Epistemic Shifts Explained in Terms of "Time Budget Deficit"

Abstract

The term "epistemic shifts" refers to a widely recognized phenomenon that knowledge ascribers would ascribe different epistemic statuses to the same beliefs held by the subjects under different internal/external conditions. Contextualists and contrastivists tend to explain this phenomenon by making the semantic connotation of "knowing" different from one context to another, while stake-based invariantists and intellectual invariantists tend to explain the same phenomenon by appealing to external factors like stakes or internal factors like "need-for-closure". I tend to argue that all the preceding theories lack either the minimal integration of the dimensions of involving parameters or the minimal elegance facilitating cognitive modelling or the minimal range of the scope of applicability. My competing model for explaining epistemic shifts is based on the notion of "time budget deficit", which refers to the deficit of the operating resources related to time that a subject incurs when she is facing a certain cognitive task. My theory has a wide scope of applicability from the bank cases to varieties of skeptic cases. Furthermore, due to the algorithmic characterization of my model, it also has the potential to be serving for the goal of AI.

Key Words and Phrases:

epistemic shifts; time-budget-deficit; stakes; intellectual invariantism; the bank case; the zebra case; skepticism

§1. Introduction: Two Missing Links in the Alliance of the Mainstream Epistemology and its Formal Counterpart

According to Vincent F. Hendricks, epistemology can be pursued in two ways: the first is the "mainstream epistemology", which seeks necessary and sufficient conditions for the possession of knowledge using largely common-sense considerations and folksy examples/counterexamples; whereas the second is "formal epistemology", which either proceed axiomatically or concentrate

on learning and knowledge acquisition using tools *from logic and computability theory*.[1] What Hendricks himself seeks for is a hybrid form of epistemology in which the minimal respect of the commonsensical intuition (which is required by the first approach) and the pursuit of logical/mathematical strictness (which is required by the second approach) can be made contributable to each other. The motivation of elaborating this new form of epistemology was foreshadowed by W. V. Quine's proposal for naturalizing epistemology[2] and more explicitly formulated in John. Pollack's philosophical manual book on "how to build a person".[3] To be more specific, according to Pollack, if token physicalism (i.e., the thesis that every mental-state-token is a physical-state-token), agent materialism (i.e., the thesis that persons are physical objects with a suitable structure) and strong AI (i.e., the thesis that one can in principle construct a silicon-based person executing appropriately designed AI programs) are all true (and they should be all true from Pollack's perspective), then there will be no reason not to believe that sooner or later we can build a silicon-based but still rational being exhibiting reasoning behaviors addressed by epistemologists. [4]

   However, although I feel sympathetic to Hendricks and Pollack's general proposal for filling the gap between the mainstream epistemology and its formal counterpart, there are a couple of missing links in their projects. One of them is related to linguistic issues (e.g., the semantic differences between English terms "knowing" , "justification", "believe", "intention", etc. and

---

[1] Vincent F. Hendricks, *Mainstream and Formal Epistemology*(Cambridge: Cambridge University Press, 2006), 4-5.

[2] W. V. Quine, "Epistemology Naturalized". In W. V. O. Quine, *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969) 69–90.

[3] John Pollock, *How to Build a Person: A Prolegomenon* (Cambridge, MA: MIT P, 1989).

[4] Ibid., 4-5.

their counterparts in non-English languages[5]), indicating the varieties of cultural backgrounds of

epistemic agents or knowledge attributors, as well as their impacts on different types of

knowledge-ascribing behaviors. Another missing link is psychology, which reveals the phenomena

such as in many occasions subjects do not tend to make judgements in accordance with logical or

probabilistic rules; instead, they tend to commit certain types of inferential fallacies .[6]

However, if inquiries into both links could be made part of the putative alliance between the

mainstream and the formal approaches, both new comers will immediately undermine each

existent approach respectively. To be more specific, the comparative linguistic considerations,

when interwoven into the web of experimental philosophy, do cast doubt on the dominance of

Anglophone subjects' epistemic intuitions, which is essentially significant for mainstream

epistemologists, whereas psychological considerations, when tangled with sympathies for the

ordinary people's thumb-rules for handling epistemic tasks, will also marginalize the status of

---

[5] Cf. Masaharu Mizumoto, "A Simple Linguistic Approach to the Knobe Effect, or the Knobe Effect without Any

Vignette", *Philosophical Studies* 175 (2018): 1613-1630. Based on investigations of intention-ascribing behaviors

of subjects who speak English, Japanese and Chinese respectively, this paper is intended to argue that there is a

legitimate role of the empirical study of concepts like "intention" (which plays a pivotal role in the formation of

the "Knobe Effect") in the investigations of cognitive processes in experimental philosophy, and such study is

impossible without correlating the concept in question with a certain natural language wherein it is shaped.

[6] There are numerous examples of this sort, such as many subject's incompetence in using the rules of Modus

Ponens and Modus Tollens in the Wason Selection Task (cf. P. C. Wason,. "Reasoning about a Rule". *Quarterly

Journal of Experimental Psychology*. 20 (1968): 273–281), or subject's incompetence in following probabilistic

rules under certain conditions (cf. Amos Tversky and Daniel Kahneman, "Extension versus Intuitive Reasoning:

The Conjunction Fallacy in Probability Judgment". *Psychological Review*. 90 (1983): 293–315).

standard formal tools that formal epistemologists appeal to.[7] Although this does not mean that both mainstream and formal approaches in epistemology have to be abandoned when linguistic/psychological factors are put on the table, the considerations of these factors have at least revealed Hendricks and Pollack's underestimation of the difficulties either for making mainstream epistemology formalized or for reconstructing an artificial agent.

But why the negligence of linguistic/psychological issues is intolerable? Why not just to build an idealized robot which is immune to both cultural diversities and cognitive biases?

