.1146/annurev-psych-122414-033729 (in press).
Richards, J.M., Plate, R.C., Ernst, M., 2013. A systematic review of fMRI reward paradigms used in studies of adolescents vs. adults: the impact of task design and implications for understanding neurodevelopment. Neurosci. Biobehav. Rev. 37 (5), 976–991.
Salmon, W.C., 1984. Scientific Evidence and the Causal Stricture of the World. Princeton University Press, Princeton, NJ.
Silverman, M.H., Jedd, K., Luciana, M., 2015. Neural networks involved in adolescent reward processing: an activation likelihood estimation meta-analysis of functional neuroimaging studies. NeuroImage 122, 427–439.
Steinberg, L., 2008. A social neuroscience perspective on adolescent risk-taking. Dev. Rev. 28 (1), 78–106.
Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., Banich, M., 2009. Age differences in future orientation and delay discounting. Child Dev. 80 (1), 28–44.
Straus, S.E., Green, M.L., Bell, D.S., Badgett, R., Davis, D., Gerrity, M., et al., Society of General Internal Medicine Evidence-Based Medicine Task Force, 2004. Evaluating the teaching of evidence based medicine: conceptual framework. BMJ 329 (7473), 1029.
Telzer, E.H., 2016. Dopaminergic reward sensitivity can promote adolescent health: a new perspective on the mechanism of ventral striatum activation. Dev. Cogn. Neurosci. 17, 57–67.
Van Duijvenvoorde, A.C.K., Achterberg, M., Braams, B.R., Peters, S., Crone, E.A., 2015. Testing a dual-systems model of adolescent brain development using resting-state connectivity analyses. NeuroImage (in press).
Wohlwill, J.F., 1970. The age variable in psychological research. Psychol. Rev. 77 (1), 49–64.
Woo, C.W., Krishnan, A., Wager, T.D., 2014. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. NeuroImage 91, 412–419.