The risk matrix
Brian Knutson¹ and Scott A Huettel²

Neuroimaging methods (e.g., functional magnetic resonance imaging or FMRI) can now resolve momentary changes in deep brain activity that not only correlate with but also predict risky choice. Accumulating evidence beginning from financial choice studies but extending into other domains indicates that risk assessment recruits activity in multiple core components which differentially promote (e.g., ventral striatum) versus inhibit (e.g., anterior insula) risky choice. Further, frontal control circuits may modulate the influence of these core components on risky choice. These findings point toward an emerging consensus about a ‘risk matrix’ whose components unite previously disparate literatures related to anticipation of reward versus pain and whose measurement can improve the prediction of risky choice.

Addresses
¹ Psychology and Neuroscience, Stanford University, United States
² Psychology and Neurosience, Duke University, United States

Corresponding author: Knutson, Brian (knutson@psych.stanford.edu)

Introduction
In the movie The Matrix, mentor Morpheus offers protagonist Neo a choice between two pills: ‘You take the blue pill — the story ends, you wake up in your bed and believe whatever you want to believe. You take the red pill — you stay in Wonderland, and I show you how deep the rabbit hole goes.’ At this point, Neo faces a classic risky choice, in which a certain option promises minimal change from the current status (i.e., small gains but also small losses), but another uncertain option offers potentially larger gains at the cost of potentially larger losses (Figure 1).

While risk can be defined in many ways [2], most risky choices require individuals to balance uncertain but significant gains against losses [3]. Risky choice predates economic and financial institutions, pervading the foraging, survival, and relational challenges faced by our forebears and other species [4]. Thus, core neural systems that support risky choice might reside not only in the most recently evolved regions of human prefrontal cortex [5], but also in more ancient and deeper affective and motivational circuits that have been conserved across evolutionary history [6].

Visualizing activity in neural systems that support risky choice thus requires methods that can resolve rapid changes in the dynamic activity of small, deep, and conserved brain regions moments before choice. The development of functional magnetic resonance imaging (FMRI) in the early 1990s provided a method for noninvasively measuring dynamic subcortical activity [7]. Since then, a rising tide of FMRI studies has identified neural correlates of risk assessment and choice — beginning with financial risk taking, but subsequently extending into other domains [3, 8, 9].

Below, we review the brief history of FMRI research on risky choice before highlighting recent findings. Borrowing from and extending the notion of a ‘pain matrix,’ [10, 11] we suggest that this work may support the existence of a ‘risk matrix,’ involving multiple interacting components that not only assess uncertain gains and losses but also shape choices across diverse risk scenarios. We conclude by considering implications for theory, research, and application.

Defining risk
Although financial incentives provide a convenient laboratory tool for eliciting risky choices, basic theoretical accounts differ on how to define risk, which has implications for experimental design and analysis. According to one of the oldest and simplest economic theories, a gamble should be chosen based on its expected value, which can be estimated as the sum of the magnitudes multiplied by the probabilities of all possible outcomes $[EV = \sum(\nu(x)\cdot p(x))]$. In a subsequent modification, a gamble’s expected utility (rather than value) is instead estimated as sum of the $utilities$ multiplied by the probabilities of all possible outcomes $[EU = \sum(u(x)\cdot p(x))]$. While objective value ($\nu$) is linear, subjective utility ($u$) instead can curve (e.g., $u = v^b$, where $b < 1$ implies concavity). Greater downward (or concave) curvature of the utility function thus implies less risk seeking [12]. Since expected utility implies that risk attitude is a feature of the utility function’s curvature, a gamble’s expected utility implicitly integrates both reward and risk attitude.

