Teoría de la Elección Social

por Christian List

Social Choice Theory, The Stanford Encyclopedia of Philosophy (Winter 2013 Edition), Edward N. Zalta (ed.).
Agradezco a los editores, sus revisores, Franz Dietrich, Iain McLean y Michael Morreau por sus comentarios.

La teoría de la elección social es el estudio de los procesos y procedimientos de decisión colectiva. No es una sola teoría, sino un conjunto de modelos y resultados relativos a la agregación de insumos individuales (por ejemplo, votos, preferencias, juicios, bienestar) en productos colectivos (p.ej., decisiones colectivas, preferencias, juicios, bienestar). Las preguntas centrales son: ¿Cómo puede un grupo de individuos elegir un resultado ganador (p.ej., política, candidato electoral) de un conjunto dado de opciones? ¿Cuáles son las propiedades de los diferentes sistemas de votación? ¿Cuándo es democrático un sistema de votación? ¿Cómo puede un colectivo (p.ej., electorado, legislatura, tribunal colegiado, panel de expertos, o comité) llegar a preferencias o juicios colectivos coherentes sobre algunas cuestiones, sobre la base de preferencias o juicios individuales de sus miembros? ¿Cómo podemos clasificar las distintas alternativas sociales en un orden de bienestar social? Los teóricos de la elección social estudian estas preguntas no sólo considerando ejemplos, sino desarrollando modelos generales y demostrando teoremas.

Teniendo como pioneros en el siglo XVIII a Nicolás de Condorcet y Jean-Charles de Borda y en el siglo XIX a Charles Dodgson (también conocido como Lewis Carroll), la teoría de la elección social despegó en el siglo XX con las obras de Kenneth Arrow, Amartya Sen, y Duncan Black. Su influencia se extiende a lo largo de la economía, la ciencia política, la filosofía, las matemáticas y, recientemente, la informática y la biología. Aparte de contribuir a nuestra comprensión de los procedimientos de decisión colectiva, la teoría de la elección social tiene aplicaciones en las áreas de diseño institucional, economía del bienestar y epistemología social.

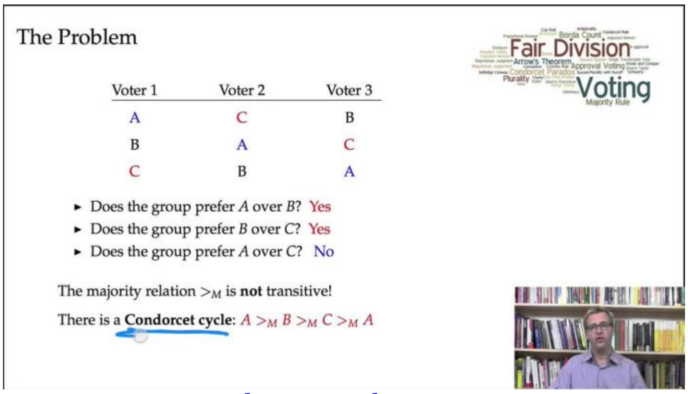
1. Historia de la teoría de la elección social

1.1 Condorcet

Nicolas de Condorcet 1743-1794
Life and times 11m 14s

Los dos estudiosos más a menudo asociados con el desarrollo de la teoría de la elección social son el francés Nicolas de Condorcet (1743-1794) y el estadounidense Kenneth Arrow (1921-2017). Marie-Jean-Antoine Nicolas de Caritat, marqués de Condorcet fue un pensador liberal en la era de la Revolución Francesa que fue perseguido por las autoridades revolucionarias por criticarlas. Después de un período de ocultación, al final fue arrestado, aunque aparentemente no identificado de inmediato, y murió en prisión (para más detalles sobre Condorcet, véase McLean y Hewitt 1994). En su *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* [Ensayo sobre la Aplicación del Análisis a la Probabilidad de las Decisiones de Mayoría] (1785), abogó por un sistema de votación particular, la mayoría de votos por parejas, y presentó sus dos ideas más destacadas. La primera, conocida como *teorema del jurado de Condorcet*, es que si cada miembro de un jurado tiene una oportunidad igual, independiente y mejor que al azar, pero peor que perfecta, de dar un juicio correcto sobre si un acusado es culpable (o alguna otra proposición factual), la mayoría de los miembros del jurado tiene más probabilidades de estar en lo correcto que cada jurado individual, y la probabilidad de que un juicio de la mayoría sea correcto se aproxima a 1 a medida que aumenta el tamaño del jurado. Por lo tanto, bajo ciertas condiciones, la regla de la mayoría es buena para "rastrear la verdad" (p.ej., Grofman, Owen y Feld, 1983; List y Goodin 2001).

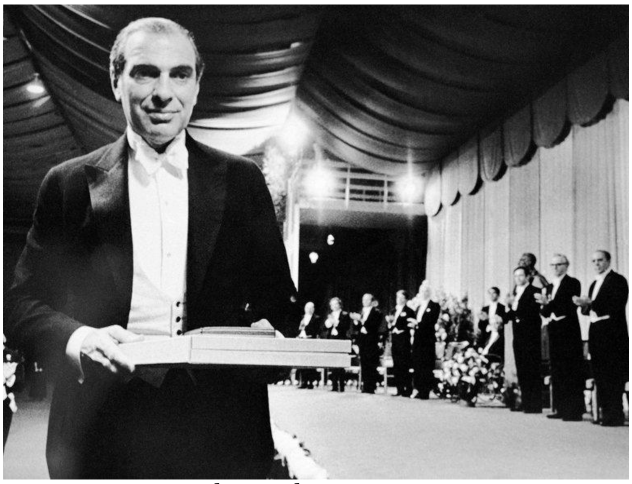
La segunda visión de Condorcet, a menudo llamada *paradoja de Condorcet*, es la observación de que las preferencias de la mayoría pueden ser "irracionales" (específicamente, intransitivas) aún cuando las preferencias individuales sean "racionales" (específicamente, transitivas). Supongan, por ejemplo, que un tercio de un grupo prefiere x a y a z, un segundo tercio prefiere y a z a x, y un tercio final prefiere z a x a y. Entonces hay mayorías (de dos tercios) para x contra y, para y contra z, y para z contra x: un 'ciclo', que viola la transitividad. Además, ninguna alternativa es un ganador de Condorcet, una alternativa que venza, o por lo menos empate con, cualquier otra alternativa en los concursos de parejas por mayoría.



Condorcet Paradox *14m 55s*

1.2 Arrow y su influencia

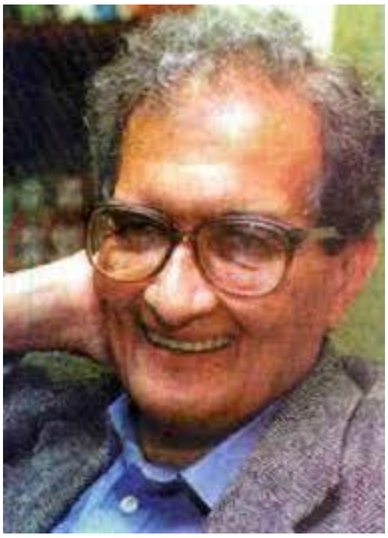
En tanto que Condorcet había investigado un *método particular de votación* (votación mayoritaria), Arrow, que ganó el Premio Nobel de Economía en 1972, introdujo un enfoque general del estudio de la agregación de preferencias, inspirado en parte por su profesor de lógica Alfred Tarski (1901- 1983), con quien había estudiado teoría de relaciones como estudiante en el City College de Nueva York (Suppes 2005). Arrow consideró una clase de *posibles* métodos de agregación, que denominó *funciones de bienestar social*, y preguntó cuál de ellos satisface ciertos axiomas o desiderata. Resultó que, sorprendentemente, no existe ningún método de agrupar preferencias de dos o más individuos sobre tres o más alternativas en las preferencias colectivas, cuando este método debe satisfacer cinco axiomas aparentemente plausibles, discutidos a continuación.



Kenneth Joseph Arrow 1921-2017
Foto de la ceremonia del Nobel 1972, Estocolmo

Este resultado, conocido como *teorema de imposibilidad de Arrow*, motivó muchos trabajos y debates en la teoría de la elección social y la economía del bienestar. William Riker (1920-1993), que inspiró la escuela de Rochester en ciencias políticas, lo interpretó como una prueba matemática de la imposibilidad de la democracia populista (p.ej., Riker 1982). Otros, en forma destacada Amartya Sen (nacido en 1933), quien ganó el Premio Nobel 1998, lo utilizó para demostrar que las preferencias ordinales son insuficientes para hacer elecciones sociales satisfactorias. Los comentaristas también cuestionaron si los desiderata de Arrow sobre un método de agregación son tan inocuos como se afirmó o si deberían ser relajados.

Las lecciones del teorema de Arrow dependen, en parte, de cómo interpretamos una función de bienestar social arroviana. El uso de las preferencias ordinales como *aggregenda* puede ser más fácil de justificar si interpretamos la regla de agregación como un *método de votación* que si lo interpretamos como un *método de evaluación del bienestar*. Sen argumentó que cuando un planificador social busca ordenar diferentes alternativas sociales en un orden de bienestar social (empleando así alguna regla de agregación como un método de evaluación del bienestar), puede justificarse el uso de información adicional más allá de las preferencias ordinales, p.ej. medidas interpersonalmente comparables de bienestar (p.ej., Sen 1982).



Amartya Kumar Sen n. 1933
Nobel 1998

El propio Arrow opinó *que la comparación interpersonal de utilidades no tiene sentido y... que no hay significado relevante en las comparaciones de bienestar de mensurabilidad de la utilidad individual.* (1951/1963: 9)

Este punto de vista estaba influido por la economía neoclásica, asociada con eruditos como Vilfredo Pareto (1848-1923), Lionel Robbins (1898-1984), John Hicks (1904-1989), coganador del Premio Nobel de Economía con Arrow y Paul Samuelson (1915-2009), otro Premio Nobel. El teorema de Arrow demuestra las crudas implicancias de los supuestos "ordinalistas" del pensamiento neoclásico.

Jean-Charles de Borda 1733-1799

Hoy en día la mayoría de los teóricos de la elección social han ido más allá de las interpretaciones negativas tempranas del teorema de Arrow y están interesados en los trade-offs implicados en hallar procedimientos de decisión satisfactorios. Sen promovió esta interpretación *'posibilista'* de la teoría de la elección social (por ejemplo, en su conferencia Nobel de 1998).

Dentro de este enfoque, el método axiomático de Arrow ha sido tal vez más influyente que su teorema de imposibilidad (sobre el método axiomático, véase Thomson 2000). El tipo paradigmático de resultado en el trabajo axiomático contemporáneo es el "*teorema de caracterización*". Aquí el objetivo es identificar un conjunto de condiciones plausibles necesarias y suficientes que caracterizan de manera única una solución particular (o clase de soluciones) a un tipo dado de problema de decisión colectiva. Un ejemplo temprano es la caracterización de Kenneth May (1952) de la regla de la mayoría, que se discute más abajo.

1.3 Borda, Carroll, Black y otros

Condorcet y Arrow no son las únicas figuras fundadoras de la teoría de la elección social. El contemporáneo y co-nacional de Condorcet Jean-Charles de Borda (1733-1799) defendió un sistema de votación que se ve a menudo como una alternativa prominente a la votación por mayoría. El recuento de Borda, formalmente definido más adelante, evita la paradoja de Condorcet pero viola una de las condiciones de Arrow, la independencia de alternativas irrelevantes. Así, el debate entre Condorcet y Borda es un precursor de algunos debates modernos sobre cómo responder al teorema de Arrow.

Los orígenes de este debate preceden a Condorcet y Borda. En la Edad Media, Ramón Llull (c 1235-1315) propuso el método de agregación de la votación por mayoría de a pares, mientras que Nicolas Cusanus (1401-1464) propuso una variante del recuento de Borda (McLean 1990). En 1672, el estadista alemán Samuel von Pufendorf (1632-1694) comparó la mayoría simple, la mayoría cualificada y las reglas de unanimidad y ofreció un análisis de la estructura de las preferencias que puede ser vista como un precursor de descubrimientos posteriores (p.ej., sobre situaciones *con punta única*, lo que se discute más adelante) (Gaertner 2005).

En el siglo XIX, el matemático y diácono anglicano británico Charles Dodgson (1832-1898), más conocido como Lewis Carroll, redescubrió de forma independiente muchas de las ideas de Condorcet y Borda y también desarrolló una teoría de la representación proporcional. En gran medida, gracias al economista escocés Duncan Black (1908-1991), las ideas de Condorcet, Borda y Dodgson sobre la teoría de la elección social fueron traídas a la atención de la comunidad de investigación moderna (McLean, McMillan y Monroe, 1995). Black también hizo varios descubrimientos relacionados con el voto por mayoría, algunos de los cuales se analizan a continuación.

En Francia, George-Théodule Guilbaud ([1952] 1966) escribió un artículo importante pero a menudo pasado por alto, revisando la teoría del voto de Condorcet desde una perspectiva lógica y provocando una literatura francesa sobre el efecto Condorcet, el problema lógico que subyace a la paradoja de Condorcet. Recientemente recibió más atención en la teoría de la elección social anglófona (Monjardet 2005). Para más contribuciones sobre la historia de la teoría de la elección social, véase McLean, McMillan y Monroe (1996), McLean y Urken (1995), McLean y Hewitt (1994) y un número especial de *Social Choice and Welfare*, editado por Salles (2005).

2. Tres argumentos formales para la regla de la mayoría

Para introducir formalmente la teoría de la elección social, es útil considerar un simple problema de decisión: una elección colectiva entre dos alternativas.

2.1 El concepto de una regla de agregación

Sea $N = \{1, 2,..., n\}$ un conjunto de individuos, donde $n \geq 2$. Los individuos tienen que elegir entre dos alternativas (candidatos, políticas, etc.). Cada individuo $i \in N$ echa un voto, denotado v_i, donde

- v_i = 1 representa un voto por la primera alternativa,
- v_i = -1 representa un voto por la segunda alternativa, y opcionalmente
- v_i = 0 representa una abstención (para simplificar, dejamos esta posibilidad a un lado).

Una combinación de votos entre los individuos, $<v_1, v_2,..., v_n>$, se denomina *perfil*. Para cualquier perfil, el grupo busca llegar a una decisión social v, donde

- v = 1 representa una decisión por la primera alternativa,
- v = -1 representa una decisión por la primera alternativa, y
- $v = 0$ representa un empate.

Una *regla de agregación* es una función f que asigna a cada perfil $<v_1, v_2,..., v_n>$ (en algún dominio de perfiles admisibles) una decisión social $v = f(v_1, v_2,..., v_n)$. Como ejemplos tenemos:

Regla de mayoría: Para cada perfil $<v_1, v_2,..., v_n>$,

$f(v_1, v_2,..., v_n)$= 1 si $v_1 + v_2 + ... + v_n > 0$ (es decir, hay más "1" que "-1");

= 0 si $v_1 + v_2 +... + v_n = 0$ (o sea, hay tantos "1" como "-1"), y

= -1 si $v_1 + v_2 +... + v_n < 0$ (es decir, hay menos "1" que "-1").

Dictadura: Para cada perfil $<v_1, v_2,..., v_n>$,

$f(v_1, v_2,..., v_n)= v_i$,

Donde $i \in N$ es un individuo fijado con anterioridad (el *dictador*).

Regla de mayoría ponderada: Para cada perfil $<v_1, v_2,..., v_n>$,

$f(v_1, v_2,..., v_n)= 1$, si $w_1v_1+w_2v_2+ ...+ w_nv_n >0$,

$= 0$, si $w_1v_1+w_2v_2+ ...+ w_nv_n = 0$,

$= -1$, si $w_1v_1+w_2v_2+ ...+ w_nv_n <0$,

donde $w_1, w_2,..., w_n$ son números reales, interpretados como "ponderaciones de votos" de los n individuos.

Vale la pena señalar dos puntos sobre el concepto de una regla de agregación. Primero, en la definición estándar, una regla de agregación se define de forma *extensiva, no intensiva*: es un *mapa* (una *relación funcional*) entre inputs individuales y outputs colectivos, *no un conjunto de instrucciones explícitas* (una regla en el sentido del lenguaje ordinario) que podría extenderse a inputs fuera del dominio formal de la función. En segundo lugar, se define una regla de agregación para un conjunto fijo de individuos N y un problema de decisión fijo, de modo que la regla de mayoría en un grupo de dos individuos es un objeto matemático diferente de la regla de mayoría en un grupo de tres.

Para ilustrar esto, las Tablas 1 y 2 muestran la regla de la mayoría para estos dos tamaños de grupo como objetos extensivos. Las filas de cada cuadro corresponden a los diferentes perfiles posibles de votos; la columna final muestra las decisiones sociales resultantes.

