Les ateliers de l'éthique
The Ethics Forum

DOSSIER :
Public Participation, Legitimate Political Decisions, and Controversial Technologies

UNE REVUE MULTIDISCIPLINAIRE SUR LES ENJEUX NORMATIFS
DES POLITIQUES PUBLIQUES ET DES PRATIQUES SOCIALES.

Les ateliers de l'éthique
The Ethics Forum

A MULTIDISCIPLINARY JOURNAL ON THE NORMATIVE CHALLENGES
OF PUBLIC POLICIES AND SOCIAL PRACTICES.

3

QUELLE VALEUR A NOTRE ENSEIGNEMENT AUX YEUX DES ÉLÈVES? PROLONGEMENT DE LA THÉORIE DE LA VALUATION DE DEWEY DANS LA RÉFLEXION PÉDAGOGIQUE

CHRISTOPHE POINT
DOCTORANT EN PHILOSOPHIE DE L'ÉDUCATION, UNIVERSITÉ DE LORRAINE ET UNIVERSITÉ LAVAL

RÉSUMÉ :

Le présent article examine la façon dont John Dewey a entrepris de poser le problème de la valuation et de ses conséquences au sein de sa théorie de l'éducation. Plus spécifiquement, nous voudrions montrer que son effort pour repenser l'articulation des moyens et des fins du processus de valuation contribue à repenser l'enquête morale. Celle-ci, si elle fait alors l'objet d'une pédagogie qui met au centre l'expérience vécue du sujet, nous oblige à concevoir à nouveaux frais les valeurs que nous accordons aux connaissances apprises. Notre analyse entend démontrer que : 1) de même que dans le domaine éthique, rien n'a de valeur *en soi* pour le sujet agissant. Mais que 2) la valuation de tout acte est déterminée en situation, alors, 3) dans le domaine de la pédagogie, aucune connaissance n'a de valeur *en soi* pour l'élève. 4) Ainsi la première tâche de l'enseignant.e est de former l'élève à des processus de valuation efficaces. Cela dans le but que celui-ci puisse à son tour, clairement, déterminer pour lui-même les connaissances qui auront de la valeur. Si notre hypothèse est juste alors cette tâche devient à la fois un point de départ important de la pédagogie pragmatiste et un impératif incontournable de l'éthique enseignante.

ABSTRACT:

The present article examines how Dewey addressed the problem of value, as well as the consequences that this solution had for his educational theory. More specifically, I aim to show that Dewey's effort to rethink means and ends within the process of valuation contributes to a recasting of moral inquiry. If moral inquiry becomes part of a pedagogy that focuses upon the lived experience of the subject, this compels us to rethink the value we attribute to acquired knowledge. My analysis will demonstrate 1) that even within the ethical domain, nothing has value in itself for the acting subject, but that 2) the value of all acts must be determined within specific situations, and therefore 3) within the domain of pedagogy, no knowledge has value in and of itself for the student. 4) Thus the first task of the teacher is to teach students how to evaluate questions of value, which will push them to determine for themselves what knowledge has value. If my hypothesis is correct, this task becomes simultaneously an important starting place for pragmatist pedagogy and a cornerstone of a new ethic of teaching.

INTRODUCTION

Les professeur.es le savent : enseigner est riche en surprises, et plus d'un peut s'étonner du fait que ses étudiant.es ne semblent montrer aucun intérêt pour la matière enseignée. Ils n'y attachent aucune valeur et semblent eux-mêmes étonnés que leur professeur.e place (ou fasse semblant de placer) sa matière au-dessus de tout. Cet étonnement est similaire aux situations éthiques où les évidences des un.es ne correspondent pas avec celles des autres; manifestement, deux personnes peuvent ne pas donner la même valeur au même acte ou à la même intention. Cette proximité entre la situation d'un.e professeur.e qui s'étonne et celle du désaccord éthique entre deux personnes montre bien qu'un processus de valuation est en cours dans le jugement éthique comme dans l'enseignement.

Nous ne définirons ici l'éthique ni comme la recherche d'un bien définitif ni comme la connaissance d'une fin ultime.[1] Nous faisons l'hypothèse selon laquelle au lieu d'être un champ de la connaissance éloigné de tous les autres, les jugements éthiques ne sont qu'une espèce de jugement de valeur parmi d'autres (comme les appréciations esthétiques, religieuses). Ainsi l'éthique entendue au sens étroit d'une étude de la conduite humaine en tant qu'elle est jugée bonne ou mauvaise est analysée ici à partir du processus plus général de la valuation, c'est-à-dire du comportement humain de donner une valeur aux choses, aux faits, aux actes, etc. Cet angle de la réflexion rend possible notre questionnement de départ : un *processus de valuation* est-il à l'œuvre chaque fois qu'un individu entre dans une démarche d'enseignement? Un choix est-il à faire entre plusieurs connaissances, méthodes, cours, et se pose-t-on la question de savoir quelle connaissance est, pour soi, plus *souhaitable* que telle autre?

Nous ouvrons ici une boîte de Pandore : si nous interrogeons la valeur que peut avoir une connaissance pour quelqu'un et non pas en elle-même, alors cette conception subjectiviste de la connaissance a pour conséquence directe un risque de clientélisation de l'élève. Les étudiant.es parcourant leur université sont-ils semblables aux consommateurs poussant leur chariot dans les allées d'un supermarché? Ce risque est réel, et se réfugier dans une conception objectiviste complète de la valuation d'une connaissance (l'étudiant.e serait naturellement intéressé.e par certaines connaissances et pas d'autres) est une façon, à notre avis tout aussi dangereuse, de se rendre aveugle au problème. Comment penser le processus de valuation du sujet quant à une connaissance? La réflexion éthique peut-elle nous éclairer sur ce problème propre à la pédagogie?

Nous souhaitons revenir ici sur la théorie de la valuation de Dewey, telle qu'il la présente dans l'*International Encyclopedia of Unified Science*[2] pour étudier les conséquences d'une telle théorie sur la conception pragmatiste de la pédagogie. En effet, pour mener à bien ce processus éthique, Dewey propose d'employer le modèle de l'enquête. Celle-ci sera ici le modèle de rationalité que nous proposons pour mener à bien des jugements éthiques. Ce modèle a le mérite de proposer une voie claire qui ne sombre ni dans un subjectivisme complet où l'individu devient un client roi et aveugle, ni un objectivisme suranné où un *deus ex machina* vient révéler la valeur objective et éternelle des connaissances.

Ainsi, 1) nous reviendrons sur l'articulation des moyens et des fins dans la théorie de Dewey pour montrer que le processus de valuation construit par le pragmatisme se distingue clairement d'une recherche aveugle du désir. Puis, 2) nous établirons aussi l'importance de la continuité des moyens et des fins qui distingue l'enquête d'une rationalité instrumentale classique. Dans un troisième temps, 3) nous verrons comment la notion d'expérience est centrale dans la valuation et nous oblige à mettre au cœur de notre réflexion pédagogique la question de la qualité de l'expérience qu'une connaissance rend possible. Enfin, 4) à partir de cette nécessité de penser la formation des valeurs dans le travail de l'enseignant.e, nous pourrons esquisser les contours de l'éthique enseignante que propose Dewey.

1. LA THÉORIE DE LA VALUATION DE DEWEY

Notre auteur étudie l'articulation des moyens et des fins-en-vue (*end-in-view*)[3] tout en essayant d'éloigner les théories confondant désirs et plaisirs au sein de ce processus, ou encore la théorie de la rationalité instrumentale développée par Max Weber[4] et Herbert Simon[5] (dont nous discuterons dans un second temps).