In contrast to this traditional economic account, finance theories have historically distinguished expected risk
from expected reward. In mean-variance theories for instance, the value of a risky investment is estimated as its potential reward (e.g., the mean of past returns) minus its potential risk (e.g., the variance of past returns), with the risk term weighted by a risk sensitivity coefficient \( [EV(x) = \bar{v}(x) - h^*r(x)] \) [13]. While both economic and financial formulations of risk were developed to describe choice rather than mechanisms that generate choice, the basic assumption of single versus multiple terms holds important implications for neuroeconomic experiments. For instance, modeling reward and risk as a single term will probably reveal unitary correlates, whereas modeling them separately increases the likelihood of detecting multiple correlates.

By resolving changing brain activity on the order of seconds, FMRI afforded investigators the opportunity not only to visualize correlates of choice but also to distinguish intermediate choice processes from those related to earlier sensory input and later motor output. Although risk assessment might recruit many brain processes (e.g., sensory attention, gain anticipation, loss anticipation, value integration, etc.), only some of these might influence or predict subsequent risky choice (Figure 2). Further, although multiple brain processes might predict risky choice (e.g., conflict resolution, action selection, motor readiness, etc.), only some of these might respond to immediate input related to the choice at hand (Figure 2). Investigators face the challenge of disentangling these processes to determine which mediate the path from input to output.

**Neural correlates of risk assessment**

Increased risk is only worth considering in the face of potentially increased reward. Setting the stage for research on risk assessment, neuroimaging studies at the beginning of the twenty-first century initially focused on reward assessment. Building from animal research [14,15], early and subsequent explorations demonstrated that anticipation of uncertain reward increased FMRI activity (or blood oxygen level dependent signal) in projection targets of midbrain dopamine neurons—which included the ventral striatum (VS, particularly including the nucleus accumbens or NAcc subregion) and medial prefrontal cortex (MPFC) [16–19]. These findings raised the question of whether this same circuit might process expected risk, as expected utility theory might imply.

Subsequent findings, however, called such a single component account into question. The anticipation of potential losses, for example, appeared not to activate circuits implicated in reward assessment like the ventral striatum as powerfully as the anterior insular cortex [16,17,20]. Researchers further demonstrated that anterior insula activity scaled with increases in anticipated risk [21] as well as changes in risk assessment [22]. Since the researchers defined risk with respect to the variance of past outcomes, their findings appeared to support the notion that risk assessment recruited more than one neural circuit. Indeed, similar to the influence of anticipated reward on ventral striatal activity, reviews suggest that the influence of anticipated risk on anterior insula activity is one of the most consistent findings in the neuroeconomic literature [3,8].

Yet, other findings seemed to challenge a multiple component account of risk assessment. For instance, findings from an influential study of loss aversion using mixed gambles implied that activity in a broad swath of regions (including the ventral striatum, MPFC, and anterior insula) increased with anticipated reward and decreased with anticipated loss [23], consistent with the notion that a single system represents both reward and risk. However, a subsequent well-powered replication featuring mixed gambles with symmetric gains and losses instead found that while anticipation of the gain component of gambles activated mesolimbic projection regions including the ventral striatum, anticipation of the loss component of gambles instead activated the medial insula [24**]. Notably, neither of these studies focused analyses only on risk assessment, instead modeling both assessment and choice together, which may have spread observed activity to connected posterior and dorsal regions (e.g., into the
dorsal striatum and premotor cortex) implicated in motor responses [25].

Together, these findings implicate multiple neural components in risk assessment. Further, they hint that recruitment of these components might also predict risky choice, but possibly in different ways. For instance, extrapolating from theories suggesting that affect can influence risky choice [26,27], an 'anticipatory affect' account might predict that although risk assessment should increase both ventral striatal and anterior insular activity, increased ventral striatal activity should promote approach toward risk (along with positive arousal), but increased anterior insula activity should instead promote avoidance of risk (along with general or negative arousal) [16]. Indeed, consistent with these predictions, meta-analyses suggest that: firstly, while risky financial options elicit both ventral striatal and anterior insular activity, the reward component most powerfully activates ventral striatum, while the risk component most powerfully activates the anterior insula; secondly, while ventral striatal activity implies a high likelihood that an article includes the term 'reward,' anterior insula activity instead implies a high likelihood that an article includes the term 'pain;' and finally resting state functional connectivity between the ventral striatum and anterior insula is localized, significant, and negative (Figure 3).