Firstly, please consider a typical topic in epistemology: "epistemic shifts", namely, the phenomenon that knowledge ascribers would ascribe different epistemic statuses to the same beliefs held by the subjects under different internal/external conditions. (For instance, we will tend to believe that a subject, say, John, knows that he has two hands in his ordinary life, and suspends this knowledge when he is indulgent in skepticism.) Is this phenomenon a linguistic one or a psychological one? Epistemic contextualists like Keith DeRose[8] and epistemic contrastivists like Jonathan Schaffer[9] tend to address this from a linguistic lens (since the "contextual factors" that

---

[7] As psychologist Gerd Gigerenzer points out, subjects tend to handle the Wason Selection Task in accordance with logical rules only when such task is reformulated in scenarios encoding information related to social contacts, since only the sensitivity of the information of this sort could be evolutionarily beneficial. Otherwise human subjects would still tend to use some logic-violating heuristics which can greatly save their operating resources. Hence, it is not hard to see that subjects' deviation from formal rules can be rationalized from an evolutionary point of view. Cf. Gerd Gigerenzer: *Adaptive Thinking: Rationality in the Real Word*(Oxford: OUP) 217-18.

[8] Keith DeRose, "Contextualism and Knowledge Attributions". *Philosophy and Phenomenological Research* 52 (1992) : 913–929. His version of contextualism is foreshadowed by David Lewis and's work (cf. David Lewis, "Elusive Knowledge", *Australasian Journal of Philosophy*, 74(1996): 549–567; Stewart Cohen, "Contextualism, Skepticism, and The Structure of Reasons", *Philosophical Perspectives 13: Epistemology*(1999): 57–89).

[9] Schaffer, J.. 2004. "From Contextualism to Contrastivism". *Philosophical Studies* 119(1–2): 73–103.

DeRose appeals to and the "semantically contrasting terms" that Schaffer appeals to both have a bite of philosophy of language), and the debate between them are even resorting to methods of experimental philosophy aimed at ascribers' linguistic behaviors. [10] By contrast, intellectual invariantists like Jennifer Nagel[11] tend to view the same phenomenon psychologically, given that the key phrase in Nagel's narrative for explaining epistemic shifts, namely, "need-for-closure", is directly borrowed from Kruglanski & Webster's psychological studies.[12] Hence, if linguistic and psychological issues were both suspended in the mainstream and formal epistemology, what sort of other resources could be used to explain epistemic shifts?

Secondly, there are practical reasons not to neglect linguistic issues in the process of reconstructing a person, given that the perspective of comparative linguistic studies has special values for certain branches of AI such as Natural Language Processing and Machine Translation, the realization of which should be fairly significant for building a socialized robot.

Thirdly, the psychological issue has to be seriously addressed as well, since many logic-violating psychological dispositions, according to psychologist Gerd Gigerenzer, are embodying the so-called "frugal rationality", namely, the evolution-favored capacity to respond to environmental challenges with very limited operating resources, like the processing time. [13]

---

[10] J. Schaffer & J. Knobe, "Contrastive Knowledge Surveyed", *Noûs* 46(2012) 675-708.

[11] Jennifer Nagel, "Knowledge Ascriptions and the Psychological Consequences of Changing Stakes". *Australasian Journal of Philosophy* 86 (2008):279–94.

[12] A. Kruglanski & D. Webster, "Motivated Closing of the Mind: 'Seizing' and 'Freezing'", *Psychological Review* 103(1996): 263–83.

[13] Cf. Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, *Simple Heuristics that Make us Smart* (Oxford: OUP, 1999); Gerd Gigerenzer: *Adaptive Thinking: Rationality in the Real Word* (Oxford: OUP, 2000); Gerd Gigerenzer, Ralph Hertwig, and Thorsten Pachur, *Heuristics: The Foundations of Adaptive Behaviors*

Since Strong AI-based robots have to tolerate the limitation of operating resources like humans, frugal-rationality-requiring psychological dispositions have to be modelled in their cognitive architectures as well. Hence, psychologists' studies of frugality has to be taken into account seriously.

Due to the limitation of space, in this article I only intend to focus on how to make psychological considerations integrated into the mainstream & formal epistemological studies, and especially how psychological considerations could lead theorists to form the most suitable formal tools for cognitive modelling, rather than doing things in an entirely reversed manner, namely, to make cognitive modelling fit the features of existing formal tools. The target phenomenon in my mind is nothing but epistemic shifts as aforementioned, and Nagel's intellectual invariantism could be accordingly viewed as an theoretic ancestor of my theory. But I still have some criticisms of her theory, as the following section will immediately show.

§2.   Nagel's Intellectual Invariantism Revisited

Epistemic shifts can be illustrated either in terms of skeptic cases or cases which are not obviously skepticism-evoking. A typical illustration of the second type of cases is the "bank case", which is composed of two scenarios:

The Bank Case:

Scenario 1: Hannah is wondering whether the bank, which is definitely open today, will still be open tomorrow, which is Saturday, and her perfect financial status allows her to deposit her paychecks in the bank a couple of days later. The question to the ascriber (whom is informed that

---

(Oxford: OUP, 2011).

the bank will open tomorrow) is: Does Hannah know that the bank will be open on the coming

Saturday if her only available evidence is that it was open on the last Saturday?

Scenario 2: Hannah is wondering whether the bank, which is definitely open today, will still

be open tomorrow, and her terrible financial status does not allow her to deposit her paychecks in

the bank a couple of days later. The question to the ascriber (whom is informed that the bank will

open tomorrow) is: Does Hannah know that the bank will be open on the coming Saturday if her

relevant evidence is that it was open on the last Saturday plus the extra piece of information that

this bank did change their serving schedule in the past?

According to Jason Stanley's account of the bank case,[14] knowledge ascribers are more

inclined to ascribe knowledge to a subject's belief *p*, say, "The local bank is open on Saturday."(if

the subject's practical interests would be in low stakes when *p* were false, as depicted by Scenario

1) than to the belief with the same content (if the subject's practical interests would be in high

stakes when such a belief were false, as depicted by Scenario 2). Stanley himself elaborates the

position of "Interest-Relative Invariantism" based on the preceding observations. Since his theory

does not predict the change of the *semantic* meaning of "knowing" in different contexts, hence, it

is not any variant of contextualism, according to which the height of the semantic bar of "knowing"

is higher in some contexts than that in others.

As to Jennifer Nagel's "intellectualist invariantism",[15] it is another form of invariantism

competing with Stanley's. According to Nagel, ascribers' epistemic shift is not going to be

explained in terms of changes of subjects' external environmental factors but by virtue of changes

---

[14]   Jason Stanley,   *Knowledge and Interest*. (Oxford: Oxford University Press, 2005): 3-4.

