

Las emociones morales en blanco y negro y en color

Ronald de Sousa

Muchas teorías psicológicas de la emoción subrayan su capacidad de motivar. Ahora bien, la capacidad de motivar corresponde primordialmente a los deseos. ¿Qué diferencia, entonces, a las emociones? Como frente a todo posible acto, sólo nos cabe decidir si lo realizamos o lo reprimimos. Al equiparar la emoción al deseo, estamos trasladando una realidad multidimensional a un plano bidimensional o, mejor, estamos proyectando en blanco y negro algo que es a todo color. Para explicar ese campo tan rico y tan coloreado de la emoción, al que voy a llamar «axiológico», estudiaré primero el papel que desempeñan las emociones en el aprendizaje. Una de las características del aprendizaje, tanto de habilidades como de conocimiento, consiste en ampliar nuestra gama de placeres. Cabe interpretar esto como una forma de adaptación biológica. Las emociones son esenciales para el aprendizaje intelectual y para el desarrollo de habilidades corporales: algunas, como la curiosidad, impulsan a investigar; otras, como el orgullo, recompensan su resultado. La educación moral implica también a nuestras emociones, pero la gama de las emociones es muy amplia, pues puede abarcar emociones intrínsecamente morales, como la compasión, la indignación o la empatía, y emociones que tendemos a considerar intrínsecamente rechazables, como es el caso del rencor y de los celos. Dado que la mayoría de estas emociones se aprenden en los estadios más tempranos del desarrollo moral, es posible que queden arraigadas en nosotros. Ahora bien, el arte y la literatura pueden educar nuestras emociones de un modo más sutil y complejo. En la buena literatura, las emociones no actúan como meros deseos. Revelan una realidad multicolor que contribuye al conocimiento axiológico y a la maduración moral. Trataré de concretar en este artículo cómo podría, a mi juicio, ser así.

I. LAS EMOCIONES EN BLANCO Y NEGRO

Si consideramos las emociones exclusivamente como *motivaciones*, concluiríamos que son fuerzas que nos impulsan y arrastran. Dominados por la totalidad de nuestras emociones, en un determinado momento, actuamos o no actuamos. Nuestra forma de *experimentar el mundo* complica las cosas, pero, en última instan-

cia, todas las emociones se reducen a una serie de decisiones singulares, en las que cada una exige obrar o no obrar, el blanco o el negro del sí o del no. Por otro lado, si consideramos las emociones como experiencias, resultan tan diversas que es imposible reducirlas a un *modelo* único. Cada una entraña una variedad de significados específicos enriquecida, además, por un conjunto de contrastes enormemente amplio. Pues bien, la *visión monocromática* se sitúa en el contexto de la acción y de la elección, mientras que la *visión a todo color* considera las emociones como modos de *percepción axiológica* multidimensional.

Basándonos en esta *visión multicolor* de las emociones, cabría objetar que el número de emociones distintas que podemos experimentar en la práctica no tiene límites, de no ser porque el número de ideas que somos capaces de concebir sí que lo tiene. Esta tesis parece aceptable a la luz de las emociones que despierta la experiencia estética. Parece claro que contemplar un ballet o escuchar un concierto, por ejemplo, suscita emociones. Ahora bien, esta afirmación puede ser tachada de trivial en dos sentidos diferentes e incluso opuestos. Por un lado, podemos pensar que las emociones suscitadas por la música o por la danza son «emociones básicas», es decir, *emociones que cabría enumerar: ira, miedo, amor, respeto, envidia, melancolía, deseo...* Pero, entonces, *¿por qué afecta a todas esta dificultad?* Si las obras de arte existen sólo para suscitar esas emociones y si no hay ninguna diferencia significativa entre dos situaciones de «miedo» o de «ira», etc., resulta muy poco probable que las obras de arte en toda su diversidad sólo nos parezcan interesantes porque representan una *emoción*, y no por otras razones. Según esta tesis, es difícil entender por qué el arte en su mayoría no es superfluo. (La mayoría del arte mediocre es superfluo sólo en este sentido: una vez que has visto una película de terror sobre una invasión de extraterrestres, ya las has visto todas.) La tesis alternativa es que cada momento diferente en música o en danza (o en cualquier arte que se considere que despierta emociones) está expresando realmente una emoción *sui generis*, y no sólo *sui generis* sino, por así decirlo, *sui ipsius*: en su propia individualidad irrepetible. Pero también en este caso trivializamos el papel del arte, pues si cada obra de arte expresa necesariamente su emoción correspondiente, ni más ni menos, entonces resulta imposible que ciertas formas de arte puedan expresar mejor o peor dicha emoción, o que las emociones suscitadas por una determinada obra de arte puedan ser más valiosas, interesantes, auténticas o más hondamente sentidas que otras. Pues todos estos juicios suponen que hay algo más allá de la expresión de una emoción en cuyos términos puede juzgarse la obra en cuestión.

Cabe superar esta última objeción si renunciamos a asumir que toda obra artística o literaria transmite un «mensaje» emocional básico. Si consideramos, por el contrario, que el arte produce y transmite por sí mismo una emoción especial, entonces podrán darse numerosas emociones, y cada obra de arte será más o menos interesante según sea la calidad de la emoción única que transmite.

Pero ahora se plantea otro tipo de problemas.

Si las emociones son, realmente, innumerables, ¿cómo entender su influencia en nuestras elecciones? La tesis parte del supuesto, que muchos hemos aprendido de Aristóteles, de que la educación moral consiste en gran medida en aprender a sentir la cosa correcta, de la manera correcta, en el momento correcto (Aristóteles,

Ética Nicomáquea, 1104b 13). A lo largo de la educación moral aprendemos, de algún modo, a sentir ciertas emociones y luego, en el curso de nuestra vida, esas emociones: a) nos permiten *ver* y *entender* las cosas correctamente y b) nos *mueven* a obrar correctamente como respuesta a ese conocimiento. ¿Cómo ejercen esa influencia mediadora?