Le premier mouvement que Dewey opère est le constat que le désir naît d'une situation où quelque chose fait problème. Quelque chose manque (un objet, une personne, une relation) et le désir que nous avons de cette chose provient du fait que nous la considérerons comme la solution au problème que nous vivons. Il ne peut donc pas y avoir de désir de *choses-en-soi*, ni non plus de *fins-en-soi*, car le processus même du désir fonctionne « en situation ». Il faut qu'un effort ait lieu, qu'une difficulté se présente pour qu'un désir naisse, et par conséquent que le processus de valuation soit engagé.

Ici c'est le schème de l'enquête[6] qui est repris en arrière-fond de la réflexion de l'auteur. Pour comprendre le processus du désir, il faut l'associer à ce schème mental. C'est parce qu'une situation de trouble, de doute, de manque s'installe chez un sujet que celui-ci cherchera une solution[7] pour sortir de cet état, l'identifiera dans une chose et ainsi la désirera en raison du problème qu'il rencontre. Il n'y a donc pas de bons ou de mauvais désirs *en-soi*, car chaque problème pourra se voir conférer un grand nombre de résolutions possibles. Il existe alors, dès le début du processus, une multiplicité de désirs possibles dans une situation, et une multitude de fins possibles (que nous pouvons pour l'instant concevoir comme des fins en tant qu'éléments de résolution du problème). Mais si on élimine de la réflexion la question des bons et des mauvais désirs *en-soi*, nous ne résolvons pas pour autant le problème de la détermination des bons ou des mauvais désirs *en situation*.

Avant d'aller plus loin, remarquons que concevoir le désir ainsi élimine également de la réflexion sur la valuation une confusion possible de celle-ci avec le processus de plaisir, car celui-ci, qu'il soit conçu comme une impulsion vitale ou comme une habitude, échappe à un processus de résolution du problème. Nous verrons pourquoi.

La notion de plaisir (*enjoyment*)[8] peut être présente sans qu'un effort de valua-tion ait eu lieu alors que le désir d'un sujet pour une chose ne peut émerger qu'au prix d'un effort, d'un travail propre au processus de valuation. Cela se constate chaque jour. Je peux éprouver du plaisir sans chercher à résoudre un problème. Par exemple, une odeur de pain chaud arrive à mon nez, car je marche près d'une boulangerie et j'aurai plaisir à sentir cette odeur indépendamment du fait que je sois rassasié ou que la faim me tiraille. Je ne peux pas donner une valeur à cette odeur, même si elle me procure du plaisir, car celle-ci n'intègre pas une enquête intérieure. Je l'apprécierai (*appraisal*) sans la valoriser (*prizing*). Notons qu'il en va de même pour la notion de déplaisir.[9] Sans une valuation, un plaisir ou un déplaisir ne peut modifier mon action, nous dirons qu'il reste sans effet et « gratuit ».[10] Un plaisir ou un déplaisir qui ne modifie pas une habitude compor-tementale est donc considéré sans valeur. Aux yeux d'une philosophie de l'ac-tion, elle reste exclue du processus de valuation.

Toutefois, à ce stade de la réflexion, on pourrait nous objecter que notre notion de désir reste vague et que de nombreuses choses peuvent être désirées sans pour autant que le sujet accorde à ces choses une valeur. Sommes-nous en train de confondre alors les notions de désir et de valeur? En effet, bien d'autres théories peuvent nous faire croire que le désir tire sa puissance et son origine[11] d'une impulsion vitale propre à l'individu même. Or si le désir est une impulsion vitale, ce n'est pas pour autant que celui-ci devient le fondement d'un processus de valuation. En effet, chez l'homme comme chez l'animal, les impulsions vitales ou « tensions organiques » donnent lieu à des enquêtes inconscientes et produi-sent des comportements que nous pouvons qualifier d'instinctifs. Mais si ces comportements sont des instincts, ils sont donc réalisés par le sujet sans réflexion sur leur valeur. Ils sont donc, de fait, exclus d'un processus de valuation.

Au contraire d'une impulsion vitale, il faut comprendre le désir comme l'étape d'un processus qui constitue la transformation d'une impulsion antérieure, d'un comportement établi ou encore d'une habitude routinière en un comportement nouveau. L'instinct, à l'inverse, conserve un comportement établi, car jugé satis-faisant dans une situation antérieure, et sera incapable de le réévaluer lors d'une nouvelle situation. Le désir possède une souplesse d'adaptation là où l'instinct se fait droit et cassant. Il a une valeur lorsqu'il modifie la résolution d'une enquête en vue d'une plus grande satisfaction d'un problème, ou de la résolution d'un nouveau problème jusqu'alors inconnu pour le sujet. Il y a donc, et c'est ce que veut montrer Dewey dans un premier temps, un élément intellectuel indis-pensable au processus de valuation :

> Une valuation n'a lieu que lorsque quelque chose fait question : quand il y a des difficultés à écarter, un besoin, un manque ou une privation à combler, un conflit entre tendances à résoudre en changeant les condi-tions existantes. Ce fait prouve à son tour qu'un élément intellectuel – un élément d'enquête – est présent chaque fois qu'il y a valuation.[12]

Le processus de valuation n'est donc absolument pas instinctif. Au contraire, il est le lieu d'une pleine réflexion intellectuelle. Ainsi, ce que propose Dewey par sa théorie de la valuation, c'est de penser le modèle de rationalité par lequel nous conduisons nos processus de valuation. Si une rationalité est à l'œuvre dans ce processus, alors il est possible de construire une éthique claire, explicite et justifiée. Celle-ci se pose les questions suivantes, par exemple : ma réflexion porte-t-elle sur le bon problème à résoudre? Ma construction de celui-ci est-elle adéquate par rapport à ma situation? La fin que je me propose d'atteindre résoudra-t-elle ce problème totalement ou partiellement? La valuation porte ainsi en partie sur les critères de justesse de la fin poursuivie, mais également sur la validité du problème construit. C'est pourquoi cette éthique doit être particulièrement attentive à deux moments de l'enquête morale.[13]

1) Le problème établi identifie-t-il clairement le manque en question? C'est-à-dire est-ce que la fin que me propose le problème que j'ai construit est *adéquate* à ma situation? C'est sur ce point que nous verrons, dans le deuxième temps de notre analyse, en quoi le modèle de l'enquête se distingue du modèle de la rationalité instrumentale. 2) Quelle est la probabilité que j'atteigne cette fin au vu de ma situation? Ou pour le dire autrement, est-ce que la fin que je me propose de poursuivre est *réalisable* au vu de ma situation?

Adéquation et *réalisation* correspondent ici aux deux critères épistémologiques de Dewey[14] pour qualifier la justesse d'une enquête. Ces deux critères s'inspirent de la conception évolutive héritée du darwinisme en biologie (pour l'adéquation d'un individu à son milieu et des évolutions qui découlent d'une inadéquation)[15] et de l'expérimentalisme hérité de l'épistémologie des sciences empiriques (pour la réalisation en tant que procès de l'expérience comme contrainte de vérification d'une hypothèse).[16] Nous comprenons ici que fins et moyens d'un désir sont à étudier par un processus de valuation qui est toujours en contexte. Un sujet portera un intérêt[17] à une chose s'il désire ce que la chose lui permettra de combler. C'est-à-dire qu'il accordera une valeur à une chose s'il pense qu'elle peut résoudre son problème. Plus le problème sera jugé important par l'individu et plus la chose en question aura une valeur à ses yeux.

Néanmoins, ces deux critères épistémologiques sont plus simples à formuler qu'à appliquer. Les enquêtes morales que tentent de réaliser les individus au quotidien peuvent être troublées par de nombreux facteurs. La pression sociale des proches, les idéologies acquises par préjugés, l'influence d'une domination symbolique politique ou religieuse… Les occurrences éloignant le processus de valuation du modèle de l'enquête font légion. Elles justifient ainsi par leur existence le travail du philosophe de chercher un moyen de les éviter ou, au moins, de diminuer leur importance. Le but poursuivi par Dewey ici est bien de soulever ces difficultés pour aider – modestement – les individus à mener un peu mieux la façon dont ils attribuent de la valeur à un acte, à un objet ou à une connaissance.