[15]   Jennifer Nagel, "Knowledge Ascriptions and the Psychological Consequences of Changing Stakes". *Australasian Journal of Philosophy* 86 (2008):279–94.

of their internal belief-forming processes, which could "internalize" environmental factors like stakes. The specific psychological mechanism most relevant to this internalizing process, according to Nagel, is "need-for-closure"(or "closure" for short), as revealed by psychologists Arie Kruglanski and Donna Webster.[16]

Psychologically speaking, according to Arie Kruglanski and Donna Webster, "need-for-closure" could be defined as "a desire for definite knowledge on some issue; it represents a dimension of stable individual differences as well as a situationally evocable state". [17] And here "a desire for definite knowledge" can be more explicitly defined as a "desire for a firm answer to a question and an aversion towards ambiguity". [18] To be even a bit more specific, the term "closure" itself can be further divided into two categories: (a) "specific closure", namely, a psychological factor inducing an agent to firmly hold a belief with some pre-fixed content (e.g., the case in which a cancer-sufferer scans the internet for convincing himself to believe that his own cancer is curable); (b) "non-specific closure", namely, a psychological factor leading an agent to firmly hold a belief the content of which is not antecedently expected (e.g., the case in which a pathologist scans the internet to know whether an anonymous patient's tumour is malignant or benign). Nagel believes that "non-specific-closure" is more useful in explaining our intuitive epistemic shifts related to the bank case. (Hence, henceforth "closure" will be only referring to "non-specific-closure".) Accordingly, as to why most ascribers do feel natural to ascribe knowledge to Hannah's belief that "The bank will be open tomorrow." when her financial status is

---

[16] A. Kruglanski & D. Webster, "Motivated Closing of the Mind: 'Seizing' and 'Freezing'", *Psychological Review* 103(1996): 263–83.

[17] Ibid., 263.

[18] Ibid., 264.

acceptable, and why they feel natural *not* to ascribe knowledge to Hannah's belief with the same

content when her financial status is much worse, Nagel's explanation goes like this:

> High-Stakes Hannah is a low need-for-closure subject: she is strongly averse to making a
>
> mistake about the banking hours, and willing to entertain hypotheses that would make her
>
> initial evidence inconclusive. Her Low-Stakes counterpart is in a neutral need-for-closure
>
> condition: Low-Stakes Hannah is not described as being under anything like the pressure for
>
> immediate decision characteristic of high-need-for-closure subjects, nor is she strongly averse
>
> to making the wrong call about Saturday banking—the scenario suggests that getting it
>
> wrong would be a mild inconvenience at worst. …Because of the perceived need-for-closure
>
> differences between the subjects, we don't expect the same information to produce the same
>
> level of belief in High and Low Stakes. [19]

Or put it differently, whether the subject's need-for-closure is low or high or neutral is the

most critical parameter that the knowledge ascriber should consider when she is invited to assess

the bank case.

I prefer Nagel's theory to Stanley's, as stakes are too diverse to be handled in an elegant

model. As a matter of fact, although Stanley attempts to clarify the issue on stakes by appealing to

the notion of "warranted expected utilities"[20], his theory gives no hint on how these utilities could

be calculated in a universal way. *Prima facie*, this problem can be solved by appealing to decision

theory or statistics. But as Jacob Ross and Mark Schroeder point out,[21] if we follow statistician

---

[19] Jennifer Nagel, "Knowledge Ascriptions and the Psychological Consequences of Changing Stakes". *Australasian Journal of Philosophy* 86 (2008):289.

[20] Jason Stanley. *Knowledge and Interest*. (Oxford: Oxford University Press, 2005): 94.

[21] J. Ross & M. Schroeder, "Belief, Credence, and Pragmatic Encroachment". *Philosophy and Phenomenological*

Leonard Savage[22] by calculating the expected utility of an agent's possible actions in terms of the actual utilities of possible consequences of these actions, the resultant computational cost would be unmanageable. In addition, in my view, how to assign proper values to the action-consequence pairs would be another salient problem when we try to handle typical skepticism-inducing epistemic scenarios, in which there is nearly no hint on how to calculate the utility of accepting /rejecting a claim like "whatever that I see is induced by a Cartesian demon". Even more troublesomely, Mark Schroeder's own refinement of Stanley's stake-based treatment of epistemic shifts cannot be of any instructing value for cognitive modelling. Please consider his following definition:

S's epistemic reasons to believe p are at least as good as S's epistemic reasons to believe $\sim p$ just in case:

$$Ev_p + Err2 + Err1_{\sim p} \geq Ev_{\sim p} + Err2 + Err1_p \quad [23]$$

The meaning of this formula is that the sum of the strength of the evidence for believing p and the cost of making "type-2 error" and the cost of making the "type-1 error concerning $\sim p$" cannot be exceeded by the sum of the strength of the evidence for believing $\sim p$ and the cost of making "type-2 error" and the cost of making the "type-1 error concerning p". (Here "type-1" error means the mistakes of making something false as true, and vice versa, and the "type-2" error means the mistakes of neglecting the content of the relevant beliefs). However, this formula is entirely incomputable due to its lack of a minimal integration of dimensions: the dimension for

measuring the weight of evidence is patently different from that for measuring the cost of making errors, and to put them on the same page is as unmanageable as computing the sum of, say, 3 *miles* of railways and 5 *tons* of Coca Cola.

With contrast to "stakes" or "costs" in Stanley& Schroeder's narrative, "need-for-closure" in Nagel's narrative is a stable type of psychological effect identified in the literature, so I am happy to take it as one of the inspirations of my own cognitive modelling concerning epistemic shifts.

But significant revisions of Nagel's position are still needed, since it still lacks an elegant explanation of how the closure is induced. Nagel has mentioned a couple of relevant factors like time pressure, fatigue and background noise,[24] but there is no unified model for explaining how these factors interact with each other, just like there is no integration of dimensions in Schroeder's model for consistently elevating both the weights of epistemic evidence and the costs of making epistemic errors. In contrast, my theory on epistemic shifts, as I will immediately show, is mainly based on the notion of "time budget deficit", namely, the "deficit" of the time-related operational resources that the subject incurs, and the basic unit of account of the "deficit" in question is represented in terms of the "steps of atomic operations" of the cognitive architecture executing the relevant epistemic tasks. It is worth nothing that the main thesis of my theory could be further algorithmically formulated rather than a claim expressed by natural language, hence, it has the potential to be employed in an AI system to intimate epistemic shifts.

