A model for multiple appearances based on Williamson's GCEL

Irene Bosco

Abstract

Human epistemic subjects cannot but employ imperfect and limited tools to gain knowledge. Even in the seemingly simple business of acquiring knowledge of the value of a physical quantity, what the instrument reads or perception tells more often that not does not correspond to real value. However, even though both our perceptual apparatus and measuring instruments are sensible to background noise, under certain conditions, collecting more information of the same quantity using the same tools leads to an improvement of the subject's epistemic condition. The aim of this paper is to formalize this intuition employing a model which extends Williamson (2013) on the arising of Gettier cases in epistemic logic. From a model where each world is univocally represented by the pair $\langle r, a \rangle$ defined by the real value and the apparent one, I will be considering a scenario wherein the subject has collected n appearances of the real parameter, so that each world is identified by a $n+1$-tuple.

On the assumptions that the n data are all independent and regarded by the subject as equally reliable, a plausible model should be able to describe three phenomena. First of all, the more coherent and closer to represent the real value the set of n appearances is, the more knowledge of the real value is gained. Secondly, the more the subject goes on acquiring "good" data, the more she knows. Thirdly, the world where the n apparent values and the real one match is the best world for the subject. Following Williamson, I shall put forth a system for multiple appearances in epistemic logic that aims at modeling these phenomena by describing the interval within which the real value is known to lie.

1 Motivation and desiderata

It is common practice in physics and the natural sciences to establish the magnitude of a given physical quantity relying not just on a single measurement but on a collection of values obtained through the same measuring process in similar conditions. For example, to gain knowledge of the temperature of a room, several measurements taken using thermometers of the same brand would be collected and their data analyzed. The result will be a value, or more realistically a range, for the temperature. Similarly, a person trying to tell the height of a tree would be willing to consider what people around her looking at the same tree would have to say, and would gain confidence in her judgment as more people agree on it. The intuitive idea behind these phenomena is that the number and the quality of the set of data affect what the subject knows about the real value. In this paper I shall attempt to provide a model that formalizes this intuition, restricted to cases where I believe our pre-theoretical intuitions to be more uniform and less controversial. However, a formal system that works for simpler cases is a valid starting point should one be willing to expand and investigate cases that are not as straightforward.

On measuring

Before digging into the technical details of the model, it is important to pause for a moment on an essential feature of the process of measuring and of gaining knowledge of the external, sensible world more in general. Our measuring tools, be them complex instruments or perceptual apparatuses, are not perfect. Assuming that there is a real value of the magnitude of a physical quantity that could be known, the value returned by a tool, the so-called apparent value, does not necessarily match the real one. Moreover, this mismatch is usually not deterministic: the process of measuring is affected by small background noises that, even in similar conditions, make it the case that whenever we perform a measurement, we could get a different value each time. Likewise, a group of people with similar perceptual capacities and abilities is likely to report different heights of the same tree.

If the instrument is well-designed, the gap between reality and appearance is rarely wide; under the assumption that the measurements are independent and that the differences in the data set are due to random errors with expected value zero, we should be confident that for a big enough number of appearances, the errors will cancel out and that the average value of the appearances will converge to the real value. However, as a matter of metaphysical possibility, we don't want to rule out that it could be possible to be in a case where not just random errors, but systematic ones or perceptual illusions may occur as well. As a result, appearance might substantially diverge from reality, and apparent values in turn might or might not diverge from each other. The case I will be considering assume that all the apparent values collected are independent of each other and that we have no reason to judge that a measuring tool is particularly worse than another. In these cases, all the apparent values are assumed to have the same weight: there is no reason to regard one more or less reliable than another. This is indeed an idealization; nonetheless, given the restricted family of cases that will be the target of the model, it is a fairly plausible one.

Three desiderata

More specifically, our model is designed to capture three phenomena that should take place in idealized scenarios where a subject is gaining knowledge of an environmental parameter just by relying on some apparent values, which are all independent and equally reliable. Intuitively, how far away reality is from appearance is given by the reality-appearance gap. To help our discussion, we shall introduce the following taxonomy for the cases I will try to model.

The "Observationally good" case (OG)**.** A case wherein the appearances collected have all identical values.

The "Extremely good case" (EG)**.** An OG case where reality matches appearance.

The "Statistically good case" (SG)**.** A case where the mean of the appearances matches reality.

Therefore, the results of this paper will be mainly limited to OG cases; although pretty rare in real life, finding a way to formalize our pre-theoretical intuitions about knowledge in these cases is a prerequisite for any more general system that aims at modeling other cases as well. Two basic conditions must be met by any such a model:

The Fitter, The Merrier In OG cases, for a fixed number of appearances, a set of values with a smaller reality-appearance gap allows the subject to acquire more knowledge than a set that displays a greater gap.

The More, The Merrier In OG cases, for a fixed reality-appearance gap, the more appearances the subject collects, the more knowledge she gains.