Una respuesta prometedora resalta la importancia de la literatura para nuestra educación moral. En este sentido, Marta Nussbaum ha hecho ver el poder de la literatura para presentarnos emociones y personajes concretos «totalmente imaginarios» frente a las descarnadas abstracciones de los ejemplos filosóficos (Nussbaum, 1990, *passim*). Con todo, esta idea plantea dos problemas. Primero, por muy viva y detallada que sea una ficción, no puede ser literalmente cierto que las ficciones nos presenten a individuos particulares. Los sujetos particulares y concretos tienen innumerables propiedades que se relacionan entre sí. Las ficciones heredan esa generalidad que es intrínseca a las palabras con las que se expresan. Pueden ser *específicas* en un sentido indefinido, pero nunca *particulares*. Segundo, si nos centramos en las características exclusivas de cada situación, en las cualidades específicas de cada individuo y en la singularidad de cada emoción, difícilmente podrán servirnos de guía. El aprendizaje requiere comparaciones, que a su vez necesitan similitudes y diferencias, clasificadas y conceptualizadas en términos que necesariamente nos llevan de lo particular a lo general. El aprendizaje, en suma, exige modelos repetibles. Literalmente hablando, la sugerencia de que cada situación concreta de la vida (lo mismo que cada episodio de una obra de arte) produce una emoción única e irrepetible resulta, pues, autocontradictoria. Los preceptos (particulares) sin conceptos (generales) no pueden guiar nuestras elecciones.

Esto no sólo es un problema para el aprendizaje individual, sino también para la evolución de nuestras capacidades emocionales mediante la selección natural. Aunque algunos científicos sociales tienden a minimizar la aportación que pueda hacer la biología al conocimiento de nosotros mismos, pocos negarían que la selección natural desempeña un papel fundamental en la determinación del potencial de nuestras emociones. Ciertas emociones, llamadas a veces «básicas», tienen una importancia evidente en determinadas funciones vitales (desempeñadas en ciertas especies por los instintos), como es el caso del hambre, el miedo, la agresividad y el deseo sexual. Es evidente que éstas no habrían sido seleccionadas si no hubieran dado lugar a modelos repetibles con propiedades causales repetibles.

Queda claro, pues, que el papel desempeñado por las emociones como factor de influencia en la conducta sobre la base de experiencias sumamente específicas debe dar por sentado que tales experiencias tienen rasgos repetibles. Ya he indicado en otras ocasiones que más allá de la media docena de modelos aproximadamente que suelen enumerar los teóricos de las emociones, como los que Ekman ha descubierto que son más o menos reconocibles universalmente en sus modos de expresión, las características comunes de las emociones más complejas están plasmadas en *escenarios paradigmáticos* que determinan no sólo los diferentes tipos de emociones, sino una forma característica de experimentarlas, al igual que sus condiciones de conveniencia. Las raíces que definen a cada emoción discriminable residen en una situación dramática específica o en un episodio típico, asociados a un *sentimiento* característico (en un sentido amplio, no en el sentido específico de los

sentimientos corporales). El drama lo representan agentes con determinados *papeles*, y la hipótesis de la que partía en otro lugar era que esos escenarios paradigmáticos, que pueden variar de una persona a otra, definen realmente el *significado* que tienen para nosotros las emociones cuyos paradigmas constituyen. Cuando experimentamos esa emoción en un momento posterior de la vida, su conveniencia para la situación viene determinada en parte por la relación existente entre la situación presente y el paradigma originario, el cual determina el *objeto formal* de esa emoción, y donde el objeto formal determina el estándar de conveniencia que debe satisfacer una emoción para ser considerada ese tipo de emoción.

Del mismo modo que podemos contar una cosa con más o menos detalle, podemos agrupar las emociones con un mayor o menor grado de especificación. De este modo, no hay ningún problema en que experiencias muy diferentes sean entendidas como miedo o como ira. Aunque el argumento de *Antígona* puede resumirse en una línea (una princesa discute con su tío y muere con su prometido), el impacto de la obra no se explica con este resumen. Lo mismo cabe decir de las emociones determinadas que experimentamos: que no las describimos adecuadamente con las palabras que solemos usar para hacer referencia a emociones determinables.

Esto indica una forma de reducir las diversas emociones que experimentamos, dado que éstas son, en gran medida, *repetitivas*, y se clasifican parcialmente en términos de un repertorio finito de *papeles*. Pero esto no basta para responder al problema planteado, esto es, ¿cómo ejercen las emociones su influencia mediadora en nuestro aprendizaje moral?

Para responder a esta pregunta, sugiero que nos dirijamos primero a la perspectiva monocromática: en lugar de considerar las emociones en toda su variedad, estudiémoslas según la doble dimensión simplificada de aversión y atracción. Como cuando se reproduce un óleo a todo color mediante un dibujo en blanco y negro, perderemos una gran riqueza de detalles, pero ganaremos determinada claridad. La versión más simple de la perspectiva monocromática es la hipótesis, tantas veces rechazada, del hedonismo psicológico, que propongo ofrecer en una versión a la que podemos llamar *hedonismo inmediato*. Luego veremos los requisitos ulteriores que nos exigen pasar a la perspectiva multicolor de la axiología.

II. EL HEDONISMO INMEDIATO

El hedonismo psicológico es la doctrina según la cual sólo obramos, de hecho, para perseguir el placer (y evitar el dolor). Naturalmente, no hay que confundir esta tesis con el *hedonismo ético*, según el cual, *debe* perseguirse el placer (y evitarse el dolor). Tradicionalmente se ha considerado que ambas teorías son incompatibles. Si frente a nuestra tendencia al placer no podemos hacer nada, no tiene sentido alguno decir que debemos seguir esa inclinación. Ahora bien, ambas tesis presentan ambigüedades tales que, si las superásemos, podríamos hacerlas compatibles.