Concevoir ainsi le processus de valuation permet également de comprendre tous les dysfonctionnements quotidiens de l'articulation entre les moyens et les fins. En effet, certains moyens sont alors *inadéquats* à telles fins et certaines fins sont alors *irréalisables*. Les moyens inadéquats peuvent être vus comme des corvées, même si la fin poursuivie reste souhaitable. Nous sommes alors dans une situation où l'information qu'a le sujet de l'objet ne suffit pas à rendre la fin poursuivie véritablement souhaitable pour le sujet lui-même. Celui-ci doit donc améliorer les moyens nécessaires à sa mise en œuvre. Certaines fins, irréalisables, deviendront des « idéaux » trop éloignés pour être possibles et décevront l'individu qui les poursuit. Le sujet doit alors modifier cette fin poursuivie pour la rendre plus accessible à la réalisation par les moyens qu'il possède. Entre Charybde et Scylla, Dewey démontre le seul lieu hospitalier au désir humain : le territoire du possible. Comment cartographier ce territoire pour les valeurs?

2. LA CONTINUITÉ DES MOYENS ET DES FINS

Le modèle de l'enquête se distingue de celui de la rationalité instrumentale, car l'enquête ne se fonde pas sur une conception formelle de la raison que l'on appliquerait à priori à la situation. Au contraire, elle prend son point de départ dans le contexte où la pensée est à l'œuvre. Ainsi la pratique de l'enquête n'est pas aussi mécanique qu'une sorte de calcul; elle se heurte à des incertitudes, des complexités, des doutes et elle rend compte des doutes et conflits de valeurs très réels au quotidien.

Pour le dire autrement, la conception de la rationalité développée par la théorie de la rationalité instrumentale reste technique : la raison est devant un problème à résoudre. Dans cette optique, les questions éthiques peuvent être traitées techniquement, de façon indépendante des fins poursuivies.[18] Or, la conception de la rationalité développée par le modèle de l'enquête fait en sorte que le problème n'est jamais donné, mais doit être déterminé.[19] C'est la situation indéterminée qui crée la qualité problématique par laquelle l'enquête se formulera par la suite. Ce travail de détermination du malaise en problème est ce qui fera réfléchir le sujet sur les moyens qu'il jugera acceptables et sur les fins qu'il se proposera d'atteindre. Mais ce travail n'est pas fixé une fois pour toutes comme dans la théorie de la rationalité instrumentale. Au contraire, pour Dewey, ce processus est dynamique et appelle les fins à être sans cesse interrogées, et non écartées en dehors de l'exercice de la rationalité.

Le problème de l'inadéquation des moyens et des fins apparaît pour Dewey à partir du moment où l'on entre dans une réflexion qui exclut la relation entre ces deux termes. Ne considérer que les moyens pour donner une valeur à telle chose et ne considérer que telle fin sans prendre en compte les moyens de sa réalisation sont deux erreurs de jugement ayant la même source : négliger la *continuité* des moyens et des fins entre eux. Cette *continuité* nous semble être le véritable idéal de la valuation pour Dewey. Elle n'est atteignable qu'au moment où le processus de valuation a pu trier entre tous les moyens possibles et toutes les fins imaginables le moyen approprié et la fin adéquate grâce à une codéter-

mination réciproque de ces deux éléments. Une continuité totale entre les deux est donc le résultat d'une enquête parfaite, idéal que le quotidien est souvent bien loin d'approcher. Combien de fois avons-nous sélectionné un moyen inadéquat pour résoudre un problème, ou bien visé une fin trop éloignée de ce qui était en notre pouvoir d'accomplir?

C'est pour répondre à ces deux types d'échecs quotidiens que notre auteur décide de rejeter la maxime populaire « la fin justifie les moyens », car celle-ci, en concentrant son attention sur la fin, risque de nous faire utiliser de mauvais moyens. Son inverse « les moyens justifient la fin » est également faux, car ne décider de la fin qu'en vue des moyens ne nous permet pas non plus de choisir la bonne fin recherchée.[20] L'erreur de ces deux formulations se construit sur la croyance que des *fins-en-soi* peuvent être attribuées (ou plutôt découvertes) au sein de certaines choses. Ce finalisme est une rupture dans la continuité de la relation moyen/fin et réduit radicalement la compréhension des processus de valuation.

C'est pourquoi, à l'inverse du modèle de la rationalité instrumentale, le modèle de l'enquête revendique une unicité de la situation. Ainsi, lors des questionnements éthiques, chaque situation est unique et l'on ne peut automatiser ou standardiser aucune procédure au sein d'une enquête. On ne peut que chercher *un* moyen en vue d'*une* fin, sans espérer pouvoir reproduire à l'identique l'opération. De plus, cette *continuité* implique un autre point essentiel dans la théorie de la valuation de Dewey : les fins recherchées ne sont jamais finales, mais peuvent devenir des moyens pour d'autres fins. Ce point permet également de justifier le point précédent sur l'inutilité et le danger dans le processus de valuation de favoriser l'un des deux éléments du binôme moyen/fin.

Dewey s'inspire des sciences empiriques pour récuser toutes fins finales et s'appuyer sur une causalité permanente des événements entre eux.[21] Cet argument éclaire d'une nouvelle façon son schème de l'enquête.[22] En effet, ce schématisme pouvait nous induire en erreur en nous faisant concevoir l'enquête sur un mode linéaire, avec un début et une fin. Or il faut davantage la comprendre comme un circuit qui ne cesse jamais. L'état de repos, d'ataraxie, qui compose le début et la fin du processus ne sont jamais que des moments, des états transitoires qu'il serait illusoire de considérer comme permanents. La fin recherchée transformera la situation du sujet, qui désirera ainsi une autre fin et ainsi de suite. Il n'est donc pas étonnant que pour Dewey, considérer une fin comme finale soit la preuve d'une immaturité, d'une routine sédimentée ou encore d'un fanatisme.[23]

Le processus de l'enquête, comme celui de la pensée ou du désir, ne se conclut jamais qu'au moment de la mort de l'individu. De même, cette circularité de l'enquête se retrouve dans celle du processus de valuation où les moyens et les fins recherchés s'enchaînent à l'infini, sans cesse relancés par la situation dans laquelle se trouve l'individu. Les temporalités respectives des moyens et des fins ne permettent pas une hiérarchisation de l'une ou de l'autre dans le proces-

sus de valuation. Cette erreur fut commise dans les morales déontologiques de l'Antiquité et on la retrouve dès que la notion de *fin-en-soi* apparaît. Mais, à ce moment de notre réflexion, un risque survient. Comment dissocier, dans une telle continuité, les moyens des fins? Ne sommes-nous pas en train de procéder à un amalgame qui réduirait la différence de ces deux éléments et en ferait des synonymes?

