

# «Confiabilismo evolucionista» y respuestas «de principio» sobre nuestras capacidades cognitivas.

Claudio Cormick.

Cita:

Claudio Cormick (2019). «*Confiabilismo evolucionista*» y respuestas «*de principio*» sobre nuestras capacidades cognitivas. *Eikasia. Revista de Filosofía*,.

Dirección estable: <https://www.aacademica.org/claudio.cormick/43>

ARK: <https://n2t.net/ark:/13683/pq15/mQy>



Esta obra está bajo una licencia de Creative Commons.  
Para ver una copia de esta licencia, visite  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

*Acta Académica* es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. *Acta Académica* fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

# «Confiabilismo evolucionista» y respuestas «de principio» sobre nuestras capacidades cognitivas

Claudio Javier Cormick

## Resumen

En el presente trabajo, intentaremos exponer la oposición entre dos acercamientos al problema de la confiabilidad global de las capacidades cognitivas humanas, y una posible solución a tal conflicto. Por un lado, como señalaremos, existe un número de acercamientos que caen bajo el amplio nombre de “confiabilismo evolucionista” y de acuerdo con los cuales las razones de las que disponemos para creer en la confiabilidad de la cognición humana son de carácter empírico. A saber, el éxito adaptativo de nuestra especie en un medio biológico caracterizado por la supervivencia del más apto nos proporciona una razón para creer que nuestros mecanismos de formación de creencias rastrean la verdad; si no lo fueran, nos habríamos extinguido. Por otro lado, sin embargo, encontramos análisis —entre los cuales nos enfocaremos en el caso de Ernest Sosa— de acuerdo con los cuales la tesis de que nuestras capacidades cognitivas son confiables no puede estar apoyada empíricamente, y como tal ser contingente; por el contrario, estos análisis señalan que necesitamos *presuponer* —sin necesidad de prueba— que tales capacidades son confiables. Dado este conflicto entre una defensa “empírica” y una defensa “de principio” de nuestras capacidades cognitivas, consideraremos dos posibles objeciones contra la propuesta de Sosa e intentaremos responderlas.

**Palabras clave:** confiabilidad, evolucionismo, cognición, epistemología, Ernest Sosa.

## Abstract

### “Evolutionary reliabilism” and “principle” answers about our cognitive abilities

In this work, we will try to state the opposition between two approaches to the problem of the overall reliability of human knowing capacities, and a possible solution to that conflict. On the one hand, as we will point out, there exist a number of approaches that fall under the broad term of “evolutionary reliabilism” and according to which the reasons that we have for believing in the reliability of human cognition are empirical in character. Namely, the adaptive success of our species in a biological environment characterised by the survival of the fittest provides us with a reason to believe that our belief-forming mechanisms are truth-tracking; if they were not, we would have gone extinct. On the other hand, nevertheless, we find analyses—of which we will focus on the case of Ernest Sosa—according to which the tenet that our knowing capacities are reliable cannot be empirically based, and thus contingent; on the contrary, they point out that we need to *presuppose*—without need of a proof—that such capacities are trustworthy. Given this conflict between an “empirical” defence and a defence “in principle” of our knowing capacities, we will consider two possible objections against Sosa’s proposal and try to answer them.

**Keywords:** reliability, evolutionism, cognition, epistemology, Ernest Sosa.

eikasía

## «Confiabilismo evolucionista» y respuestas «de principio» sobre nuestras capacidades cognitivas

Claudio Javier Cormick

### 1. Presentación general: dos tipos de argumentos sobre la fiabilidad de nuestras capacidades cognitivas

Como es sabido, un cierto número de elaboraciones en el seno de la epistemología anglosajona ha buscado establecer que las leyes de la selección natural nos brindan un argumento de carácter empírico contra las dudas escépticas que podrían poner en cuestión el ajuste de nuestras capacidades cognitivas con respecto a la realidad objetiva. En una palabra, para este “confiabilismo evolucionista”, el hecho mismo de que no nos hayamos extinguido habla a favor de la fiabilidad de tales capacidades, entendida como tendencia a producir, a grandes rasgos, creencias verdaderas. Las raíces de esta propuesta pueden encontrarse, en una forma temprana, en Antony Flew, quien señaló, en el marco de un debate con Ernest Gellner, que la biología contemporánea puede proveernos un argumento para garantizar que de conjunto nuestras creencias han de ser verdaderas, puesto que, “si no lo fueran, los hombres no habrían podido sobrevivir mucho tiempo” (Flew, 1958: 45); en Popper (1972: 261), quien, aunque centrado en describir en realidad el crecimiento y progreso del conocimiento humano que en ofrecer garantías para él, ha señalado -y, según insiste, no a modo metafórico- que “el conocimiento animal [...] crece mediante la eliminación de aquellos que sostienen las hipótesis inadecuadas”; esto es, la selección natural escoge a aquellos individuos menos proclives a incurrir en la falsedad; en Fodor (1981: 121) quien declara que “la selección darwiniana garantiza que los organismos conozcan los principios de la lógica, o fallezcan”, o más recientemente en Boulter (2007: 32-52), para quien las posiciones filosóficas que