Tabla 1 Regla de mayoría con dos individuos

Voto del Individuo 1	*Voto del Individuo 2*	*Decisión Colectiva*
1	*1*	*1*
1	*-1*	*0*
-1	*1*	*0*
-1	*-1*	*-1*

Tabla 1 Regla de mayoría con tres individuos

Voto del Individuo 1	*Voto del Individuo 2*	*Voto del Individuo 3*	*Decisión Colectiva*
1	*1*	*1*	*1*
1	*1*	*-1*	*1*
1	*-1*	*1*	*1*
1	*-1*	*-1*	*-1*
-1	*1*	*1*	*1*
-1	*1*	*-1*	*-1*
-1	*-1*	*1*	*-1*
-1	*-1*	*-1*	*-1*

Esta forma de representar una regla de agregación nos ayuda a ver cuántas posibles reglas de agregación existen (véase también List 2011). Supongan que hay k perfiles en el dominio de insumos admisibles (en el ejemplo presente, $k = 2^n$, ya que cada uno de los n individuos tiene dos opciones, con la abstención descartada). Supongan, además, que hay L decisiones sociales posibles para cada perfil (en el ejemplo, $L = 3$, permitiendo empates). Entonces hay L^k reglas de agregación posibles: la tabla relevante tiene k filas, y en cada fila, hay L formas posibles de especificar la entrada final (la decisión colectiva). Luego, el número de posibles reglas de agregación crece *exponencialmente* con el número de perfiles admisibles y el número de posibles resultados de la decisión.

Para seleccionar una regla de agregación en forma no arbitraria de esta clase enorme de posibilidades, se requieren algunas restricciones. Ahora consideraré tres argumentos formales para la regla de la mayoría.

2.2 Un argumento de procedimiento para la regla de la mayoría

El primer argumento consiste en imponer algunos requisitos «de procedimiento» sobre la relación entre votos individuales y decisiones sociales y mostrar que la regla de la mayoría es la única regla de agregación que los satisface. May (1952) introdujo cuatro requisitos de este tipo:

Dominio universal: El dominio de inputs admisibles de la regla de agregación consiste de todos los perfiles de votos lógicamente posibles $<v_1, v_2,..., v_n>$, donde cada $v_i \in \{-1,1\}$.

Anonimato: Para cualesquier perfiles admisibles $<v_1, v_2,..., v_n>$ y $<w_1, w_2,..., w_n>$ que son permutaciones entre sí (es decir, uno puede obtenerse del otro reordenando las entradas) la decisión social es la misma, es decir, $f(v_1, v_2,..., v_n) = f(w_1, w_2,..., w_n)$.

Neutralidad: Para cualquier perfil admisible $< v_1, v_2,..., v_n >$, si los votos por las dos alternativas son invertidos, la decisión social también se invierte, es decir, $f(-v1, -v2,..., -vn) = -f(v_1, v_2,..., v_n)$.

Capacidad de Respuesta Positiva: Para cualquier perfil admisible $< v_1, v_2,..., v_n >$, si algunos votantes cambian sus votos a favor de una alternativa (digamos la primera) y todos los demás votos permanecen iguales, la decisión social no cambia en la dirección opuesta; si la decisión social era un empate antes del cambio, el empate se rompe en la dirección del cambio, es decir, si [$w_i > v_i$ para algún i y $w_j = v_j$ para todos los demás j] y $f(v_1, v_2, ..., v_n) = 0$ o 1, entonces $f(w_1, w_2, ..., w_n) = 1$.

Christian List Profesor de Ciencia Política y Filosofía, London School of Economics
Web Page

El dominio universal requiere que la regla de agregación haga frente a cualquier nivel de "pluralismo" en sus inputs; el anonimato, que se trate a todos los votantes por igual; la neutralidad, que se trate a todas las alternativas por igual; y la capacidad de respuesta positiva exige que la decisión social sea una función positiva de la manera en que vota la gente. May demostró lo siguiente:

> *Teorema* (May 1952): Una regla de agregación satisface los supuestos de dominio universal, anonimato, neutralidad y capacidad de respuesta positiva si y sólo si es la regla de la mayoría.

Además de proporcionar un argumento para la regla de la mayoría basada en cuatro desiderata plausibles de procedimiento, el teorema nos ayuda a caracterizar otras reglas de agregación en términos de desiderata violados. La dictadura y las reglas de mayoría ponderada con ponderaciones individuales desiguales violan el anonimato. Las reglas de las súper mayorías asimétricas (en las que se requiere una súper mayoría de votos, como dos tercios o tres cuartos, para tomar una decisión a favor de una de las alternativas, cuando la otra alternativa es la opción por defecto) violan la neutralidad. Esto a veces puede justificarse, por ejemplo cuando hay una presunción a favor de una alternativa, como una presunción de inocencia en una decisión del jurado. Las reglas de las súper mayorías simétricas (en las cuales no se elige ninguna alternativa a menos que esté sostenida por una súper mayoría suficientemente grande) violan la respuesta positiva. Un ejemplo más inverosímil de una regla de agregación que viola la capacidad de respuesta positiva es la regla inversa de la mayoría (aquí gana la alternativa rechazada por una mayoría).

2.3 Un argumento epistémico para la regla de la mayoría

El teorema del jurado de Condorcet proporciona un argumento consecuencialista para la regla de la mayoría. El argumento es "epistémico", en la medida que la regla de agregación se interpreta como un dispositivo de rastreo de la verdad (p.ej., Grofman, Owen y Feld 1983, List y Goodin 2001).

Supongan que el objetivo es hacer un juicio sobre algún hecho o estado del mundo independiente del procedimiento, denotado X. En una decisión del jurado, el acusado es culpable (X = 1) o inocente (X = -1). En una decisión de un panel de expertos sobre la seguridad de alguna tecnología, la tecnología puede ser segura (X = 1) o no (X = -1). El voto de cada individuo expresa un juicio sobre ese hecho o estado, y la decisión social representa el juicio colectivo. El objetivo es alcanzar un juicio colectivo factualmente correcto. La regla de agregación que se desempeña mejor en "rastrear la verdad" depende de la relación entre los votos individuales y el hecho o estado relevante del mundo.

Condorcet supuso que al hacer un juicio correcto el de cada individuo resulta mejor que hacerlo en forma aleatoria (supuesto de *competencia*) y que los juicios de diferentes individuos son estocásticamente independientes, dado el estado del mundo (supuesto de *independencia*). Formalmente, $V_1, V_2,..., V_n$ (mayúsculas) denotan las variables aleatorias que generan los votos individuales específicos $v_1, v_2,..., v_n$ (minúsculas) y hacemos que $V= f(V_1, V_2,..., V_n)$ denote la variable aleatoria resultante que representa la decisión social $v = f(v_1, v_2,..., v_n)$ bajo una regla de agregación f, como p.ej. la regla de la mayoría. Los supuestos de Condorcet se pueden enunciar de la siguiente manera:

> *Competencia*: Para cada individuo $i \in N$ y cada estado del mundo $x \in \{-1,1\}$, $Pr(V_i = x \mid X = x) = p > ½$, donde p es la misma entre individuos y estados.
>
> *Independencia*: Los votos de diferentes individuos $V_1, V_2,..., V_n$ son independientes entre sí, condicionados a cada valor $x \in \{-1,1\}$ de X.

Con estos supuestos, la votación por mayoría es un buen rastreador de la verdad:

> *Teorema* (teorema del jurado de Condorcet): Para todo estado del mundo x ∈ {-1,1}, la probabilidad de una decisión correcta de la mayoría, $Pr(V = x \mid X = x)$, es mayor que la probabilidad de un voto correcto de cada individuo $Pr(Vi = x \mid X = x)$, y converge a 1, a medida que aumenta el número de individuos n.[1]

El primer enunciado ("es mayor que la probabilidad de un voto correcto de cada individuo") es la *conclusión no asintótica*, el segundo ("converge a 1") la *conclusión asintótica*. Se puede

[1] Cuando n es par, la primera parte del teorema sólo es válida para tamaños de grupo n por encima de cierto límite inferior (que depende de p), debido a la posibilidad de empates. Si n es impar, es válido para cualquier n> 1.

probar además que, si los dos estados del mundo tienen una probabilidad previa igual (es decir, *Pr (X = 1) = Pr (X = -1) = ½*), la regla de la mayoría es la más fiable de todas las reglas de agregación, maximizando Pr (V = X) (p.ej., Ben-Yashar y Nitzan 1997).

Aunque el teorema del jurado se invoca a menudo para establecer los méritos epistémicos de la democracia, sus supuestos son altamente idealistas. El supuesto de competencia no es un enunciado conceptual sino empírico y depende del problema de decisión dado. Aunque una competencia individual media (no necesariamente igual) por encima de ½ puede bastar para la conclusión de Condorcet (por ejemplo, Grofman, Owen y Feld 1983, Boland 1989, Kanazawa 1998),[2] el teorema deja de tener validez si los individuos randomizan (no son mejores ni peores que arrojar una moneda) o si son peores que una decisión aleatoria ($p < ½$). En este último caso, la probabilidad de una decisión correcta de la mayoría es menor que la probabilidad de un voto correcto de cada individuo y converge a 0, a medida que aumenta el tamaño del jurado. La conclusión del teorema también puede ser socavada en casos menos extremos (Berend y Paroush 1998), p.ej., cuando la fiabilidad de cada individuo, aunque superior a ½, sea una función exponencialmente decreciente que se aproxima a ½ al aumentar el tamaño del jurado (List 2003a).

De modo similar, que el supuesto de independencia sea verdadero depende del problema de decisión en cuestión. Aunque la conclusión de Condorcet sea robusta a la presencia de algunas interdependencias entre votos individuales, la estructura de esta interdependencia sí importa (p.ej., Boland 1989, Ladha 1992, Estlund 1994, Dietrich y List 2004, Berend y Sapir 2007, Dietrich y Spiekermann 2013). Si todos los votos de los individuos están perfectamente correlacionados entre sí o imitan a un pequeño número de líderes de opinión, el juicio colectivo no es más confiable que el juicio de un pequeño número de individuos independientes.

Las redes bayesianas, empleadas en el trabajo de Pearl sobre causalidad (2000), se han utilizado para modelar los efectos de las dependencias de los votantes sobre el teorema del jurado y para distinguir entre variantes más fuertes y más débiles de independencia condicional (Dietrich y List 2004; Dietrich y Spiekermann 2013). Dietrich (2008) ha sostenido que los dos supuestos de Condorcet nunca están justificados simultáneamente, en el sentido de que, aun cuando ambos sean verdaderos, no se puede obtener evidencia que apoye a ambos a la vez.

Finalmente, en teoría de los juegos se desafía un supuesto implícito del teorema del jurado, a saber que los votantes siempre revelarán sus juicios en forma veraz. Aún si todos los votantes prefieren un juicio colectivo correcto a uno incorrecto, pueden tener incentivos para falsificar sus juicios individuales. Esto puede suceder cuando, a condición de que el evento sea fundamental para el resultado, un votante espera una mayor probabilidad de lograr un juicio colectivo correcto votando en contra de su propio juicio privado y no en línea con él (Austin-Smith y Banks 1996; Feddersen y Pesendorfer 1998).

2.4 Un argumento utilitario para la regla de la mayoría

[2] Si diferentes individuos tienen diferentes niveles conocidos de confiabilidad, el voto por mayoría ponderada supera al voto por mayoría simple para maximizar la probabilidad de una decisión correcta, con el peso de cada voto individual proporcional a $log\ (p\ /\ (1-p))$, donde p es la confiabilidad del individuo como se definió anteriormente (Shapley y Grofman 1984, Grofman, Owen y Feld 1983, Ben-Yashar y Nitzan 1997).

Otro argumento consecuencialista para la regla de la mayoría es utilitario en lugar de epistémico. No requiere la existencia de un hecho o estado independiente del mundo que la decisión colectiva supuestamente rastree. Supongan que cada votante obtiene alguna utilidad de la decisión colectiva, que depende de si la decisión coincide con su voto (preferencia): específicamente, cada votante obtiene una utilidad de 1 de una coincidencia entre su voto y el resultado colectivo y una utilidad de 0 de un desajuste.[3] El teorema de Rae-Taylor enuncia entonces que si cada individuo tiene una probabilidad previa igual de preferir cada una de las dos alternativas, la regla de la mayoría maximiza la utilidad esperada de cada individuo (véase, por ejemplo, Mueller 2003).

En forma relacionada, la regla de la mayoría minimiza el número de votantes frustrados (definidos como votantes del lado perdedor) y maximiza la utilidad total entre los votantes. Brighouse y Fleurbaey (2010) generalizan este resultado. Definen la apuesta del votante i en la decisión, d_i, como la diferencia de utilidad entre su resultado preferido y su resultado no deseado. Brighouse y Fleurbaey muestran que cuando se permite que las apuestas varíen entre los votantes, la utilidad total se maximiza no con la regla de la mayoría, sino con una regla de mayoría ponderada, en la que la ponderación del voto de cada individuo i sea proporcional a su apuesta d_i.

3. Agregación de preferencias

En el corazón de la teoría de la elección social está el análisis de la agregación de preferencias, entendida como la agregación de varios rankings de preferencias de dos o más alternativas sociales en un solo ranking de preferencias (o decisiones) colectivo sobre estas alternativas.

El modelo básico es como sigue. De nuevo, consideren un conjunto N = {1, 2,..., n} de individuos (n ≥ 2). Sea $X = \{x, y, z,...\}$ un conjunto de alternativas sociales, p.ej. mundos posibles, plataformas de políticas, candidatos a elecciones o asignaciones de bienes. Cada individuo $i \in N$ tiene un *orden de preferencias* R_i sobre estas alternativas: una relación binaria completa y transitiva sobre X.[4] Para cualquier $x, y \in X$, xR_iy significa que el individuo i prefiere débilmente x a y. Escribimos xP_iy si xR_iy y no yR_ix ("individuo i prefiere estrictamente x a y"), y xI_iy si xR_iy e yR_ix ("individuo i está indiferente entre x e y").

Una combinación de órdenes de preferencia entre los individuos, $<R_1, R_2,..., R_n>$, se denomina un *perfil*. Una *regla de agregación de preferencias*, F, es una función que asigna a cada perfil $<R_1, R_2,..., R_n>$ (en algún dominio de perfiles admisibles) una relación de preferencia social $R = F(R_1, R_2,..., R_n)$ en X. Cuando F resulta clara por el contexto, simplemente escribimos R para la relación de preferencia social correspondiente a $<R_1, R_2,..., R_n>$.

Para cualquier $x, y \in X$, xRy significa que x es socialmente débilmente preferido a y. También escribimos xPy si xRy y no yRx ('x es estrictamente socialmente preferido a y'), y xIy si xRy e yRx ('x e y están socialmente empatados o indiferentes'). Para mayor generalidad, el requisito de que R sea completo y transitivo no está incorporado en la definición de una regla de agregación de preferencias.

---

[3] Opcionalmente, se puede estipular que la utilidad de un empate es ½.
[4] La completitud requiere que, para cualquier $x, y \in X$, xR_iy o yR_ix, y la transitividad requiere que, para cualquier x, y, z ∈ X, si xR_iy e yR_iz, luego xR_iz.

El ejemplo paradigmático de una regla de agregación de preferencias es la votación mayoritaria por parejas, tal y como la discutió Condorcet. En tal caso, para cualquier perfil $<R_1, R_2,..., R_n>$ y cualesquiera $x, y \in X$, xRy si y sólo si al menos tantos individuos tienen xR_iy como tienen yR_ix, formalmente $| \{i \in N: xR_iy\} | \geq | \{i \in N: yR_ix\} |$. Como hemos visto, esto no garantiza preferencias sociales transitivas.[5]

¿Cuán frecuentes son las preferencias de mayoría intransitivas? Se puede demostrar que la proporción de perfiles de preferencia (entre todos los posibles) que conducen a preferencias cíclicas mayoritarias aumenta con el número de individuos (n) y el número de alternativas (| X |). Si todos los perfiles de preferencias posibles tienen la misma probabilidad de ocurrir (el supuesto escenario de "cultura imparcial"), los ciclos mayoritarios deberían ser probables en los grandes electorados (Gehrlein 1983). (Los trabajos técnicos distinguen aún más entre "ciclos-top" y ciclos por debajo de una posible alternativa ganadora-Condorcet.) Sin embargo, la probabilidad de ciclos puede ser significativamente menor bajo ciertas desvíos sistemáticos, incluso pequeños, de una cultura imparcial (List y Goodin, 2001, Apéndice 3, Tsetlin, Regenwetter y Grofman 2003, Regenwetter et al., 2006).