If the first condition is not peculiar of a multiple-appearance model, the other is meant to capture the point of collecting more measurements or asking for the opinion of peers, keeping in mind that quantity is not always the key to knowledge. To illustrate the above points more precisely, let us consider the following scenarios.

Suppose we have collected three measurements of a physical quantity, all with the same value, e.g. $a_1 = 5, a_2 = 5, a_3 = 5$ and suppose that the real value is actually 5: we are in an EG case. Then consider the case where the real value is 5, however the apparent values are $a_1 = 3, a_2 = 3, a_3 = 3$. Our intuition is that more knowledge of the real value is gained in the former case than in the latter. This intuition is conveyed by *The Fitter, The Merrier* desideratum.

Now imagine that an extra appearance is collected. If the appearance-reality gap has not changed, viz. if the value of the extra appearance is a 5 or a 3 respectively, more knowledge has been gained; this is the point of the *The More, The Merrier* requirement. Lastly, a third desideratum could be added as well, which concerns the best case for the subject who has collected n data, not necessarily all matching:

The Best-Case Principle For a fixed number of appearances, the EG case is the world wherein the subject knows the most.

The subject knows more when she sees only 5s than in any other possible case with the same number of appearances. The aim of this paper is to design a model that satisfies these conditions.

2 Starting point: Williamson's model

The model that I shall advance will be a generalization to multiple appearances of the model that Williamson designed to show the possibility of Gettier cases (Williamson, 2013); as such, it is helpful for our discussion to recall that model and its details, especially regarding the constraints. We are asked to imagine an ideal subject who is gaining knowledge of an environmental parameter just by learning about its apparent value, which is maximally specific: exactly one value is displayed. To formalize this, Williamson designs a model in epistemic and doxastic logic which consists in a frame $\langle W, R, S\rangle$ where each possible world in W can be uniquely identified by the pair $\langle r, a\rangle$ where r, a are the real and apparent value respectively. For ease of exposition, I shall follow Williamson[1] and take $r, a \in \mathbb{R}$, so that each possible world could be univocally identified with a pair of real numbers. Propositions are subsets of W and R and S denote the epistemic and doxastic accessibility relation respectively; $R(w)$ and $S(w)$ refer to the strongest proposition that a subject knows and is justified in believing[2] at a world w.

Before defining the accessibility relations, thus establishing what is known or believed at a world, it is important to pause on the constraints of this model, for our model will rely heavily on them in order to get the desiderata right.

Luminosity The subject knows the apparent value, so that any world with a different apparent value is not epistemically accessible. Note that sameness of appearances is not a sufficient condition for

[1] I shall be using Williamson's notation as well, so that $w, u, v...$ range over worlds. The only difference is in the choice of letters to denote the reality-appearance pair, which for me will be $\langle r, a\rangle$

[2] Williamson needs the doxastic accessibility relation to track not merely the belief operator but the stronger one of "justified belief". This stronger sense is plausible in the multiple appearance model as well, but it is not an essential feature of the model

accessibility; if that were the case, for all we know the real value could be any value on the real line, yielding to crass skepticism.

Actuality The real value r is always epistemically accessible.

Inexactness Even when the apparent value matches the real one, for all that the subject knows, she could be in a world wherein the real value is different from the apparent one.

Convexity If a world is accessible from another, then all in-between worlds are as well: by *Inexactness*, for all I know I could be in a world where r is different than a, say it is a bit higher; *Convexity* implies that I could also be in a world where r' is greater than a and less than r[3]

Distance For a given apparent value, if the reality-appearance gap in w_1 is wider than the gap in w_2, then the subject knows less in w_1 than in w_2.

Appearance-centering Knowledge of the real value is cashed out as an interval centered on the apparent value.

Distance imposes that a subject in a world where reality and appearance diverge substantially knows less than what she would know in world wherein the gap is narrower. Since in principle there are no metaphysical limits to how much reality and appearance could come apart, the wider the gap grows, the the less the subject will know; however, unless the gap were infinitely large, scenario that we can rule out a priori, the subject will still come to know something non-trivial about the real value r. Through the use of the metric $d : \mathbb{R}^2 \to \mathbb{R}$, the reality-appearance gap could be easily quantified, for example as the absolute value $|r - a|$.

Given two arbitrary worlds $w = \langle r, a\rangle, w' = \langle r', a'\rangle$ the epistemic accessibility relation R envisaged by Williamson is:

$$wRw' \Leftrightarrow d(r', a') \leq d(a, r) + c \wedge (a = a')$$

where c is a positive, non-zero constant real number; I shall call c the "inexactness value".

It is straightforward to see that this definition satisfies all the above constraints. As a result, what the subject will know is that the real value falls within a certain range. This is crucial, for it follows that gaining knowledge will be cashed out in terms of coming to know an interval $I \subset \mathbb{R}$, which I shall dub

[3]Note that *Convexity* is an idealization that is more plausible in scenarios where the set of the real/apparent values is continuous, like $\mathbb{R}$. Indeed, in a case of a discrete set, *Convexity* loses its appeal: suppose that the subject wants to know on which day of the week a particular event is, if a Thursday-world is accessible from a Tuesday-world, it does not make much sense to allow for the Wednesday-world to be accessible as well.

the “accessibility interval”.