reivindican el sentido común pueden encontrar un soporte “neutral”, externo a las disputas de la disciplina, apelando al testimonio de la biología.

Estas propuestas son ampliamente conocidas. Nos interesa subrayar, sin embargo, una particularidad que ellas presentan y que conducirá al conflicto -que desarrollaremos- entre este confiabilismo y otra estrategia de respuesta al escepticismo. A saber, este enfoque empíricamente informado, como todo otro intento de “fundamentar” de conjunto la fiabilidad de nuestras capacidades cognitivas sobre la base de evidencias científicas, presupone que, al menos en principio, nuestras facultades cognitivas *podrían* ser descritas, inteligiblemente, como *no* confiables, y que en consecuencia si consideramos que ellas *son* confiables esto se deberá a que existen evidencias empíricas para hacerlo (en particular, evidencias empíricas, susceptibles de ser desmentidas en este mismo terreno, que parecen conectar la capacidad de supervivencia de un organismo con la de formar creencias verdaderas sobre su entorno). El punto que debemos subrayar, a riesgo de cometer una perogrullada, es que buscar un aval *empírico* para una tesis implica negar que esta pueda apoyarse en consideraciones *de principio*.

Así, pues, un abordaje confiabilista evolucionista no solamente se opondrá al escepticismo: también tendrá un adversario, menos evidente, en la forma de las epistemologías para las cuales la fiabilidad de nuestras capacidades cognitivas puede reivindicarse sobre la base de argumentos de principio, independientes de cualquier dato empírico que las ciencias puedan proporcionarnos. En efecto, la respuesta al escepticismo aparece de esta manera como “sobre-determinada”: si las tesis confiabilistas evolucionistas están en lo cierto, entonces una defensa de principio de nuestras capacidades cognitivas será injustificadamente fuerte, puesto que considerará que estas pueden ser declaradas confiables al margen de los datos concretos que tenemos sobre el modo en que el mundo incide sobre nuestros modos de formación de creencias. A la inversa, si son las argumentaciones de principio las que resultan verdaderas, entonces los argumentos del confiabilismo evolucionista serán excesivamente débiles, puesto que hacen depender de datos contingentes una fiabilidad que podemos establecer sin necesidad de apoyarnos en aquellos. Al menos uno de estos dos tipos de enfoques tiene, entonces, que ser rechazado.

Ahora bien, ¿en qué puede consistir una defensa *de principio*, no empírica, de nuestras capacidades cognitivas, incompatible del modo señalado con el confiabilismo evolucionista? Creemos que hay un ejemplo interesantes para considerar aquí: el llamado “argumento trascendental” de Ernest Sosa en su *Knowing Full Well*, de 2011. Sosa busca responder por medio de este argumento al problema del carácter “accidental” que, desde el punto de vista evolutivo, parece tener el “éxito de nuestras facultades cognitivas” (Sosa 2011: 153ss). De acuerdo al autor norteamericano, el aval de la fiabilidad de tales facultades proviene de un argumento *de principio*, no de uno empírico que podamos tomar de las ciencias. Por medio del experimento mental de comparar nuestra situación con la de un sujeto que tuviera que determinar si está o no bajo el influjo de una píldora que lo incapacita cognitivamente, llamada “Disablex”, Sosa arguye que no tenemos más opción que presuponer -como lo haría, a título individual, tal sujeto- que nuestras facultades *son* fiables; debemos presuponerlo, porque son *las nuestras*, y por tanto las únicas a que podemos apelar. Incluso si, según admite, “hay escenarios concebibles en los que uno adquiere evidencia considerable de que *ha* tomado una píldora tal”, el sujeto que creyera que efectivamente está bajo el flujo de tal droga se comprometería con una posición “epistémicamente auto-derrotadora” (*self-defeating*; Sosa 2011: 155). Del mismo modo, incluso una evaluación de nuestra situación epistémica que se enfocase en las posibles evidencias sobre la *falta* de fiabilidad de tales facultades no podría más que evaluar tales evidencias *utilizando* las facultades mismas que se supone que están puestas en cuestión. Si la tesis según la cual nuestras propias capacidades cognitivas *son* fiables incurre, de este modo, en cierto tipo de circularidad, sin embargo *la única alternativa disponible* (cf. *ibíd.*: 156) es la posición autocontradictoria consistente en apoyarnos en evidencias *disponibles para nosotros* a efectos de concluir que nuestras evidencias no son dignas de crédito. El argumento, pues, no es de índole empírica, dado que señala que tendremos que comprometernos con tal fiabilidad de nuestras capacidades *independientemente de cualquier dato del que dispongamos acerca de ellas*; en rigor, *cualquier* evaluación empíricamente informada que quisiésemos hacer de tales capacidades cognitivas se encuentra en una relación de *dependencia evidencial* con respecto al supuesto de que ellas son, de hecho,