**Neural prediction of risky choice**

Although both ventral striatal and anterior insula components are activated during risk assessment, mounting evidence suggests that before risky choice, anterior insula activity may serve as a motivational counterpoint to ventral striatal activity. In an initial study using gambles, although risk assessment elevated insula activity, insular activity correlated with individual differences in risk aversion rather than risk seeking [20]. Subsequent research using an investing task further established that while ventral striatal activity predicted optimal as well as excessive risk seeking within subjects on a trial-to-trial basis, anterior insula activity instead predicted optimal as well as excessive risk aversion [29], a pattern supported by later findings [30–32].

Converging research on neural correlates of risk assessment and neural predictors of risky choice therefore implies the existence of a ‘risk matrix’ consisting of multiple components that activate in response to a risky option (involving a mix of uncertain but significant gains and losses), but then promote subsequent approach toward or avoidance of that option. These findings further point toward candidate components for the approach circuit in the ventral striatum (particularly the nucleus accumbens) and for the avoidance circuit in the anterior insula. More sophisticated parametric experimental designs and multivariate analyses will doubtless illuminate additional aspects of these circuits, and so could expand the precision and predictive power of such a risk matrix account [33]. Although the risk matrix alludes to at least two components, it differs from other popular ‘dual system’ accounts of decision making that typically juxtapose motivational versus control elements, since it instead invokes two opposing motivational components [34]. Control circuits, accordingly, may modulate either or both of the risk matrix components, depending on the choice context (e.g., gains, losses, probability, ambiguity, time, etc.) — a possibility that we consider next.

**Neural control of risky choice**

Risky choice minimally involves choosing whether or not to accept one risky option, and classically involves
choosing between low risk versus high risk options (e.g., the blue versus the red pill in the Matrix). While these simplified choice scenarios may confer experimental control, they may at the same time sacrifice much of the complexity commonly encountered in real-world choices. Thus, recent research has focused on how decision makers calibrate risky choice to accommodate multiple options and emphasize different decision features. These studies have adapted a welter of tasks and dynamic processes, but there has been one point of convergence. Much of this research now points to a particular role for the dorsomedial prefrontal cortex (or DMPFC) — alternatively labeled as the anterior cingulate cortex (or ACC) — in exerting cognitive control over risky choice.

Anterior cingulate cortical activity has traditionally been linked to notification of anticipated or obtained errors [35,36] as well as environmental unpredictability [37], suggesting that it supports cognitive control processes capable of altering currently active behavioral tendencies. Research has linked ACC activity to increased demand for control in complex choice scenarios, such as when people choose against a typical decision frame [38] or adopt a decision strategy that runs counter to their usual tendency [31,39]. More recent work has further linked dynamic ACC activity to changes in risk preference based on the current environmental context [40]. These and related findings have been codified into a ‘cost of control’ framework which postulates that ACC activity tracks tradeoffs between potential rewards and the cognitive and physical effort necessary for obtaining them [41] — which can then can alter activity in the risk matrix to shape upcoming choices [32,40]. Despite this convergence of findings and theory, exactly how these prefrontal cortical circuits modulate the risk matrix remains to be determined.

Implications and extensions
Core components of the risk matrix have also been implicated in anticipation of diverse gains (ventral striatum) and diverse losses (including pain anticipation; anterior insula) [10], and so might bridge literatures focusing on reward circuitry and the pain matrix [11,25] (Figure 3). Particularly when choices are complex, prefrontal value integration and control circuits may modulate activity in both components of the risk matrix. Physiologically, these findings cohere with recent neuroanatomical accounts of risky choice in which ascending frontostratal and insular circuits enlist anticipatory afferent to inform subsequent value integration and motivation (e.g., the AIM framework; [42]).