Pour contrer ce problème d'identification des moyens et des fins, Dewey propose de considérer les fins comme les éléments d'organisation des moyens. Ainsi la fin est ce qui coordonne (et veille sur) l'ensemble du processus où les moyens travaillent. Ce rôle de coordination n'en fait pas un élément supérieur à celui qu'il coordonne, mais en livre la particularité. Et c'est celle-ci qui empêchera cette circularité de se figer dans une succession absurde de moyens sans but. Considérer les fins ainsi permet d'écarter le problème du fondement de la valeur si souvent rencontré dans les réflexions éthiques. Ce problème, formalisé par Hans Albert, montre que chercher une *fin-en-soi* ou un principe suffisant à la fondation d'une valeur revient à se confronter à une triple aporie appelée « le *trilemme* de Munchausen ».[24]

Ce *trilemme* menace toute théorie de la valuation qui cherche à fonder la valeur dans une *fin-en-soi* et l'oblige à se confronter à 1) l'apparition d'un cercle vicieux, ou 2) une régression à l'infini, ou encore 3) un arrêt arbitraire et dogmatique du processus de justification. Ce à quoi échappe la théorie de Dewey, car :

1) Pour la première aporie du cercle vicieux : ce n'est jamais le même processus qui se reproduit à l'infini dans la valuation réalisée par le sujet, étant donné que la situation de celui-ci évolue sans cesse, ainsi les *coordonnées* du problème (moyens + fins) évoluent également et l'enquête ne se réalise jamais à l'identique. De fait, il n'y a pas de répétition du même processus, car chaque enquête « nourrit » la suivante de ses acquis. En ce sens, la théorie de la valuation de Dewey s'écarte une fois de plus de la théorie de la rationalité instrumentale pour se rapprocher de la théorie wittgensteinienne du « voir comme » et de l'« agir comme » où la valuation d'une situation peut être influencée par une autre suivant son degré de familiarité avec une situation antérieure.[25]

2) Pour la seconde aporie de la régression à l'infini : le sujet ne cherche pas à l'infini une fin finale qui fonderait la chaîne des moyens. L'éthique de Dewey n'est pas la quête d'un bonheur à partir d'un principe qui détermine hors contexte la valeur de toutes choses. Au contraire d'une régression à l'infini, c'est l'évolution permanente d'une satisfaction contextualisée qui est prônée. C'est la continuité des éléments qui permettra une progression dans la valuation. Ainsi, dans le processus de l'enquête, il ne s'agit pas de distinguer la pensée et l'action comme deux phases séparées. Là aussi, à rebours de l'agir instrumental, le modèle de l'enquête peut s'apparenter à une conversation réflexive où agir et penser ont une co-extension qui donne lieu à une transaction incessante entre le sujet et le contexte, le moyen et la fin. L'enquêteuse ou l'enquêteur ne sont pas des joueurs

d'échecs ayant déjà prévu tous leurs coups avant le début de la partie. Au contraire, chacun de leurs coups est une expérimentation qui fait évoluer leur jeu.[26]

3) Enfin, pour la troisième aporie de l'arrêt arbitraire et dogmatique du processus de valuation : celle-ci est sans cesse reportée par l'activité même du sujet. Cette activité correspond à la vie du sujet et ne s'arrête qu'à sa mort, ce qui est donc un arrêt qui ne dépend pas de la valuation même. Toutefois, cet arrêt peut se produire également par un recours à l'instinct ou à l'habitude routinière, mais nous avons vu que ces deux éléments sont des obstacles externes au processus de valuation. Un dogmatisme éthique n'aurait aucun sens au sein de la théorie de la valuation de Dewey, car l'objet de la valuation est contemporain et immanent à son processus. Ce que je juge bon ou mauvais pour moi peut l'être avec justesse, mais seulement si cet objet est devant moi, compris dans ma situation. Sinon mon jugement éthique formera une enquête artificielle avec un problème factice.[27]

Pour revenir à cette continuité des moyens et des fins, nous pouvons dès lors fixer un objectif à nos pratiques éthiques où un processus de valuation est nécessaire. Cet objectif sera celui de la coordination réussie des moyens et des fins au sein de chacune de nos activités. Cette coordination est une réussite lorsqu'elle est à la fois un *rétablissement* et un *déploiement* d'une situation. Si l'usage d'un moyen est si coûteux moralement qu'il ne permet pas au sujet de viser, une fois son problème résolu, d'autres fins, alors la valuation sera déclarée défectueuse, car elle n'a permis de rétablir une situation qu'au prix d'un enfermement de l'individu sur soi. À l'inverse, si une fin une fois réalisée permet de déployer une grande diversité de nouvelles situations possibles, mais sans pouvoir rétablir la situation initiale, celle qui pose problème, alors le processus de valuation sera également défectueux.

Ainsi chacun des résultats successifs a un contenu ou un objet qui diffère de celui de ses prédécesseurs, car, si chacun d'eux est le rétablissement (*reinstatement*) d'un cours d'activité unifié, interrompu un temps par le conflit et le manque, il est aussi le déploiement (*enactment*) d'un nouvel état de choses.[28] De même qu'en peinture, chaque geste du peintre est réalisé sous la coordination du tableau que le peintre a en tête, chaque coup de pinceau à la fois réaffirme aux yeux du peintre le tableau imaginé et déploie dans son imagination de nouvelles possibilités au tableau commencé. Le modèle de l'enquête deweyenne pour penser les processus de valuation est donc plus proche de la métaphore de la « conversation réflexive » de Schön que de celle du « calcul automatique » de Simon. Ce qui distingue ces deux modèles de rationalité tient dans l'importance donnée à la notion d'expérience dans la valuation.

3. UNE PÉDAGOGIE DE L'EXPÉRIENCE

Nous faisons ici l'hypothèse selon laquelle l'importance de la continuité dans un jugement de valeur permet de déterminer la façon dont les pédagogues doivent proposer des expériences à leurs élèves. Mais pour cela, il nous faut dire à quel

point l'expérience est primordiale chez Dewey. En effet si celle-ci est centrale dans la pensée de Dewey, il faut alors, pour réfléchir à une bonne éducation, déterminer la nature de ces expériences et pouvoir les distinguer entre elles.

L'individu fera de multiples expériences tout au long de sa vie, mais toutes ne seront pas « authentiques » ou fructueuses pour lui. Dans ces conditions, construire un milieu où se développe l'enfant, c'est rendre possibles certaines expériences et en interdire d'autres; un règlement intérieur d'établissement scolaire ne fait explicitement pas autre chose que de déterminer les expériences « autorisées » et « proscrites ». L'une des tâches des éducatrices et éducateurs sera donc de faire le tri, de choisir les expériences à favoriser chez l'enfant. C'est ce que Dewey exprime ainsi :

> D'où il apparaît que le problème central d'une éducation basée sur l'expérience consiste à choisir la nature des expériences présentes capables de demeurer fécondes et créatrices dans les expériences suivantes.[29]

Le choix des expériences doit donc se faire à la lumière de critères solides au sein d'un processus de valuation de celles-ci. Dans cet extrait, ces critères semblent être la « fécondité » et la « créativité ». Qu'est-ce qu'une expérience féconde et créatrice? Cette question se pose au vu de la notion de continuum expérimental qui permet de sélectionner les expériences selon leur « valeur éducative ».[30] Mais quelle est donc cette continuité que l'on recherche entre les expériences pour qu'elles forment un continuum expérimental entendu ici en son sens le plus étendu et non plus seulement éthique? S'agit-il simplement d'une relation cumulative ou d'une succession temporelle d'expériences?

> En voici la caractéristique fondamentale : chaque expérience faite modifie le sujet et cette modification, à son tour, affecte – que nous le voulions ou non – la qualité des expériences suivantes, le sujet étant un peu différent après chaque expérience de ce qu'il était auparavant. […] De ce point de vue, le principe de continuité de l'expérience signifie que chaque expérience, d'une part, emprunte quelque chose aux expériences antérieures[31] et, d'autre part, modifie de quelque manière la qualité des expériences ultérieures.[32]

Voici ce qu'est la continuité expérimentale en tant que principe : une expérience sera jugée « positive » si elle permet au sujet de réaliser d'autres expériences. Téter le sein de sa mère pour un nourrisson sera une expérience enrichissante, car cela lui permettra de croître pour ensuite se nourrir par lui-même d'aliments autres que le lait maternel. La continuité des expériences alimentaires du nourrisson forme l'accroissement du nombre d'objets susceptibles de le nourrir. Chaque étape de ce processus expérimental, c'est-à-dire chaque expérience prise séparément, a une valeur dont la réussite est relative aux expériences réalisées – en amont et en aval.[33] Le but de l'éducation est donc de favoriser cette continuité des expériences pour que le sujet soit toujours plus enclin à faire des expériences qui l'autorisent à leur tour à en faire d'autres. Ainsi, mener un processus

de valuation sur nos expériences est non seulement possible, mais prouve également une performativité de celle-ci. Cette enquête démontre notre intelligence. C'est pourquoi Dewey affirme que :