fiables<sup>1</sup>. De hecho, lo interesante del argumento de Sosa es que elige discutir a partir de un caso límite: tan lejos se encuentra de argumentar que debemos apoyar *empíricamente* la confianza en nuestras propias capacidades que el escenario que propone es uno en el cual la evidencia a favor de tal fiabilidad sería *la peor posible*. Si incluso en el escenario de “considerable evidencia” empírica contra la idea de que podemos confiar en nuestras capacidades, la actitud racional radica en afirmar que sí podemos, entonces ciertamente el tipo de aval que tenemos para la tesis de la fiabilidad de aquellas no es, en absoluto, empírico.

Así es que concluye:

¿Por qué es que no tenemos más opción racional que responder por la afirmativa a la pregunta sobre nuestra propia confiabilidad? Bueno, consideremos las alternativas. Supongamos que decimos que no. ¿Cómo entonces podríamos coherentemente confiar en nuestras facultades al apoyar esa misma respuesta? En efecto, supongamos que siquiera suspendemos el juicio sobre la pregunta, encogiéndonos de hombros con un “tal vez sí, tal vez” no, o incluso simplemente evitando juzgar [...]. Incluso aquí, ¿cómo podríamos coherentemente comprometernos con *esta* actitud al mismo tiempo que decimos que no podemos realmente saber si, procediendo así, estamos procediendo de un modo cognitivamente correcto? [...] [A]nte esta pregunta, *solo la respuesta confiadamente afirmativa puede ser plenamente coherente*. Una vez que vemos que esta actitud está racionalmente requerida para ser enteramente coherentes, ello nos da razones para extraer sus consecuencias deductivas, incluyendo (a) que nunca hemos tomado ninguna píldora incapacitante y (b) que nuestras facultades no tienen orígenes *incapacitantes* (por ejemplo, orígenes tales que involucren un engaño poderoso y sistemático) (ibíd.: 156-157).

## 2. Precisando el argumento de Sosa. Una primera objeción a la propuesta: sobre el problema del rechazo de evidencia

Habiendo introducido la propuesta de Sosa como alternativa “de principio” a la clase de argumento antiescético que encontramos bajo la forma de un confiabilismo evolucionista, debemos dar un paso más e intentar reconstruir con algo más de

<sup>1</sup> En respuesta a Alvin Plantinga, Sosa (2002: 102) reconoce que el naturalista que recurriese a sus propias capacidades cognitivas para reivindicar el valor de estas estaría cometiendo una circularidad, pero remarca que este tipo de circularidad no es diferente de la que aquejaba a un *no* naturalista como Descartes. Para Sosa este no es, por lo demás, el tipo de círculo que quepa razonablemente desacreditar como vicioso (1997, 2003, 2009).

precisión en qué consiste el argumento. Cabe desglosar el argumento de la siguiente manera:

1. Para que podamos *razonablemente* inferir que nuestras capacidades cognitivas no son confiables, sería necesario que podamos contar con evidencia de que efectivamente no lo son.

2. La única forma en que podemos obtener tal evidencia, al igual que evidencia a favor de cualquier otra creencia, es *asumiendo la confiabilidad* de nuestras capacidades cognitivas.

3. La actitud descrita en (2) implica adoptar creencias contradictorias sobre tales capacidades.

4. No es posible inferir razonablemente, por lo tanto, que nuestras capacidades cognitivas no son confiables, ni siquiera cuando *aparentemente* tenemos evidencia de que no lo son.