3.1 Teorema de Arrow

Haciendo abstracción de la votación mayoritaria por parejas, Arrow presentó las siguientes condiciones de una regla de agregación de preferencias, F:

> _Dominio universal_: El dominio de F es el conjunto de todos los perfiles lógicamente posibles de ordenamientos de preferencias individuales completas y transitivas.
>
> _Ordenamiento_: Para todo perfil $<R_1, R_2,..., R_n>$ en el dominio de F, la relación de preferencia social R es completa y transitiva.
>
> _Principio débil de Pareto_: Para todo perfil $<R_1, R_2,..., R_n>$ en el dominio de F, si para todo $i \in N$ xP_iy, entonces xPy.
>
> _Independencia de alternativas irrelevantes_: Para cualesquiera dos perfiles $<R_1, R_2,..., R_n>$ y $<R^*_1, R^*_2,..., R^*_n>$ en el dominio de F y cualquier $x, y \in X$, si para todos $i \in N$ el ordenamiento de R_i entre x e y coincide con el ordenamiento de R^*_i entre x e y, luego xRy si y sólo si xR^*y.
>
> _No-dictadura_: No existe un individuo $i \in N$ tal que, para todo $<R_1, R_2,..., R_n>$ en el dominio de F y todo $x, y \in X$, xP_iy implica xPy.

El dominio universal requiere que la regla de agregación haga frente a cualquier nivel de "pluralismo" en sus inputs. El ordenamiento requiere producir preferencias sociales "racionales", evitando los ciclos de Condorcet. El principio débil de Pareto requiere que cuando todos los individuos prefieran estrictamente la alternativa x a la alternativa y, también lo haga la sociedad. La independencia de alternativas irrelevantes requiere que la preferencia social entre dos alternativas x e y dependa sólo de las preferencias individuales entre x e y, no de las preferencias de los individuos sobre otras alternativas. La no-dictadura requiere

---

[5] En el ejemplo clásico, hay tres individuos con ordenamientos de preferencia xP_1yP_1z, yP_2zP_2x y zP_3xP_3y sobre tres alternativas x, y, y z. Las preferencias mayoritarias resultantes son cíclicas: tenemos xPy, yPz, y sin embargo zPx.

que no haya "dictador", que determine siempre la preferencia social, en forma independiente de las preferencias de los demás. (Tengan en cuenta que la votación mayoritaria por parejas cumple todas estas condiciones, excepto el ordenamiento.)

Teorema (Arrow 1951/1963): Si | X | > 2, no existe ninguna regla de agregación de preferencias que satisfaga el dominio universal, el ordenamiento, el principio débil de Pareto, la independencia de alternativas irrelevantes y la no-dictadura.

Es evidente que este resultado se traslada a la agregación de otros tipos de ordenamientos, a diferencia de los ordenamientos de preferencias, tales como (i) los ordenamientos de creencias sobre varias hipótesis (credenciales ordinales), (ii) los criterios múltiples que un único tomador de decisiones puede usar para generar un ordenamiento de todas las cosas de varias opciones de decisión, y (iii) las ordenaciones de valores en conflictos a reconciliar.

Ejemplos de otros problemas de agregación a los que se ha aplicado el teorema de Arrow son: problemas de agregación intrapersonal (por ejemplo, May 1954, Hurley 1985), agregación de restricciones en la teoría de la optimalidad en lingüística (p.ej., Harbour y List 2000), elección de teorías (p.ej. Okasha 2011, cf. Morreau próximamente), evidencia de fusión (p.ej., Stegenga 2013) y la agregación de ordenamientos de similitud múltiples en un ordenamiento de similitud todo-considerado (por ejemplo, Morreau 2010, Kroedel y Huber 2013). En cada caso, la plausibilidad del teorema de Arrow depende de la plausibilidad casuística específica del marco ordinalista de Arrow y de las condiciones del teorema.

Generalmente, si consideramos apropiado el marco de Arrow e indispensables sus condiciones, el teorema de Arrow plantea un serio desafío. Para evitarlo, debemos relajar al menos una de las cinco condiciones o abandonar la restricción de los inputs de la regla de agregación a ordenamientos y defender el uso de inputs más ricos, como se discute en la Sección 4.

3.2 Reglas de agregación de preferencias no dictatoriales

3.2.1 Relajación del dominio universal

Una forma de evitar el teorema de Arrow es relajar el dominio universal. Si se requiere que la regla de agregación acepte como inputs sólo perfiles de preferencia que satisfagan ciertas condiciones de "cohesión", las reglas de agregación como la votación mayoritaria por pares producirán preferencias sociales completas y transitivas. La condición de cohesión más conocida es la de *punta única* (Black 1948).

Un perfil $<R_1, R_2, \ldots, R_n>$ tiene punta única si las alternativas pueden alinearse de "izquierda" a "derecha" (por ejemplo, en alguna dimensión cognitiva o ideológica) de manera que cada individuo tenga una posición preferida en esa alineación con preferencia decreciente a medida que las alternativas se alejan más (en cualquier dirección) desde la posición más preferida. Formalmente, esto requiere la existencia de un ordenamiento lineal Ω en X tal que, para cada triple de alternativas $x, y, z \in X$, si y está entre x y z con respecto a Ω, no sea el caso que xR_iy y zR_ix (Esto descarta una "cueva" entre x y z, en y). En algunos contextos democráticos es plausible la punta única. Si las alternativas en X son tasas de impuestos diferentes, por ejemplo, cada individuo puede tener una tasa de impuesto más preferida (que será más baja para un individuo libertario que para un socialista) y prefiere otras tasas de impuestos menos a medida que se alejan del ideal.

Black (1948) demostró que si el dominio de la regla de agregación está restringido al conjunto de todos los perfiles de ordenamientos de preferencias individuales que satisfagan la condición de punta única, no pueden ocurrir ciclos mayoritarios y la alternativa más preferida del individuo *mediano* con respecto a la alineación izquierda-derecha es un ganador de Condorcet (suponiendo que n sea impar). La votación mayoritaria por pares satisface entonces el resto de las condiciones de Arrow.

Otras condiciones de restricción de dominio con implicancias similares incluyen la condición de *cavidad única*, una imagen especular geométrica de la condición de punta única (Inada 1964), *separabilidad en dos grupos* (ibid.) y "*latin-squarelessness*" (Ward 1965), siendo las dos últimas condiciones combinatorias más complicadas (para una revisión, véase Gaertner 2001). Sen (1966) mostró que todas estas condiciones implican una condición más débil, la *restricción de valor triple*. Exige que para cada triple de alternativas $x, y, z \in X$, exista una alternativa en $\{x, y, z\}$ y un rango $r \in \{1, 2, 3\}$ tal que ningún individuo clasifique esa alternativa en el lugar r-ésimo entre x, y, y z. Por ejemplo, todos los individuos pueden estar de acuerdo en que y no es la peor parte ($r = 3$) entre x, y, y z. La restricción de valor triple es suficiente para preferencias mayoritarias transitivas (para una simple prueba del teorema de Sen, véase Elsholtz y List 2005).

Ha habido mucha discusión sobre si, y bajo qué condiciones, las preferencias en el mundo real caen en un dominio tan restringido. Se sugirió, por ejemplo, que la deliberación de grupo puede inducir preferencias de punta única, llevando a los participantes a enfocarse en una dimensión cognitiva o ideológica compartida (Miller, 1992, Knight y Johnson, 1994; Dryzek y List 2003). La evidencia experimental de las encuestas de opinión deliberativas es consistente con esta hipótesis (List, Luskin, Fishkin y McLean, 2013), aunque se necesita más trabajo empírico.

3.2.2 Relajación del ordenamiento

Normalmente se espera que las reglas de agregación de preferencias tengan como outputs ordenamientos, pero a veces sólo podemos requerir ordenamientos parciales o relaciones binarias no plenamente transitivas. Una regla de agregación que genera preferencias sociales transitivas pero a menudo incompletas es el procedimiento de dominancia de Pareto: aquí, para cualquier perfil $\langle R_1, R_2, \ldots, R_n \rangle$ y cualquier $x, y \in X$, xRy si y sólo si, para todo i ∈ N, XP_iy. Una regla de agregación que produce preferencias sociales completas pero a menudo intransitivas es el procedimiento de extensión de Pareto: aquí, para cualquier perfil $\langle R_1, R_2, \ldots, R_n \rangle$ y cualquier $x, y \in X$, xRy si y sólo si no es el caso que, para todo $i \in N$, yP_ix. Ambas reglas tienen un espíritu unánime, otorgando a cada individuo *poder de veto* o bien contra la presencia de una débil preferencia social por x sobre y o contra su ausencia.

Gibbard (1969) demostró que incluso si reemplazamos el requisito de transitividad por lo que llamó *cuasi-transitividad*, las posibilidades de agregación resultantes son todavía muy limitadas. Llamamos a una relación de preferencia R *cuasi-transitiva* si la relación estricta inducida P es transitiva (mientras que la relación de indiferencia no requiere serlo). Llamamos a una regla de agregación *oligárquica* si existe un subconjunto $M \subseteq N$ (los "oligarcas") tal que (i) si, para todo $i \in M$, xP_iy, entonces xPy, y (ii) si para algún $i \in M$, xP_iy, entonces xRy. El procedimiento de extensión de Pareto es un ejemplo de una regla de agregación oligárquica con $M = N$. En una oligarquía, los oligarcas son conjuntamente decisivos y tienen poder de veto individual. Gibbard demostró lo siguiente:

Teorema (Gibbard 1969): Si $|X| > 2$, no existe una regla de agregación de preferencias que sea cuasi-transitiva, y satisfaga el dominio universal, la completitud de las preferencias sociales, el principio débil de Pareto, la independencia de las alternativas irrelevantes y la no-oligarquía.

3.2.3 Relajación del principio débil de Pareto

Es sin duda difícil renunciar al principio débil de Pareto. Un caso en el que podemos levantarlo es el de unanimidad espuria, cuando una preferencia unánime para x sobre y se basa en razones mutuamente inconsistentes (por ejemplo, Mongin 1997, Gilboa, Samet y Schmeidler 2004). Dos hombres pueden cada uno preferir luchar un duelo (alternativa x) a no combatir (alternativa y) porque cada uno sobre-estima sus posibilidades de ganar. Puede no existir una asignación de probabilidad mutuamente aceptable sobre posibles resultados del duelo (es decir, quién ganaría) que "racionalice" la preferencia unánime por x sobre y. En este caso, la preferencia unánime es un mal indicador de la preferencia social. Este ejemplo, sin embargo, depende del hecho de que las alternativas de luchar y no combatir no son resultados completamente especificados sino perspectivas inciertas. Se puede argumentar que el principio débil de Pareto es más plausible en casos sin incertidumbre.

Una regla de agregación que se torna posible cuando el débil principio de Pareto es abandonado es un *procedimiento impuesto*, donde, para cualquier perfil $<R_1, R_2,..., R_n>$, la relación de preferencia social R es un ordenamiento prefijado ("impuesto") $R_{impuesto}$ de alternativas. Aunque sea completamente insensible a las preferencias individuales, esta regla de agregación satisface el resto de las condiciones de Arrow.

Sen (1970a) ofreció otra crítica del principio débil de Pareto, mostrando que está en conflicto con un principio "liberal". Aquí se interpreta la regla de agregación como un método que un planificador social puede usar para clasificar las alternativas sociales en un orden de bienestar social. Supongan que cada individuo de la sociedad tiene ciertos derechos básicos, en el sentido de que su preferencia es a veces socialmente decisiva (es decir, no puede ser invalidada por las preferencias de los demás). Tanto Lewd como Prude, por ejemplo, deberían ser decisivos sobre si cada uno lee un libro en particular, *Lady Chatterley's Lover*.

Mínimo liberalismo : Hay *al menos* dos individuos distintos $i, j \in N$ que son cada uno decisivos en al menos un par de alternativas; es decir, hay al menos un par de alternativas $x, y \in X$ tal que, para cada perfil $<R_1, R_2,..., R_n>$, xP_iy implica xPy, e yP_ix implica yPx, y al menos un par de alternativas $x^*, y^* \in X$ tal que, para cada perfil $< R_1, R_2,..., R_n >$, $x^*P_jy^*$ implica x^*Py^*, e $y^*P_jx^*$ implica y^*Px^*.

Sen nos pidió que imagináramos que lo que más prefiere Lewd es que Prude lea el libro (alternativa x), lo segundo es que prefiere leer él mismo el libro (alternativa y), y lo que menos prefiere es que ninguno lea el libro (z). Lo que más prefiere Prude es que ninguno lea el libro (z), en segundo lugar, que él mismo lea el libro (x), y lo que menos prefiere es que Lewd lea el libro (y). Suponiendo que Lewd es decisivo sobre el par y y z, la sociedad debe preferir y a z. Suponiendo que Prude es decisivo sobre el par x y z, la sociedad debe preferir z a x. Pero como Lewd y Prude ambos prefieren x a y, el principio débil de Pareto (aplicado a N = {Lewd, Prude}) implica que la sociedad debe preferir x a y. Por lo tanto, nos enfrentamos a un ciclo de preferencia social. Sen llamó a este problema la "paradoja liberal" y lo generalizó de la siguiente manera.

Teorema (Sen 1970a): No existe una regla de agregación de preferencias que satisfaga el dominio universal, la aciclicidad de las preferencias sociales, el principio débil de Pareto y el mínimo liberalismo.

El resultado sugiere que si queremos respetar los derechos individuales, a veces tenemos que sacrificar la eficiencia paretiana. Una conclusión alternativa es que el principio débil de Pareto puede hacerse compatible con el liberalismo mínimo sólo cuando el dominio de perfiles de preferencias admisibles está adecuadamente restringido, por ejemplo, a preferencias que sean "tolerantes" o no "entrometidas" (Blau 1975, Craven 1982, Gigliotti 1986, Sen 1983). Las preferencias de Lewd y Prude en el ejemplo de Sen son "entrometidas".

Varios autores han cuestionado la relevancia del resultado de Sen, sin embargo, criticando su formalización de los derechos (p.ej., Gaertner, Pattanaik y Suzumura 1992; Dowding y van Hees 2003).

3.2.4 Relajación de la independencia de alternativas irrelevantes

Una forma común de obtener posibles reglas de agregación de preferencias es renunciar a la independencia de alternativas irrelevantes. Casi todos los métodos de votación familiares sobre tres o más alternativas que involucran algún tipo de voto preferencial (con votantes a los que se pide expresar ordenamientos de preferencias plenos o parciales) violan esta condición.

Un ejemplo estándar es la *regla de pluralidad*: aquí, para cualquier perfil $<R_1, R_2,..., R_n>$ y cualquier $x, y \in X$, xRy si y sólo si $|\{i \in N: \text{para todo } z \neq x, xP_iz\}| \geq |\{i \in N: \text{para todo } z \neq y, yP_iz\}|$. Dicho informalmente, las alternativas se clasifican socialmente en el orden de cuántos individuos prefieren cada uno de ellas. La regla de la pluralidad evita la paradoja de Condorcet, pero se tropieza con otros problemas. El más notable es que una alternativa que es dispreferida por la mayoría a cualquier otra alternativa puede ganar bajo la regla de la pluralidad: si el 34% de los votantes ordena a x por encima de y por encima de z, 33% ordenan a y por encima de z por encima de x, y 33% ordenan a z por encima de y por encima de x, la regla de pluralidad ordena a x por encima de cada uno de y y z, mientras que la votación mayoritaria por parejas clasificaría a y por encima de z por encima de x (y es el ganador de Condorcet). Al ignorar las alternativas de menor rango, la regla de pluralidad también viola el principio débil de Pareto. Sin embargo, la regla de la pluralidad puede ser plausible en "ambientes restringidos de información", donde el procedimiento de votación recopila información solamente sobre las preferencias principales de los votantes, no sobre su clasificación completa de preferencias. Aquí, la regla de pluralidad satisface variantes generalizadas de las cuatro condiciones de May introducidas previamente (Goodin y List 2006).

Un segundo ejemplo de una regla de agregación de preferencias que viola la independencia de alternativas irrelevantes es el *recuento de Borda* (por ejemplo, Saari 1990). Aquí, para cualquier perfil $<R_1, R_2,..., R_n>$ y cualquier $x, y \in X$, xRy si y sólo si $\Sigma_{i \in N} | \{z \in X: xR_iz\} | \geq \Sigma_{i \in N} | \{z \in X: yR_iz\} |$. Informalmente, cada votante asigna una puntuación a cada alternativa, que depende de su rango en su clasificación de preferencia. La alternativa más preferida obtiene una puntuación de k (donde $k = | X |$), la segunda alternativa más preferida, una puntuación de $k - 1$, la tercera alternativa más preferida, una puntuación de $k - 2$, y así sucesivamente. Las alternativas son entonces ordenadas socialmente en términos de las sumas de sus calificaciones entre los votantes: la alternativa con la mayor suma total es la superior, la alternativa con la segunda mayor suma total es la siguiente, y así sucesivamente.