A nice upshot is that the phenomenon of gaining more knowledge, i.e. of improving one’s epistemic condition, could be neatly quantified. Acquiring knowledge amounts to reducing the length of the accessibility interval, which is uncontroversially represented by the absolute value of the difference of the two endpoints.

When the subject is at world w, reads a and the real value of the environmental parameter is $r = a$, for all that she knows she could be in a world w' where $r' = a \pm c$. Therefore, the strongest proposition known by a subject at w is $R(w) = r \in [a - c, a + c]$, whereas the strongest proposition known at a generic world $w = \langle r, a \rangle$ where $r \neq a$ is that $r \in [a - |r - a| - c, a + |r - a| + c]$. The length of the accessibility interval I at world w is:

$$|I(w)| = |a + |r - a| + c - (a - |r - a| - c)| = 2(|r - a| + c)$$

As such, it depends entirely on two factors: the inexactness value c and the reality-appearance gap $|r - a|$, which I shall be denoting by the function $G(r, a)$.

3 Moving on: multiple appearances

Recall that the scenario we are trying to model is that of a subject or a group who is acquiring knowledge of an external parameter with value r by collecting a finite number n of data, the appearances, which are supposed to be independent from one another but caused by the same environmental parameter. For the sake of simplicity, I shall assume that the case is that of a single epistemic subject who has access to a finite number n of appearances. For example, the value of the temperature as read using n different thermometers the subject has no reason to trust one more than another. A further constraint that we might want to add at this point is that the order in which the subject has been presented with the n appearances does not matter: we have no reason to consider a particular order of appearances more significant than another. Thus, we want the sample of appearances to form a multiset S of cardinality n; for the sake of simplicity we shall be talking about the *set* of appearances, with the understanding that multiple instances of the same elements are allowed. Formally, a world is defined as an equivalence class of $n + 1$-tuple: $\langle r, a_1, a_2, ..., a_n \rangle$ where the first element is the real value of the physical quantity and the following n are the apparent values. The equivalence relation $\sim$ is defined on the set of ordered $n+1$-tuples as $\langle r, a_1, a_2, ..., a_n \rangle \sim \langle r', a'_1, a'_2, ..., a'_n \rangle$ if and only if the real values are the same, $r = r'$, and the respective sets of appearances S and S' are the same. This means that, given the ordered $n+1$-tuples

$w = \langle r, a_1, a_2, ..., a_n \rangle$ and $w' = \langle r, a'_1, a'_2, ..., a'_n \rangle$, there is a permutation $\tau : [1, n] \to [1, n]$ such that $a_j = a'_{\tau(j)} \quad \forall j \in [1, n] \quad \forall a_j \in S$.

For ease of exposition, henceforth the term "world" and the letter w will denote an equivalence class of ordered tuples, so that the accessibility relations are to be understood as holding for $n+1$-tuples in different equivalence classes as well.

Now let's turn to the implications of Williamson's constraints for our model. Some requirements will naturally carry over:

Luminosity For n appearances, it implies that a necessary condition for a world to access another is that the sets of the n appearances are exactly the same. If this were a sufficient condition as well, then for all the subject knows, the temperature could be any real value: this consequence is obviously unwanted.

Moreover, since our measuring tools are not perfectly discriminating, we need also:

Actuality The actual value is always epistemically accessible.

Inexactness Given a set of n appearances, for all that the subject knows she could be in a world where the real value is not exactly the one conveyed, somehow, by such a set.

Convexity Given that our scale to measure the environmental parameter is the real line, it seems natural to assume that all those real values that fall between the real value at the subject's world and that of one accessible from there are epistemically possible.

Moreover, since the real value of the temperature is given by a single value, when we have several appearances, possibly distinct, the natural analog to Williamson's model is to take their arithmetic mean as central point.

Mean-centering The accessibility interval is centered on the mean value of the appearances.

Distance is the constraint that is the least easy to generalize to a case of multiple appearances. The underlying idea that more knowledge is lost as reality and appearance diverge remains, yet now the reality-appearance distance itself is a strange gap between a single value and a set of apparent ones. In OG cases, the set could be easily represented by the repeated apparent values, so we would have to take into account the fact that it is repeated only. The matter becomes much more complex for arbitrary cases; in the next section I shall provide a somewhat general discussion of how much the subject knows could respect *Distance* in multiple-appearance scenarios.

4 Describing the accessibility interval

Recall that the initial desiderata regard a change in the amount of knowledge; in the light of the above considerations, knowledge is cashed out as an interval centered on the mean. According to Williamson's model, the total width of the accessibility interval $|I|$ is a function of c and G. Recall that for OG cases, all that we need is to have c and G depend on the real value r, the apparent value a and the number of appearances n. For less good cases, more inventive work has to be done. As we shall see in this section, the definition of the accessibility interval must take into account not only the number of appearances, but also their variance σ and their mean $\bar{a}$. A viable mathematical definition of both functions c and G will be advanced in the last section; here I shall limit myself to justifying my choice in less formal terms.