Ante todo, para evaluar la aceptabilidad de este argumento debemos intentar retraducir los términos del experimento mental a la problemática inicial a propósito de la cual el autor lo presenta. Sosa nos dice que dudar de la fiabilidad de las capacidades cognitivas de nuestra especie a la luz de sus orígenes sería, en algún sentido a precisar, comparable con dudar de nuestras capacidades cognitivas individuales a la luz de la evidencia que sugiere que hemos tomado “Disablex”: en uno y otro casos, se supone, el ejercicio de duda no puede realmente estar fundamentado, puesto que la fuente de evidencias para tal duda es precisamente el mismo conjunto de capacidades sobre las cuales incidiría la duda. Ahora bien, ¿qué significaría exactamente esto respecto de los problemas epistemológicos suscitados por una explicación evolucionista de nuestros orígenes? Si tenemos que tomarnos la comparación con el “Disablex” al pie de la letra, la conclusión a extraer parecería ser la siguiente: (i) efectivamente tenemos, en la forma de los resultados de la biología evolutiva, evidencia que al menos *prima facie* parecería desacreditar la fiabilidad de nuestras capacidades cognitivas -evidencia que funcionaría, siguiendo la comparación de Sosa, como el descubrimiento de que probablemente hayamos tomado “Disablex”-; sin embargo (ii) tenemos que resignificar el valor que tiene esa evidencia, puesto que, si efectivamente concluimos lo que ella parece indicar, entonces nos estaremos comprometiendo con una posición epistémicamente



autoderrotadora. En otras palabras: nuestros fundamentos para creer en la proposición “Nuestras capacidades no son confiables” presupondrán, justamente, la proposición “Nuestras capacidades son confiables”.

Ahora bien, estos corolarios son todavía excesivamente abstractos; restan problemas que el defensor de un abordaje empírico puede, en consecuencia, aducir contra la alternativa propugnada por Sosa. En efecto, ¿qué significaría, más exactamente, decir que, al igual que en el caso del “Disablex”, no podemos concluir lo que la evidencia *parece* indicar? Esto parece conducir a un problema adicional, a saber: ¿cómo puede exactamente ser racional la actitud de simplemente desconocer cierta evidencia fáctica?

Veamos estos problemas de a uno. Ante todo, la cuestión de rechazar cierta conclusión aparentemente justificada. El experimento mental es analizado por Sosa en términos tales que queda claro que la conclusión racional, de acuerdo con el autor, es asumir que no hemos tomado la droga en cuestión, pese a las apariencias en contrario (la “evidencia considerable” de la que nos habla Sosa). No se trata de que podamos aceptar la información fáctica que parecen brindarnos nuestra memoria y sentidos (a saber, que sí hemos tomado la droga) pero interpretar de modo no autoderrotador las implicaciones de tal hecho (por caso, asumir que la droga en cuestión no era realmente incapacitante), sino que lisa y llanamente *tenemos que negar que ese hecho haya ocurrido*. ¿Qué significaría esto en el caso que Sosa desea iluminar, el de nuestros orígenes evolutivos? ¿Debemos asumir, por analogía con el experimento mental, que es falsa la información *fáctica* que podría indicar que el origen de nuestras capacidades cognitivas es conducente al error? No resulta claro qué es lo que tendríamos que negar en este caso, puesto que no es claro cuál de las diferentes opciones de obstáculos epistemológicos surgidos de la teoría de la evolución es la que tiene en mente Sosa. ¿Debemos negar por razones de principio la aparente evidencia fáctica de que, por ejemplo, existe una conexión puramente accidental entre el valor adaptativo de nuestras creencias y su valor de verdad, y asumir, por el contrario, que si nuestras capacidades cognitivas se han mostrado adaptativamente valiosas ellas son cognitivamente confiables, pese a que la evidencia empírica parece indicar lo contrario?

Si la solución que tiene en mente Sosa es algo de este tenor, y ciertamente lo es, si no para el problema de nuestros orígenes evolutivos, sí al menos para el problema paralelo de determinar si hemos o no tomado “Disablex”, esto nos conduce necesariamente a la segunda de nuestras preguntas; a saber: ¿cómo es exactamente que podemos simplemente desconocer la evidencia empírica? La situación es particularmente irónica puesto que parece constituir exactamente el tipo de contradicción que Sosa nos invita a evitar, solo que vuelta del revés: si, según Sosa, es contradictorio confiar en nuestras capacidades cognitivas (los datos de nuestros sentidos y memoria) para declarar que no podemos confiar en nuestras capacidades cognitivas (declarar, en efecto, que hemos tomado “Disablex” y que este tiene un efecto incapacitante), esta es, invertida, la situación en la que nos encontraríamos si simplemente declaráramos *que podemos ignorar* los datos proporcionados por nuestras capacidades cognitivas (los datos que nos indican que hemos tomado “Disablex”) en el mismo momento en que decimos que podemos confiar en esas capacidades (puesto que, según afirmamos, *no* hemos tomado “Disablex”). Si concluimos, por las razones de principio que menciona Sosa, que no hemos tomado la droga incapacitante, parecería ser que tenemos que explicar *de algún modo* el hecho de que tenemos datos mnémicos y sensoriales que nos indican justamente lo contrario.