La primera ambigüedad es la siguiente. El hedonismo psicológico dice que nuestra motivación última es siempre el placer. Pero, ¿el placer de quién? Se supo-

ne que se quiere decir el mío, el de cada agente. El hedonismo ético consideraría el valor del placer más ampliamente, desembocando, así, en la teoría utilitarista según la cual debemos tratar de maximizar el placer para el mayor número. Esto puede ser incompatible, sin duda alguna, con la persecución exclusiva de mi propio placer. De ahí que quepa la posibilidad de proponer un ideal para el que todo agente se hallaría mal dotado. Ahora bien, estar mal dotado para aspirar a un ideal no es lo mismo que obedecerlo o seguirlo de modo trivial. La segunda ambigüedad afecta sólo al hedonismo psicológico. ¿Se trata de maximizar mi placer en general, o de perseguir constantemente el placer que me sea más accesible en cada momento concreto? Esta última alternativa equivale a la política de «perseguir el placer de cada momento», que Aristóteles atribuye al *akólastos* o «autoindulgente». Ofrece una visión del hedonismo ético que ni siquiera tiene en cuenta a los demás; me exige, simplemente, que vea mi placer desde una perspectiva más amplia.

Entre las críticas que ha recibido el hedonismo psicológico figura la que describe el experimento mental de la *máquina de experiencias* (Nozick, 1974, 42-45).

Se pidió a los sujetos que se imaginaran que podían elegir entre continuar su vida actual con sus incertidumbres, sus riesgos y sus dolores, o estar conectados a una máquina que sólo les suministrara placer durante el resto de su vida —o mejor, ya que eso último parece en exceso aburrido, que les suministrara una combinación apropiada de momentos de anticipación, frustración y éxito final, de momentos de esperanza, previsión y logro definitivo, como los que caracterizan a la existencia—. La mayoría de los sujetos no deseaba conectarse a una máquina así, rechazaba la idea. Al margen de que muchos no creían probablemente en las virtualidades de dicha máquina, el resultado de este experimento cuestiona el hedonismo psicológico. Sin embargo, Sober y Wilson han señalado recientemente que este argumento no tiene fuerza frente a una versión del hedonismo psicológico que lo interprete en términos del placer presente y no de los resultados de cálculos que traten de maximizar el placer en el futuro o en su totalidad (Sober y Wilson, 1998, 275-291). Llamaré *hedonismo inmediato* a esta versión del hedonismo psicológico.

El hedonismo inmediato otorga un poder causal decisivo al placer inmediato y especialmente a la experiencia real de hacer una determinada elección, en lugar de hacer hincapié en las consecuencias de esa elección para el futuro. El problema de la máquina de experiencias puede afrontarse, así, con facilidad, pues sólo se necesita postular que el acto de elegir la vida real es en sí más placentero que el acto de elegir la máquina de experiencias, por muy agradables que puedan parecer las consecuencias de esta última elección.

Esta hipótesis tiene la ventaja de admitir una elección cuyas consecuencias serían menos placenteras que las consecuencias de la elección no realizada, pues no es contradictoria una proposición que afirme:

1. *A* es preferida a *B*, pero *elegir B* es preferido a *elegir A*.

De este modo el hedonismo inmediato es inmune a los mejores argumentos contra el hedonismo psicológico. Esto no quiere decir que la tesis sea indiscutible.

Puede calificarse también de trivial. Pero antes de caer en la tentación de rechazarla por ese motivo, observemos que nos permite decir que 1. es coherente, pero también es *irracional*.

III. EL PRINCIPIO DEL FILEBO

La razón de que 1. sea irracional es que viola un principio plausible de racionalidad que cabría llamar *el principio del Filebo*.

A lo largo de su investigación sobre el placer, Platón examina en el *Filebo* la posibilidad de concebir el placer como una *especie de percepción* y no sólo como una especie de *sensación*. Esta distinción significa que la percepción, a diferencia de la sensación, parece proporcionar información sobre algo que está más allá de ella misma. Los placeres de la anticipación, en especial, que dicen relación a un placer futuro, pueden suministrar información respecto al alcance de dicho placer. Es probable que esto no baste para considerarlos verdaderos o falsos literalmente, como afirma Platón, pero al parecer confiere racionalidad al siguiente principio:

Principio del Filebo (PP): Un placer de anticipación ha de tener una intensidad proporcional al placer anticipado que concibe.

Desde un punto de vista biológico, este principio es razonable. Si el hedonismo inmediato es verdadero, la intensidad del placer anticipador presente determinará la tenacidad de mi motivación actual para perseguir el placer anticipado. Pero si la intensidad de mi placer actual no guarda relación con la cantidad de placer futuro que me motiva a perseguir, entonces mi motivación presente no guardará ninguna relación establecida con la satisfacción futura. Nuestras vidas se parecerían siempre a la de una persona que queriendo organizar una comida en el campo, en una fecha y un lugar concretos, hiciera una meticulosa previsión del tiempo para una fecha y un lugar posiblemente distintos y elegidos al azar.

IV. LAS EMOCIONES Y EL APRENDIZAJE

Ahora podemos distinguir algunas de las formas en que podrían ser factores motivacionales de nuestro progreso en distintos campos las emociones en tanto que proyectadas en la dimensión «monocromática» del placer y del dolor. De hecho, intervienen de tres maneras al menos.

La más evidente corresponde a la observación de Aristóteles de que, por naturaleza, los seres humanos desean saber (*Metafísica*, 980a 21). Es esta una observación empírica correcta que cabe constatar desde los primeros días de la infancia. El ejercicio de cualquier capacidad y el aprendizaje dirigido a aumentar nuestras aptitudes implica placer. Todo niño da muestras de experimentar este tipo de placer, al menos hasta que los colegios frustran su curiosidad y la convierten en rechazo. Evolutivamente hablando, el placer de aprender no es más sorprendente que el placer sexual.

Segundo, el modelo nos brinda también una forma de entender la *akrasía*. La *akrasía* puede ser entendida como la elección de A, en lugar de B, cuando estamos dando por sentado que B es preferible a A. Como digo, eso instanciaría el modelo, pues, en el momento de elegir, considero la perspectiva de B preferible —esto es, según el esquema bidimensional, en blanco y negro, al que nos ceñimos ahora, B proporciona una mayor cantidad de placer que A—; pero, debido a lo que Platón llamaría una anticipación de placer falsa, experimento más placer en el acto real de elegir A que en el acto de elegir B.