> Le progrès de la biologie a accoutumé nos esprits à l'idée que l'intelligence n'est pas un pouvoir extérieur présidant de manière suprême, et statique aux désirs et aux efforts des hommes, mais une méthode d'ajustement des capacités et des conditions à l'intérieur de situations spécifiques.[34]

Ainsi l'enquête est la méthode de l'intelligence et penser devient la capacité à choisir les expériences les plus propices pour *l'adaptation* à l'environnement. S'adapter signifie alors choisir l'expérience qui rendra plus sûre la continuité entre mes expériences passées et mes expériences futures. Concevoir l'adaptation comme un ajustement entre une situation et un individu c'est avant tout donner un espace de manœuvre, un pouvoir d'agir à l'individu. Ou pour le dire autrement avec Jean Piaget,[35] l'adaptation chez Dewey est l'expression d'une liberté de l'individu sur sa situation.[36] Cette liberté n'est réelle que parce qu'elle est située dans un milieu (c'est-à-dire une situation), et donc également limitée par la constitution de celui-ci. L'avantage conceptuel d'une telle pensée est de ne plus concevoir le mouvement de l'adaptation comme une lutte pour la survie, mais comme la progression d'une vie grâce à l'intelligence. L'adaptation devient l'expression d'une intelligence et non plus le mouvement de forces aveugles.[37] Ainsi, au sens propre, réaliser le processus de l'enquête consiste à enquêter pour déterminer quelles sont les expériences favorables à l'individu, et que l'on pourrait, pour cette raison, souhaiter lui enseigner. On peut en conclure que la méthode de l'enquête correspond pour Dewey à une attention intelligente au continuum expérimental. L'éducation doit mener l'individu à acquérir la compétence de l'enquête grâce à une attention extrême à la continuité de ses expériences.

C'est pourquoi il faut maintenant reconstruire le gain de l'éducation à partir de cette notion d'expérience. L'éducation doit permettre à l'individu d'accroître son discernement sur la valeur éducative des expériences qu'il peut effectuer. Celle-ci nous apparaît dans sa capacité à soutenir une croissance naturelle chez l'élève. En liant des expériences entre elles, l'enquête nous permet de dépasser nos préjugés, nos problèmes éthiques et nos contradictions morales pour soutenir la croissance de l'élève en élargissant son horizon d'expériences possibles. Dewey ira même jusqu'à dire que la valeur de l'école se mesure à cette capacité d'entreprendre ce processus de valuation des connaissances à enseigner :

> La valeur de toute institution : sociale, économique, domestique, politique, juridique, religieuse, se mesure à l'effet qu'elle exerce sur l'élargissement et l'amélioration de l'expérience.[38]

Le critère de la valuation qui permet d'évaluer une connaissance est l'« élargissement et l'amélioration » de l'expérience. C'est alors au regard de la croissance du continuum expérimental des individus que l'on jugera du bénéfice des institutions de son pays. On juge donc cette école suivant la qualité des expériences qu'elle a permises à ses élèves et celles qu'elle leur permet à l'avenir. La théorie de la valuation de Dewey peut dès lors s'exprimer pleinement dans sa théorie de l'éducation. L'enseignant.e peut à sa façon rendre possible la réalisation de ce processus de valuation.

4. ESQUISSE D'UNE ÉTHIQUE ENSEIGNANTE

À partir de ce qui a été développé dans les sections précédentes, est-il possible de mettre à profit les implications de la théorie de la valuation de Dewey pour l'éthique enseignante? Cette interrogation n'a rien de rhétorique, car si c'est le sujet qui est le lieu de la valuation de ses actes et de ses pensées, alors on ne peut lui imposer frontalement une autre valuation en lui disant : « tu te trompes, voici la vraie valeur des choses! ». Cela signifie-t-il que les professeur.es ne peuvent donner tort aux élèves, et que si ces derniers n'accordent aucune valeur aux théorèmes géométriques, par exemple, leur professeur.e ne doive pas les détromper? De fait, une position subjectiviste de la valuation contient le risque de rendre relatif à chacun.e la valuation des connaissances. Or cette relativité peut entraîner chez l'élève un comportement de clientélisme : « telle connaissance m'intéresse, mais pas telle autre ». Cet.te élève n'aurait alors qu'une relation marchande avec l'éducation et risquerait de tomber dans tous les pièges de la consommation. Un tel comportement transformerait l'éducation en un service et la desservirait gravement dans son organisation et ses objectifs.[39]

C'est pour éviter un tel clientélisme de la part des élèves qu'une éthique enseignante est nécessaire. Celle-ci se fonde sur une nécessité indéniable : éviter à l'élève les erreurs dans sa valuation des connaissances. En effet, sans formation, le sujet risque de se laisser guider par ses intérêts immédiats, des influences extérieures de domination, et le choix opéré à la fin du processus de valuation risque non pas d'accroître, mais de faire décroître la qualité de son continuum expérimental. C'est en raison de l'immaturité naturelle des élèves que leurs professeur.es ont pour tâche première de guider ceux-ci dans leur valuation. Loin de les abandonner à eux-mêmes comme des consommateurs dans un supermarché, les enseignant.es doivent d'abord tâcher de les former à des processus de valuation efficaces. Cela dans le but que les élèves puissent à leur tour, clairement, déterminer pour eux-mêmes les connaissances qui auront de la valeur.

Pour mener à bien cette formation, Dewey propose deux objectifs qui forment les premiers jalons de l'éthique enseignante recherchée. Le premier objectif relève de façon évidente de l'intérêt de la formation elle-même, c'est-à-dire que le sujet doit, à la fin de celle-ci, être autonome dans la formulation de ses propres jugements. Faire de l'autonomie de l'individu dans ses processus de valuation un objectif et non une prémisse montre bien que les élèves ne peuvent décider seul.es de ce qu'ils doivent apprendre. Elles/ils ne savent pas, au début, quelles

connaissances leur seront nécessaires, donc ils ne peuvent estimer avec justesse les connaissances qui auront de la valeur pour elles/eux. Les professeur.es deviennent des guides pour que les élèves maîtrisent de mieux en mieux la valuation en la rendant explicite, claire et révisable. Il faudra interroger les élèves sur leurs choix, leur permettre de formaliser ces choix et d'en discuter pour qu'ils ne se laissent pas mener aveuglément par leurs désirs, mais qu'ils soient capables d'enquêter sur ce qu'elles et ils jugent bon ou mauvais. Ce premier objectif de l'éthique enseignante est de rechercher l'autonomie des élèves dans leurs jugements éthiques, c'est-à-dire de faire d'eux des enquêteurs aussi lucides que possible sur leurs processus de valuation.