Una solución parecería ser declarar que somos víctimas de algún tipo de alucinación y de falsos recuerdos, pero esto, precisamente, no es concluir que nuestras capacidades cognitivas son confiables. En todo caso, de lo que se tratará, para adoptar una actitud epistémicamente racional, es de concluir que nuestras capacidades cognitivas son confiables *en cierto grado* -el grado suficiente como para admitir errores *puntuales*, como los datos que nos convencen de que hemos tomado “Disablex” cuando no lo hemos hecho-, pero que, paradójicamente, resultan poco fiables cuando parecen indicarnos que ellas son *radicalmente* no fiables (es decir, cuando parecen indicarnos que hemos tomado “Disablex”). La respuesta a un interrogante como el de si hemos tomado una droga incapacitante, cuando tenemos datos que lo sugieren, solo puede consistir en algún tipo de solución de compromiso, de “trade-off”, entre necesidades epistemológicas de principio -a saber, la tesis de que nuestras capacidades cognitivas tienen que ser confiables- y datos empíricos particulares -a saber, el dato de que *al menos en apariencia* hemos efectivamente

tomado “Disablex”-. Aplicando este esquema al problema evolutivo que Sosa desea modelar por medio de su experimento, la única respuesta a la que llegamos es *que de algún modo*, para evitar caer en una posición autoderrotadora, se ha de poder encontrar un equilibrio, aquí también, entre lo que los datos empíricos parecen indicar y lo que necesitamos por principio presuponer. Sin embargo, esta solución es indeterminada; no se nos dice exactamente cómo es que debemos evitar un análisis epistemológicamente autoderrotador acerca de nuestras capacidades cognitivas. En este sentido, la clase de presuposición de confiabilidad con la que Sosa se compromete es comparable con la posición suscrita por C. S. Lewis en contra de lo que consideraba una “idiotez autocontradictoria” (Lewis 2004: 303) bajo la forma de la crítica radical de la razón que habría sido emprendida por el freudismo y el marxismo: según Lewis, para evitar caer en la paradoja por la cual apelamos a la razón para desacreditar a la razón, debemos oponerle una “creencia tenaz en nuestro poder de razonamiento” (ibíd.). Pero hablar de “tenacidad” implica reconocer que esta reivindicación de la razón necesita *enfrentarse* a algo; reconocer que hay motivos legítimos de duda acerca de nuestra propia racionalidad -como, en el caso estudiado por Lewis, los determinantes psicológicos y económicos que inciden sobre nuestra formación de creencias (cf. ibíd.: 299-300)- y que, *de algún modo*, el reconocimiento de estos motivos (puesto que Lewis nunca señala que el freudismo y el marxismo estén *empíricamente* equivocados) habrá de poder equilibrarse con la suficiente confianza en nuestro propio raciocinio para no desembocar en una posición autodestructiva.

Análogamente, la respuesta de Sosa al problema del carácter accidental del surgimiento de nuestras capacidades cognitivas es entonces, en realidad, algo así como *el bosquejo, el esquema*, de una respuesta, compatible con una variedad concebible de soluciones en concreto, soluciones que incluyen, claro está, las propuestas específicas que pueden provenir de las distintas variedades del confiabilismo evolucionista. ¿En qué consiste, entonces, el aporte específico de una solución “de principio” como la de Sosa? En establecer el marco dentro del cual podrá proponerse racionalmente un relato acerca de nuestras capacidades cognitivas, relato que tendrá, de un modo u otro, que presentar a estas como lo suficientemente confiables como para no constituir un discurso autoderrotador. Pero incluso un aporte en apariencia tan mínimo como este tiene la consecuencia de que los análisis

empíricos sobre el conocimiento humano no pueden considerar la fiabilidad de nuestras capacidades como una tesis susceptible de ser falsada; no pueden asumir que existen condiciones, lógicamente concebibles, bajo las cuales *quedaría empíricamente probado* que nuestro conocimiento es radicalmente poco fiable.