Para ver cómo es violada la independencia de alternativas irrelevantes, consideren los dos perfiles de ordenamientos de preferencia individuales sobre cuatro alternativas (x, y, z, w) en las Tablas 3 y 4.

Tabla 3. Un perfil de ordenamiento de preferencias individuales

	Individuo 1	*Individuos 2 a 7*	*Individuos 8 a 15*
1ra preferencia	*y*	*x*	*z*
2da preferencia	*x*	*z*	*x*
3ra preferencia	*z*	*w*	*y*
4ta preferencia	*w*	*y*	*w*

Tabla 4. Perfil ligeramente modificado de preferencias individuales

	Individuo 1	*Individuos 2 a 7*	*Individuos 8 a 15*
1ra preferencia	*x*	*x*	*z*
2da preferencia	*y*	*z*	*x*
3ra preferencia	*w*	*w*	*y*
4ta preferencia	*z*	*y*	*w*

En la Tabla 3, las puntuaciones de Borda de las cuatro alternativas son:

- *x: 9*3 + 6*4 = 51,*
- *y: 1*4 + 6*1 + 8*2 = 26,*
- *z: 1*2 + 6*3 + 8*4 = 52,*
- *w: 1*1 + 6*2 + 8*1 = 21,*

conduciendo a una preferencia social por z sobre x sobre y sobre w. En la Tabla 4 las puntuaciones de Borda son:

- *x: 7*4 + 8*3 = 52,*
- *y: 1*3 + 6*1 + 8*2 = 25,*
- *z: 1*1 + 6*3 + 8*4 = 51,*
- *w: 7*2 + 8*1 = 22,*

conduciendo a una preferencia social por *x* sobre *z* sobre *y* sobre *w*. La única diferencia entre los dos perfiles radica en el orden de preferencias del individuo 1, e incluso aquí no hay cambio en el ranking relativo de *x* y *z*. A pesar de preferencias individuales idénticas entre *x* y *z* en las Tablas 3 y 4, se invierte la preferencia social entre x y z, una violación de la independencia de alternativas irrelevantes.

Tales violaciones son comunes en las reglas de votación del mundo real, y hacen que la agregación de preferencias sea potencialmente vulnerable a la votación estratégica y / o a la fijación estratégica de una agenda. Se ilustra esto en el caso del voto estratégico.

3.3 El teorema de Gibbard-Satterthwaite

Hasta ahora se han discutido *reglas de agregación de preferencias*, que mapean perfiles de preferencias individuales a relaciones de preferencia social. Ahora consideramos *reglas de elección social*, cuyo resultado, en cambio, es una o varias alternativas ganadoras. Formalmente, una regla de elección social, f, es una función que asigna a cada perfil $<R_1, R_2,..., R_n>$ (en algún dominio de perfiles admisibles) un conjunto de opciones sociales $f(R_1, R_2,..., R_n) \subseteq X$. Una regla de elección social f puede derivarse de una regla de agregación de preferencias F, definiendo $f(R_1, R_2,..., R_n) = \{x \in X$: para todo $y \in X, xRy\}$ donde $R=F(R_1, R_2,..., R_n)$. Lo inverso no es generalmente válido. Llamamos al conjunto de alternativas a veces elegidas el *rango* de f.[6]

El criterio del ganador de Condorcet define una regla de elección social, donde, para cada perfil $<R_1, R_2,..., R_n>$, $f(R_1, R_2,..., R_n)$ contiene cada alternativa en X que gana o al menos empata con cualquier otra alternativa en votación por mayoría de a pares. Como muestra la paradoja de Condorcet, esto puede producir un conjunto vacío de opciones. Por el contrario, la regla de pluralidad y el conteo de Borda inducen reglas de elección social que siempre producen conjuntos de elección no vacíos. También satisfacen las siguientes condiciones básicas (la última para $|X| \geq 3$):

> *Dominio universal*: El dominio de f es el conjunto de todos los perfiles lógicamente posibles de ordenamientos de preferencias individuales completos y transitivos.
>
> *No-dictadura*: No existe un individuo $i \in N$ tal que, para todo $<R_1, R_2,..., R_n>$ en el dominio de f y todo x en el rango de f, yR_ix donde $y \in f(R_1, R_2,..., R_n)$.[7]
>
> *Restricción de rango*: El rango de f contiene al menos tres alternativas distintas (e idealmente todas las alternativas en X).

Cuando se las complementa con un criterio de desempate apropiado, las reglas de pluralidad y de Borda pueden hacerse más "resueltas":

> *Resolución*: La regla de elección social f siempre produce una única alternativa ganadora (un conjunto de elección singleton). (Luego escribimos $x = f(R_1, R_2,..., R_n)$ para denotar la alternativa ganadora de perfil $< R_1, R_2,..., R_n >$.)

Sorprendentemente, esta lista de condiciones entra en conflicto con el siguiente requisito adicional.

> *Inmunidad Estratégica*: No existe un perfil $< R_1, R_2,..., R_n >$ en el dominio de f en el que f sea manipulable por algún individuo $i \in N$, donde *manipulabilidad* significa lo

[6] Formalmente, $\{x \in X: x \in f(R_1, R_2,..., R_n)$ para algunos $< R_1, R_2,..., R_n >$ en el dominio de $f\}$.
[7] A los propósitos actuales, se puede estipular que la última cláusula (para todo x en el rango de f, yR_ix donde $y \in f(R_1, R_2,..., R_n)$) es violada si $f(R_1, R_2,..., R_n)$ es vacío.

siguiente: si i somete un ordenamiento de preferencias falso R'_i ($\neq R_i$), la ganadora es una alternativa y' que i prefiere estrictamente (de acuerdo a R_i) a la alternativa y que hubiera ganado si i presentara el verdadero ordenamiento de preferencias R_i.[8]

<u>*Teorema*</u> (Gibbard 1973; Satterthwaite 1975): No existe una regla de elección social que satisfaga el dominio universal, la no-dictadura, la restricción de rango, la resolución y la inmunidad estratégica.

Allan Gibbard n. 1942
Web Page

Este resultado plantea cuestiones importantes sobre los trade-offs entre diferentes requisitos en una regla de elección social. Una dictadura, que siempre escoge la alternativa más preferida del dictador, trivialmente tiene inmunidad estratégica. Obviamente, el dictador no tiene ningún incentivo para votar estratégicamente, y nadie más lo hace tampoco, ya que el resultado depende solamente del dictador.

Para ver que el recuento de Borda viola la inmunidad estratégica, recuerden el ejemplo de las Tablas 3 y 4 anteriores. Si el Individuo 1 de la Tabla 3 informa verdaderamente el ordenamiento de preferencias $yP_1xP_1zP_1w$, el ganador de Borda es z, como hemos visto. Si el mismo Individuo 1 reporta falsamente el ordenamiento de preferencias $xP_1yP_1wP_1z$, como en la Tabla 4, el ganador de Borda es x. Pero el Individuo 1 prefiere x a z de acuerdo a su orden de preferencias real (en la Tabla 3), y por lo tanto él o ella tendrán un incentivo a votar estratégicamente.

Moulin (1980) ha demostrado que cuando el dominio de la regla de la elección social es restringido a perfiles de preferencias de punta única, el voto mayoritario por parejas y otros sistemas denominados de "votación por la mediana" pueden satisfacer el resto de las condiciones del teorema de Gibbard-Satterthwaite. Del mismo modo, cuando las decisiones colectivas se limitan a opciones binarias solas, lo que equivale a dejar de lado la restricción de rango, el voto mayoritario satisface el resto de las condiciones. Otras rutas de escape posibles del teorema se abren si se deja de lado la resolución. En el caso limitante en el que siempre se eligen todas las alternativas, las demás condiciones se satisfacen vacuamente.

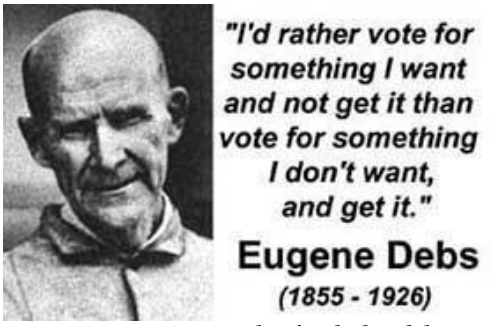
Mark Satterthwaite Web Page

El requisito de inmunidad estratégica también ha sido cuestionado. Una línea de argumento es que, aun cuando existan incentivos estratégicos en el sentido técnico del teorema de Gibbard-Satterthwaite, los individuos no necesariamente actuarán sobre ellos. Requerirían información detallada sobre las preferencias de otros y suficiente poder computacional para averiguar cuáles serían las preferencias óptimas estratégicamente modificadas. Ninguna de estas exigencias se satis-

[8] Formalmente, $y'P_iy$, donde $y' = f(R_1,..., R'_i,..., R_n)$ e $y = f(R_1,..., R_i, ..., R_n)$, suponiendo que $<R_1, ..., R'_i, ..., R_n>$ está en el dominio de f. La definición presupone que los conjuntos de selección social para los perfiles $< R_1,..., R_i, ..., R_n >$ y $< R_1, ..., R'_i, ..., R_n>$ son singleton.

face generalmente. Bartholdi, Tovey y Trick (1989) mostraron que, debido a la complejidad computacional, algunas reglas de elección social son resistentes a la manipulación estratégica: puede ser un problema muy complejo para un votante determinar cómo votar estratégicamente. En tal sentido, Harrison y McDaniel (2008) proporcionan evidencia experimental que sugiere que la "regla de Kemeny", una extensión de la votación mayoritaria por pares diseñada para evitar los ciclos de Condorcet, es "compatible con incentivos de comportamiento": es decir, la manipulación estratégica es computacionalmente difícil.

Dowding y van Hees (2008) han argumentado que no todas las formas de votación estratégica son normativamente problemáticas. Distinguen entre formas de manipulación "*sinceras*" e "*insinceras*" y argumentan que sólo las últimas, pero no las anteriores, son normativamente problemáticas. La manipulación sincera ocurre cuando un votante (i) vota por una alternativa de transacción cuyas posibilidades de ganar son aumentadas y (ii) realmente prefiere esa alternativa de compromiso a la alternativa que de otra manera ganaría. Por ejemplo, en la elección presidencial de Estados Unidos de 2000, los partidarios de Ralph Nader (un candidato de terceros con pocas posibilidades de ganar) que votaron por Al Gore para aumentar sus posibilidades de derrotar a George W. Bush se dedicaron a una manipulación sincera en el sentido de i) y (ii). La regla de pluralidad es susceptible de manipulación sincera, pero no es vulnerable a la manipulación insincera.



Eugene Victor Debs, fundador del Partido Social-Demócrata de USA 1897

4. Agregación del bienestar

Un supuesto implícito hasta ahora ha sido que las preferencias son ordinales y que no son comparables interpersonalmente: los ordenamientos de preferencias no contienen información sobre la fuerza de la preferencia de cada individuo o sobre cómo comparar las preferencias de diferentes personas entre sí. Enunciados como "el Individuo 1 prefiere la alternativa x más que lo que el Individuo 2 prefiere la alternativa y" o "el Individuo l prefiere un cambio de x a y más que lo que el Individuo 2 prefiere un cambio de x^* a y^*" se consideran sin sentido.

En los contextos de votación, este supuesto puede ser plausible, pero en contextos de evaluación del bienestar - cuando un planificador social busca clasificar diferentes alternativas sociales en un orden de bienestar social - puede justificarse el uso de información más rica. Sen (1970b) generalizó el modelo de Arrow para incorporar tal información más rica.

Como antes, consideren un conjunto $N = \{1, 2,..., n\}$ de individuos ($n \geq 2$) y un conjunto $X = \{x, y, z,...\}$ de alternativas sociales. Ahora cada individuo $i \in N$ tiene una *función de bienestar* W_i sobre estas alternativas, que asigna un número real $W_i (x)$ a cada alternativa $x \in X$, interpretada como una medida del bienestar de i en la alternativa x. Cualquier función de bienestar en X induce un ordenamiento en X, pero lo contrario no es cierto: las funciones de bienestar codifican más información. Una combinación de funciones de bienestar a través de individuos, $<W_1, W_2,..., W_n>$, es denominada un *perfil*.

Un *funcional de bienestar social* (SWFL), también denotado F, es una función que asigna a cada perfil ($<W_1, W_2,..., W_n>$) una relación de preferencia social $R = F (W_1, W_2,..., W_n)$ en X, con la interpretación familiar. De nuevo, cuando F resulta claro en el contexto, escribimos R para la relación de preferencia social correspondiente a $< W_1, W_2,..., W_n >$. El output de un SWFL es similar al de una regla de agregación de preferencias (de nuevo, no incorporamos completitud o transitividad de R en la definición[9]), pero su _input_ es más rico.

Lo que ganamos de esto depende de la cantidad de información enriquecida que nos permitamos utilizar para determinar las preferencias de la sociedad: técnicamente, depende de nuestro supuesto sobre la mensurabilidad y la comparabilidad interpersonal del bienestar.

4.1 Mensurabilidad y comparabilidad interpersonal del bienestar

Al asignar números reales a alternativas, los perfiles de bienestar contienen mucha información por encima de los perfiles de ordenamientos sobre X que inducen. En particular, muchas asignaciones diferentes de _números_ a alternativas pueden dar lugar a los mismos _ordenamientos_. Pero podemos no considerar que toda esta información sea significativa. Parte de ella podría ser un artefacto de representación numérica. Por ejemplo, la diferencia entre el perfil $<W_1, W_2,..., W_n>$ y su versión escalada $<10 * W_1, 10 * W_2,..., 10 * W_n>$, donde todo es igual en términos proporcionales, podría ser como la diferencia entre medidas de longitud en centímetros y en pulgadas. Los dos perfiles pueden ser vistos como representaciones alternativas de la misma información exacta, sólo que en escalas diferentes.

Para expresar diferentes supuestos acerca de qué información es verdaderamente codificada por un perfil de funciones de bienestar y qué información no (y debería ser visto, en el mejor de los casos, como un artefacto de representación numérica), es útil introducir la noción de _enunciados significativos_. Algunos ejemplos de enunciados sobre bienestar individual que son candidatos como enunciados significativos son los siguientes (List 2003b, véase también Bossert y Weymark 1996: Sección 5):

> _Una comparación de nivel_: El bienestar de la persona i en la alternativa x es al menos tan grande como el bienestar de la persona j en la alternativa y, formalmente $W_i (x) \geq W_j (y)$. (La comparación es intrapersonal si i = j, e interpersonal si i ≠ j.)
>
> _Una comparación unitaria_: La proporción de [ganancia o pérdida de bienestar de un individuo i si cambiamos de la alternativa y_1 a la alternativa x_1] a [ganancia o pérdida de bienestar del individuo j si cambiamos de la alternativa y_2 a la alternativa x_2] es λ, donde λ es algún número real, formalmente $(x_1 - y_1) / (x_2 - y_2) = \lambda$. (De nuevo, la comparación es intrapersonal si i = j, e interpersonal si i ≠ j.)
>
> _Una comparación de cero_: El bienestar de la persona i en la alternativa x es mayor que / igual a / menor que cero, formalmente $signo\ (W_i (x)) = \lambda$, donde $\lambda \in \{-1, 0, 1\}$ y _signo_ es una función de valor real que asigna números estrictamente negativos a -1, cero a 0 y números estrictamente positivos a +1.

El punto de vista de Arrow, como se observó, es que sólo las comparaciones de nivel intrapersonal son significativas, mientras que todas las otras clases de comparaciones no lo son. Sen (1970b) formalizó varios supuestos sobre la mensurabilidad y la comparabilidad interpersonal del bienestar: (i) definiendo una relación de equivalencia en los perfiles de bienestar que especifica cuándo dos perfiles cuentan como "conteniendo la misma información", y

[9] Sen, como Arrow en su definición de las funciones de bienestar social (en oposición a los funcionales), requirió que R fuera un ordenamiento por definición.

(ii) requiriendo que cualquier perfil en la misma clase de equivalencia genere el mismo orden de preferencia social. De los tres tipos de afirmaciones de comparación presentadas anteriormente, las significativas son aquellas que son invariantes en cada clase de equivalencia. El supuesto ordinalista de Arrow puede expresarse de la siguiente manera:

> *Mensurabilidad ordinaria sin comparabilidad interpersonal (ONC)*: Dos perfiles $< W_1, W_2, \ldots, W_n >$ y $< W^*_1, W^*_2, \ldots, W^*_n >$ contienen la misma información siempre que, para cada $i \in N$, $W^*_i = \varphi_i (W_i)$, donde φ_i es una transformación monótona positiva, posiblemente diferente para diferentes individuos.

Por lo tanto, las funciones individuales de bienestar en cualquier perfil pueden ser transformadas arbitrariamente monotónicamente ("estiradas o contraídas") sin pérdida de información, descartando así cualquier comparación interpersonal o incluso comparaciones intrapersonales unitarias.