4.1 The reality-appearance gap G

The purpose of this value is to represent of how far away the set of appearances, somehow regarded as a whole, is from reality. In Williamson's model, G was simply the distance between two points, r and a. The natural extension of this definition to a case of n apparent values is to take the distance of the real value from the mean: $G = d(\bar{a}, r)$. Thus, if we follow Williamson and keep the inexactness value c constant, the accessibility interval at a world $w(r, \bar{a}) \in W_n$ is defined as:

$$|I| = 2(|\bar{a} - r| + c)$$

Unfortunately this model would not satisfy the desiderata. Suppose the subject is in an OG world where the real temperature is $82F$ and she has collected 2 appearances: $a_1 = a_2 = 80F$. Since $\bar{a} = 80F$, the reality-appearance gap is $G = 2$. Then she keeps measuring and her data sample grows into 100 appearances, all with value $80F$. I think that the intuitive response is that the subject has gained knowledge by performing 98 measurements more. However, if the accessibility interval depended only on G and a constant c, in both worlds the strongest proposition known would be that $r \in [78 - c, 82 + c]$. The reality-appearance gap has not varied, thus this model clearly fails to meet the *The More, The Merrier* condition; as such, it is not a viable one.

Neither would do to consider the total distance of r from the set of apparent values, viz. the total sum of each distance $|r - a_i|$, so that $G = \sum_{i \in [1,n]} |r - a_i|$. Then $G_2 = 4$ and $G_{100} = 200$, so that we would reach the absurd conclusion that the subject's knowledge has decreased. To conclude, so far the best way to describe the reality-appearance gap G is to cash it out as the distance between the real value and the mean of the apparent ones.

4.2 The inexactness value c

Once established a viable definition of G and learned that, contrarily to Williamson's model, c cannot be a constant value, we are left with the more creative task of providing a description of the inexactness value that lets the model satisfy the desiderata. Arguably, in the case of a single apparent value, the choice of the inexactness value has to do with the sensitivity of the measuring tool or of the subject's perceptual apparatus. For example, if we are measuring the temperature of a room, we expect c to have an order of magnitude of one; on the contrary, if the temperature of a star is the physical quantity at hand, then the order of magnitude of c would be at least twice higher. In brief, we can think of the inexactness value as depending on a given value C that is meant to subsume all these features, which depends on a given scenario but that is constant at every world. Moreover, the inexactness value has to decrease, as n gets larger: one could be tempted to define c as:

$$c = \frac{C}{n}$$

However, this option is not feasible. Imagine a scenario where the real value $r = 7$ and the subject has collected four perfectly matching appearances of 7 at world w. Then consider a world u with the same $r = 7$ where the sample is comprised of four appearances, $a_1 = 4, a_2 = 5, a_3 = 8, a_4 = 11$. Then

$$c_w = \frac{C}{4} = c_u$$

This identity still obtains when we take the reality-appearance gap $G = |\bar{a} - r|$ also into account, for both are SG cases, so that $G_w = G_u = 0$. Thus, the total lengths of the accessibility intervals is $|I_w| = \frac{C}{2} = |I_u|$. Thus, according to this model, the subject would not know more in w, which is an EG case: this model fails to meet *The Best-Case Principle.*

As a consequence, our definition of c has to take into account another feature of the sample of data, which helps discriminating among cases with same number of appearances and identical mean.

The variance

Intuitively, the closer a single appearance is to the real value, the better. In statistics, the variance represents how far a set of numbers are spread out from their average value; likewise, we can define a variance with respect to the real value, which represents the spreading out of appearances from the real value. This value, denoted by σ, encodes the difference between the two cases in the example above: given the same number of appearances and identical mean value, we can say that we know more in an EG case for the variance is smaller, actually null. The mathematical definitions will be advanced in the last section; for the time being, we could informally conclude that, other factors being equal, a smaller variance is better than a greater one. Moreover, in OG cases the variance is exactly the value of G, so that the reality-appearance gap and the number of appearances are all that matters, in accordance with what claimed by the first two desiderata. Any OG world could be univocally characterized just by the real value, the apparent one and the number of appearances.

It is natural to wonder whether we could design a model where the accessibility interval depends just on variance and number of appearances, viz. on the inexactness value $c(n, \sigma, r)$. Thus we could have $c = |I|$, disposing of the G-factor altogether, and rephrase the initial desiderata in terms of variance instead of the reality-appearance gap. This would be tempting, but a model that disregards G would be feasible for OG cases only. Yet, OG cases in themselves are not that interesting: they serve us as a starting point, but a model that could naturally be expanded to cover also more general cases should be our final goal. For such cases, we need G.