### 3. Una segunda objeción al argumento de Sosa desde una perspectiva naturalista

Hasta aquí, pues, puede hablarse de un primer sentido en que una respuesta “de principio” como la de Sosa entra en tensión con la orientación empírica propia del confiabilismo evolucionista: en efecto, si la pretensión de estos es exhibir la fiabilidad de nuestras capacidades cognitivas como un dato contingente, la propuesta del autor norteamericano excluye de plano, como señalábamos, esta posibilidad. Pero hay también, por otra parte, un segundo modo en que cabe delimitar la propuesta de Sosa con respecto a estas formas de confiabilismo, y concierne a la relación con lo que puede entenderse como cierta exigencia específica que debe satisfacer una epistemología de inspiración naturalista. Nos referimos a la que plantea Howard Sankey al señalar que este tipo de epistemología -como la que desarrolla, en su caso, para responder al relativismo- “se aplica a este mundo, el que de hecho habitamos, no a todos los mundos posibles. En buena vena naturalista, es un acercamiento [...] que se adecua a *nuestras* circunstancias, a criaturas constituidas como *nosotros* lo estamos, en el mundo constituido como es” (Sankey 2010, 14). Es interesante notar que, justamente, la estrategia de Sosa *no* se enfoca en absoluto en lo que tienen de específico nuestros mecanismos cognitivos humanos. El experimento mental consistente en preguntarse “¿Son mis capacidades cognitivas radicalmente poco fiables, como lo serían si hubiera tomado una droga incapacitante?” podría en principio ser efectuado por *cualquier* sujeto capaz de operaciones reflexivas de este tenor, y conduciría al mismo resultado -esto es: que se debe *presuponer* que ellas son fiables- al margen de si, en lo que concierne al resto de sus rendimientos cognitivos, este sujeto es radicalmente diferente de nosotros, al punto de que, en presencia del tipo de creencias que forma, lo declararíamos *no* confiable.

En efecto, las únicas condiciones que se requieren para poder efectuar el experimento mental del “Disablex” consisten, aparentemente, en (a) ser un sujeto de

conocimiento dotado de las capacidades de hacer inferencias y de obtener “inputs” cognitivos a partir de la percepción o la memoria, (b) poder objetivar reflexivamente estas capacidades a efectos de interrogarse sobre su fiabilidad. No hay nada de específicamente *humano* en tales condiciones, y ellas bien podrían venir acompañadas de una propensión a formar creencias que, desde nuestra perspectiva, aparecerían como disparatadas. ¿No quiere esto decir que la propuesta “de principio”, según la cual debemos *presuponer* el valor de nuestras capacidades más que probarlo, *trivializa* la idea misma de “capacidades cognitivas fiables”, puesto que, en su carácter excesivamente “democrático”, el argumento indica que *cualquier* sujeto cognitivo capaz de reflexión está justificado en considerarse a sí mismo confiable? En particular, ¿implica el argumento de Sosa que, frente a cualquier sujeto reflexivo S, *nosotros* (nosotros los seres humanos) deberíamos concebirlo como “fiable” porque este es el modo en que S se presenta *a sí mismo*? Supongamos que hubiéramos evolucionado de un modo radicalmente diferente a como lo hicimos, y que nuestras contrapartes en el mundo  $W_2$  no tienen nuestras capacidades cognitivas sino otras que los vuelven, por ejemplo, sistemáticamente inclinados a creer en contradicciones (pese a que, en todo lo demás, el mundo  $W_2$  es como el nuestro). Parecería ser que un argumento confiabilista puede describir en qué sentido estas contrapartes tendrían sistemas radicalmente poco fiables de formación de creencias -más aun, que como consecuencia de ello no podrían sobrevivir demasiado-, mientras que, por el contrario, una estrategia argumental como la de Sosa, que recurre a la reflexión, es incapaz de conducir a este resultado. ¿Qué decir sobre esta objeción?

En rigor, la pregunta por el riesgo de trivialización nos conduce a la necesidad de distinguir tajantemente, a la luz del argumento de Sosa, entre lo que cabe evaluar desde una perspectiva de primera persona y lo que solamente podríamos determinar razonablemente desde una perspectiva de tercera. El corolario es doble: *por un lado*, nada nos obliga a considerar, desde nuestra perspectiva de tercera persona, que otros sujetos cualesquiera son “cognitivamente fiables” solo porque ese es el modo en que ellos se perciben a sí mismos; desde *nuestra* perspectiva, podemos juzgar a estos sujetos, reales o hipotéticos, a partir de lo que *nosotros* sabemos del mundo, y es por eso que la calificación de “confiable” no queda trivializada. Que nuestras contrapartes evolutivas hipotéticas que, en otro mundo posible, desarrollaron una