Si el bienestar es cardinalmente mensurable pero aún interpersonalmente no comparable, tenemos:

> *Mensurabilidad cardinal sin comparabilidad interpersonal (CNC)*: Dos perfiles $< W_1, W_2, \ldots, W_n >$ y $< W^*_1, W^*_2, \ldots, W^*_n >$ contienen la misma información siempre que, para cada $i \in N$, $W^*_i = a_i W_i + b_i$, donde los a_is y b_is son números reales (con $a_i > 0$), posiblemente diferentes para diferentes individuos.

Aquí, la función de bienestar de cada individuo es única hasta transformaciones afines positivas ("cambio de escala y traslación"), pero todavía no hay una escala común entre individuos. Esto hace que las comparaciones del nivel intrapersonal y de unidades sean significativas, pero descarta comparaciones interpersonales y comparaciones de cero.

La comparabilidad interpersonal se logra bajo la siguiente variante enriquecida de mensurabilidad ordinal:

> *Mensurabilidad ordinal con comparabilidad de nivel interpersonal (OLC)*: Dos perfiles $< W_1, W_2, \ldots, W_n >$ y $< W^*_1, W^*_2, \ldots, W^*_n >$ contienen la misma información siempre que, para cada $i \in N$, $W^*_i = \varphi (W_i)$, donde φ es la misma transformación monótona positiva para todos los individuos.

Aquí, un perfil de funciones de bienestar individuales puede ser arbitrariamente transformado monotónicamente ("estirado o contraído") sin pérdida de información, pero la misma transformación debe ser utilizada para todos los individuos, haciendo así que las comparaciones de nivel interpersonal sean significativas.

La comparabilidad de unidades interpersonales se logra bajo la siguiente variante enriquecida de mensurabilidad cardinal:

> *Mensurabilidad cardinal con comparabilidad de unidades interpersonales (CUC)*: Dos perfiles $< W_1, W_2, \ldots, W_n >$ y $< W^*_1, W^*_2, \ldots, W^*_n >$ contienen la misma información siempre que, para cada $i \in N$, $W^*_i = aW_i + b_i$, donde a es el mismo número real para todos los individuos ($a > 0$) y los b_is son números reales.

En este caso, las funciones de bienestar en cada perfil pueden ser reescaladas y trasladadas sin pérdida de información, pero el mismo múltiplo escalar (aunque no necesariamente la misma constante de cambio) debe aplicarse a todos los individuos, haciendo que las comparaciones de unidades interpersonales sean significativas.

Las comparaciones cero, finalmente, se vuelven significativas bajo la siguiente variante enriquecida de mensurabilidad ordinal (List 2001):

> *Mensurabilidad ordinal con comparabilidad de cero (ONC+0)*: Dos perfiles < W_1, W_2,..., W_n > y < W^*_1, W^*_2,..., W^*_n > contienen la misma información siempre que, para cada $i \in N$, $W^*_i = \varphi_i (W_i)$, donde φ_i es una transformación monótona positiva y que conserva el cero, posiblemente diferente para diferentes individuos. (Aquí conserva el cero significa que $\varphi_i(0) = 0$.)

Esto permite el estiramiento y la compresión arbitrarios de las funciones individuales de bienestar sin pérdida de información, *siempre que el nivel de bienestar de cero permanezca fijo*, garantizando así la comparabilidad cero.

Varias otros supuestos de mensurabilidad y comparabilidad interpersonal han sido discutidos en la literatura. El siguiente asegura que las comparaciones interpersonales de tanto los niveles como las unidades sean significativas:

> *Mensurabilidad cardinal con plena comparabilidad interpersonal (CFC)*: Dos perfiles < W_1, W_2,..., W_n > y < W^*_1, W^*_2,..., W^*_n > contienen la misma información siempre que, para cada $i \in N$, $W^*_i = a W_i + b$, donde a, b son los mismos números reales para todos los individuos ($a > 0$).

Por último, las comparaciones intra e interpersonales de los tres tipos (nivel, unidad y cero) son significativas si aceptamos lo siguiente:

> *Escala-razón de mensurabilidad con plena comparabilidad interpersonal (RFC)*: Dos perfiles < W_1, W_2,..., W_n > y < W^*_1, W^*_2,..., W^*_n > contienen la misma información siempre que, para cada $i \in N$, $W^*_i = a W_i$, donde a es el mismo número real para todos los individuos ($a > O$).

Cuál hipótesis se justifica depende de cómo se interprete el bienestar. Si el bienestar es la utilidad hedónica, que sólo puede experimentarse desde una perspectiva en primera persona, las comparaciones interpersonales son más difíciles de justificar que si el bienestar es la satisfacción objetiva de preferencias o deseos subjetivos (la visión de la satisfacción del deseo) o un bien o estado objetivo (una vista de una lista de objetivos) (p.ej., Hausman 1995, List 2003b). El punto de vista de la satisfacción de deseos puede hacer que las comparaciones interpersonales sean empíricamente significativas (vinculando los niveles máximos y mínimos de bienestar interpersonalmente significativos para cada individuo con el logro de sus alternativas más y menos preferidas), pero a expensas de encontrarse con problemas de gustos costosos o preferencias adaptativas (Hausman 1995). Por el contrario, los enfoques de bienestar basados en recursos, en el funcionamiento o en los bienes primarios pueden permitir comparaciones interpersonales empíricamente significativas y moralmente menos problemáticas.

4.2 La posibilidad de agregación del bienestar

Una vez que introducimos comparaciones interpersonales de niveles o unidades de bienestar, o comparaciones cero, existen SWFLs posibles que satisfacen los análogos de las condiciones de Arrow, así como desiderata más fuertes. En un contexto de agregación del bienestar, la imposibilidad de Arrow se puede atribuir a una falta de comparabilidad interpersonal.

Como se ha indicado, un SWFL respeta un supuesto dado sobre mensurabilidad y comparabilidad interpersonal si, para dos perfiles $< W_1, W_2,..., W_n >$ y $< W^*_1, W^*_2,..., W^*_n >$ que se considera que contienen la misma información, se tiene $F (W_1, W_2,..., W_n) = F (W^*_1, W^*_2,..., W^*_n)$. Las condiciones y el teorema de Arrow pueden ser reformulados como sigue:

> <u>Dominio universal</u>: El dominio de F es el conjunto de todos los perfiles lógicamente posibles de funciones de bienestar individuales.
>
> <u>Ordenamiento</u>: Para cualquier perfil $< W_1, W_2,..., W_n >$ en el dominio de F, la relación de preferencia social R es completa y transitiva.
>
> <u>Principio débil de Pareto</u>: Para cualquier perfil $< W_1, W_2,..., W_n >$ en el dominio de F, si para todo $i \in N$ $W_i (x) > W_i (y)$, entonces xPy.
>
> <u>Independencia de alternativas irrelevantes</u>: Para cualesquiera dos perfiles $< W_1, W_2,..., W_n >$ y $< W^*_1, W^*_2,..., W^*_n >$ en el dominio de F y cualquier $x, y \in X$, si para todos $i \in N$ $W_i (x) = W^*_i (x)$ y $W_i (y) = W^*_i (y)$, entonces xRy si y sólo si xR^*y.
>
> <u>No-dictadura</u>: No existe un individuo $i \in N$ tal que, para todo $< W_1, W_2,..., W_n >$ en el dominio de F y todo $x, y \in X$, $W_i (x) > W_i (y)$ implica xPy.
>
> <u>Teorema</u>: Bajo ONC (o CNC, como Sen 1970b ha demostrado), si $|X| > 2$, no existe SWFL que satisfaga el dominio universal, el ordenamiento, el principio débil de Pareto, la independencia de alternativas irrelevantes y la no-dictadura.

Sin embargo, de modo crucial, cada OLC, CUC y ONC+0 es suficiente para la existencia de SWFLs que satisfagan todas las demás condiciones:

> <u>Teorema</u> (combinando varios resultados de la literatura, como se ilustra a continuación): Bajo cada OLC, CUC y ONC+0, existen SWFLs que satisfacen el dominio universal, el ordenamiento, el principio débil de Pareto, la independencia de alternativas irrelevantes y la no-dictadura (así como condiciones más fuertes).

Algunos ejemplos de tales SWFLs provienen de la filosofía política y la economía del bienestar. Una posible SWFL bajo OLC es una versión del *principio de la diferencia* de Rawls (1971).

> <u>Maximin</u>: Para cualquier perfil $< W_1, W_2,..., W_n >$ y cualesquiera $x, y \in X$, xRy si y sólo si $min_{i \in N} (W_i (x)) \geq min_{i \in N} (W_i (y))$.

Mientras que el maximin clasifica las alternativas sociales en términos del nivel de bienestar del individuo más desfavorecido solo, su extensión lexicográfica (leximin), que fue respaldada por el mismo Rawls, usa el nivel de bienestar del segundo peor individuo para romper el empate cuando hay empate en el nivel de los peores, el nivel de bienestar de la tercera persona en peor situación como desempate cuando hay empate en la segunda etapa, y así sucesivamente. (Observen, sin embargo, que Rawls se centró en los bienes primarios, en lugar del bienestar, como la "moneda" relevante). Esto satisface el principio fuerte (no sólo débil) de Pareto, que requiere que si para todo $i \in N$ $W_i (x) \geq W_i (Y)$, entonces xRy, y si además para algún $i \in N$ $W_i (x) > W_i (y)$, entonces xPy.

Un ejemplo de una posible SWFL bajo CUC es el utilitarismo clásico.

> *Utilitarismo*: Para cualquier perfil $< W_1, W_2,..., W_n >$ y cualquier $x, y \in X$, xRy si y sólo si $W_1(x) + W_2(x) +... + W_n(x) \geq W_1(y) + W_2(Y) + ... + W_n(y)$.

Por último, un ejemplo de una posible SWFL bajo ONC+0 es una variante de una medida de pobreza de uso frecuente, aunque bastante simplista.

> *Regla de recuento*: Para cualquier perfil $< W_1, W_2,..., W_n >$ y cualquier $x, y \in X$, xRy si y sólo si $| \{i \in N: W_i (x)<0\}| < |\{i \in N: W_i (y) <0\}|$ o $[|\{i \in N: W_i(x)<0\}|=|\{i \in N: W_i(y)<0\}|$ y $xR_jy]$, donde $j \in N$ es un individuo que desempata previamente determinado.

Aunque sustancialmente menos atractivas que las reglas maximin o utilitaristas, las reglas de recuento sólo requieren comparabilidad cero del bienestar (List 2001).

Una conclusión importante, por lo tanto, es que tanto el principio de la diferencia de Rawls, el principio utilitario clásico, como incluso el método de medición de recuento de la pobreza, pueden ser vistos como soluciones al problema de agregación de Arrow que son posibles una vez que vamos más allá del marco de las preferencias ordinales, no comparables interpersonalmente.

Bajo el CFC, se puede proporcionar una caracterización simultánea del maximin y del utilitarismo rawlsiano (Deschamps y Gevers 1978). Utiliza dos axiomas adicionales. Uno, *equidad mínima*, requiere (en palabras de Sen 1977: 1548) 'que una persona que va a estar mejor de todas formas no siempre lo hace estrictamente a su manera', y otro, *separabilidad*, requiere que dos perfiles de bienestar que coinciden en algún subconjunto $M \subseteq N$ cuando todo el mundo en $N \setminus M$ está indiferente entre todas las alternativas en X conduzcan al mismo ordenamiento social.

> *Teorema* (Deschamps y Gevers 1978): Bajo CFC, cualquier SWFL que satisfaga el dominio universal, el ordenamiento, el principio fuerte de Pareto, la independencia de alternativas irrelevantes, el anonimato (como en el teorema de May), la mínima equidad y la separabilidad es leximin o de tipo utilitario (lo cual significa que, salvo posiblemente cuando hay empates en la suma de bienestar total, coincide con la SWFL utilitaria definida anteriormente).

Por último, la información adicional disponible en RFC hace posibles las SWFLs "prioritarias".[10] Al igual que las SWFL utilitarias, ordenan las alternativas sociales en base a las sumas de bienestar de los individuos en N, pero en lugar de resumir directamente el bienestar, resumen el bienestar transformado de manera cóncava, otorgando mayor ponderación marginal a los niveles inferiores de bienestar.

> *Prioritarismo*: Para cualquier perfil $< W_1, W_2,..., W_n >$ y cualquier $x, y \in X$, xRy si y sólo si $W_1r(x) + W_2r(x) +... + W_n r(x) \geq W_1r(y) + W_2r(y) + ... + W_nr(y)$, con $0 < r < 1$.

El prioritarismo requiere RFC y no meramente CFC porque, por diseño, el ordenamiento social prioritario para cualquier perfil de bienestar no es invariante a los cambios en los niveles de bienestar (traslación).

[10] Técnicamente, esto requiere una restricción de dominio a perfiles positivos de bienestar.

4.3 Otros temas

El actual marco de agregación del bienestar se ha aplicado a varias áreas adicionales. Se ha generalizado a problemas de elección de población variable, con el fin de formalizar la ética de la población en la tradición de Parfit (1984). Aquí, se deben clasificar las alternativas sociales (por ejemplo, mundos posibles) en las que existen diferentes individuos. Sea *N (x)* el conjunto de individuos existentes bajo la alternativa x. Por ejemplo, el conjunto *N (x)* podría diferir del conjunto *N (y),* cuando *x* e *y* son alternativas distintas (esto generaliza nuestro supuesto anterior de un conjunto fijo *N*). El caso de población variable plantea preguntas tales como si un mundo con un número menor de individuos en mejores condiciones es mejor que, igual de bueno o peor que un mundo con un número mayor de individuos en peores condiciones. (El enfoque aquí es sobre cuestiones axiológicas acerca de la bondad relativa de tales mundos, no preguntas normativas sobre la corrección o incorrección de llevarlos a cabo).

Parfit (1984) y otros argumentaron que el utilitarismo clásico está sujeto a la *conclusión repugnante*: un mundo con un número muy grande de individuos cuyos niveles de bienestar están situados apenas por encima de cero podría tener una suma mayor de bienestar y por lo tanto contar como mejor, que otro mundo con un número menor de individuos muy acomodados.

Blackorby, Donaldson y Bossert (p.ej., 2005) han caracterizado axiomáticamente distintos métodos de agregación de bienestar de población variable que evitan la conclusión repugnante y satisfacen algunos otros desiderata. Una solución es la siguiente:

> *Utilitarismo de nivel crítico*: Para cualquier perfil $< W_1, W_2,..., W_n >$ y cualquier $x, y \in X$, xRy si y sólo si $\Sigma_{i \in N}(x)[W_i(x) - c] \geq \Sigma_{i \in N}(y)[W_i(y) - c]$, donde $c \geq 0$ es un "nivel crítico" de bienestar por encima del cual la calidad de vida cuenta como "decente/buena".

El utilitarismo de nivel crítico evita la conclusión repugnante cuando el parámetro c es suficientemente grande. Requiere una medida del bienestar más sólida que el utilitarismo clásico, ya que genera un ordenamiento social R que no es generalmente invariante bajo el cambio de escala de las unidades de bienestar o cambios en los niveles de bienestar. Incluso el rico marco de RFC obligaría al nivel crítico c a ser cero, colapsando así el utilitarismo de nivel crítico en el utilitarismo clásico y volviéndolo vulnerable a la conclusión repugnante. Como señalan Blackorby, Bossert y Donaldson (1999: 420),

> *Algunos entornos de información que son éticamente adecuados en marcos de población fija tienen consecuencias éticamente poco atractivas en marcos de población variable.*

Por lo tanto, en el caso de población variable, se necesita una desviación más significativa del marco informativo limitado del modelo original de Arrow para evitar resultados de imposibilidad.

El enfoque SWFL ha sido generalizado al caso en que cada individuo tiene múltiples funciones de bienestar (por ejemplo, una k-tupla de ellas), capturando (i) opiniones múltiples sobre el bienestar de cada individuo (por ejemplo, Roberts 1995; Ooghe y Lauwers 2005) o (ii) múltiples dimensiones del bienestar (por ejemplo, List 2004a). En este caso, enfrentamos no sólo cuestiones de mensurabilidad y comparabilidad interpersonal, sino también cuestiones de inter-opinión o comparabilidad inter-dimensional. Para obtener resultados de posibilidad convincentes, es necesaria la comparabilidad entre los individuos y las dimensiones/opiniones. Una literatura relacionada aborda la medición de la desigualdad multidimensional (para una revisión introductoria, ver Weymark 2006).

Por último, en filosofía de la biología, se han utilizado los marcos SWFL unidimensionales y multidimensionales (Okasha 2009 y Bossert, Qi y Weymark, 2013) para analizar la noción de aptitud grupal, definida como una función de los indicadores de aptitud individual.