De ser cierto todo lo que vengo diciendo, esa pauta sería no sólo posible sino también irracional. Ahora bien, donde hay irracionalidad, hay esperanza de racionalidad. El lado bueno de la *akrasía* es que permite el progreso emocional y moral. Todo progreso entraña conflicto: presupone que en el terreno de la conducta moral y de la experiencia de emociones, es posible tanto cometer errores como lamentarlos. Como en el futuro cabe corregir esos errores, éstos anticipan una mejora. (Aristóteles dice con acierto que su «akólastos», al «perseguir el placer que tiene delante» y hacer el mal sin experimentar conflicto interior alguno, no se arrepiente y es, por eso, «incurable» (*Ética Nicomáquea*, 1150b 29). Se trata de una especie de idiota moral feliz, incapaz de reformarse.)

De este modo, nuestra visión restringida y bicolor del hedonismo inmediato nos permite explicar en qué sentido facilitan las emociones determinado aprendizaje. Existe otro camino puramente psicológico, más directo y más neutral entre diversas teorías de la motivación, porque su efecto es meramente causal y no motivacional. La estimulación intensa del sistema límbico constituye un factor crucial para que se fije en la memoria un acontecimiento importante (Rosenfield, 1992, 165), como sabe todo americano que recuerde dónde estaba exactamente cuando se enteró de la muerte de John Kennedy. (Cuando los asesinatos se convierten en una rutina, ya no son memorables.)

El cuadro escueto, en blanco y negro, de la motivación como placer nos permite hablar también de una serie de emociones que influyen en la conducta y el aprendizaje mediante el cálculo instrumental. Las emociones que son capaces de intervenir en este campo incluyen la rivalidad, la admiración, el orgullo, e incluso un cierto número de emociones que en otro terreno nos resistiríamos a admitir, pero que aquí nos hacen, quizás, ser mejores: la codicia, la envidia, la vanidad, los celos, la humillación o el miedo pueden ser estímulos para que nos esforcemos por ampliar, tal vez, el ámbito de nuestras competencias. En este sentido, pueden extender nuestro aprendizaje más allá de las cosas que aprendemos a lo largo de un proceso que comporte un intenso placer. Lamentablemente los procesos de aprendizaje donde pueden intervenir estos estímulos son muy diversos, empezando por los métodos de aprendizaje de la lectura que más de un niño y más de un padre habrán considerado alguna vez una forma de tortura, pero que el padre más preocupado por el bienestar de su hijo será el menos motivado para ahorrárselos a éste. Con todo, estas emociones son incidentales en la promoción de estos fines respetables; nada en ellas parece servir intrínsecamente al fin del progreso cognitivo, emocional o moral.

V. LAS EMOCIONES MORALES

Hay, en cambio, emociones que tradicionalmente se ha venido considerando que sí fomentan ese avance, emociones que nos sentimos tentados a llamar intrínsecamente morales. Entre ellas, ocupan un lugar importante la benevolencia y la compasión, aunque también podrían incluirse en esa lista los sentimientos de culpa o de vergüenza y la indignación. Esta última, en especial, fue considerada la emoción moral básica por François Itard, el médico que en el siglo XVIII redactó un informe sobre el niño salvaje que había encontrado en los bosques de Aveyron. Itard llevó a cabo un experimento diseñado para comprobar si el niño salvaje había adquirido el sentido de la justicia. Este experimento consistió en «castigarle» cuando hacía algo correcto. Itard quedó muy satisfecho al comprobar que, en contraste con su aceptación habitual del castigo, el niño reaccionó con ira e indignación (Itard, 1964, sección 54).

El aspecto más interesante de este experimento es que, como el niño seguía sin saber hablar, al interpretar su conducta, Itard supuso que sabía distinguir con claridad la expresión de indignación o de ira, de otras como las de vergüenza, tristeza o frustración, que esas emociones y sus expresiones eran naturales y prelingüísticas. Itard estaba probablemente en lo cierto, según las observaciones posteriores de Darwin y más recientemente de Paul Ekman (Darwin, 1998), que parecen dar por sentado que una media docena de emociones, incluyendo la ira y el miedo, tienen expresiones faciales que las hacen claramente identificables en todas las culturas. Pese a todo, es difícil saber si partiendo de la actitud de un niño hacia el castigo podemos generalizar y referirnos a otros aspectos de la justicia. Especialmente la falta de lenguaje impide determinar si el niño extendería también su indignación ante ultrajes cometidos contra otras personas distintas a él o si sólo reacciona de este modo cuando es él quien sale perdiendo por una distribución injusta. Lo más probable es que, cuando intervine el lenguaje, puedan utilizarse estos sentimientos «naturales» para configurar ese sentido más universal de la justicia que descubrimos en los estudios sobre el altruismo.

El altruismo es una característica que originalmente se aplica a la motivación, aunque no hay argumentos definitivos que permitan concluir si los sujetos son siempre altruistas o si ayudan a otros, en última instancia, para asegurarse la satisfacción que obtienen al hacerlo. Los argumentos no son convincentes porque, aun en el caso de que deseemos realmente el bien de otro, nada excluye que el placer experimentado se deba al logro de un bien que nos interesaba. La tesis del hedonismo inmediato que he usado hasta ahora para basar mi argumento acepta claramente la idea de que el placer obtenido por el agente en el acto es la causa directa que motiva las conductas benevolentes, pero esto es plenamente compatible con el hecho de que mi interés por la otra persona sea verdaderamente sincero. Esto resta mucho interés a la discusión.

Desde hace unos años, el altruismo ha sido analizado en el terreno de la biología, donde ha sido definido como un tipo de conducta que acarrea una pérdida clara en el bienestar de un agente, vinculada a una obtención de beneficio por parte de otros. Así entendido, esta concepción representa una paradoja para el darwinismo, planteando, así, cuestiones mucho más sustanciales. Efectivamente, el reto

que el darwinismo tiene planteado (en suma, que toda tendencia heredable a reducir nuestra conveniencia en beneficio de la conveniencia de otro ha de ser eliminada necesariamente de la existencia) exige que detectemos mecanismos capaces de explicar cómo los organismos obran de forma altruista. De ahí que se apele a la selección familiar, al altruismo recíproco y a los modelos de selección grupal, que han suscitado feroces disputas en las publicaciones de las dos últimas décadas. Pues cabe preguntar si la evolución puede habernos inculcado ciertas disposiciones emocionales que quepa interpretar en términos altruistas y no sólo en términos de la sociabilidad que promueven (Sober y Wilson, 1998).