These developments have important implications for neuroscientific and behavioral theory. Neuro- and psychologically, they suggest that affect and motivation play a significant role in risky choice, and further imply that symbolic numerical representation of costs and benefits may influence choice through affective circuits [3,26,27]. Economically and financially, the findings suggest that accounts of risky choice that incorporate multiple components (e.g., mean-variance theory) more closely reflect the underlying mechanics of the risk matrix than do accounts that include only one component (e.g., expected utility theory) [43].

Beyond the surprisingly close fit of brain activity to mean-variance theory, these advances have also generated novel predictions that extend beyond the scope of traditional finance theories, and which researchers have just begun to explore. First, by definition, mean-variance theory does not consider higher-order sources of variability involving asymmetric (skewed) or extreme (kurtotic) outcomes. But researchers have begun to demonstrate that asymmetric outcomes [44,45], and lottery-like (or positive-skewed) gambles in particular, can drive ventral striatal activity and preference [46]. Second, traditional finance theories do not typically account for the origin or dynamic adjustment of reward and risk expectations. A growing literature on reward and risk learning, however, suggests not only that updating these expectations recruits risk matrix components [22,47], but also that people may learn in biased ways reflected in brain activity [48], such as updating expectations with heavier weights on gain than loss outcomes [49]. Third, a surprising implication of proposed affective contributions is that even incidental activation of risk matrix components might bias risky choice [16]. In fact, researchers have found that presentation of irrelevant positive images can increase risky choice by activating the ventral striatum [50], while negative images and threats of shock can instead decrease risky choice [49], possibly by activating the anterior insula [51].

Along with input from prefrontal circuits in response to demands for value integration and control, other circuits doubtless modulate the risk matrix and its output, depending on the choice context. For instance, socially risky choices powerfully activate risk matrix components in ways that can promote cooperation or competition [52]. Further, in investment settings involving others, modulatory control from circuits that support social inference (e.g., DMPFC) can increase performance [53].

Even more remarkably, new findings suggest that group risk matrix activity may allow researchers to forecast choice at the aggregate level of markets. For example, in a neuroimaging study of financial market bubble formation, group ventral striatal (i.e., specifically in the nucleus accumbens) activity tracked market bubble formation, while individual differences in anterior insula activity predicted who would bail out most rapidly and so minimize their losses after a crash [54]. If deciding to share resources with a stranger represents a socially risky choice, emerging research further suggests that ventral
striatal activity predicts individual choices to donate or lend [55], and further, that group ventral striatal activity provides a forecast of the success of loan requests on the internet — even better than group choice itself [56]. These discoveries imply that risk matrix activity not only can predict socially risky choices in individuals, but further that in some cases, group brain activity might forecast aggregate choice better than the behavior of that group.

In retrospect, neuroimaging research on risky choice has advanced over the span of a decade with unexpected speed. Findings have progressed from documenting initial correlates of risk assessment, to identifying neural predictors of risky choice in individuals, to exploring the potential for the neural activity of groups to forecast aggregate choice. Despite this rapid progress, much work remains to be done on several fronts. Parallel animal models of risky choice could be combined with targeted neurochemical probes (e.g., optogenetics) to causally test neuroscientific accounts of risky choice and to help explain observed patterns of brain activity in humans. Improved experimental designs and multivariate analyses might optimize generalizable models of neural predictors of risky choice. Further research should explore the extent to which risk matrix activity predicts risky choice in nonmonetary domains, since risk preferences can vary across domains (e.g., financial, physical, social) [57]. Future studies will undoubtedly also explore which components of the risk matrix support forecasting choice at larger scales (e.g., on the internet and in markets). Though presently more potential than reality, in the near future, the risk matrix may transform from science fiction into scientific fact.

Conflict of interest statement
Nothing declared.

Acknowledgements
We thank Colin Camerer, Russell Poldrack, and spanlab for contributions to earlier drafts. During manuscript preparation, BK was supported by the Stanford Neuroscience Institute’s NeuroChoice Initiative and SAH was supported by an Incubator Award from the Duke Institute for Brain Sciences.

References