Le deuxième objectif prévient un danger du premier objectif énoncé : si les professeur.es recherchent l'autonomie des élèves, ne risquent-t-ils pas de les enfermer dans une position solipsiste quant à leurs choix et, *in fine*, quant à leurs connaissances? Enfermé.es dans une telle position, les élèves risquent d'être influencé.es inconsciemment par leur culture, les mœurs en vigueur et les institutions en place. Autant d'obstacles à l'enquête qui nuisent au sujet s'il ne les rend pas explicites pour lui-même. Or la meilleure façon de rendre claires les influences extérieures aux yeux d'un sujet est de les partager avec autrui. Ce partage est indispensable aux jugements éthiques :

> La capacité à supporter la publicité et la communication est le test par lequel on décide si un bien supposé est authentique ou sans fondement.[40]

Ainsi les professeur.es auront comme principe éthique de permettre aux élèves de partager leurs processus de valuation pour les confronter à ceux des autres. Les valeurs pourront alors être publiques et discutées par l'ensemble des élèves.[41] Cette confrontation contribue à ce que tou.tes participent à la formation des valeurs et révèle du coup le caractère démocratique de l'éthique enseignante prônée par Dewey.[42]

Nous pouvons, pour conclure, formaliser une esquisse de l'éthique enseignante que rend possible une telle théorie de la valuation en trois propositions. Celles-ci seraient moins là pour imposer par principe tel ou tel comportement que pour orienter la pratique des professeur.es. Une éthique minimale a plus de chance d'être efficiente, car en étant plus souple, son adaptation aux différences de situations pédagogiques permet davantage sa mobilisation.[43]

1) La première proposition est de ne pas faire croire aux élèves que les connaissances ont une valeur en soi. Il s'agit de ne pas projeter nos propres valeurs en donnant l'impression qu'elles sont vraies ou bonnes pour tou.tes. Cela peut sembler être un geste contre-intuitif, mais il est primordial de faire comprendre aux élèves qu'il n'y a pas, pour les valeurs, de « réalité antérieure »[44] au processus de valuation. Leur réalité appartient toujours à la situation actuelle des personnes;[45] c'est donc à elles d'engager ce processus et non aux professeur.es de l'imposer normativement.[46]

2) Sensibiliser les élèves au processus de valuation pour qu'ils puissent peu à peu maîtriser de mieux en mieux celui-ci est notre deuxième proposition. Elle constitue le cœur de l'exigence de formation de notre éthique enseignante. Nous ne pouvons ici interroger les moyens didactiques d'une telle formation, mais nous souhaitons souligner la nécessité que ce processus de valuation devienne pour eux une habitude à la fin de leur éducation. L'éthique enseignante de Dewey, si elle veut atteindre son objectif d'autonomie des élèves, doit prendre en compte très tôt cette formation à la valuation dans leur éducation.[47]

3) Notre troisième proposition suggère de sensibiliser les élèves à l'aspect public de la valuation. Cela signifie qu'il faut, d'une part, exposer l'influence de leur milieu social sur leurs jugements moraux, et d'autre part leur montrer l'intérêt du partage de ces jugements entre eux pour parvenir à un partage conscient et explicite de leurs valeurs. Ce partage de l'enquête morale est indispensable à la théorie de la démocratie de Dewey, car elle conditionne la possibilité d'un partage de l'enquête sociale sans domination des un.es sur les autres.[48] Ainsi, en formant les élèves à échanger et à partager leurs processus de valuation, nous commençons le processus de formation qui permettra de les éduquer à la citoyenneté.[49]

NOTES

1 Dewey, John, *Reconstruction in Philosophy*, The Middle Works of John Dewey, vol. 12, *The Collected Works of John Dewey*, 1920, p. 172.

2 Dewey, John, « Theory of valuation », *International Encyclopedia of Unified Science*, vol. II, no. 4, Chicago, The University of Chicago Press, 1939, p. I-67. Dewey, John, « La théorie de la valuation », trad. A. Bidet, *Tracés. Revue de Sciences humaines*, no. 15, 2008, pp. 217-228.

3 *Ibid.*, p. 217.

4 Weber, Max, *L'éthique protestante et l'esprit du capitalisme*, trad. Jacques Chavy, Paris, Plon, 1964 [1904].

5 Simon, Herbert, « The Logic of Rational Decision » et « The Logic of Heuristic Decision Making », *Models of Discovery*, Dordrecht, Riedel, 1977, p. 137-175.

6 Dewey, John, « Le schème de l'enquête » (chap. 6), in *Logic: The Theory of Inquiry*, The Later Works, vol. 12, 1938, *The Collected Works of John Dewey*, p.101. Dewey, John, *Logique. La théorie de l'enquête*, trad. G. Deledalle, Paris, PUF, 1967, p. 165-185.

7 Cette « solution » peut s'entendre ici au sens d'une « hypothèse de résolution du problème ». Madelrieux, Stéphane, « La logique de l'enquête expérimentale », *La philosophie de John Dewey*, Paris, Vrin, « Repères », 2016, p. 96-156.

8 Dewey, « Theory of valuation », p. 220.

9 En effet, si un fait provoque chez moi un déplaisir (imaginons une mauvaise odeur), mais si je n'en modifie pas mon comportement en la considérant comme un problème et en cherchant donc une façon de m'en débarrasser, alors elle restera sans valeur à mes yeux. Si je continue mon chemin sans changer mon itinéraire pour éviter cette mauvaise odeur, alors je peux dire que cette odeur m'est désagréable, mais cela ne suffit pas pour donner une mauvaise valeur à celle-ci.

10 Ce terme est entendu ici au sens où nous ne serons pas prêts à mettre un prix sur la chose en question. Nous ne l'estimerons pas, elle n'aura pas de valeur, et nous ne modifierons pas notre comportement pour l'atteindre ou la sauvegarder.

11 Il est en effet pratique d'associer la puissance d'une chose avec le mystère de son origine pour laisser inexpliqués ces deux points. Ce diallèle est présent au sein de nombreuses métaphysiques de la volonté, par exemple.

12 Dewey, « Theory of valuation », p. 218.

13 Dewey, John, « Moral Judgment and knowledge », *in* John Dewey, *Ethics*, USA, Henry Holt and Company, 1936, p. 288-314.

14 Mais une étude approfondie des textes éthiques de Charles Sanders Peirce ou de William James pourrait également partager ces critères. Frega, Roberto, *John Dewey et la philosophie comme épistémologie de la pratique*, chapitre 2, Paris, L'Harmattan, 2006.

15 Sur l'influence du darwinisme sur la pensée de Dewey on pourra consulter Dewey, John, *L'influence de Darwin sur la philosophie et autres essais de philosophie contemporaine*, trad. L. Chataigné Pouteyo, S. Madelrieux et E. Renault, Paris, Gallimard, 2016.

16 Dewey, John, « Expérience et pensée » (chap. 11), *Démocratie et éducation*, suivi de *Expérience et éducation*, intro. D. Meuret et J. Zask, Paris, Armand Colin, 2011 [1916; 1968], p. 223-236.

17 Dewey associe étroitement les notions d'intérêt et de désir jusqu'à les associer à de nombreuses reprises : « La valuation comme *désir-intérêt* », « sans que n'interviennent désir et intention », « l'accomplissement d'un désir ou de la réalisation d'un intérêt ». Dewey, « Theory of valuation », p. 220-221.

18 Ce point est soutenu également par Hans-Georg Gadamer dans sa théorie de l'agir expert. Gadamer, Hans-Georg, *Das Erbe Europas*, Frankfurt am M., Suhrkamp Verlag, 1989.

19 Schön, D., *The Reflective Practitioner*, New-York, Basic Book, 1983, p. 37.

20 Dewey, « Theory of valuation »*, p 224.

21 « Dans toutes les sciences physiques, on considère aujourd'hui que tout "effet" est aussi une "cause" ou, plus exactement, qu'il n'arrive rien qui soit final, c'est-à-dire ne participant pas d'un flux continu d'événements. » *Ibid.*, p. 225.

[22] Dewey, « Le schème de l'enquête », p. 101.

[23] Dewey, « Theory of valuation », p 226.

[24] Albert, Hans, « The problem of foundation » (chap. 1), *Treatise on Critical Reason*, New Jersey, Princeton University Press, 1985, p. 18.

[25] Schön, D., « The Theory of Inquiry: Dewey's Legacy to Education », *Curriculum Inquiry*, no. 22, 1992. Waks, L. J., « Donald Schön's Philosophy of Design and Design Education », *International Journal of Technology and Design Education*, no. 11, p. 37-51.

[26] Schön, D., *The Reflective Practitioner*, p. 94.