La perspectiva evolucionista, como han observado Sober y Wilson, «nos obliga a considerar la mente humana como un mecanismo próximo a la hora de causar que algunos organismos produzcan conductas adaptativas» (Sober y Wilson, 1998, 200). Los deseos intrínsecos o últimos, en especial, al margen de si son o no invariablemente egocéntricos, parecen hallarse psicológicamente lejos de los medios concretos con los que se satisfacen, aunque son los medios próximos por los que la selección asegura su «finalidad». Desde este ángulo, mi lista de «emociones morales» admisibles podría haber sido, en parte, seleccionada con ese objetivo.

Como ha defendido Martha Nussbaum (1996), la compasión es una de esas pocas emociones que tienen un valor especial por su carácter moral, esto es, emociones que por sí mismas tienden a generar comportamientos morales. La perspectiva sociobiológica que he estado adoptando hasta aquí no considera que sea un vicio metodológico tener en cuenta el efecto global de un determinado tipo de conducta como una clave para detectar su función. Pero la lección principal que podemos aprender de Nussbaum es que hemos de comprobar cómo actúa la compasión en un contexto mucho más amplio. Necesitamos verla a todo color, esto es, en su verdadero contexto multidimensional. Lo que sigue explica por qué.

VI. AVANZANDO HACIA LA VISIÓN A TODO COLOR

Nussbaum conjetura que «sólo será objeto de mi interés el dolor de otro si reconozco que existe una cierta comunidad ente el otro y yo, y si entiendo lo que representaría para mí enfrentarme a ese dolor» (Nussbaum, 1996, 35). E insiste: «El juicio sobre las posibles similitudes forma parte de una construcción que salva el vacío existente ente el interés prudente y el altruismo» (36). Me pregunto, sin embargo, si Nussbaum no lo entiende al revés; esto es, si una determinada capacidad natural para experimentar el dolor de otro, de un modo que nos mueva realmente a tenerlo en cuenta, no podría ser en parte la causa de esa capacidad nuestra que realiza ese acto mental mucho más complicado de «reconocer una comunidad». No veo por qué no dar crédito a los informes que apuntan a la existencia de una capacidad en algunos de nuestros primos mamíferos, como las ballenas, las marsopas o los elefantes, que les permite confraternizar con otros miembros de su especie o incluso con otras especies de mamíferos (Masson, 1995, 79-82). Pero dar por válida esta observación directa no parece que nos exija atribuir a las ballenas algo tan elaborado como el reconocimiento de una comunidad a la que pertenecerían ellas y nosotros.

Esta sugerencia responde aproximadamente a la idea de Robert Gordon y otros de que simular el estado mental de otro puede ser la forma más práctica de acabar conociendo la mente ajena (Gordon, 1986; Davies y Stone, 1995). Aunque este supuesto es teórico, no carece enteramente de fundamento práctico. Hay pruebas de que esta simulación podría estar realmente mediada por unas neuronas especiales que, ya sea por aprendizaje o por maduración, se activan precisamente al observar los actos motores relativos a ciertos fines del mismo modo que se activan durante el proceso de realizar esos mismos actos. Esta conclusión sobre las «neuronas-espejo» ha llevado a conjeturar que compartimos con los primates no humanos una capacidad de «atribución de un estado intencional». Tal atribución no llega a ser una comprensión auténtica y genuina de estados mentales ajenos, pero «constituye un estadio filogenético necesario en la senda de la evolución que conduce al desarrollo de la capacidad de lectura mental que es propia de los seres humanos» (Gallese y Goldman, 1998, 500). En esta investigación, no se abordan las emociones; pero no parece inverosímil ampliar la especulación a la captación de estados emocionales de otros. Para otorgarle al menos cierta credibilidad, basta decir que las expresiones de las emociones implican actividad motora, y que entre las fascinantes conclusiones de Paul Ekman y su equipo de investigadores sobre la expresión emocional, se encuentra el hecho de que una configuración de los músculos de la cara ocasionada por un control directamente aprendido de esos mismos músculos, cuando coincide plenamente con la expresión de un tipo de emoción, es la causa que predispone a que se sienta esa emoción (Ekman y Friesen, 1975; Ekman, 1999).

A la luz de estas consideraciones, si concebimos la compasión como una capacidad algo más primitiva de conocer emociones, podríamos incluso conceder a Nussbaum que la compasión es la «emoción social básica», mientras evitamos, en cierta medida, el debate sobre la «evolución del altruismo». Pues la capacidad para la compasión podría ser en realidad un efecto colateral de una capacidad claramente adaptativa como es el poder leer la mente de otro, lo que S. J. Gould ha llamado un «spandrel». Es evidente que interpretar correctamente el estado mental de alguien que puede influir en nuestro bienestar constituye una ventaja fundamental (Humphrey, 1984). Pero puede darse también el caso de que no se pueda llevar a cabo esto sin que exista un cierto grado de empatía. Para aceptar esta idea no se requiere realmente que resolvamos el problema de saber en qué consiste: («¿Pensamos realmente (...) que somos la víctima? ¿Imaginamos que nuestras respuestas se confunden de un modo misterioso con las de la víctima?») (Nussbaum, 1996, 34). Todo lo que se requiere es que nuestra captación del estado mental de otro sea correcta, y que dicha corrección tenga el efecto colateral de poder mover la mente del intérprete. A veces, este efecto colateral puede tener un coste, si conduce a una conducta altruista, pero puede no ser tan costoso cuando se sopesa el beneficio general que reporta.