Consider two worlds, w and u, where the real value is $r = 4$; at w the subject has collected six apparent values, $a_1 = 3, a_2 = 3, a_3 = 3, a_4 = 5, a_5 = 5, a_6 = 5$, whereas at u she had 6 appearances of $a' = 3$. Computing the variances, it turns out that $\sigma_w = \sigma_u = 1$. Since the number $n = 6$ is they same, we have that $c_w(6, 1, 4) = c_u(6, 1, 4)$. However, intuitively the subject knows more in w than in u. In order to achieve this result we would need to add the respective reality-appearance components $G_w = 0, G_u = 1$, for the mean values are $\bar{a} = 4$ and $\bar{a'} = 3$ respectively.

To conclude, in light of the introduction of variance, we can put forth a fourth desideratum that goes a little beyond OG cases:

Supremacy of Statistics For a fixed number of appearances, for a fixed variance, the SG case is the world wherein the subject knows the most.

Note that if we limit ourselves to OG cases, then the only SG case is the EG one, so *The Best-Case Principle* is exactly a special case of the above.

5 A model for multiple appearances

Once we have described the components that make up the length of the accessibility interval, it is time to describe how this interval is actually generated, i.e. to see why it corresponds to the strongest proposition known by the subject at a specific world. In order to do that, we need a definition of an epistemic accessibility relation for a model of multiple appearances. I will now advance the general relation, then discuss how it fares with respect to single-appearance cases.

5.1 Generalized R

For any $n \in \mathbb{N}$, let $\langle W_n, R_n \rangle$ be a frame, with $W_n \subseteq \mathbb{R}^{n+1}$ a non-empty set of possible worlds univocally represented by $n+1$-tuples where the first element is the real value r, and the other $a_1, ..., a_n$ are the values of the appearances: $w = \langle r, a_1, a_2, a_3, ..., a_n \rangle \in W_n$.
Let $w = \langle r^*, a_1, a_2, a_3, ..., a_N \rangle, w' = \langle r', a'_1, a'_2, a'_3, ..., a'_N \rangle$. For a world to be accessible from another, the sets of multiple appearances $S_1 = \{a_1, ...a_N\}$ and $S_2 = \{a'_1, ..., a'_N\}$ have to be the same.
Let σ^*, σ' be the variances of the sets S_1, S_2 respectively, and let $\bar{a^*}, \bar{a}'$, be the respective mean values. Note that $\sigma = \sigma'$ and $\bar{a^*} = \bar{a}'$; let $G(r, a)$ be the measure of the reality-appearance gap, and $c(n, \sigma, r)$ be the inexactness value as described above.
For any $N \in \mathbb{N}_0$, the epistemic accessibility relation R_N is a binary relation on W_N defined as:

$$wR_Nw' \Leftrightarrow G(r', \bar{a^*}) \leq G(r^*, \bar{a^*}) + c(N, \sigma^*, r^*)$$

This definition of epistemic accessibility is a natural generalization of Williamson's, to the extent that it is straightforward to prove that all formal properties of R carry over into R_N. For any natural N, R_N is reflexive, yet neither symmetric nor transitive. The logic is closed under entailment, and the strongest proposition known at w, $R_N(w) = \{u : wR_Nu\}$, is that r lies within an interval centered on the mean value $\bar{a^*}$, whose length is

$$|I_N(w)| = 2(G(r^*, \bar{a^*}) + c(N, \sigma^*, r^*))$$

5.2 Generalized S

We can expand the model with a doxastic accessibility relation that is a natural generalization of Williamson's. In his model, Williamson defines the doxastic accessibility relation in terms of the epistemic one and of the Good Case w_G:

$$wSw' \Leftrightarrow w_G R w'$$

I follow suit, and for any $w, w' \in W_N$ as above, for any $N \in \mathbb{N}_0$ the generalized doxastic accessibility relation S_N is defined as:

$$wS_N w' \Leftrightarrow w_{SG} R_N w'$$

What the subject at w is justified in believing is what she would know were she in the corresponding Statistically Good case, where $r = \bar{a^*}$. Hence, since $G_{SG} = 0$, the strongest justifiedly believed proposition at w is that

$$r \in [\bar{a^*} - c(N, \sigma^*, \bar{a^*}), \bar{a^*} + c(N, \sigma^*, \bar{a^*})]$$

Note that the value of the inexactness value in this world could be know: c depends on the variance computed using the mean of appearances and their number, indexes that are available to the subject. Also, if a world is doxastically accessible, then it is epistemically accessible from one with the same set of appearances. The principle that Knowledge entails Belief also follows, as well as the principles of Weak Positive and Negative Introspection.

Therefore, in either case the subject knows that she is justified in believing the real value to be within an interval centered at the mean of the appearances, whose length is given by a certain inexactness value which depends on the quality of her data sample. No matter in what particular world she currently is in, a skeptical case or an EG one, the strongest proposition believed is across all worlds with the same appearances. In the rest of the section I will show that R_n and S_n are a continuous extension of Williamson's R and S. In the next one we shall see the model at work.