5. Agregación de juicios

Una rama más reciente de la teoría de la elección social es la teoría de la agregación de juicios. Se puede motivar observando que los votos, ordenamientos o funciones de bienestar sobre múltiples alternativas no son los únicos objetos que podemos agregar de un individuo a un nivel colectivo. Muchos órganos de toma de decisiones, como legislaturas, tribunales colegiados, grupos de expertos y otros comités, enfrentan agregados más complejos. En particular, pueden tener que agrupar conjuntos individuales de juicios sobre múltiples proposiciones conectadas lógicamente en conjuntos colectivos de juicios.

Un tribunal puede tener que juzgar si un acusado es responsable de incumplimiento de contrato sobre la base de si había un contrato válido y si hubo una violación. Un grupo de expertos puede tener que juzgar si las concentraciones atmosféricas de gases de efecto invernadero excederán un umbral determinado en 2050, si hay una cadena causal de mayores concentraciones de gases de efecto invernadero a los aumentos de temperatura y si la temperatura aumentará. Los legisladores pueden tener que juzgar si un fin en particular es socialmente deseable, si una política propuesta es el mejor medio para lograr ese fin, y si seguir esa política.

Estos problemas no pueden formalizarse en modelos de agregación de preferencias estándar, ya que los *agreganda* no son ordenamientos sino conjuntos de juicios sobre proposiciones múltiples. La teoría de la agregación de juicios representa estos agreganda en la lógica proposicional (u otra lógica adecuada). El campo fue inspirado por la "paradoja doctrinal" en la jurisprudencia, con la que comenzamos.

5.1 La "paradoja doctrinal" y el "dilema discursivo"

Kornhauser y Sager (1986) describieron el siguiente problema. (Un problema estructuralmente similar fue descubierto por Vacca 1921 y, como Elster 2013 señala, por Poisson 1837.) Un tribunal de tres jueces tiene que emitir juicios sobre las siguientes proposiciones:

- *p*: El demandado estaba obligado contractualmente a no realizar la acción *X*.
- *q*: El acusado hizo la acción *X*.
- *r*: El demandado es responsable por incumplimiento de contrato.

Según la doctrina legal, las *premisas* p y q son conjuntamente necesarias y suficientes para la *conclusión* r. Supongan que los jueces individuales tienen las opiniones mostradas en la Tabla 5.

Tabla 5: un ejemplo de la "paradoja doctrinal"

	p (obligación)	*q (acción)*	*r (responsabilidad)*
Juez 1	*Verdad*	*Verdad*	*Verdad*
Juez 2	*Falso*	*Verdad*	*Falso*
Juez 3	*Verdad*	*Falso*	*Falso*
Mayoría	*Verdad*	*Verdad*	*Falso*

Aunque cada juez individual respeta la doctrina legal relevante, hay una mayoría para p, una mayoría para q, y sin embargo una mayoría contra r - en violación de la doctrina legal. El tribunal se enfrenta a un dilema: puede ir con los juicios de la mayoría en las premisas (p y q) y llegar a un veredicto "responsable" por inferencia lógica (el *enfoque problema por problema* o *basado en las premisas*); o ir con el juicio de la mayoría sobre la conclusión (r) y llegar a un veredicto de 'no responsable', ignorando los juicios de la mayoría sobre las premisas (el *enfoque caso por caso* o *basado en la conclusión*). La "paradoja doctrinal" de Kornhauser y Sager consiste en el hecho de que estos dos enfoques pueden conducir a resultados opuestos.

Podemos aprender otra lección de este ejemplo. Con relación a la doctrina legal, los juicios mayoritarios son lógicamente inconsistentes. Formalmente expresado, el conjunto de proposiciones aceptadas por la mayoría, {p, q, no r}, es inconsistente con respecto a la restricción r si y sólo si (p y q). Esta observación fue el punto de partida de la literatura más reciente basada en la lógica formal sobre la agregación de juicios (comenzando con un modelo y el resultado de imposibilidad en List y Pettit, 2002).

La posibilidad de juicios de la mayoría inconsistentes no está vinculada a la presencia de una doctrina legal u otra restricción lateral explícita (como lo señaló Pettit 2001, quien llamó a este fenómeno el "dilema discursivo"). Supongan, por ejemplo, que un panel de expertos tenga que emitir juicios sobre tres proposiciones (y sus negaciones):

- *p*: El CO_2 atmosférico superará 600ppm en 2050.
- *si p entonces q*: Si el CO2 atmosférico sobrepasa este nivel en 2050, habrá un aumento de temperatura de más de 3,5° en 2010.
- *q*: En 2010 habrá un aumento de temperatura de más de 3,5°.

Si los juicios individuales son como se muestra en la Tabla 6, los juicios de la mayoría son inconsistentes: a pesar de juicios que son individualmente consistentes, el conjunto de proposiciones aceptadas por la mayoría, {p, si p entonces q, no q}, es lógicamente inconsistente.

Tabla 6: Inconsistencia mayoritaria

	p	*si p entonces q*	*q*
Experto 1	*Verdad*	*Verdad*	*Verdad*
Experto 2	*Falso*	*Verdad*	*Falso*
Experto 3	*Verdad*	*Falso*	*Falso*
Mayoría	*Verdad*	*Verdad*	*Falso*

5.2 Condiciones para juicios de mayoría inconsistentes

Los patrones de juicios de las Tablas 5 y 6 son estructuralmente equivalentes al patrón de preferencias que conduce a la paradoja de Condorcet, cuando reinterpretamos esas preferencias como juicios sobre proposiciones de la forma x es preferible a y, y es preferible a z, y así sucesivamente, como se muestra en la Tabla 7 (List y Pettit, 2004; una interpretación anterior de las preferencias a lo largo de estas líneas se puede encontrar en Guilbaud [1952] 1966). Aquí, el conjunto de propuestas aceptadas por la mayoría es inconsistente con respecto a la restricción de transitividad.

Tabla 7: La paradoja de Condorcet, proposicionalmente reinterpretada

	"x es preferible a y"	*"y es preferible a z"*	*"x es preferible a z"*
Individuo 1 (prefiere x a y a z)	*Verdad*	*Verdad*	*Verdad*
Individuo 2 (prefiere y a z a x)	*Falso*	*Verdad*	*Falso*
Individuo 3 (prefiere z a x a y)	*Verdad*	*Falso*	*Falso*
Mayoría (prefiere x a y a z a x, un "ciclo")	*Verdad*	*Verdad*	*Falso*

Un resultado combinatorio general abarca todos estos fenómenos. Llamemos a un conjunto de proposiciones *mínimamente inconsistente* si es un conjunto lógicamente inconsistente, pero todos sus subconjuntos propios son consistentes.

> *Proposición* (Dietrich y List 2007a: Nehring y Puppe 2007): La votación de enunciados por mayoría puede generar juicios colectivos inconsistentes si y sólo si el conjunto de enunciados (y sus negaciones) sobre los que se deben emitir juicios tiene un subconjunto mínimamente inconsistente de tres o más enunciados.

En los ejemplos de las Tablas 6, 5 y 7, los conjuntos mínimamente inconsistentes relevantes de tamaño (al menos) tres son: {p, si p entonces q, no q}, que es mínimamente inconsistente *simpliciter*; {p, q, no r}, que es mínimamente inconsistente con respecto a la restricción lateral r si y sólo si (p y q); y {'*x es preferible a y*', '*y es preferible a z*', '*z es preferible a x*'}, el cual es mínimamente inconsistente con relación a una restricción de transitividad sobre las preferencias.

5.3 Un modelo general y un resultado simple de imposibilidad

El modelo básico de agregación de juicios puede definirse de la siguiente manera (List y Pettit 2002). Sea $N = \{1, 2,..., n\}$ un conjunto de individuos ($n \geq 2$). Las proposiciones sobre las que se deben hacer los juicios están representadas por frases de la lógica proposicional (o alguna otra lógica expresivamente más rica, como una lógica predicada, modal o condicional, véase Dietrich 2007). Definimos la *agenda*, X, como un conjunto finito de proposiciones, cerrado bajo la sola negación.[11] Por ejemplo, X podría ser $\{p, \neg p, p \rightarrow q, \neg (p \rightarrow q), q, \neg q\}$, como en el caso del panel de expertos.

Cada individuo $i \in N$ tiene un conjunto de juicios J_i, definido como un subconjunto $J_i \subseteq X$ e interpretado como el conjunto de proposiciones que el individuo i acepta. Un conjunto de juicios es consistente si es un conjunto lógicamente *consistente* de proposiciones[12] y *completo* (relativo a X) si contiene un miembro de cada par proposición-negación $p, \neg p \in X$.

Una combinación de conjuntos de juicios entre individuos, $<J_1, J_2,..., J_n>$, se denomina *perfil*. Una *regla de agregación de juicios*, F, es una función que asigna a cada perfil $<J_1, J_2,..., J_n>$ (en algún dominio de perfiles admisibles) un conjunto de juicios colectivos $J = F(J_1, J_2,..., J_n)$, interpretado como el conjunto de proposiciones aceptadas por el grupo como un todo. Como antes, cuando F resulta claro por el contexto, escribimos J para el conjunto de juicios colectivos correspondiente a $<J_1, J_2,..., J_n>$. De nuevo, para mayor generalidad, no fijamos ningún requisito de racionalidad sobre J (como la consistencia o la completitud) en la definición de una regla de agregación de juicios.

El ejemplo más simple de una regla de agregación de juicios es la *votación por mayoría de las proposiciones*. Aquí, para cualquier perfil $<J_1, J_2,..., J_n>$, $J=\{p \in X: | \{i \in N: p \in J_i\} | > n/2\}$. Como hemos visto, esto puede producir juicios colectivos inconsistentes.

Consideren las siguientes condiciones en una regla de agregación:

> *Dominio universal*: El dominio de F es el conjunto de todos los perfiles lógicamente posibles de conjuntos de juicios individuales coherentes y completos.
>
> *Racionalidad colectiva*: Para cualquier perfil $<J_1, J_2,..., J_n>$ en el dominio de F, el conjunto de juicios colectivos J es consistente y completo.
>
> *Anonimato*: Para cualesquiera dos perfiles $<J_1, J_2,..., J_n>$ y $<J^*_1, J^*_2,..., J^*_n>$ que son permutaciones entre sí, $F(J_1, J_2,..., J_n) = F(J^*_1, J^*_2,..., J^*_n)$.
>
> *Sistematicidad*: Para cualesquiera dos perfiles $<J_1, J_2,..., J_n>$ y $<J^*_1, J^*_2,..., J^*_n>$ en el dominio de F y cualquier $p, q \in X$, si para todo $i \in N$ $p \in J_i$ si y sólo si $q \in J^*_i$, entonces $p \in J$ si y sólo si $q \in J^*$.

---

[11] Formalmente, $X = \{p, \neg p: p \in X+\}$, donde $X+$ es un conjunto de proposiciones no negadas. Para evitar los tecnicismos, se supone que X no contiene proposiciones contradictorias o tautológicas.

[12] En principio, la coherencia puede definirse con relación a alguna restricción lateral, como la doctrina jurídica en el ejemplo de la «paradoja doctrinal».

Las tres primeras condiciones son análogas al dominio universal, al ordenamiento y al anonimato en la agregación de preferencias. La última es la contrapartida de la independencia de las alternativas irrelevantes, aunque más fuerte: requiere que el juicio colectivo sobre cualquier proposición $p \in X$ (de la cual una proposición de ordenamiento binario como *'x es preferible a y'* es un caso especial) dependa sólo de proposiciones individuales sobre p (la parte de *independencia*), y (ii) el patrón de dependencia entre juicios individuales y colectivos sea el mismo en todas las proposiciones en X (la parte de *neutralidad*). Formalmente, *independencia* es el caso especial con la cuantificación restringida a $p = q$. La votación por mayoría de proposiciones satisface todas estas condiciones, excepto la parte consistente de la racionalidad colectiva.

> *Teorema* (List y Pettit 2002): Si $\{p, q, p \wedge q\} \subseteq X$ (donde p y q son proposiciones mutuamente independientes y '$\wedge$' también puede ser reemplazado por '$\vee$' o '$\rightarrow$'), no existe regla de agregación del juicio que satisfaga el dominio universal, la racionalidad colectiva, el anonimato y la sistematicidad.

Al igual que otros teoremas de imposibilidad, este resultado se interpreta mejor como describiendo los trade-offs entre diferentes condiciones en una regla de agregación. El resultado ha sido generalizado y fortalecido de varias maneras, comenzando con la prueba de Pauly y van Hees (2006) de que la imposibilidad persiste si el anonimato se debilita a la no dictadura (para otras generalizaciones, ver Dietrich 2006 y Mongin 2008).

5.4 Teoremas de imposibilidad y posibilidad más generales

Como hemos visto, en la agregación de preferencias, la "frontera" entre los resultados de posibilidad e imposibilidad es fácil de trazar: cuando sólo hay dos alternativas de decisión, se pueden satisfacer todos los desiderata de una regla de agregación de preferencias analizada anteriormente (y la regla mayoritaria funciona); cuando hay tres o más alternativas, hay resultados de imposibilidad. En la agregación de juicios, por el contrario, el cuadro es más complicado. Lo que importa no es el número de proposiciones en X sino la naturaleza de las interconexiones lógicas entre ellas.

Los resultados de imposibilidad en la agregación de juicios tienen la siguiente forma genérica: para una clase dada de agendas, las reglas de agregación que satisfacen un conjunto particular de condiciones (usualmente una condición de dominio, una condición de racionalidad y algunas condiciones de respuesta) son inexistentes o degeneradas (p.ej., dictatoriales). Diferentes tipos de agendas dan lugar a diferentes instancias de este esquema, con condiciones más fuertes o más débiles impuestas a la regla de agregación dependiendo de las propiedades de esas agendas (para una revisión más detallada, véase List 2012). La importancia de las propiedades combinatorias de la agenda fue descubierta por Nehring y Puppe (2002) en un marco matemáticamente relacionado pero interpretativamente distinto (elección social a prueba de estrategias sobre los llamados espacios de propiedades). Se destacan tres tipos de agenda:

> *Una agenda no simple*: X tiene un subconjunto mínimamente inconsistente de tres o más proposiciones.

> *Una agenda de pares negativos*: X tiene un subconjunto Y mínimamente inconsistente que puede ser hecho consistente negando un par de proposiciones en él. (En forma

equivalente, X no es isomorfo a un conjunto de proposiciones cuyos únicos conectivos son ¬ y ↔, véase Dokow y Holzman 2010a).

Una agenda conectada (o *totalmente bloqueada*, en Nehring y Puppe 2002): Para cualquier $p, q \in X$, hay una secuencia $p_1, p_2, \ldots, p_k \in X$ con $p_1 = p$ y $p_k = q$ tales que p_1 implica condicionalmente p_2, p_2 implica condicionalmente $p_3, \ldots$, y p_{k-1} implica condicionalmente p_k. (Aquí, *condicionalidad de p_i implica* p_j si $p_i \cup Y$ implica p_j para algunos $Y \subseteq X$ consistente con cada p_i y $\neg p_j$.)

Algunas agendas tienen dos o más de estas propiedades. Las agendas en nuestros ejemplos de "paradoja doctrinal" y "dilema discursivo" son no-simples y par-negables. La agenda de preferencias, X = {'x es preferible a y', 'y es preferible a x', 'x es preferible a z', 'z es preferible a x',...}, es no-simple, par-negable, y conectada por caminos (suponiendo que la preferibilidad sea transitiva y completa). Se cumple el siguiente resultado:

Teorema (Dietrich y List 2007b; Dokow y Holzman 2010a; elaborando sobre Nehring y Puppe 2002): Si X es no-simple, par-negable y conectada por caminos, no existe regla de agregación de juicios que satisfaga el dominio universal, la racionalidad colectiva, la independencia, la preservación de la unanimidad (que requiere que, para cualquier perfil unánime $\langle J, J, \ldots, J \rangle$, $F(J, J, \ldots, J) = J$) y la no-dictadura.[13]

Aplicado a la agenda de preferencias, este resultado da como resultado el teorema de Arrow (para órdenes de preferencia estrictos) (precedentes de este resultado se pueden hallar en List y Pettit 2004 y Nehring 2003).[14] Por lo tanto, la agregación de preferencias arrovianas puede reinterpretarse como un caso especial de agregación de juicios.

La literatura contiene muchas variantes de este teorema. Una variante deja caer la propiedad de agenda de conexión de caminos y fortalece la independencia y la sistematicidad. Una segunda variante deja caer la propiedad de agenda de pares negativos e impone una condición de monotonicidad en la regla de agregación (requiriendo que el soporte adicional nunca dañe una proposición aceptada) (Nehring y Puppe 2010, el último resultado fue probado por primera vez en el marco mencionado de Nehring y Puppe 2002). Una variante final deja caer tanto la conexión de caminos como la negatividad de pares, imponiendo tanto la sistematicidad como la monotonicidad (ibid.).