El proceso por el que esa capacidad de «leer» emocionalmente la mente de otro evolucionó podría ilustrar el interesante mecanismo conocido como el «efecto Baldwin» (Baldwin, 1986; Hall, 2001; Deacon, 1997, 322 y siguientes). El efecto Baldwin fue concebido originalmente como un modo de explicar ciertos aspectos de la teoría de Lamarck en términos estrictamente darwinistas. En su forma más

general, consiste en un proceso de realimentación en bucle que empieza con una conducta no instintiva, que produce un cambio en el medio ambiente, cambio que lleva, a su vez, a nuevas urgencias selectivas que favorecen disposiciones genéticas a ciertos tipos de conducta. Cuando la conducta que aparece al final del proceso es del mismo tipo que la del principio, podemos obtener algo que se parece bastante a un episodio de la evolución lamarckiana, puesto que se dará realmente el caso de que un organismo particular (un grupo de organismos), que elija una determinada forma de conducta, podrá, en última instancia, hacer que quede genéticamente codificada la predisposición a dicha conducta. Ahora bien, esto sólo se produce en casos especiales; en toda su generalidad, el mecanismo identificado por Baldwin sólo tiene en cuenta el hecho de que la conducta puede influir en los genes; el carácter indirecto de la influencia garantizará normalmente que los resultados sean por entero impredecibles, de forma que el efecto Baldwin lo que hace, en realidad, es dirigir la evolución de los organismos capaces de realizar conductas en direcciones enteramente nuevas.

En nuestro caso, cabe sugerir que el «sentido de comunidad» podría no ser la condición sino el efecto lejano e impredecible de un cambio en la forma de vida de nuestros antepasados que aumentó su dependencia mutua. Las cosas podrían haberse originado así: supongamos un cambio de conducta en una población que conlleve un comienzo de sociabilidad. Una vez producido, aumentará naturalmente la ventaja de quienes son capaces de «leer la mente» de los demás. En tales circunstancias, puede que el método más eficaz de leer la mente sea catalogar las emociones. (De otro modo, como observa con acierto Nussbaum, no distinguiríamos «entre aceptar realmente una proposición y limitarse a pronunciar las palabras». De ahí que comente Nussbaum a propósito del aserto de Rousseau «Ver algo sin sentirlo es no conocerlo», que ello «significa algo muy concreto: que el sufrimiento de otros no ha llegado a formar parte del repertorio cognitivo de Emilio (...)» (Nussbaum, 1996, 38).) Cuando se da el caso que venimos mencionando, alcanzamos un nivel de sociabilidad más pleno, un estadio donde, como dice Nussbaum del espectador de una tragedia, uno se sienta arrastrado por la compasión que suscita el arte del trágico o del moralista a «captar las desigualdades respecto a la vulnerabilidad: así tendremos razones para elevar el nivel de seguridad para todos» (45). Y, en este sentido, nuestra capacidad innata para la compasión empática hará que nos acerquemos al ideal estoico de desarrollo moral consistente en «desplazar los círculos (concéntricos de intereses que están más alejados de los puramente egoístas) cada vez más y de modo progresivo hacia el centro, de forma que nuestros padres lleguen a ser como nosotros mismos, nuestros familiares como nuestros padres, los desconocidos como familiares, y así sucesivamente» (Nussbaum, 1996, 48).

Recordemos que, al comienzo de mi exposición, hablaba de la necesidad de pasar del ámbito de lo que he llamado la proyección en blanco y negro de nuestras emociones a la visión a todo color. Mi exposición empezaba con la simple instrumentalidad de la lectura de otras mentes; y al formular mi postura, no muy alejada de la mantenida por Nussbaum, hablaba de la posibilidad de abordar la compasión en su forma más simple, de forma que la pregunta no sería: «¿Cómo puede aprenderse un sentimiento tan complejo?» (Nussbaum, 1996, 39), sino: «¿Cómo

pudo evolucionar este instinto tan simple y tan fuerte?». Ahora bien, para dar paso a la imagen completa de lo que implica semejante lectura, tenemos que ver más allá de la necesidad unidimensional de hacer elecciones: es evidente que mi conocimiento de otras mentes me importa ante todo y sobre todo por fines prácticos, pero si la vida emocional es como Nussbaum la describe (y como yo he tratado de captar con la noción de escenarios paradigmáticos), entonces conocer la mente de otro es conocer algo que no puede representarse adecuadamente proyectándolo en blanco y negro, en los términos bidimensionales de elegir entre actuar y no actuar. Se trata de experiencias axiológicas ricas y de dimensiones inconmensurables, que la compasión o empatía se esfuerzan realmente en captar.

VII. EL RETO DE LAS EMOCIONES RECHAZABLES

Entre los argumentos que esgrime Nussbaum para defender la importancia moral de la compasión, hay algunos destinados a responder a la queja de que tolerarla supone dar pie a la aparición de pasiones mucho menos aceptables. Quien acepta los valores que implica considerar que algunas cosas son realmente dignas de compasión «admite ser prisionero del azar. Y ser prisionero del azar supone dar rienda suelta no sólo a la compasión, sino también al miedo, a la ansiedad, a la aflicción (...), así como a la cólera y a la disposición retributiva» (Nussbaum, 1996, 43). Nussbaum responde a esta queja diciendo que el defensor de la importancia moral de la emoción «no está obligado a admitir que cualquier perjuicio, ofensa o insulto sea una ocasión apropiada para la cólera retributiva. Pues el mayor número de los males sociales causados por la venganza ocasionan daños a la riqueza, la posición social, el poder y el honor, a los que el defensor de la compasión no concede mucho valor» (Nussbaum, 1996, 49). Esta respuesta me parece bastante aceptable, pero me gustaría situarla en una perspectiva algo más amplia.

Cabe esperar que los defensores de la importancia de las emociones para la vida racional y moral sean acusados de ignorar el hecho de que algunas emociones son *rechazables*. Una emoción es característicamente rechazable si se dan todas las circunstancias que siguen:

- (i) experimentarla es muy repugnante;
- (ii) está expuesta a una fuerte desaprobación;
- (iii) va siempre o normalmente seguida de consecuencias malas.

Las emociones rechazables representan un reto para quien defiende que las emociones son percepciones de valor. ¿Cómo se puede compaginar la perspectiva axiológica con la existencia de emociones rechazables? ¿Afecta negativamente a la idea de que las emociones son captaciones de valores el hecho de que existan emociones rechazables?