5.3 Case $n = 1$

For our model to be continuous with Williamson, R_n reduced to the case of a single appearance must be equal to R, ditto for S_n. When we have a singleton $S = \{a\}$, where $a \neq r$, then $\sigma = (r - a) \neq 0$ and $\bar{a} = a$. Let $w = \langle r, a \rangle, w' = \langle r^*, a' \rangle$ and $a = a'$. $G(r, a) = |r - a|$, then

$$wR_1 w' \Leftrightarrow |r^* - a| \leq |r - a| + c(1, r - a, r)$$

The only difference regards the inexactness-value component. Recall that Williamson envisaged a constant inexactness value, whereas now c depends on the distance of the apparent value from reality: the inexactness value is different at each world. Williamson's initial constraints would be met anyway; however, the difference in knowledge required by *Distance* will grow faster. The strongest proposition know is that

$$r \in [a - |r - a| - c(1, |r - a|, r), a + |r - a| + c(1, |r - a|, r)]$$

and the interval has length $|I_1(w)| = 2(|r - a| + c(1, r - a, r))$. Clearly, as a gets closer to r, the interval will get shorter and shorter.

The "Good case". When appearance matches reality, at the world $w_G = \langle a, a \rangle$, we have that $\sigma = 0$ and $\bar{a} = a$: the average corresponds to the real value, and we are trivially in an EG case. It follows that $c(1, 0, a) = C$, thus the value of c, being independent from r, could be known by the subject. Since only those worlds wherein $G(r, a) \leq C$ are accessible, the accessibility interval is $|I_1(w_G)| = 0 + 2C = 2C$. This is the shortest interval for $n = 1$, exactly as it is the case in Williamson's model.

Moreover, given that S_n is defined in terms of R_n and the SG case, it follows that the generalized doxastic relation restricted to a case $n = 1$ is defined in terms of R_1 and w_G, yielding exactly Williamson's relation S.

6 A satisfactory model?

To get a better idea of whether the one advanced above is a viable model, I shall put it to work, and asses how it deals with more concrete examples. To work out the exact numeric values I will be using the function for the inexactness value $c(n, \sigma, r)$ and the reality-appearance gap G given in the following section. For ease of exposition, in line with the order of magnitude of the data in the examples, I will be assuming that the constant $C = 1$.

Some concrete examples

Recall the first two examples given in section one to arise our intuitive responses. In this section I shall formalize them, to see if the model conveys the original intuitions, hence if it meets the initial desiderata. In addition, I will consider a world v where the appearances are all distinct. This world is helpful to test the potential of the model with respect to cases that are not so intuitively settled.

Same n

Let's start by analyzing the three worlds when the number of appearances is kept fixed. As before, assume $n = 3$, $r = 5$; apparent values for w, u, v and relative statistics are reported in the tables below, where $\mathbf{K}|I|, \mathbf{B}|I|$ denote the width of the accessibility and doxastic interval respectively.

3-appearance case

	w	u	v
r	5	5	5
a_1	5	3	5.5
a_2	5	3	3.5
a_3	5	3	5

3-appearance case Statistics

	w	u	v
σ^2	0	4	0.83
$\bar{a}$	5	3	4.66
c	0.33	0.39	0.34
G	0	2	0.33
$\mathbf{K}\vert I\vert$	0.66	4.78	1.34
$\mathbf{B}\vert I\vert$	0.66	0.66	0.68

Strongest Proposition **K** and **B**

world	**K**	**B**
w	[4.66,5.33]	[4.66,5.33]
u	[0.60,5.38]	[2.67,3.33]
v	[3.98,5.32]	[4.32,5.00]

It's no surprise that the world showing the smallest variance and the smallest reality-appearance gap is the EG case w. On the contrary, u has the highest variance, and also the widest reality-appearance gap. Since $\mathbf{K}|I(w)| < \mathbf{K}|I(u)|$, the first desideratum of *The Fitter, The Merrier* is met. Unbeknown to the subject, w is a fitter world, for the appearances match reality: she knows more of the environmental parameter by being in w than by being in u. Yet, the width of the interval of what the subject is justified in believing is the same in both worlds: this is the case because in the eyes of the subject the two worlds are exactly equally good.

Even though v is not an OG case, it could be beneficial to compare it with the other worlds. Again, $\mathbf{K}|I(w)| < \mathbf{K}|I(v)|$, so that the *The Best-Case Principle* is satisfied. Indeed, we did not have any initial expectations regarding which case between u and v would give the subject more knowledge; our intuitions have been regarded as silent on this sort of arbitrary cases. However, our model seems to suggest that the appearances collected at v are not much worse than those of w. The variance is not much higher, and the value for the inexactness value differ by a 0.01. That said, in w the reality-appearance gap is null, whereas in v it is a fifteenth of the real value. As a result, the accessibility interval in v is clearly shorter than in u, for both the values of c and G are: appearances available in u, as a whole, are further away from reality than in v. Therefore, if one is willing to carry over the idea of *The Fitter, The Merrier* desideratum beyond OG cases, then our model seems to be on the right track.