Now it's the right time to unpack my theory.

---

[24] Jennifer Nagel, "Knowledge Ascriptions and the Psychological Consequences of Changing Stakes". *Australasian Journal of Philosophy* 86 (2008):287-288.

§2. The Notion of "Time Budget Deficit" and the Bank Case again

Before algorithmically reconstructing my model on time budget deficit, it may be helpful to provide an intuitive illustration of what the very notion of it is.

> **The first approximation to "time budget deficit"**: A subject handling an epistemic task has incurred certain quantity of time budget deficit , when and only when according to her estimation, the time budget required for completing the epistemic task does exceed her actually available time budget, or in other words, the quantity of such a deficit is numerically the difference between the requisite time budget and the available time budget.

But why is this definition useful in explaining epistemic shifts? Here comes a supplemental definition:

> **The relevance of time budget deficit to epistemic shifts**: The quantity of the time deficit is negatively correlated with the degree of the difficulty of ascribing knowledge: the bigger the deficit is felt by the subject, the more reluctance is also felt for delivering the label of "knowledge" to the target belief.

I have three clarifications for the preceding definitions.

Firstly, the metaphoric term "epistemic shifts" indicates the possibility of talking about the quantitative transitions of some variables' values related to knowledge ascription, and "time budget deficit" is precisely such a variable allowing its values to be graded in a continuous spectrum. Hence, such treatment has certain built-in instructing values for cognitive modelling, which has to be resorting to algorithmic characterizations.

Secondly, it is obvious that the preceding definitions do not explicitly involve

multiple-dimension-assuming parameters like "closures" or "stakes", and exclusively relies on the parameter on time. Hence, it will not evoke the troublesome problem related to the integration of different dimensions from different parameters. Hence, my treatment of epistemic shifts in the light of the preceding definitions will possess the merit of elegance and unification.

Thirdly, the "time" mentioned in this definition is surely not objective time measured by physical devices, but psychological time which could be only subjectively estimated. This treatment is introduced to catch the intuition that the same epistemic task may bring about different time pressures to different subjects with different cognitive talents. However, as I will explain later, subjectivity of this type will not make it an obstacle for cognitive modelling, since there is an algorithmic characterization of this subjectivity in terms of "steps of atomic operations".

Though it may be still a bit too early to give an algorithmic account of my model, the currently available definitions do suffice for the purpose for showing how it can be used to handle a wide range of epistemic scenarios from the bank case to the skeptic cases.

Let us begin with the bank case as aforementioned. Most people will feel more reluctant to ascribe knowledge to the High-Stakes Hannah's belief than to the Low-Stakes Hannah's belief, and my time-deficit-based model can explain this easily. Actually this explanation is not substantially relevant to stakes. To be more specific, the High-Stakes Hannah is facing a more complicated epistemic task than her Low-stakes counterpart, given that the Low-Stakes Hannah only needs to induce the schedule of this bank in this month from her stored belief on what the schedule was in last month, whereas the High-Stakes Hannah needs to negotiate between the preceding stored belief and a piece of additional information that this bank would change the

schedule irregularly. Other things being equal, the more complicated the cognitive task is felt, the more requisite time budget will also be estimated, and hence, the more time deficit will be represented as well. Consequently, more reluctance of ascribing knowledge to the relevant belief will be perceived.

However, this treatment may be confronted with the following objection from the perspective of stake-based invariantism: Why is the High-Stakes Hannah more likely to consider the additional information that this bank would change the schedule irregularly? Just because she is a High-Stakes Hannah, to whom the cost of making the schedule wrong would be intolerable. Hence, your time-pressure-based model has simply assumed a stake-based model.

My response: Surely in some cases the awareness of stakes related to p-based actions may trigger further beliefs encoding negative evidence against p, but these negative-evidence-encoding beliefs may also be triggered due to other causes unrelated to stakes. For instance, even the Low-Stakes Hannah may be aware of the information that this bank would change the schedule irregularly due to an accidentally recalled image of the difference between this bank's schedule in last month and that in the month before last month, and in this case, according to my model, her time deficit for fixing the epistemic status of the target belief will still be huge.

The preceding response may seem to be counterintuitive for many. Many would feel that even if the Low-Stakes Hannah were able to visualize the possibility for the bank to change its schedule, the degree of the stress felt by her would still be much smaller than the degree of the stress felt by her High-Stakes counterpart, and only the stake-based model, rather than mine, can explain this. But I think here critics have confused a subject's psychological stress felt by her when she is practically engaged in a stake-related task with her time deficit for fixing the truth-value of a

14

belief (which is not necessarily stake-involving), since the two kinds of things do not always go hand in hand with each other. As for an instance, although to check the veracity of Goldbach Conjecture is obviously time-budget-deficit-requiring for a non-mathematician, in most cases, it would not bring about psychological stress to her, given that the pursuit of abstract mathematical truth is irrelevant to most people's practical concerns. But this simply does not mean that the time deficit felt by non-mathematicians could be accordingly ignored. Rather, it is still tremendous on the table, despite the subject's practical indifference to it. Another illustrating case is as the follows: a prisoner under death sentence is surely under tremendous psychological stress when the executing day is coming. But this does not imply that the epistemic task of figuring out which day is the executing day would bring about big time budget for him. Rather, this task is just a cake work, despite the psychological stress produced by the connotation of the relevant belief is still huge. Paralleling to both the Goldbach-conjecture case and the prisoner case, in the bank case, though the bad financial status would usually bring about psychological stress to Hannah, this does not imply that the complexity of the relevant epistemic task is positively correlated with the strength of the stress. Hence, it is possible for high stress, which may be substantially stake-related, be accompanied with a small time deficit of completing the epistemic task for fixing the content of the bank schedule, and vice versa. And Stanley's stake-based invariantist explanation of epistemic shifts, in my view, is precisely based on a systematic confusion of stake-related stress from the practical perspective with the complexity-related time deficit from the epistemological perspective.[25]