La respuesta consiste en atender a la diferencia entre lo que he llamado la perspectiva de las emociones en blanco y negro y a todo color. En un mundo exclusivamente en blanco y negro, es decir, en un mundo donde sólo se tiene en cuenta la dimensión práctica inmediata (perspectiva donde, como he defendido, las simplificaciones del hedonismo inmediato pueden aportar una sorprendente

claridad), las emociones rechazables serían condenadas por el simple hecho de que su existencia ejerce, en conjunto, una influencia negativa en la utilidad, bien directamente o bien combinadas con sus consecuencias. Ya sea que «los daños a la riqueza, la posición social, el poder y el honor» sean considerados importantes o no, constituyen un perjuicio y son una consecuencia típica de las emociones dirigidas a la venganza. Por el contrario, desde el punto de vista axiológico el problema no es si tienen consecuencias malas, del mismo modo que no puede ser un problema epistemológico el que una determinada creencia pueda tener consecuencias malas. Las creencias verdaderas pueden ser desalentadoras o incluso deprimentes (y disponemos también de interesantes pruebas de que la depresión, precisamente, puede favorecer creencias verdaderas (Pacini y otros, 1998)); pero esto no es importante epistemológicamente hablando: las consecuencias de tener una creencia no constituyen su corrección. De igual modo, tener un deseo «correcto», es decir, desear algo que es «bueno» o realmente deseable, puede acarrear consecuencias sumamente indeseables; pero las consecuencias de tener el deseo no son importantes para la cuestión de si el objeto del deseo era, efectivamente, como lo representaba el deseo, esto es, bueno o deseable. Lo que determina que un estado específico sea «correcto», en sus propios términos, es si la meta que persigue se acomoda o no se acomoda al objeto formal del estado. En el caso de la creencia, la dimensión análoga es la verdad; en el caso del deseo, lo es la bondad o la deseabilidad. En el caso de la emoción, esto viene determinado por el escenario paradigmático que la define: ese objeto formal establece una dimensión axiológica. Ahora bien, hemos de resistirnos a pensar que existe una sola dimensión que constituiría, por así decirlo, la conveniencia o propiedad emocional, pues cada emoción tiene, en verdad, su propia pauta de conveniencia o propiedad. Otra cuestión distinta y separable de cualquier problema epistemológico específico sobre una creencia determinada sería si esa creencia concreta constituye o no un componente probable de una vida buena en un contexto específico. De igual modo, en cada emoción se plantean dos cuestiones separables: una de ellas se refiere al grado en que la emoción se ajusta a su objeto en términos de la relación de dicho objeto con el objeto formal de la emoción. La segunda cuestión es cómo podría contribuir, en el contexto presente, una emoción así a la vida buena.

Aquí se hace necesario recurrir a lo que en otro lugar he llamado holismo axiológico (De Sousa, 2000). En nuestro mundo a todo color de la dimensionalidad emocional indefinida, sólo podemos juzgar el valor de una emoción en particular dentro del contexto global en el que se experimenta. Cada emoción ha surgido de un modo específico, constituyendo su propio objeto formal específico. Ahora bien, ni en la selección natural ni en la biografía del individuo hay nada que proporcione alguna forma de teleología intrínseca y que pueda dictaminar que todas las emociones serán útiles, placenteras, altruistas, o recomendables por alguna otra razón. Con todo, la sola existencia de una emoción tiene, seguramente, su peso en nuestras vidas o, mejor, en la vida que queremos llevar. Ese peso puede verse contrarrestado por el de otras emociones; y esto es verdad incluso si se trata de emociones coherentes con una concepción de la buena vida razonable. Así, en algunas ocasiones, emociones socialmente útiles como la compasión o la benevolencia podrán confundirnos moralmente, pero, de igual modo, la indudable utilidad

de la facultad de ver no nos garantiza que nunca se producirán ilusiones ópticas. El hecho de que las emociones en general constituyan captaciones de realidad axiológica no implica que todas las emociones sean fidedignas por igual. Esto vale para los dos niveles de apreciación que hemos distinguido: algunas emociones pueden ser engañosas en su contexto, aunque sean adecuadas en sí mismas; otras pueden ser sistemáticamente engañosas o rechazables. En ambos casos, cabría hablar de errores axiológicos. Pero ello no demuestra que no haya algo que cada emoción capta correctamente; ni demuestra tampoco que el hecho de que una emoción capte algo correcto sea garantía suficiente de que esa emoción es algo bueno.

Todavía debemos abordar otro problema. Aunque muchos filósofos han tendido a pensar que las emociones son más un producto de la naturaleza que de la educación, yo estaría más de acuerdo en que las emociones más importantes para el desarrollo moral son fruto del aprendizaje. Esto podría ser cierto en gran medida incluso respecto a las emociones «básicas» universales, pero vale indudablemente al menos para todas las demás. No estoy queriendo decir que sean uniformes necesariamente dentro de todo grupo humano, sino que se producen en conformidad con la presión social o como una reacción contra ésta: en cualquier caso, es muy improbable que el repertorio de emociones de un niño escape a las influencias culturales que las han configurado, aunque haya cambios en un pequeño número de vectores. La ley de hierro de la mayoría de las sociedades es la conformidad. Esto se aprecia claramente en el informe de Sober y Wilson sobre culturas de distintas partes del mundo: son extraordinariamente raras las sociedades como la nuestra, donde la tiranía del grupo no fuerza a los individuos a responder a las expectativas sociales, sobre todo en aquellos aspectos que podrían afectar, aunque fuese remotamente, a la reproducción, el parentesco y la sexualidad (Sober y Wilson, 1998, capítulo 5). Es fácil encontrar también ejemplos contemporáneos de lo que vengo diciendo. Hace poco presencié una entrevista a varios hombres que habían asesinado a una hermana o a una hija como una cuestión de «honor», es decir, porque la víctima había sido violada. Lo más asombroso de las actitudes de esos hombres era que no parecían sufrir ni el más mínimo asomo de arrepentimiento, ni siquiera de culpa, respecto a lo que su sentido del «honor» les había movido a hacer. Hasta tal punto llega la fuerza del condicionamiento social.