En cada caso, las propiedades de agenda no sólo son suficientes sino también (si $n \geq 3$) necesarias para el resultado (Nehring y Puppe 2002, 2010; Dokow y Holzman 2010a). Obsérvese también que la conexión de caminos implica no simplicidad. Por lo tanto, la no simplicidad no necesita ser enumerada entre las condiciones del teorema, aunque es necesaria en las variantes que dejan caer la conexión del camino.

5.5 Reglas no dictatoriales de agregación de juicios

5.5.1 Relajación del dominio universal

Como en la agregación de preferencias, una manera de evitar los resultados de imposibilidad es relajar el dominio universal. Si el dominio de los perfiles admisibles de juicios individuales

[13] Véase también la observación sobre la relación entre la conexión por caminos y la no simplicidad al final de esta subsección.

[14] Una contribución anterior relacionada matemáticamente, aunque interpretativamente distinta, es el trabajo de Wilson sobre agregación abstracta (1975).

se limita a aquellos que satisfacen condiciones específicas de "cohesión", la votación por mayoría de enunciados sugiere juicios colectivos consistentes.

La condición de cohesión más simple es la *alineación unidimensional* (List 2003c). Un perfil <$J_1, J_2,..., J_n$> está *unidimensionalmente alineado* si los individuos en N pueden ser ordenados de izquierda a derecha (por ejemplo, en alguna dimensión cognitiva o ideológica) de modo que, para cada proposición $p \in X$, los individuos que aceptan p (es decir, aquellos con $p \in J_i$) están todos a la izquierda o a la derecha de los que rechazan p (es decir, aquellos con $p \notin J_i$), como se ilustra en la Tabla 8. Para tales perfiles, los juicios de la mayoría son consistentes: prevalecerá el conjunto de juicios del individuo mediano relativo al ordenamiento de izquierda-derecha (con n impar). Este conjunto de juicios heredará su consistencia del individuo mediano, suponiendo que los juicios individuales sean consistentes. Por implicancia, en los dominios unidimensionalmente alineados, la votación por mayoría de enunciados satisfará el resto de las condiciones sobre las reglas de agregación de juicios analizadas previamente.

Tabla 8: Alineación unidimensional

	Individuo 1	*Individuo 2*	*Individuo 3*	*Individuo 4*	*Individuo 5*
p	*Verdad*	*Verdad*	*Falso*	*Falso*	*Falso*
q	*Verdad*	*Verdad*	*Verdad*	*Verdad*	*Falso*
r	*Falso*	*Falso*	*Falso*	*Verdad*	*Verdad*
$p \wedge q \wedge r$	*Falso*	*Falso*	*Falso*	*Falso*	*Falso*

En analogía con el caso de punta única en la agregación de preferencias, varias condiciones menos restrictivas ya son suficientes para juicios de mayoría consistentes. Una condición de este tipo (introducida en Dietrich y List 2010a, donde se proporciona una encuesta), generaliza la restricción de valor triple de Sen. Un perfil <$J_1, J_2,..., J_n$> es de *valor restringido* si cada subconjunto mínimamente inconsistente $Y \subseteq X$ tiene un par de elementos p, q tal que ningún individuo $i \in N$ tenga $\{p, q\} \subseteq J_i$. La restricción de valor evita que cualquier subconjunto mínimamente incoherente de X se acepte por mayoría y, por lo tanto, garantiza juicios de mayoría consistentes. Aplicada a la agenda de preferencias, la restricción de valor se reduce a la condición igualmente nombrada de Sen.

5.5.2 Relajación de la racionalidad colectiva

Si bien el requisito de que los juicios colectivos sean consistentes es ampliamente aceptado, la exigencia de que los juicios colectivos sean completos (en X) es más polémica. En apoyo de la completitud, se podría decir que una proposición dada no se incluiría en X a menos que se suponga que se juzga colectivamente. Contra la completitud, se podría decir que hay circunstancias en las que el nivel de desacuerdo sobre una proposición particular (o conjunto de proposiciones) es tan grande que formar una opinión colectiva sobre él es indeseable o contraproducente. Varios trabajos ofrecen resultados de posibilidad o de imposibilidad sobre relajaciones completas (p.ej., List y Pettit 2002, Gärdenfors 2006, Dietrich y List 2007a, 2008, Dokow y Holzman 2010b).

Reglas de agregación de juicios que violan la completitud colectiva mientras que satisfacen (la totalidad o la mayor parte) de las otras condiciones introducidas anteriormente incluyen: la *regla de unanimidad*, donde, para cualquier perfil $<J_1, J_2, ..., J_n>$, $J = \{p \in X: p \in J_i \text{ para todo } i \in N\}$; *reglas de súper-mayorías*, donde, para cualquier perfil $< J_1, J_2, ..., J_n >$, $J = \{p \in X: | \{i \in N: p \in J_i\} | > qn\}$ para una cuota de aceptación adecuada $q \in (0.5,1)$; y las *reglas basadas en la conclusión*, donde un subconjunto $Y \subseteq X$ de proposiciones lógicamente independientes (y sus negaciones) es designado como un conjunto de conclusiones y $J=\{p\in Y: | \{i \in N: p \in Ji\} | > N / 2\}$. En el ejemplo de la Tabla 5 de varios miembros del tribunal, el conjunto de conclusiones es simplemente $Y = \{r, \neg r\}$.

Dado conjuntos de juicios individuales consistentes, la regla de unanimidad garantiza conjuntos de juicios colectivos consistentes, porque la intersección de varios conjuntos consistentes de proposiciones es siempre consistente. Las reglas de la súper-mayoría también garantizan conjuntos de juicios colectivos consistentes, siempre que la cuota q se elija como mínimo $(k-1) / k$, donde k es el tamaño del subconjunto mínimamente inconsistente de X. La razón es combinatoria: cualquier k súper-mayorías distintas del tamaño pertinente siempre tendrán al menos un individuo en común. Por lo tanto, para que cualquier conjunto de proposiciones mínimamente incoherentes (que sea a lo sumo de tamaño k) sea aceptado mayoritariamente, al menos un individuo tendría que aceptar todas las proposiciones del conjunto, contradiciendo la consistencia de este individuo (Dietrich y List 2007a; y Pettit, 2002). Las reglas basadas en la conclusión, por último, producen conjuntos de juicios colectivos consistentes con la construcción, pero siempre dejan sin decidir las no conclusiones.

Gärdenfors (2006) y, más en general, Dietrich y List (2008) y Dokow y Holzman (2010b) han demostrado que si se requiere - al relajarse la completitud - que los conjuntos de juicios colectivos sean deductivamente cerrados (es decir, para cualquier $p\in X$ implicado por J, debe darse que $p\in J$), nos enfrentamos de nuevo a un resultado de imposibilidad. Para las mismas agendas que llevan al resultado de imposibilidad revisado en la Sección 5.4, no existe regla de agregación de juicios que satisfaga el dominio universal, la consistencia colectiva y la clausura deductiva, la independencia, la preservación de la unanimidad y la no-oligarquía. Una regla de agregación se denomina *oligárquica* si existe un subconjunto previamente fijado $M \subseteq N$ (los "oligarcas") tal que, para cualquier perfil $<J_1, J_2,..., J_n>$, $J = \{p \in X: p\in J_i \text{ para todos } i\in M\}$. La regla de unanimidad y las dictaduras son casos especiales con M=N y M = {i} para algún i ∈ N, respectivamente.

La desventaja de las reglas de agregación oligárquicas es que o bien caen en la dictadura o bien conducen a un punto muerto, con los desacuerdos más leves entre los oligarcas resultando en indecisión (ya que cada oligarca tiene poder de veto sobre cada proposición).

5.5.3 Relajando la sistematicidad / independencia

Una variedad de reglas de agregación de juicios son posibles cuando relajamos la sistematicidad / independencia. Recordemos que la sistematicidad combina una exigencia de independencia y de neutralidad. Relajar sólo la neutralidad no nos lleva muy lejos, ya que para muchas agendas hay resultados de imposibilidad sólo con independencia, como se ilustró en la Sección 5.4.

Una clase muy discutida de reglas de agregación violatorias de la independencia está dada por las reglas basadas en la premisa. Aquí, un subconjunto $Y \subseteq X$ de proposiciones lógicamente independientes (y sus negaciones) es designado como un conjunto de premisas, como en el ejemplo del tribunal. Para todo perfil $<J_1, J_2,..., J_n>$, $J= \{p \in X: JY$ implica $p\}$ donde JY son las proposiciones entre las premisas aceptadas por la mayoría, formalmente $\{p\in Y:|\{i\in N: p\in J_i\}|>N/2\}$. Informalmente, los votos de la mayoría son tomados en las premisas, y los juicios colectivos sobre todas las demás proposiciones son determinados por implicancia lógica. Si las premisas constituyen una base lógica para toda la agenda, una regla basada en premisas garantiza la coherencia y (si no hay empates) conjuntos de juicios colectivos completos. (La presente definición sigue a List y Pettit 2002; para generalizaciones, ver Dietrich y Mongin 2010. Las propiedades procesales y epistémicas de las reglas basadas en premisas se discuten en Pettit 2001, Chapman 2002, Bovens y Rabinowicz 2006, Dietrich 2006, List 2006).

Una generalización viene dada por las *reglas de prioridad secuenciales* (Lista 2004b, Dietrich y List 2007a). Aquí, para cada perfil $<J_1, J_2,..., J_n>$, las proposiciones en X son juzgadas colectivamente en un orden de prioridad fijo, por ejemplo, temporal o epistémico. El juicio colectivo sobre cada proposición $p\in X$ se realiza de la siguiente manera. Si el juicio de la mayoría sobre p es consistente con juicios colectivos sobre proposiciones anteriores, prevalece este juicio mayoritario; de lo contrario el juicio colectivo sobre p es determinado por las implicancias de los juicios anteriores. Por construcción, esto garantiza juicios consistentes y (si no hay empates), juicios colectivos completos. Sin embargo, el procedimiento es *dependiente del camino*: el orden en que se consideran las proposiciones puede afectar el resultado, específicamente cuando los juicios mayoritarios subyacentes son inconsistentes. Por ejemplo, cuando esta regla de agregación se aplica a los perfiles de las Tablas 5, 6 y 7 (pero no la 8), los juicios colectivos dependen del orden en que se consideran las proposiciones. Por lo tanto, las reglas de prioridad secuenciales son vulnerables a *manipulación de la agenda*. Fenómenos similares ocurren en el voto por mayoría secuencial de a pares en la agregación de preferencias (por ejemplo, Riker 1982).

Otra clase prominente de reglas de agregación que violan la independencia está dada por las reglas basadas en la distancia (Pigozzi 2006, basándose en Konieczny y Pino Pérez 2002, véase también Miller y Osherson 2009). Una regla basada en la distancia se define en términos de alguna métrica de distancia entre conjuntos de juicio, por ejemplo la *distancia de Hamming*, donde, para cualesquiera dos conjuntos de juicios $J, J' \subseteq X$, $d(J, J') = |\{p \in X: No [p \in J \Leftrightarrow p \in J']\}$. Cada perfil $<J_1, J_2,..., J_n>$ es correlacionado con un conjunto de juicios consistente y completo J que minimiza la suma total de las distancias de cada uno de los J_i. Las reglas basadas en la distancia pueden ser interpretadas como capturando la idea de identificar juicios de compromiso. A diferencia de las reglas basadas en premisas o las reglas de prioridad secuenciales, no requieren una distinción entre premisas y conclusiones o cualquier otro orden de prioridad entre las proposiciones.

Como en la agregación de preferencias, el costo de relajar la independencia es la pérdida del carácter "a prueba de estrategias". La conjunción de independencia y monotonicidad es necesaria y suficiente para la no manipulación de una regla de agregación de juicios por voto estratégico (Dietrich y List 2007c, para resultados relacionados, véase Nehring y Puppe 2002). Por lo tanto, generalmente no podemos lograr el carácter a prueba de estrategias sin relajar o el dominio universal, o la racionalidad colectiva, o la preservación de la unanimidad, o la no-dictadura. En la práctica, debemos buscar formas de minimizar las oportunidades de manipulación estratégica.

6. Otros temas

Como debe resultar evidente, la teoría de la elección social es un campo extenso. Áreas no cubiertas en este documento, o mencionadas sólo de pasada, incluyen: las teorías de la división equitativa (cómo dividir uno o varios bienes divisibles o indivisibles, como pasteles o casas, entre varios demandantes, p. ej., Brams y Taylor 1996 y Moulin 2004); teoría de la elección social conductista (analiza la evidencia empírica de comportamiento del voto bajo varias reglas de agregación, p. ej., Regenwetter et al., 2006, List, Luskin, Fishkin y McLean 2013); teoría de la elección social empírica (analiza encuestas y experimentos sobre las intuiciones de la gente acerca de la justicia distributiva, p. ej., Gaertner y Schokkaert 2012); teoría de la elección social computacional (análisis de las propiedades computacionales de las reglas de agregación, incluyendo su complejidad computacional, p. ej., Bartholdi, Tovey y Trick 1989; Brandt, Conitzer y Endriss 2013); teorías de agregación de probabilidades (estudia la agregación de funciones de probabilidad o credibilidad, p. ej., Lehrer y Wagner 1981, McConway 1981, Genest y Zidek 1986, Mongin 1995, Dietrich y List 2007d); teorías de agregación de actitud general (generaliza agregación de juicios de dos valores, agregación de probabilidad / credibilidad y agregación de preferencias, p. ej., Dietrich y List 2010b; Dokow y Holzman 2010c); estudio de la toma de decisiones colectiva en animales no humanos (estudia las decisiones de grupo en una variedad de especies animales de insectos sociales a primates, según lo encuestado en Conradt y List 2009 y el número especial que introduce); y aplicaciones a la epistemología social (análisis de estados doxásticos de grupo y su relación con estados doxásticos individuales, p.ej., Goldman 2004, 2010).

Bibliografía

- Arrow, K., 1951/1963, *Social Choice and Individual Values*. New York: Wiley.
- Austen-Smith, D. and J. S. Banks, 1996, "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review,* 90: 34–45.
- Bartholdi, J. J., C. A. Tovey, and M. A. Trick, 1989, "The computational difficulty of manipulating an election." *Social Choice and Welfare,* 6: 227–241.
- Ben-Yashar, R. and S. Nitzan, 1997, "The optimal decision rule for fixed-size committees in dichotomous choice situations: the general result." *International Economic Review,* 38: 175–186.
- Berend, D. and J. Paroush, 1998, "When is Condorcet's Jury Theorem valid?" *Social Choice and Welfare,* 15: 481–488.
- Berend, D. and L. Sapir., 2007, "Monotonicity in Condorcet's Jury Theorem with dependent voters." *Social Choice and Welfare,*28: 507–528.
- Black, D., 1948, "On the Rationale of Group Decision-Making."*Journal of Political Economy,* 56: 23–34.
- Blackorby, C., W. Bossert, and D. Donaldson, 1999, "Information Invariance in Variable-Population Social-Choice Problems."*International Economic Review,* 40: 403–422.
- –––, 2005, *Population Issues in Social Choice Theory, Welfare Economics, and Ethics,*. Cambridge: Cambridge University Press.
- Blau, J. H., 1975, "Liberal Values and Independence." *Review of Economic Studies,* 42: 395–401.
- Boland, P. J., 1989, "Majority systems and the Condorcet jury theorem." *Statistician,* 38: 181–189.
- Bossert, W. and J. A. Weymark, 1996, "Utility in social choice."*Handbook of Utility Theory, Volume 2*. S. Barberà, P. J. Hammond and C. Seidel, (eds.). Boston: Kluwer.
- Bossert, W., C. X. Qi, and J. A. Weymark, 2013, "Extensive social choice and the measurement of group fitness in biological hierarchies." *Biology and Philosophy,* 28: 75–98.
- Bovens, L. and W. Rabinowicz, 2006, "Democratic Answers to Complex Questions: An Epistemic Perspective." *Synthese,*150: 131–153.
- Brams, S. J. and A. D. Taylor, 1996, *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge: Cambridge University Press.
- Brandt, F., V. Conitzer, and U. Endriss, 2013, "Computational Social Choice." *Multiagent Systems*. G. Weiss (ed.). Cambridge, MA: MIT Press, pp. 213–283.
- Brighouse, H. and M. Fleurbaey, 2010, "Democracy and Proportionality." *Journal of Political Philosophy,* 18: 137–155.
- Chapman, B., 2002, "Rational Aggregation." *Politics, Philosophy and Economics,* 1: 337–354.
- Condorcet, Nicolas de, 1785, *Essay sur l'Application de l'Analyse à la Probabilité des Décisions Rendue à la Pluralité des Voix*. Paris.