propensión a creer contradicciones, no puedan más que considerarse a sí mismas “fiables” solo nos informa sobre lo que resulta justificado desde *su* perspectiva, no desde la nuestra. Nosotros seguimos teniendo que usar *nuestras* capacidades cognitivas, no las de ellos, y *nuestras* capacidades nos indican que, entre otras cosas, el mundo es no-contradictorio, de manera que no podemos más que asumir que ellos están equivocados. Sin embargo, *por otro lado*, la distinción entre perspectivas de primera y tercera persona quiere decir también que, en lo que hace a la fiabilidad que nosotros podemos atribuirnos *a nosotros mismos*, ella es del mismo tipo que la que *cualquier* sujeto reflexivo podría arrogarse, como se desprende del argumento de Sosa. La clase de reivindicación que podemos hacer de nuestras propias capacidades presupone circularmente el valor de estas, tanto como lo presupondría respecto de sus capacidades (capacidades que *para nosotros* pueden ser poco fiables) cualquier otro sujeto reflexivo. En otras palabras, el argumento en apariencia excesivamente “democrático” de Sosa prueba, sencillamente, que el confiabilismo no es capaz de lograr el milagro de permitirnos escapar de nosotros mismos y garantizar el valor de nuestro conocimiento desde una perspectiva del ojo de Dios. El *factum* de ser quienes somos, de no poder juzgarnos desde fuera, es un dato al que la teoría del conocimiento no puede más que resignarse.

#### **4. Una consecuencia de la argumentación de principio: ausencia de un sentido neutral para hablar de “confiabilidad”**

Sin embargo, precisamente por ello, lo interesante de la argumentación de principio con la que se compromete Sosa es qué preguntas ella vuelve *irracional* formularse. Si no podemos más que garantizar nuestra confiabilidad circularmente, entonces no tiene mayor sentido preguntarse qué mecanismos -si, por ejemplo, una improbable gran dosis de suerte, o un mecanismo evolutivo orientado hacia la confiabilidad epistémica- han hecho que los seres humanos desarrolláramos capacidades cognitivas confiables *habiendo podido* desarrollar capacidades *no* confiables. Consideremos por ejemplo la objeción de Alvin Plantinga contra la improbabilidad de que hayamos “ganado la lotería evolutiva” (Plantinga 2002b: 246); esto es, que hayamos tenido, como especie, una implausible dosis de suerte



epistémica. Este argumento presupone, en efecto, que podemos distinguir, por un lado, lo que es conocer confiablemente y, por otro lado, lo que es conocer como *nosotros los humanos* lo hacemos; solo así cabe preguntarnos, luego, por qué justamente nuestras capacidades son de hecho confiables.

Por el contrario, si nos tomamos el argumento de Sosa en serio, este parece implicar que si hubiéramos desarrollado *cualesquiera otras* capacidades entonces ellas también habrían sido “confiables” *para nosotros mismos*, en la medida en que también en ese caso la evaluación de tales capacidades habría necesitado ser autorreferencial: estas otras capacidades serían evaluadas a la luz del uso de esas capacidades. Si llamamos C1 al conjunto de las capacidades cognitivas que tenemos en el mundo real (cierta percepción sensorial, cierta capacidad de memoria, ciertas capacidades para las inferencias deductivas e inductivas, etcétera), y consideramos que C1 es un conjunto de capacidades confiables, *podemos* preguntarnos por qué tenemos ese conjunto de capacidades C1 y no ciertas capacidades distintas, llamémoslas C2. En otras palabras, *es* contingente que las tengamos a ellas en vez de tener otras. Pero, una vez más, no es contingente que las capacidades C1, las que tenemos en el mundo real, *sean confiables*, para nosotros que no podemos juzgarlas más que a la luz de sí mismas; de allí que problemas como el de la “suerte epistémica” que inquietan a Plantinga desaparezcan. Simplemente, no hay un sentido *neutral* -neutral entre nosotros y nuestras contrapartes que tendrían las capacidades C2- para decir que *las nuestras* son confiables. Al no haber tal sentido neutral, no estamos en absoluto haciendo un enunciado sintético cuando decimos que ellas son confiables. Así, preguntarse si no sería posible que *también nuestras* creencias sean en su mayoría falsas parece, en consecuencia, tan irracional como preguntarse si no sería posible que el metro patrón no mida realmente un metro. En otras palabras: ¿con qué criterios, salvo *los nuestros*, podríamos llegar a *saber* que un conjunto de capacidades cognitivas no es realmente confiable?