La moraleja es que el holismo axiológico no puede meramente limitarse a examinar, más allá de las emociones que experimenta un individuo, las emociones aprobadas por su sociedad; pues nuestro entendimiento nos hace ver que muchas de las emociones promovidas socialmente son rechazables en grado sumo. Podemos juzgar si una emoción contribuye a la realización del bien o del mal de la misma manera que comprobamos la veracidad de una información sensorial: observando lo mismo desde distintos puntos de vista, lo que significa, naturalmente, recurrir a otras percepciones o, salvando las diferencias, a otras emociones. El método del equilibrio reflexivo viene a ser un lugar común tanto en la ciencia como en la ética. Aunque con frecuencia no se aprecia algo que Nussbaum ha dejado sumamente claro: que en el caso de la ética los datos que se han de equilibrar no son de ordinario hechos empíricos sino respuestas emocionales. Tales respuestas emo-

cionales se entienden mejor como percepciones axiológicas, y para dar cuenta de ellas suficientemente no basta con atender a sus proyecciones bidimensionales en el terreno del placer y del dolor, o a la media docena de dimensiones que suministran las emociones estandarizadas. Trasladar a unas pocas dimensiones un mundo multidimensional puede ser no sólo engañoso sino positivamente incoherente, como sucede con esos dibujos bidimensionales de Escher que representarían algo imposible en las tres dimensiones que pretenden captar. Esta es la razón de que, en el mundo multicolor de las emociones reales, el elevado nivel de especificidad que puede alcanzar el arte y la literatura de ficción no sea sólo un instrumento sino un elemento básico del aprendizaje moral y emocional.

Traducción: *Enrique López Castellón*

REFERENCIAS BIBLIOGRÁFICAS

- Baldwin, J. M. (1896): «Consciousness and evolution», en *Psychological Review*, nº 3, pp. 300-308.
- Campbell, S. (1998): *Interpreting the Personal: Expression and the Formation of Feeling*, Ithaca, Cornell University.
- Clynes, M. (1977): *Sentics: The Touch of Emotions*, Garden City, Anchor Press.
- D'Arms, J. y Jacobson, D. (2001): «The moralistic fallacy: on the 'appropriateness' of emotions», en *Philosophy and Phenomenological Research*, nº 61, pp. 65-90.
- Darwin, Charles (1998) (1872, 1889): *The Expression of the Emotions in Man and Animals*, Ed. with intro, afterword and commentaries by Paul Ekman, London, Harper Collins.
- Davies, M. y Stone, T. (eds.) (1995): *Mental Simulation: Evaluations and Applications*. Readings in Mind and Language, Oxford, Blackwell.
- De Sousa, R. (2000): «Learning to be natural», en Roughley, N. (ed.), *Being Humans*, Berlin-New York, De Gruyter, pp. 87-307.
- (1987): *The Rationality of Emotion*, MIT Press, Bradford Book.
- (1974): «The good and the true», en *Mind*, nº 83, pp. 534-551.
- Deacon, T. W. (1997): *The Symbolic Species: The Coevolution of Language and the Brain*, New York, Norton.
- Ekman, P. (1999): «Emotional and conversational nonverbal signals». Conferencia. San Sebastián.
- Ekman, P. y Friesen, W. (1975): *Unmasking the Face: A guide to Recognizing Emotions from Facial Expressions*, Englewood Cliffs (NJ), Prentice-Hall.
- Gallese, V. y Goldman, A. (1998): «Mirror neurons and the simulation theory of mind-reading», en *Trends in Cognitive Sciences*, vol. 2, nº 12, pp. 493-501.
- Gallop, D. (1960): «True and false pleasures», en *Philosophical Quarterly*.
- Gordon, R. M. (1986): «Folk psychology and simulation», en *Mind and Language*, nº 1, pp. 158-171.
- Gould, S. J. (1979): *Proceedings of the Royal Society of London*, nº 205, pp. 581-598. *The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme*.
- Gould, S. J. y Vrba, E. (1982): *Paleobiology*, nº 8, pp. 4-15 y *Exaptation: A Missing Term in the Science of Form*.
- Griffiths, P. E. (1996): «The historical turn in the study of adaptation», en *British Journal*

for the *Philosophy of Science*, n° 47, pp. 511-532.

Hall, Brian K. (2001): «Organic selection: proximate environmental effects on the evolution of morphology and behaviour», en *Biology and Philosophy*, n° 16, pp. 215-237.

Humphrey, N. (1984): «Nature's psychologists», en *Consciousness Regained: Chapters in the Development of Mind*, Oxford: Oxford University Press, pp. 29-41.

Itard, J. (1964): «Mémoire et rapport sur Victor de l'Aveyron», en L. Malson, *Les enfants sauvages: mythe et réalité*, Paris, Union Générale d'éditions.

Masson, J. M. (1995): *When Elephants Weep: The Emotional Lives of Animals*, New York, Delacorte Press.

Nozick, R. (1974): *Anarchy, State, and Utopia*, New York, Basic Books.

Nussbaum, M. (1997): *Cultivating Humanity: a Classical Defense of Reform in Liberal Education*, Cambridge (MA), Harvard University Press.

— (1996): «Compassion: The basic social emotion», en *Social Philosophy and Policy*, n° 13, pp. 27-58.

— (1994): *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*, Princeton, Princeton University Press.

— (1990): *Love's Knowledge*, Oxford-New York-Toronto, Oxford University Press.

Pacini, R., Muir, F. y Epstein, S. (1998): «Depressive realism from the perspective of cognitive-experiential self-theory», en *Journal of Personality and Social Psychology*, n° 74, pp. 1056-1068.

Penner, T. (1970): «False anticipatory pleasures in the *Philebus* 36b13-41a6», en *Phronesis*, pp. 166-178.

Rosenfield, I. (1992): *The Strange, Familiar and Forgotten*, New York, Alfred A. Knopf.

Sober, E. y Wilson, D. S. (1998): *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge (MA), Harvard University Press.