- Conradt, L. and C. List, 2009, “Group decisions in humans and animals: a survey.” *Philosophical Transactions of the Royal Society,* B 364: 719–742.
- Craven, J., 1982, “Liberalism and Individual Preferences.” *Theory and Decision,* 14: 351–360.
- Deschamps, R. and L. Gevers, 1978, “Leximin and utilitarian rules: A joint characterization.” *Journal of Economic Theory,* 17: 143–163.
- Dietrich, F., 2006, “Judgment Aggregation: (Im)Possibility Theorems.” *Journal of Economic Theory,* 126: 286–298.
- –––, 2007, “A generalised model of judgment aggregation.” *Social Choice and Welfare,* 28: 529–565.
- –––, 2008, “The premises of Condorcet's jury theorem are not simultaneously justified.” *Episteme,* 5: 56–73.
- Dietrich, F. and C. List, 2004, “A Model of Jury Decisions Where All Jurors Have the Same Evidence.” *Synthese,* 142: 175–202.
- –––, 2007a, “Judgment aggregation by quota rules: majority voting generalized.” *Journal of Theoretical Politics,* 19: 391–424.
- –––, 2007b, “Arrow's theorem in judgment aggregation.” *Social Choice and Welfare,* 29: 19–33.
- –––, 2007c, “Strategy-proof judgment aggregation.” *Economics and Philosophy,* 23: 269–300.
- –––, 2007d, “Opinion pooling on general agendas.” [Dietrich and List 2007d available online (pdf)]
- –––, 2008, “Judgment aggregation without full rationality.” *Social Choice and Welfare,* 31: 15–39.
- –––, 2010a, “Majority voting on restricted domains.” *Journal of Economic Theory,* 145: 512–543.
- –––, 2010b, “The aggregation of propositional attitudes: towards a general theory.” *Oxford Studies in Epistemology,* 3: 215–234.
- Dietrich, F. and P. Mongin, 2010, “The premise-based approach to judgment aggregation.” *Journal of Economic Theory,* 145: 562–582.
- Dietrich, F. and K. Spiekermann, 2013, “Epistemic Democracy with Defensible Premises.” *Economics and Philosophy,*29(1): 87–120.
- Dokow, E. and R. Holzman, 2010a, “Aggregation of binary evaluations.” *Journal of Economic Theory,* 145: 495–511.
- –––, 2010b, “Aggregation of binary evaluations with abstentions.”*Journal of Economic Theory,* 145: 544–561.
- –––, 2010c, “Aggregation of non-binary evaluations.” *Advances in Applied Mathematics,* 45: 487–504.
- Dowding, K. and M. van Hees, 2003, “The Construction of Rights.”*American Political Science Review,* 97: 281–293.
- –––, 2008, “In Praise of Manipulation.” *British Journal of Political Science,* 38: 1–15.

- Dryzek, J. and C. List, 2003, “Social Choice Theory and Deliberative Democracy: A Reconciliation.” *British Journal of Political Science,* 33: 1–28.
- Elsholtz, C. and C. List, 2005, “A Simple Proof of Sen's Possibility Theorem on Majority Decisions.” *Elemente der Mathematik,* 60: 45–56.
- Elster, J., 2013, “Excessive Ambitions (II).” *Capitalism and Society,* 8, Issue 1, Article 1.
- Estlund, D., 1994, “Opinion Leaders, Independence, and Condorcet's Jury Theorem.” *Theory and Decision,* 36: 131–162.
- Feddersen, T. J. and W. Pesendorfer, 1998, “Convicting the Innocent.” *American Political Science Review,* 92: 23–35.
- Gaertner, W., 2001, *Domain Conditions in Social Choice Theory*. Cambridge: Cambridge University Press.
- –––, 2005, “De jure naturae et gentium: Samuel von Pufendorf's contribution to social choice theory and economics.” *Social Choice and Welfare,* 25: 231–241.
- Gaertner, W., P. K. Pattanaik, and K. Suzumura, 1992, “Individual Rights Revisited.” *Economica,* 59: 161–177.
- Gaertner, W. and E. Schokkaert, 2012, *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*. Cambridge: Cambridge University Press.
- Gärdenfors, P., 2006, “An Arrow-like theorem for voting with logical consequences.” *Economics and Philosophy,* 22: 181–190.
- Gehrlein, W. V., 1983, “Condorcet's Paradox.” *Theory and Decision,* 15: 161–197.
- Genest, C. and J. V. Zidek, 1986, “Combining Probability Distributions: A Critique and Annotated Bibliography.”*Statistical Science,* 1: 113–135.
- Gibbard, A., 1969, “Social Choice and the Arrow Conditions.” Unpublished manuscript. [Gibbard 1969 available online (pdf)]
- –––, 1973, “Manipulation of voting schemes: a general result.”*Econometrica,* 41: 587–601.
- Gigliotti, G. A., 1986, “Comment on Craven.” *Theory and Decision,* 21: 89–95.
- Gilboa, I., D. Samet, and D. Schmeidler, 2004, “Utilitarian Aggregation of Beliefs and Tastes.” *Journal of Political Economy,* 112: 932–938.
- Goldman, A., 2004, “Group Knowledge versus Group Rationality: Two Approaches to Social Epistemology.” *Episteme, A Journal of Social Epistemology,* 1: 11–22.
- –––, 2010, “Why Social Epistemology Is Real Epistemology.”*Social Epistemology*. A. Haddock, A. Millar, and D. Pritchard (eds.), Oxford: Oxford University Press.
- Goodin, R. E. and C. List, 2006, “A Conditional Defense of Plurality Rule: Generalizing May's Theorem in a Restricted Informational Environment.” *American Journal of Political Science,* 50: 940–949.
- Grofman, B., G. Owen, and S. L. Feld, 1983, “Thirteen theorems in search of the truth.” *Theory and Decision,* 15: 261–278.
- Guilbaud, G. T., [1952] 1966, “Theories of the General Interest, and the Logical Problem of Aggregation.” *Readings in Mathematical Social Science*. P. F. Lazarsfeld and N. W. Henry (eds.). Cambridge/MA: MIT Press, pp. 262–307.

- Harbour, D. and C. List, 2000, "Optimality Theory and the problem of constraint aggregation." *MIT Working Papers in Linguistics and Philosophy 1: The Linguistics/Philosophy Interface*. R. Bhatt, P. Hawley, M. Hackl and I. Maitra (eds.). Cambridge, MA: MITWIPL, pp. 175–213.
- Harrison, G. W. and T. McDaniel, 2008, "Voting Games and Computational Complexity." *Oxford Economic Papers,* 60: 546–565.
- Hausman, D., 1995, "The Impossibility of Interpersonal Utility Comparisons." *Mind,* 104: 473–490.
- Hurley, S., 1985, "Supervenience and the Possibility of Coherence." *Mind,* 94: 501–525.
- Inada, K.-I., 1964, "A Note on the Simple Majority Decision Rule."*Econometrica,* 32: 525–531.
- Kanazawa, S., 1998, "A brief note on a further refinement of the Condorcet Jury Theorem for heterogeneous groups."*Mathematical Social Sciences,* 35: 69–73.
- Knight, J. and J. Johnson, 1994, "Aggregation and Deliberation: On the Possibility of Democratic Legitimacy." *Political Theory,*22: 277–296.
- Konieczny, S. and R. Pino Pérez, 2002, "Merging Information Under Constraints: A Logical Framework." *Journal of Logic and Computation,* 12: 773–808.
- Kornhauser, L. A. and L. G. Sager, 1986, "Unpacking the Court."*Yale Law Journal,* 96: 82–117.
- Kroedel, T. and F. Huber, 2013, "Counterfactual Dependence and Arrow." *Noûs,* 47(3): 453–466.
- Ladha, K., 1992, "The Condorcet Jury Theorem, Free Speech and Correlated Votes." *American Journal of Political Science,*36: 617–634.
- Lehrer, K. and C. Wagner, 1981, *Rational Consensus in Science and Society*. Dordrecht/Boston: Reidel.
- List, C., 2001, "A Note on Introducing a 'Zero-Line' of Welfare as an Escape-Route from Arrow's Theorem." *Pacific Economic Review,* 6, special section in honour of Amartya Sen, 223–238.
- –––, 2003a, "The epistemology of special majority voting." Working paper, London School of Economics. [List 2003a available online (pdf)]
- –––, 2003b, "Are Interpersonal Comparisons of Utility Indeterminate?" *Erkenntnis,* 58: 229–260.
- –––, 2003c, "A Possibility Theorem on Aggregation over Multiple Interconnected Propositions." *Mathematical Social Sciences,*45: 1–13 (with Corrigendum in *Mathematical Social Sciences,* 52: 109–110).
- –––, 2004a, "Multidimensional Welfare Aggregation." *Public Choice,* 119: 119–142.
- –––, 2004b, "A Model of Path-Dependence in Decisions over Multiple Propositions." *American Political Science Review,*98: 495–513.
- –––, 2006, "The Discursive Dilemma and Public Reason." *Ethics,*116: 362–402.
- –––, 2011, "The Logical Space of Democracy." *Philosophy and Public Affairs,* 39: 262–297.

- ———, 2012, "The Theory of Judgment Aggregation: An Introductory Review." *Synthese,* 187: 179–207.
- List, C. and R. E. Goodin, 2001, "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy,* 9: 277–306.
- List, C., R. C. Luskin, J. S. Fishkin, and I. McLean, 2013, "Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls."*Journal of Politics,* 75: 80–95.
- List, C. and P. Pettit, 2002, "Aggregating Sets of Judgments: An Impossibility Result." *Economics and Philosophy,* 18(1): 89–110.
- ———, 2004, "Aggregating Sets of Judgments: Two Impossibility Results Compared." *Synthese,* 140: 207–235.

May, K. O., 1952, "A set of independent, necessary and sufficient conditions for simple majority decision." *Econometrica,* 20: 680–684.

———, 1954, "Intransitivity, Utility, and the Aggregation of Preference Patterns." *Econometrica,* 22: 1–13.

McConway, K. J., 1981, "Marginalization and Linear Opinion Pools." *Journal of the American Statistical Association,* 76: 410–414.

McLean, I., 1990, "The Borda and Condorcet principles: Three medieval applications." *Social Choice and Welfare,* 7: 99–108.

McLean, I. and F. Hewitt (eds.), 1994, *Condorcet: Foundations of Social Choice and Political Theory*. Cheltenham: Edward Elgar Publishing.

McLean, I. S., A. McMillan, and B. L. Monroe, 1995, "Duncan Black and Lewis Carroll." *Journal of Theoretical Politics,* 7: 107–123.

——— (eds.), 1996, *A Mathematical Approach to Proportional Representation: Duncan Black on Lewis Carroll*. Dordrecht: Kluwer.

McLean, I. and A. B. Urken (eds.), 1995, *Classics of Social Choice*. Ann Arbor: University of Michigan Press.

Miller, D., 1992, "Deliberative Democracy and Social Choice."*Political Studies,* 40 (special issue): 54–67.

Miller, M. K. and D. Osherson, 2009, "Methods for distance-based judgment aggregation." *Social Choice and Welfare,* 32: 575–601.

Monjardet, B., 2005, "Social choice theory and the 'Centre de Mathématique Sociale': some historical notes." *Social Choice and Welfare,* 25: 433–456.

Mongin, P., 1995, "Consistent Bayesian aggregation." *Journal of Economic Theory,* 66: 313–351.

———, 1997, "Spurious Unanimity and the Pareto Principle." Paper presented at the Conference on Utilitarianism, New Orleans, March 1997. [Mongin 1997 available online (pdf)]

———, 2008, "Factoring Out the Impossibility of Logical Aggregation." *Journal of Economic Theory,* 141: 100–113.

Morreau, M., 2010, "It simply does not add up: Trouble with overall similarity." *Journal of Philosophy,* 107: 469–490.

———, forthcoming, "Theory Choice and Social Choice: Kuhn Vindicated." *Mind.*

Moulin, H., 1980, "On Strategy-Proofness and Single Peakedness."*Public Choice,* 35: 437–455.

———, 2004, *Fair Division And Collective Welfare*. Cambridge, MA: MIT Press.

Mueller, D. C., 2003, *Public Choice III*. Cambridge: Cambridge University Press.

Nehring, K., 2003, "Arrow's theorem as a corollary." *Economics Letters,* 80: 379–382.

Nehring, K. and C. Puppe, 2002, "Strategyproof Social Choice on Single-Peaked Domains: Possibility, Impossibility and the Space Between." Unpublished manuscript, University of California at Davis.

———, 2007, "The structure of strategy-proof social choice—Part I: General characterization and possibility results on median spaces." *Journal of Economic Theory,* 135: 269–305.

———, 2010, "Abstract Arrovian Aggregation." *Journal of Economic Theory,* 145: 467–494.

Okasha, S., 2009, "Individuals, groups, fitness and utility: multi-level selection meets social choice theory." *Biology and Philosophy,* 24: 561–584.

———, 2011, "Theory Choice and Social Choice: Kuhn versus Arrow." *Mind,* 120: 83–115.

Ooghe, E. and L. Lauwers, 2005, "Non-dictatorial extensive social choice." *Economic Theory,* 25: 721–743.

Parfit, D., 1984, *Reasons and Persons*. Oxford: Oxford University Press.

Pauly, M. and M. van Hees, 2006, "Logical Constraints on Judgment Aggregation." *Journal of Philosophical Logic,* 35: 569–585.

Pearl, J., 2000, *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press.

Pettit, P., 2001, "Deliberative Democracy and the Discursive Dilemma." *Philosophical Issues,* 11: 268–299.

Pigozzi, G., 2006, "Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation." *Synthese,* 152: 285–298.

Poisson, S. D., 1837, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile: précédées des règles générales du calcul des probabilités*.

Rawls, J., 1971, *A Theory of Justice*. Cambridge/MA: Harvard University Press.

Regenwetter, M., B. Grofman, A. A. J. Marley, and I. Tsetlin, 2006,*Behavioral Social Choice*. Cambridge: Cambridge University Press.

Riker, W., 1982, *Liberalism against Populism*. San Francisco: W.H. Freeman and Co.

Roberts, K., 1995, "Valued Opinions or Opinionated Values: The Double Aggregation Problem." *Choice, Welfare and Development: A Festschrift in Honour of Amartya Sen*. K. Basu, P. K. Pattanaik, and K. Suzumura (eds.), Oxford: Oxford University Press, pp. 141–165.

Saari, D. G., 1990, "The Borda dictionary." *Social Choice and Welfare,* 7: 279–317.

Salles, M. (edited with introduction), 2005, "The history of Social Choice." Special issue, *Social Choice and Welfare,* 25: 229–564.

Satterthwaite, M., 1975, "Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions." *Journal of Economic Theory,* 10: 187–217.

Sen, A. K., 1966, "A Possibility Theorem on Majority Decisions."*Econometrica,* 34: 491–499.

–––, 1970a, "The Impossibility of a Paretian Liberal." *Journal of Political Economy,* 78: 152–157.

–––, 1970b, *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

–––, 1977, "On weights and measures: informational constraints in social welfare analysis." *Econometrica,* 45: 1539–1572.

–––, 1982, *Choice, Welfare and Measurement*. Oxford: Blackwell.

–––, 1983, "Liberty and social choice." *Journal of Philosophy,*80: 5–28.

–––, 1998, "The Possibility of Social Choice." Nobel lecture, December 8, 1998, Stockholm. [Sen 1998 available online (pdf)]

Shapley, L. and B. Grofman, 1984, "Optimizing group judgment accuracy in the presence of interdependencies." *Public Choice,* 43: 329–343.

Stegenga, J., 2013, "An impossibility theorem for amalgamating evidence." *Synthese,* 190(2): 2391–2411.

Suppes, P., 2005, "The pre-history of Kenneth Arrow's social choice and individual values." *Social Choice and Welfare,* 25: 319–226.

Thomson, W., 2000, "On the axiomatic method and its recent applications to game theory and resource allocation." *Social Choice and Welfare,* 18: 327–386.

Tsetlin, I., M. Regenwetter, and B. Grofman, 2003, "The Impartial Culture Maximizes the Probability of Majority Cycles."*Social Choice and Welfare,* 21: 387–398.

Vacca, R., 1921, "Opinioni Individuali e Deliberazioni Collettive."*Rivista Internazionale di Filosofia del Diritto,*: 52–59.

Ward, B., 1965, "Majority Voting and Alternative Forms of Public Enterprises." *The Public Economy of Urban Communities*. J. Margolis (ed.). Baltimore: Johns Hopkins Press.

Weymark, J., 2006, "The Normative Approach to the Measurement of Multidimensional Inequality." *Inequality and Economic Integration*. F. Farina and E. Savaglio (eds.), London: Routledge, pp. 303–328.

Wilson, R., 1975, "On the Theory of Aggregation." *Journal of Economic Theory,* 10: 89–99.