[27] Pour Habermas, il s'agit d'un des problèmes des « dilemmes moraux » développés par Lawrence Kohlberg. Habermas, Jürgen, *Morale et Communication. Conscience morale et activité communicationnelle*, trad. C. Bouchindhomme, Paris, Cerf, 1986.

[28] Dewey, « Theory of valuation », p. 228.

[29] Dewey, « Utilité d'une théorie de l'expérience », p. 467.

[30] Dewey, « Les critères de l'expérience », p. 471.

[31] Nous corrigeons ici une faute de traduction de la version française qui écrit « expériences intérieures » là où Dewey écrit « those which have gone before ».

[32] *Ibid.*, p. 472-473.

[33] C'est précisément pourquoi notre conception de la situation ne peut se définir de façon structuraliste ni existentialiste, car il s'agit d'une situation subjective d'expériences propres à chaque individu et à la fois déterminée par les expériences qui précèdent et suivent celle-ci.

[34] Dewey, John, « Intelligence and morals », Essays*, Moral Principles in Education*, The Middle Works of John Dewey, vol. 4, *The Collected Works of John Dewey*, 1907-1909.

[35] Piaget, Jean, « L'évolution intellectuelle entre l'adolescence et l'âge adulte », in *3rd International Convention and Awarding of FONEME prizes 1970*, Milan, 1970.

[36] Cette importance de l'adaptation dans la théorie de Dewey a été comparée à (et rapprochée de) celle de Piaget pour la première fois par Kolb en 1984. Kolb, D. A., *Experiential Learning: Experience as the Source of Learning and Development*, chap. 2, Englewood Cliffs, New Jersey, Prentice-Hall, 1984.

[37] Dewey, John, « The Development of American Pragmatism », *Essays*, The Later Works of John Dewey, vol. 2, *The Collected Works of John Dewey*, 1925-1927.

[38] Dewey, « L'éducation, nécessité biologique », p. 85.

[39] Freitag, Michel, *Le naufrage de l'université et autres essais d'épistémologie politique*, Paris, Éd. La Découverte, 1995. Martin, Éric et Maxime Ouellet, *La gouvernance des universités dans l'économie du savoir*, rapport de recherche, Institut de recherche et d'informations socio économiques (IRIS), Montréal, 2010.

[40] Dewey, John, *Reconstruction in Philosophy*, Boston, Beacon Press, 1959, p. 205. On notera ici la proximité de cette idée avec celle de Kant qui, dans *Vers la paix perpétuelle*, l'applique à la notion de droit public juste. Kant, Emmanuel, *Vers la paix perpétuelle*, trad. J.-F. Poirier, Paris, Flammarion, 1991, p. 124.

[41] Sur ce point, l'éthique développée par Dewey est assez proche de celle d'Habermas. Habermas, Jürgen, *L'éthique de la discussion*, trad. M. Hunyadi, Paris, Cerf, 1992, p. 104.

[42] La qualité démocratique de l'éducation se mesure à « la participation de tout être humain adulte à la formation des valeurs qui régulent le vivre ensemble ». Dewey, John, « Democracy and Educational Administration », *The Later Works of John Dewey,* vol. 11, *The Collected Works of John Dewey*, 1937, p. 217.

[43] Prairat, Eirick, *De la déontologie enseignante*, Paris, PUF, 2014, p. 124-126.

[44] Gouinlock, James, *John Dewey's Philosophy of Value*, New York, Humanities Press, 1972, p. 308.

[45] C'est ce qui fait dire à Judith Green, qui poursuit la réflexion de Dewey dans la philosophie politique (nous traduisons) : « L'idéal démocratique de Dewey n'a pas de contenu fixe […] De la même manière, le contenu des processus éducatifs qui servent de moyens d'atteindre et de mesurer cet idéal doit changer en réponse aux expériences passées et aux nouveaux besoins qui surgissent dans le contexte des situations actuelles. » Green, Judith, *Deep Democracy*, Rowman & Littlefield Press, New York, 1999, p. 62.

[46] Prévenons ici un risque de confusion. Dewey fait partie des philosophes, avec Putnam, qui ne distinguent pas radicalement les normes et les valeurs, contrairement à Habermas. Et c'est justement parce que les valeurs peuvent, autant que les normes, acquérir une objectivité par l'enquête morale, que la confusion entre les normes et les valeurs est possible. Or cette confusion ne doit pas faire perdre de vue que les professeur.es ne doivent pas donner aux valeurs une connotation normative les rendant étrangères à l'expérimentation et à la révision. Habermas, Jürgen, *Droit et démocratie*, section 2, chap. 6, Paris, Gallimard, 1992. Putnam, Hilary, *Fait/Valeur : la fin d'un dogme et autres essais*, Paris, Éditions de l'Éclat, 2002.

[47] Lipman, Matthew, « Renforcer le raisonnement et le jugement par la philosophie », *in* Leleux, Catherine (dir.), *La philosophie pour enfants. Le modèle de Matthew Lipman en discussion*, Bruxelles, de Boeck, 2005 [1989].

[48] Dewey, John, « Creative Democracy – The Task Before Us », *The Later Works*, vol. 14, 1938, *The Collected Works of John Dewey*, p. 229.

[49] Zask, Joëlle, *L'opinion publique et son double*, livre 2, *John Dewey, philosophe du public*, Paris, L'Harmattan, 2000.

DOSSIER

PUBLIC PARTICIPATION, LEGITIMATE POLITICAL DECISIONS, AND CONTROVERSIAL TECHNOLOGIES

XAVIER LANDES
STOCKHOLM SCHOOL OF ECONOMICS IN RIGA

MARTIN MARCHMAN ANDERSEN
TECHNICAL UNIVERSITY OF DENMARK

KLEMENS KAPPEL
UNIVERSITY OF COPENHAGEN

INTRODUCTION

How should a diversity of public opinions and perceptions of controversial policies or technologies regarding, for example, food, health, and medicine should be accommodated or respected in the overall legal, administrative, and political frameworks? What is required to enhance or preserve the democratic legitimacy of such a range of decisions? What sort of public participation should we want or require in designing the legal, administrative, and political frameworks? In particular, what weight should public participation have compared to other requirements of justice and legitimacy?

This special issue concerns the above questions, and is one of the outcomes of a multidisciplinary research project undertaken at the University of Copenhagen titled "Plants for Changing World." The project involved researchers in plant and environmental sciences, pharmacology, law, food and resources economics, and philosophy. This project aimed at underlining the scientific and social challenges raised by a variety of agricultural developments, in particular in plant design. These developments included genetically modified organisms (GMOs) and rewilding of crops (e.g., Marchman et al. 2015, Palmgren et al. 2015). As is familiar, GMOs create a lot of public resistance, often despite contrary scientific evidence that they may be beneficial and involve no significant risk to health or the environment.

At least in the European context, what we might call the public-participation paradigm has been influential in the regulation of GMOs. This is the idea that public participation—as well as some degree of public acceptance—is a precon-

dition for legitimate decision making in the domain of novel plant technologies. Numerous surveys have found a considerable public resistance in all European countries regarding GMOs and their development (European Commission 2010, p. 18), and in part this may explain the costly regulatory system and the lack of will among producers to engage in the development of GMOs.

Yet the suspicion is often voiced that this resistance is illegitimate because it is not sufficiently scientifically informed, or because it is not based on sufficiently cohesive and justified moral values. An additional concern is that one major potential of novel plant technology touches climate change and food supply, issues that mainly affect humans in developing countries and future generations. But it is unclear how benefits to people other than those who are involved in the political decision-making process should be reflected in the public-participation paradigm. So, there is considerable reason to rethink and refine the public-partic-ipation paradigm as it applies to novel plant technology.