## 5. Recapitulación

1. Hemos visto en primer lugar, a partir del ejemplo de Ernest Sosa en *Knowing Full Well*, que los abordajes confiabilistas evolucionistas acerca de las capacidades

cognitivas humanas pueden ser vistos como contrapuestos no solo a las dudas escépticas que ellos contribuyen a enfrentar, sino también a la clase de enfoque antiescético que buscaría avalar tales capacidades no desde evidencias empíricas sino a partir de argumentos de principio. Que existan respuestas empíricas y de principio a la misma cuestión indica que la solución al problema está *sobredeterminada*, y es necesario en consecuencia cotejar una contra otra.

2. A continuación, sin embargo, hemos tenido que problematizar la efectiva aceptabilidad de esta propuesta, en la medida en que la tesis de que debemos *por principio* considerar fiables nuestras capacidades cognitivas suscita la dificultad de determinar cómo se vincula este presupuesto, por un lado, con los datos empíricos que podrían contradecirlo, por otro lado (situación que es efectivamente escenificada por Sosa a partir del experimento mental del “Disablex”): si, como señala Sosa, es contradictorio apoyarnos en nuestras capacidades para concluir que no son fiables, también lo es, a la inversa, declarar que confiamos en nuestras capacidades a la vez que elegimos desconocer justamente evidencias que ellas arrojan. Puede establecerse, en cualquier caso, que la clase de solución propuesta por Sosa requiere algún tipo de “*trade-off*” entre los presupuestos de principio y las evidencias empíricas, y que, en consecuencia, ella constituye menos una respuesta que un esquema general compatible con diferentes respuestas concretas.

3. Un problema adicional para la solución “de principio” propuesta por Sosa proviene de la circunstancia de que ella parece hacer abstracción de las características *específicas* de las capacidades cognitivas humanas: la clase de interrogante que el autor nos invita a responder puede ser efectuada por *cualquier* sujeto cognoscente capaz de reflexión, lo cual es en principio excesivamente abarcador y *trivializa* la caracterización de “capacidad cognitiva confiable” atribuyéndosela a potenciales sujetos cognoscentes cuyas creencias diferirían radicalmente de las que podemos sentirnos inclinados a considerar confiables. Sin embargo, esta consecuencia no se sigue: del argumento de Sosa solo se desprende, en rigor, la necesidad de distinguir tajantemente entre una perspectiva de primera y de tercera personas.

4. Por último, en virtud de esta misma distinción, el argumento prueba que, si bien no queda trivializada la caracterización de “confiable”, tampoco podemos realmente privilegiar nuestra propia auto-evaluación global por sobre la que podrían



realizar sobre sí mismos cualesquiera otros sujetos cognoscentes reflexivos: unos y otros estamos igualmente condenados a no poder apelar a un “punto de vista del ojo de Dios”.

## 6. Bibliografía

- Flew, Antony (1958), “Determinism and Validity Again”, *The Rationalist Annual*, 1958.
- Fodor, J. (1981) *Representations. Philosophical Essays on the Foundations of Cognitive Science*. Brighton, The Harvester Press.
- Lewis, C. S. (2014 [1970]), “‘Bulverism’, or, the Foundation of Twentieth Century Thought”, en *God in the Dock*, Cambridge, Eerdmann Publishing.
- Nagel, Thomas (2012), *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. Oxford, Oxford University Press.
- Plantinga, Alvin (1991), “An Evolutionary Argument Against Naturalism”, *Logos* 12.
- Plantinga, Alvin (1993), *Warrant and Proper Function*, Nueva York y Oxford, Oxford University Press.
- Plantinga, Alvin (2002), “Introduction: The Evolutionary Argument Against Naturalism” y “Reply to Beilby’s Cohorts”, en Beilby, James (ed.), *Naturalism defeated? Essays on Plantinga’s Evolutionary Argument Against Naturalism*, Ithaca, Nueva York y Londres, Cornell University Press.
- Popper, Karl, *Objective Knowledge: An Evolutionary Approach*, Oxford, Clarendon Press, 1972.
- Sosa, Ernest (1997), “How to Resolve the Pyrrhonian Problematic: A Lesson from Descartes”, *Philosophical Studies* 85 (2-3):229-249.
- Sosa, Ernest (2002), “Plantinga’s Evolutionary Meditations”, en *Naturalism Defeated?*, *op. cit.*
- Sosa, Ernest (2003), “Reflective Knowledge in the Best Circles”, en Steven Luper (ed.), *Essential Knowledge: Readings in Epistemology*, New York, Longman.
- Sosa, Ernest (2009), *Reflective Knowledge*, Oxford, Oxford University Press.
- Sosa, Ernest (2011), *Knowing Full Well*, Princeton, Princeton University Press.
- Sosa, Ernest (2017), *Epistemology*, Princeton, Princeton University Press.