---

[25] My position concerning the stake-knowledge relationship is not as radical as Wesley Buckwalter and Janathan Schaffer's position that stakes do not affect knowledge affection at all (cf. W. Buckwalter & J. Schaffer, "Knowledge, Stakes, and Mistakes". *Noûs* 49(2015): 201-234). I only intend to claim that stakes may be somehow

The same criticism can be even applied to Nagel's intellectual invariantism as well. Although the need-for-closure is a psychologically robust phenomenon, from the perspective of cognitive modelling, it may be both further underpinned by stake-evaluating mechanisms and time-deficit-evaluating mechanisms, since both the increase of the stakes and that of the perceived complexity of the relevant epistemic task will decrease the need-for-closure. For instance, when a pathologist scans the internet to know whether an anonymous patient's tumor is malignant or benign, she will feel reluctant to make a final conclusion too quickly if both the following pieces of additional information are delivered to her: firstly, a misdiagnose will produce serious legal responsibility for her; secondly, the available diagnosing time is less than what this type of work normally requires. In this sense, even Nagel's closure-based invariantism has involved the risk of systematically confusing stake-related closure from the practical perspective with the time-deficit-related closure from the epistemological perspective.

But why is the temptation of confusing the issues related to stakes with those related to complexity so irresistible for many? I believe that this may be due to a bad phenomenology assumed by stake/closure-based theorists. Phenomenally speaking, the feelings of the stake-inducing stress is undistinguishable from the feelings of the complexity-inducing time pressure for completing an epistemic task, and even neurologically speaking, the two types of feelings may also be unluckily related to the same type of physical events like the activation of the hypothalamic-pituitary-adrenal axis and the secretion of cortisol/catecholamines.[26] However, on a

related to need-for-closure (and it may be not), but it is definitely not a proximate cause of knowledge ascription and hence does not deserve a niche in my model.

[26] A comprehensive empirical account of the functioning of psychological stressors and their neural basis can be found in S. S. Dickens & M. E. Kemeny, "Acute Stressors and Cortisol Responses: a Theoretical Integration and

theoretical level, it is still necessary to distinguish stake-inducing stress from the complexity-inducing time deficit, otherwise it would be hopeless to handle epistemic cases which are not obviously stake-involving.

§3. The Zebra Case Revisited

As even beginners of epistemologists know, typical epistemic cases, like the skeptic ones, are highly speculative and remote from practical concerns about stakes. And a qualified model on epistemic shifts needs to handle these cases with appropriate priority, although it is also expected to handle those stake-involving ones. The zebra case [27] is such a skeptic case which deserves a serious treatment. This case is composed of two contrasting scenarios:

The Zebra Case

Scenario 3: Hannah is watching a zebra wandering in a zoo, and her father, who is standing by her side, points to her that it is possible that what she sees is not a zebra but a mule. Hence, Does Hannah know that what she is watching is a zebra?

Scenario 4: Hannah is watching a zebra wandering in a zoo, and her father, who is standing by her side, points to her that it is possible that what she sees is not a zebra but a cleverly painted mule, which is perceptually indistinguishable from a real zebra from a distance. Hence, Does Hannah know that what she is watching is a zebra?

Synthesis of Laboratory Research". *Psychol Bull*, 130(2004):355-91. According to their study, tasks containing both uncontrollable and social-evaluative elements were associated with the largest cortisol and adrenocorticotropin hormone changes. Hence, the neurological link, which is the last link in the causal chain leading to phenomenal feelings about stress, is insensitive to high-level content which initially evokes stress.

[27] The original form of the case can be found in Fred Dreske, "Epistemic Operators," *Journal of Philosophy*, 67(1970): 1007–1023.

Most ascribers would give a positive answer to the query raised in scenario 3 but a negative one to the query in scenario 4, and my theory can tell why. In the zebra case, the intuitive epistemic shifts from scenario 3 to 4 is to be explained in terms of the difference between the time-deficit-value in scenario 3 and that in 4, and this difference further hinges on the difference between the complexity of the task in scenario 3 (namely, "to distinguish a zebra from a mule", which seems to be an easy job) and the complexity of the task in scenario 4 (namely, "to distinguish a zebra from a cleverly-painted-zebra-alike-mule", which seems to be a very challenging job). Hence, the more complicated the identification task is, the higher time-deficit-value will be, and consequently, the less naturally the ascriber will feel to ascribe knowledge to the target belief. Therefore, it should be easier to ascribe knowledge to the target belief in scenario 3.

The analysis of the zebra case can be easily applied to more typical skeptical scenarios like the Cartesian demon case. Since to distinguish a perceptual object induced by normal conditions and one induced by a Cartesian demon should be unpredictably complicated (and this is due to a further fact that any piece of evidence for the existence of the normal conditions can be reinterpreted to be fitting the putative existence of a Cartesian demon), this will result in a tremendously huge time-deficit-value, which makes the ascription of knowledge extremely hard.

It is worth noting that the preceding skeptic cases are not so favored by stake/closure-based invarianists due to their irrelevance to practical concerns, and the applicability of my time-deficit-based narrative to both skeptic cases and ordinary cases like the bank case simply exhibits the explanatory power of my theory. And I even believe that my theory is more powerful than some epistemological stance deliberately designed for handling skeptic cases, e.g., Schaffer's

epistemic contrastivism. Please consider an updated form of the zebra case:

The Updated Zebra Case:

Scenario 5: Hannah is watching a zebra wandering in a zoo, and her father, who is standing by her side, points to her that it is possible that what she sees is not a zebra but a cleverly painted mule. Hence, Does Hannah's know that what she is watching is a zebra?

Scenario 6: Hannah is watching a zebra wandering in a zoo, and her philosophy teacher, who is standing by her side, points to her that it is possible that everything that she sees, the zebra included, is ultimately induced by a Cartesian demon. Hence, Does Hannah's know that what she is watching is a zebra?

I believe that most readers will tend to say that Hannah is ignorant in both scenarios, and the Hannah in scenario 4 is even *more* ignorant. Now a new question comes: why do we have this intuition?

My explanation is quite simple: It is part of our common sense that to discriminate a zebra from a cleverly painted mule is not entirely impossible, while the confirmation or disconfirmation of the existence of a Cartesian demon will bring about a tremendously complicated epistemic task. Hence, the epistemic task in scenario 5 will involve a smaller amount of time deficit than its counterpart in scenario 6, and the difference between the two cases may quickly lead the ascriber to conclude that the Hannah in scenario 6 is even more ignorant.