Gaining an extra appearance

Now suppose, that the subject in each of the worlds above is presented with another appearance; the numbers are the following:

4-appearance case			
	w	u	v
r	5	5	5
a_1	5	3	5.5
a_2	5	3	3.5
a_3	5	3	5
a_4	5	3	4

4-appearance case Statistics			
	w	u	v
σ^2	0	4	0.875
$\bar{a}$	5	3	4.5
c	0.25	0.29	0.26
G	0	2	0.50
$\mathbf{K}\|I\|$	0.50	4.58	1.50
$\mathbf{B}\|I\|$	0.50	0.50	0.52

Strongest Proposition **K** and **B**		
world	**K**	**B**
w	[4.75,5.25]	[4.75,5.25]
u	[0.71,5.29]	[2.75,3.25]
v	[3.75,5.25]	[4.24,4.76]

At world w, gaining another matching appearance resulted in gaining more knowledge: $\mathbf{K}|I_4| < \mathbf{K}|I_3|$. This is the case since $c_4 < c_3$, although the variance has not changed. Note that the variance has not changed in u either: here as well more knowledge has been gained thanks to the extra 3, and indeed the subject is more justified in believing the real value to be closer to 3 than before. It is clear that the more coherent appearances one has, the smaller both intervals become, until we reach a limit point. Suppose the subject in u has an infinite number of appearances of 3: the strongest proposition she could ever come to know if she keeps collecting 3s is that $r \in [1, 5]$ and her strongest justified belief is that $r = 3$. To sum up, the *The More, The Merrier* condition is satisfied in both worlds w and u, where the distance reality-appearance has not varied.

Discussing what happens in v is not so straightforward. First of all, compared to the case of just 3 appearances, the variance has slightly increased, yet c has decreased anyway. This is due to the fact that a minor change in σ cannot make up for a change of the factor n. Admittedly, a constant $C = 1$ seems a bit too high: after-all it is a fifth of the real value, it is not very realistic; if we were adopting a smaller C, $c(4)$, although less than $c(3)$ would be much smaller than G: a smaller inexactness-value component would not reduce the width of the interval. However, the average as well has moved further away from the real value. Thus, it looks like the subject's epistemic condition has slightly impaired since the 3-appearance case. The intuitions that I have is that she does not know more, even if she has collected more appearances. Thus if one is willing to retain the intuition behind the *The Fitter, The Merrier* principle for a case of distinct appearances and different numbers, since $\mathbf{K}|I_4| > \mathbf{K}|I_3|$ in v, our model would satisfy it.

The power of Statistics

Lastly, let us consider also a world s of four distinct appearances where the mean matches the real value $r = 5$ and the variance is the same as in v.

s and v		
	s	v
r	5	5
a_1	5.6	5.5
a_2	6.2	3.5
a_3	4.3	5
a_4	3.9	4

s and v Statistics		
	s	v
σ^2	0.875	0.875
$\bar{a}$	5	4.5
c	0.26	0.26
G	0	0.50
$\mathbf{K}\|I\|$	0.52	1.50
$\mathbf{B}\|I\|$	0.52	0.52

Strongest Proposition **K** and **B**		
world	**K**	**B**
s	[4.74,5.26]	[4.74,5.26]
v	[3.75,5.25]	[4.24,4.76]

Clearly, s is a SG case, hence *Supremacy of Statistics* is satisfied for $\mathbf{K}|I(s)| < \mathbf{K}|I(v)|$. Finally note that $\mathbf{K}|I(s)| \sim \mathbf{K}|I(w)|$: it follows that under the same conditions as above, as the number of appearances increases, $\mathbf{K}|I(s)|$ approaches to $\mathbf{K}|I(w)|$, so that knowledge in a SG case is equal to knowledge in a EG case. However, the SG case is the case one would expect to represent a scenario where the difference of appearances is due to small errors randomly distributed, for example when all thermometers are actually equally good at measuring the temperature. Thus, under suitable, though realistic, conditions, the subject's knowledge derived from a set of multiple and distinct appearances is the same as that conveyed by a realistically more improbable set of all-matching apparent values.

7 Mathematical details of the model

In the previous two sections I have been working with a description of the accessibility interval that depended on the inexactness value and the reality-appearance gap, on the understanding that there was a mathematically description of them. Finally, in this conclusive section I shall introduce two functions meant to describe G and c. They are by no means the only possibilities, nor maybe the best one. However, I will defend my choice by showing that the resulting model represents a natural extension of Williamson's and that, as informally anticipated already in the previous section, the desiderata mathematically follow. Let $S = \{a_1, ..., a_n\}$ be a set of n appearances available to the subject, and assume as above that set

$S \subset \mathbb{R}$ and that $r \in \mathbb{R}$; we can define the square of variance as $\sigma^2 = \frac{\sum_{i \in n}(a_i - r)^2}{n}$ and the mean $\bar{a} = \frac{\sum_{i \in n} a_i}{n}$.

7.1 The reality-appearance gap

Let $d : \mathbb{R}^2 \to \mathbb{R}^+$ be the metric described by Williamson, viz. $d(a, b) = |a - b|$. Let $\bar{a}$ be the mean of the data in S_n; if we regards $\bar{a}$ as representative of S, then a natural option would be to define the distance as $G = d(r, \bar{a}) = |r - \bar{a}|$. If $\bar{a} = r$, then $G = 0$.

7.2 The inexactness value

My option for the inexactness value c is:

$$c(n, \sigma, r) = \frac{Ce^{\frac{\sigma^2}{r^2}}}{n}$$

Where C is the value specific for each scenario, as discussed in an earlier section. Note that in order to get a pure multiplicative factor for C, we need to divide the variance by the real value. Taking the square of this ratio ensures that the exponent is going to be a positive number, regardless of the sign of r. Clearly, $c(n, \sigma, r) > 0$ for any n, σ, r, and $c \to 0$ as $n \to \infty$

7.3 The accessibility intervals in simple cases

Now I shall analyze what happens to c and G as we move across worlds with the same real value but letting the number of appearances and their variance vary.

Case $n = 1$

If $a \neq r$, $\sigma = (r - a) \neq 0$, and $c(1, (r - a), r) = Ce^{\frac{(a-r)^2}{r^2}}$. $|I_1(w)| = 2(|r - a| + c(1, r - a, r))$

The "Good case". If $w_G = \langle a, a \rangle$, we have that $\sigma = 0$; it follows that $c(1, 0, a) = C$, $|I_1(w_G)| = 2(|r - a| + C)$. Note that $Ce^{\frac{(a-r)^2}{r^2}} \geq C$

Case $n > 1$

The "Observationally good" case (OG)**.** All values of multiple appearances match: $a_1 = a_2 = a_3 = ... = a_N$; call this value α. Suppose $\alpha \neq r$: clearly, $\bar{a} = \alpha$ and $G = |r - \alpha|$, and $\sigma^2 = \frac{N(\alpha - r)^2}{N} = (\alpha - r)^2$. Then $c(N, (\alpha - r), r) = \frac{Ce^{\frac{(\alpha-r)^2}{r^2}}}{N} = \frac{Ce^{(\frac{\alpha-r}{r})^2}}{N}$.
Thus, $|I_N(w_{OG})| = 2(|r - \alpha| + \frac{Ce^{\frac{(\alpha-r)^2}{r^2}}}{N})$.

The "Extremely good case" (EG)**.** $r = \alpha = a_1 = ... = a_N$. In this case $\sigma = 0$, so $c(N, 0, \alpha) = \frac{C}{N}$ and $G = 0$. $|I_N(w_{EG})| = \frac{2C}{N}$. This is the shortest interval for $n = N$. Also, $|I_N(w_{EG})| < |I_{N+1}(w_{EG})|$: the more appearances α the agent collects, the more she knows.

The "Statistically good case" (SG)**.** The N data available to the subject are not all the same, so that $\sigma \neq 0$, yet $\bar{a} = r$. Then $c(N, \sigma, \bar{a}) = \frac{Ce^{\frac{\sigma^2}{\bar{a}^2}}}{N}$. For a fix set S_n of not-all-matching appearances, this case has the smallest value of c. $G = d(r, \alpha) = 0$, thus $|I_N(w_{SG})| = \frac{Ce^{\frac{\sigma^2}{\alpha^2}}}{N}$

Meeting the desiderata

In conclusion, the above definitions of G and c are such that the resulting model satisfies all the fourth desiderata I listed.

- *The Fitter, The Merrier*: Let $n = N$, then $I(|r - \alpha|) = 2|r - \alpha| + 2\frac{Ce^{\frac{(\alpha - r)^2}{r^2}}}{N}$ it's a monotonically increasing function of $|r - \alpha|$: as this reduces, so does the length of $|I|$
- *The More, The Merrier*: Let $|r - \alpha| = k$, then $I(n) = 2k + 2\frac{Ce^{\frac{k^2}{r^2}}}{n}$ it's a monotonically decreasing function of n: as n increases, the length of $|I|$ reduces
- *The Best-Case Principle* Let $n = N$, then $I(\sigma, \bar{a}, r)_N = 2(|r - \bar{a}| + \frac{Ce^{\frac{\sigma^2}{r^2}}}{N}) \geq \frac{2C}{N} = I_N(EG)(|r - \alpha|)$ for any $|r - \alpha|$ and for any N, $I_{EG} = I \Leftrightarrow r = \alpha = a_1 = ... = a_N$
- *Supremacy of Statistics* Let $n = N$, $\sigma = s$, then $I_{SG} = \frac{Ce^{\frac{s^2}{r^2}}}{N} \leq 2|r - \bar{a}| + 2\frac{Ce^{\frac{s^2}{r^2}}}{N}$ for any $|r - \bar{a}|$ and for any N, s we have that $I_{SG} = I \Leftrightarrow r = \bar{a}$

References

T. Williamson. Gettier Cases in Epistemic Logic. *Inquiry (United Kingdom)*, 56(1):1–14, 2013. ISSN 0020174X. doi: 10.1080/0020174X.2013.775010.