However, I cannot see how Schaffer's contrastivism can explain this. The basic idea of his theory can be generalized as the follows: "knows" denotes a ternary relationship among the subject (S), the semantic content of her belief (p) and a contrast proposition (q), hence, expressions like "S knows that p." have to be *always* unpacked as "S knows that p rather than q." Accordingly,

the subject cannot attribute knowledge to *p* when *q*, which is *p*'s contrast proposition, shares the same evidence-basis with *p* itself, since her preference of *p* to *q* has to be based on the relatively bigger evidential power of *p* over that of *q*. This treatment surely can explain why Hannah is both ignorant in scenarios 5&6, given that the contrast proposition in scenario 5 (i.e., "This is a cleverly painted mule") and that in scenario 6 (i.e., "This is a Cartesian-demon-inducing zebra") both share the same evidence-basis with the target belief (i.e., "This is a real zebra".). However, there is no theoretic resources in epistemic contrastivism for explaining why Hannah is more ignorant in scenario 6 than in scenario 5, unless a contrastivist could appeal to the estimations of the complexity of the tasks involved in the two scenarios respectively.   But I simply cannot see how a contrastivist could do this estimation without assuming my time-deficit-based model.

Now it is the right time for explaining how my model could be algorithmically reconstructed, as I have promised.

§4. An Algorithmic Account of Time Deficit

As aforementioned, the most salient obstacle for algorithmically building my time-deficit-based model on epistemic shifts is the seemingly huge gap between the subjectivity of the notion of "time" involved in the compound expression "time deficit" and the objectivity of the constructing rules for any cognitive modelling project. My recipe for filling the gap is to resort to the notion of "steps of atomic operations".

My conception of this notion is motivated by chess game. Here is an illustrating case: Within 1 second after making a certain move on the chess board, say, move A, a beginner of the game of Go may be only capable of predicting 2 steps of further moves that her opponent would make for

responding to A, whereas an experienced player may be capable of predicting 5 moves within 1 second in the similar situation. Hence, the information processing speed of the second player is 2.5 times of that of the first player. Thereby, one can easily transfer the calculation of the time-deficit for completing a task into that of the step-deficit of doing the same thing. As for instance, if the maximal temporal duration of the task is both 2-second long for both players, and if the given task is expected to be completed only when the putative player could be able to successfully predict at least 10 steps of further moves that her opponent would make, then for the aforementioned unexperienced player, the step-deficit that she incurs will be 6 steps (=10 steps-2 step/sec × 2 sec), whereas as to the aforementioned experienced player, the step-deficit that she incurs will be 0 step (=10 steps-5 step/sec ×2 sec). It is worth noting that the notion of "step-deficit" is both subjective and objective here: it is subjective since it can catch the differences among different players' different talents, and it is also objective since it can catch the preceding differences by using a single dimension, namely, the number of steps.

To be more specific, the "steps" in this case can be construed as "steps of atomic operations of the chess game", wherein the adjective "atomic" is used to modify a unit of operation the further analysis of which is meaningless according to the chess rules. Surely there are numerous epistemic tasks which are neither constrained by chess-rule-like explicit rules nor analyzable into discrete units like a single act of moving a piece of chess. But even for these epistemic tasks, in which nearly all typical epistemological vignettes may be included, it is still theoretically possible to talk about atomic operations. Here goes the recipe for doing this.

Firstly, no matter what an atomic operation is, the number steps of them (N) is obviously the quotient of the temporal duration of the processing of a task (T) divided by the temporal duration

of a single atomic operation (Ts). Or a bit more formally,

Formula-1: $N=\frac{T}{Ts}$

It is obviously that the T-value is measurable. And the Ts-value is at least theoretically measurable. To be more specific, Psychologically speaking, it makes sense to measure the minimal temporal duration of subject's execution of a simple task like "naming a color"; [28] neurologically speaking, especially under the umbrella term "nerve conduction velocity studies", it also makes sense to measure the average temporal duration of "the propagation of nerve impulses from a neuron to a nearby one". [29] And machine-functionalistically speaking, it still makes sense to talk about the average temporal duration for a human-cognition-intimating Universal Turing Machine to execute a simplest task like erasing a "0" or printing a "1" on the tape. Surely different methodologies for empirically measuring the Ts-value will produce different results, and even when we are talking about the integrated Ts-value of a single person, complicated statistics has to be involved to negotiate the sub-Ts-values of different psychological modules or neurological micro-structures. But once a convention of how to handle this problem can be made and consistently abided, the arbitrariness of such a convention will not significantly influence the behaviors of a epistemic-shifts-predicting model, just like an arbitrary choice between the British System and the Metric System will not significantly influence the functioning of an electronic balance. The reason is quite simple: our epistemic-shifts-predicting machine is

---

[28] Empirical study of this sort is well discussed in M. B. Denckla, & L. E. Cutting, "History and Significance of Rapid Automatized Naming". *Annals of Dyslexia*. 49 (1999): 29–42.

[29] Empirical study of this sort is well discussed in D. C. Preson, & B. E. Shapiro, 2012. *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations* (3rd Edition) (London: Elsevier, 2012) 97-124.

only expected to give some high-level information like "there is more reluctance in ascribing knowledge in some cases than in others", just like an electronic balance is only expected to give some high-level information like "a normal baby is heavier than a normal snail". Hence, no big worry is needed to have here concerning what kind of low-level story is wanted to underpin the high-level one. Rather, the Ts-value in formula-1 is just a place-holder, and the representation of it is mainly for the sake of the theoretical accountability of the notion of "steps".

It is also worth noting that the introduction of formula-1 can help the modelers avert the trouble of directly representing the steps of atomic operations. But why is this a trouble? It is a trouble, because it will be very resource-consuming for any cognitive system to store the step-consuming data related to each task, whereas to store the T-values related to each task looks much easier. As for instance, it is usually hard for anyone to recall how many footsteps that she used to walk from her office room to the nearby bank, but it is much easier to recall how many minutes that she used to do the same thing yesterday. As to the Ts-value, though the estimation of it may be empirically complicated for humans, it can be easily determined in the AI-context by virtue of the performance of corresponding hardware.

From formula-1, it is not hard to derive the formulas for computing the *requisite* number of operating steps of a given task (Nr), as well as the *actually available* number of operating steps of a given task (Na). It is also easy to perceive that the difference between Nr and Na is mainly based on the difference between the temporal duration of the requisite time (Tr) and that of the available time (Ta). Here goes the two new formulas:

Formula-2: $Nr = \frac{Tr}{Ts}$

Formula-3: $Na = \frac{Ta}{Ts}$

Now if we combine the intuitive formulation of the time deficit (D) in section 2 with the preceding formulas, we can immediately have a formal definition of deficit:

Formula-4: $D = k(Nr - Na) = k\left(\frac{Tr - Ta}{Ts}\right)$

(Here "k" refers to a constant characterizing the specific-agent-related disposition of underestimating/overestimating the time deficit. Surely this can be omitted in a more simplified model.)

According to Formula-4, the subject will feel the existence of time deficit insofar as the computing result of the right side of the formula is a positive number. By contrast, if it is a negative number, the subject will feel the existence of time budget surplus. However, such cases are rare in skepticism-inspired epistemological scenarios.

We can even reconstruct the epistemic notion of "need-for-closure" out of formula-4. (The modifier "epistemic" indicates that the stake-related version of closure is not my primary concern here, although Nagel's notion of "closure" seems to have covered both versions.) To be more specific, the strength of the closure is negatively correlated to the D-value: the more deficit the subject incurs, the weaker the need for freezing a belief as knowledge will be represented. Thereby we can have formula-5 as the follows:

Formula-5: $NC = \frac{1}{D} = \frac{1}{k\left(\frac{Tr - Ta}{Ts}\right)} = \frac{Ts}{k(Tr - Ta)}$

And the derivability of formula-5 from formula-4, not vice versa, additionally explains why my deficit-based model on epistemic shifts is more fundamental than Nagel's closure-based counterpart.

Now we can use the preceding constructions to re-explain epistemic shifts in both the bank case and the zebra case. Insofar as the bank case is concerned, the High-stakes Hannah's Tr-value is higher than the Low-stakes Hannah's Tr-value, and when both the Ta-value and the Ts-value are invariant, the difference between them is basically related to two subjects' different estimations of

the number of operational steps of relevant tasks.(However, as I have explained in section 3, such difference is not necessarily stake-related.) Hence, according to formula-4, it is not hard to predict that the High-stakes Hannah will perceive a higher D-value and hence be more reluctant to freeze the relevant belief as knowledge. Insofar as the zebra case is concerned, since the task of discriminating zebras from mules is less step-demanding than the task of discriminating zebras from cleverly painted mules, the Tr-value has to be lower in the first case than in the second, hence, according to formula-4, other things being equal, the subject handling the mule-zebra discrimination task incurs a smaller amount of time deficit, and hence, she will be less reluctant to freeze the relevant belief as knowledge.

Now I think I need respond to some possible puzzles or objections related to my theory in the form of Q&A.

§5. Q&A

Q1: According to the mainstream understanding of the term "knowing", it has to be ungradable, and hence, there is no gradually fading spectrum from knowledge to ignorance: if you know something, then it is clearly not the case that you are ignorant of it. But your notion of time deficit is apparently gradable. Then, how to make the duality of knowledge accountable in your theory?

A1: It is not hard to so do. A modeler can further set a parameter to intimate the knowledge-triggering threshold, namely, the Kn-value: if the D-value is beneath the Kn-value, then the whole system will attribute the label of knowledge to the target belief, otherwise it would not do that. The height of the Kn-value-bar can be variant from one cognitive system to another, thereby we can easily characterize the difference between the dispositions of bolder knowledge ascribers, who are endowed with a higher Kn-value, and those of more prudent knowledge ascribers, who are endowed with a lower Kn-value. However, in typical epistemological cases, Kn-value-based interpersonal differences could be ignored, given that the subject in these vignettes, like Hannah, are basically fictional persons deprived of certain personal traits.

Q2: However, putting the treatment of fictional subjects in epistemological vignettes aside, according to your theory, different ascribers endowed with different Kn-values will deliver

different results of knowledge ascriptions. Hence, it seems that your theory is still on the side of intellectual invariantism, according to which the semantic level of "knowing" is invariant, while epistemic shifts are only to be explained psychologically, rather than semantically. Is this a correct description of your position?

A2: Not so correctly. I would rather say that it is very natural to view my theory as a species of intellectual invariantism. However, it can be also made to be consistent with epistemic contextualism or epistemic contrastivism as well. Nonetheless, if our purpose is only to explain epistemic shifts, intellectual invariantism looks like a more humble and hence less-risk-taking stance, since it only requires a theoretical move linking shifts to the operating principles of some internal mental mechanism rather than a bolder move to adjust the whole semantic web concerning "knowing" on the interpersonal level.

And moreover, from the perspective of AI, to take either the contextualist or the contrastivist stance (as two typical stances for formulating semantic variantism) looks simply infeasible. To be more specific, given that any algorithmic construction of semantic web requires the minimal cross-contextual stability of the inferential roles played by web-constituting lexicons, it will bring about a huge programing burden to realize the contextualist idea that the inferential role of "knowing" is different from context to context, or the contrastivist idea that "S knows q" has to be always unpacked as "S knows that p rather than q"(wherein the issue on how to saturate the q-value is still sensitive to contextual factors). By contrast, the intellectual-inspired modeling of internal cognitive mechanism is definitely free from the risk of overcomplicating the existent human semantic networks.

Q4. In your treatment of the bank/zebra case, the attributions of the Tr-values are still hinging on the humans' intuitions. But how can an AI system get these values automatically?

A4. A simplified algorithm for realizing this function will be unpacked as the follows:

Step 1. To identify the given epistemic task as a new token of an old type of tasks that were completed in the system's operating history, and assign the average consuming time of the task-type to the Tr-slot of this new token. And if the system cannot identify this token as something subject to a known type, then it will go to execute step. 2.

Step 2. To identify the given epistemic task as a new token of new type of task, insofar as this type is composed of sub-types of tasks that were completed in the system's operating history, and assign the sum of the average consuming time of each sub-task-type to the Tr-slot of this new task. If the system cannot identify this new task as a compound of known types, then it will go to execute step 3.

Step 3. To view the new task as something entirely beyond the system's capacity of analyzing it, and hence, the system will automatically assign a very high Tr-value to the task. And this move will automatically result in a huge deficit number.(By the way, albeit in a somehow disappointing style, step 3 catches our intuition that entirely unfamiliar problems will bring tremendous stress to problem-solvers)

Q5. Your model can at most explain how epistemic shifts occur rather than provide certain normative principles for telling what type of epistemic shifts are rational. Hence, your model lacks the normative dimension required by typical epistemological inquiries. Isn't this a problem?

A5. This accusation is too general in content, since it can be used to criticize any naturalized approach in epistemology. Moreover, as a matter of fact, my model does have niches to accommodate normativity-based considerations. One of the niches could be found in the K-/Kn-values. To recall, the k-value characterizes subjects' dispositions of overestimating/underestimating the deficit-value, whereas the Kn-value characterizes subjects' thresholds facilitating the transformations of the quantitative representations of the deficit-value into the duality of the conceptual switch of "knowing". Hence, there should be a huge space for probing into meta-rules on how to fine-tune the k-/Kn-values in order to make the resultant agent behave rationally in knowledge ascription. Given that k-/Kn-values are critical for shaping the designed agents' virtues, the elaboration of preceding metal-rules should be also potentially related to the algorithmic realization of virtue epistemology, which is a decent branch of mainstream epistemology. However, how to actualize this potentiality is simply requiring another project.

Q6. Stakes are not explicitly mentioned in your model, but some epistemic scenarios do involve stakes. (Just to think about the High-Stakes Hannah again.). Hence, does your time-deficit-based theory have the potential to account for stake-related factors in order to widen

its own scope of applicability?

A6.   I still believe that stake-related factors do not deserve an independent niche in my model even when its scope of applicability is expected to be widened. The reason is quite simple: stakes with respect to a certain aspect (say, the economic aspect), if disconnected to subjects' desires to minimalize their risks in the same aspect, will not bring stress to them. As for instance, if the High-Stakes Hannah were not caring about financial responsibilities and related social credits, even her awareness of her own bad financial conditions would not lead her to pay enough attention to the epistemic task of figuring out the correct date of depositing the paychecks. Hence, the most relevant factor here should be the subject's willingness of consuming a certain amount of time budgets for completing the relevant task, and wills of this type are definitely not necessarily triggered by stakes. Hence, it should be more appropriate to explicitly mention the parameter characterizing the amount of time intended by the subject to solve the given task, or the "Ti-value" as its abbreviation. Thereby, if the size of the overlapping part between Ta and Ti could be defined as |Ta -Ti|, then we can update Formula-4 as Formula-6:

Formula-6: $D = k\left(\frac{Tr - |Ta - Ti|}{Ts}\right)$

Formula-6 could be used to handle some of the more complicated forms of the bank case, like the following one: [30]

Scenario 7: Hannah is wondering whether the bank, which is definitely open today, will still be open tomorrow, which is Saturday, and her perfect financial status allows her to deposit her paychecks in the bank a couple of days later. And her car dies right beside the bank, and she just keeps on waiting her tow truck to come instead of entering the bank to deposit her paychecks, and finally she kills 2 hours' time in her car without doing anything. The question to the ascriber (whom is informed that the bank will be open tomorrow) is: Does Hannah know that the bank will be open on the coming Saturday if her only available evidence is that it was open on the last

---

[30] This case is originally formulated by J. Schaffer, "The Irrelevance of Subjects: against Subjective-Sensitive Invariantism", *Philosophical Studies*, 127 (2006): 87–107.

Saturday?

Intuitively speaking, Hannah does not know in scenario 7, and my explanation of this is: If the ascriber is supposing she is in the similar epistemic conditions as Hannah, she will feel a relatively high D-value, which is not knowledge-triggering. Some readers may feel puzzled by this diagnosis, since it seems that the Hannah in this scenario has much time at hand to increase the Ta-value and hence to decrease the D-value. But according to Formula-6, a certain quantity of available time, say, *2 hours*, cannot substantially help in decreasing the D-value if there is nearly no perceivable overlapping part between the Tr-value and the Ti-value, or in other words, two hours' time means nothing if the subject does *not intend* to efficiently use it. (However, why Hannah is so idle in this scenario is a relatively trivial issue.)

§6. Concluding Remarks

Metaphilosophically speaking, the "original sin" committed by post-Gettier Anglophone epistemology is theorists' overreliance on both the intuitions triggered by deliberately designed thought-experiments and the formal tools inherited from the Fregean tradition. But the problem is: ordinary people's intuitions are not directly triggered by rationalized reasoning favored by professional epistemologists. Rather, they are merely deliveries of their cognitive architectures, which are nothing but information-processing mechanisms evolved to be adaptive to environmental conditions. Hence, a clever management of time budget has to be a critical sub-function of these mechanisms, given that the time resource is one of the most important resources that an adaptivity-oriented cognitive system has to efficiently exploit. And the efficiency of such exploitation looks hopeless when there is no time-budget-deficit perceiving module to monitor it on a higher level, just like the efficiency of using, say, the domestic gas, would be hopeless when the gas valve or the gas meter were not doing its job. And such a time-resource-monitoring module is indispensable even for knowledge ascriptions, given that the

29

competence in appropriately freezing the process of fixing truth-values of certain beliefs would be of tremendous evolutionary values. Hence, a plausible account of epistemic shifts has to be based on a time-deficit-monitoring module, the putative algorithmic principles of which have been provided by this paper.

However, the preceding observation does not imply that my epistemic-shifts-predicting model cannot be used in purely speculative (and hence evolutionarily insignificant) scenarios like those on the Cartesian demons or the Brain-in-the-Vat and so on. Rather, as I have shown, my model can easily handle skeptic cases, since the widely perceived difficulties of responding to skeptic challenges, according to my theory, are based on nothing but the salient difference between the time budget required to respond to them and the time budget actually available at hand. In this sense, my theory has a wide scope of applicability from the bank case to varieties of skeptic cases. Furthermore, due to the algorithmic characterization of my model, it also has the potential to be serving for the goal of AI.