This special issue relates to current discussions in political theory about legiti-macy, deliberative democracy, epistemic conceptions of democracy, delibera-tive failures, public understanding of science, and collective decisions in contexts of uncertainty. Moreover, its main focus is the public-participation paradigm (i.e., the idea that public decision making needs citizens' involvement in order to be legitimate). According to the public-participation paradigm, affected indi-viduals should give their qualified consent to a given policy in some sort of deliberative process.

The public-participation paradigm raises all sorts of questions that are central for political theory—for example:

1) What counts as being affected by a given policy? Are you affected by a change in the production system in another country? Are you affected in the relevant sense merely by a product being available on the market? Are you affected if your moral views or religious views are not followed by others? Are you affected only if you could be harmed in a specific way?

2) What counts as a reasonable level of factually correct information upon which views should be based? What about cases in which part of the public debate is about what information is factually correct? Which level of understanding of probabilities should be expected/required from participants in order to count opinions as properly informed? What sort of idealized version of the actual expressed views should be permitted, or required?

3) What sort of consent is required? Actual consent might seem clearly too demanding. A commonly mentioned possibility for defining consent as legitimate in political theory is to consider decisions as being legitimate

when they are supported by reasons that no one can reasonably reject. How should that be interpreted? Is the possibility of contesting or the possibility of being heard enough?

4) How should the views of populations of the developing world and future generations be accommodated in the public-participation paradigm?

This special issue is obviously too short to deal with all these questions. Nonetheless, these questions indicate the richness and the depth of the issues related to the participation of citizens in public decision making, issues that need to be tackled by political theory.

The task undertaken by the authors of this special issue of *Les Ateliers de l'Éthique/The Ethics Forum* is double pronged. On the one hand, it is to discuss the implications of the public-participation paradigm for decision making that bears on scientific activities and advances that are perceived as risky by citizens. On the other hand, it is to question the very role played by consensus and deliberation in contemporary theories of legitimacy.

The first dimension is covered by Andreas Christiansen, Karin Joench-Clausen, and Klemens Kappel's article "Does Controversial Science Call for Public Participation? The Case of GMO Skepticism" and Andreas Christiansen and Björn Gunnar Hallsson's article "Democratic Decision Making and the Psychology of Risk."

In "Does Controversial Science Call for Public Participation? The Case of GMO Skepticism," Christiansen, Joench-Clausen, and Kappel challenge the assumption widely shared in public policy and science communication that public participation could overcome citizens' suspicion towards controversial technological advances. The main justification for what the authors label as the "Public Participation Paradigm" in a narrow sense (i.e., the view positing that controversial science and technology require public participation in the policy-making process) would be to bridge the gap between citizens and experts. However, despite the desirability of such ambition, the authors claim that theories of political legitimacy do not firmly support the paradigm. So, while widely endorsed, the Public Participation Paradigm is actually not well supported in the case of GMOs.

In "Democratic Decision Making and the Psychology of Risk," Christiansen and Hallsson expose the tensions between people's preferences and beliefs, on the one hand, and, on the other, scientific expertise on activities that are perceived by the public as risky. To do so, the authors undertake two tasks. First, they mobilize the resources drawn from psychological studies in order to better understand the reasons why citizens sometimes oppose activities that scientists do not judge particularly risky. Second, they reflect on the lessons that psychology can teach in relation to democratic decision making, in particular with regards to how public policy should answer to citizens' perceptions of risk.

As they explain in detail in the paper,

> for a substantial number of risks, lay opinion is divided along cultural lines. In these cases, agreement with experts is not correlated with scientific literacy or deliberate, careful reasoning—rather the opposite is true. Instead, an individual's beliefs about the riskiness of some phenomenon largely depends on whether that phenomenon is good or bad according to her basic cultural worldview—her basic values. Furthermore, cases where risk debates have become culturally charged are overrepresented among the risks that exhibit the conflict between experts and (some) citizens

A main question of the paper, then, is in what way liberal-democratic decision making should be responsive to values and preferences of citizens that in this way are subject to cultural cognition.

In "Consensus and Liberal Legitimacy: From First to Second Best?," Xavier Landes claims that public participation plays a prominent role in part of the liberal theory of political legitimacy through consensus. Consensus acts as what economists call the "first best"—that is, a set of conditions that, if they cannot be fulfilled, should be nonetheless approximated because they lead to welfare improvement. This centrality of consensus as a first best would be present in the liberal theory as well as in democratic practices. Landes's paper offers a review of the reasons why factual and epistemological disagreements may create second-best issues. He points to the importance of stability in liberal thought. However, the conclusion is that more work is required for "importing" the theorem of the second best into political theory, especially considering the apparent contextualism embodied in the original, economic, formulation of the theorem.

The special issue closes on a contribution titled "New Trouble for Deliberative Democracy," where Robert Talisse takes a critical stance on the deliberative turn experienced by political theory and practices for the last two decades. Deliberative conceptions of democracy explicitly value exchange of reasons among citizens during the elaboration of public policies. According to Talisse, the same factors that facilitate the deliberative turn nurture deliberative pathologies such as group polarization, dialectical fallacies, and deliberativization of democracy.

REFERENCES

European Commission, *Eurobarometer 73.1: Biotechnology*, Brussels, TNS and European Commission, 2010.

Marchman Andersen, Martin, Xavier Landes, Wen Xiang, Artem Anyshchenko, Janus Falhof, Jeppe Thulin Østerberg, Lene Irene Olsen, Anna Kristina Edenbrandt, Suzanne Elizabeth Vedel, Bo Jellesmark Thorsen, Peter Sandøe, Christian Gamborg, Klemens Kappel, and Michael G. Palmgren, "Feasibility of New Breeding Techniques for Organic Farming", *Trends in Plant Science*, vol. 20, no. 7, 2015, p. 426-434.

Palmgren, Michael G., Anna Kristina Edenbrandt, Suzanne Elizabeth Vedel, Martin Marchman Andersen, Xavier Landes, Jeppe Thulin Østerberg, Janus Falhof, Lene Irene Olsen, Søren Brøgger Christensen, Peter Sandøe, Christian Gamborg, Klemens Kappel, Bo Jellesmark Thorsen and Peter Pagh, "Are We Ready for Back-to-Nature Crop Breeding?" *Trends in Plant Science,* vol. 20, no. 3, 2015, p. 155-164.

DOES CONTROVERSIAL SCIENCE CALL FOR PUBLIC PARTICIPATION? THE CASE OF GMO SKEPTICISM

ANDREAS CHRISTIANSEN
POSTDOCTORAL FELLOW, DEPARTMENT OF MEDIA COGNITION AND COMMUNICATION, UNIVERSITY OF COPENHAGEN

KARIN JONCH-CLAUSEN
PH.D. IN PHILOSOPHY, UNIVERSITY OF COPENHAGEN & STUDENT, UNIVERSITY OF WISCONSIN LAW SCHOOL

KLEMENS KAPPEL
UNIVERSITY OF COPENHAGEN

ABSTRACT:

Many instances of new and emerging science and technology are controversial. Although a number of people, including scientific experts, welcome these developments, a considerable skepticism exists among members of the public. The use of genetically modified organisms (GMOs) is a case in point. In science policy and in science communication, it is widely assumed that such controversial science and technology require public participation in the policy-making process. We examine this view, which we call the Public Participation Paradigm, using the case of GMOs as an example. We suggest that a prominent reason behind the call for public participation is the belief that such participation is required for democratic legitimacy. We then show that the most prominent accounts of democratic legitimacy do not, in fact, entail that public participation is required in cases of controversial science in general, or in the case of GMOs in particular.

RÉSUMÉ :

Beaucoup d'avancées scientifiques et de technologies émergentes sont controversées. Bien qu'un certain nombre de personnes, incluant des experts scientifiques, sont favorables à ces développements, la population demeure largement sceptique. Le recours aux organismes génétiquement modifiés (OGM) illustre une telle situation. Dans les politiques et communications scientifiques, il est largement tenu pour acquis que de telles controverses scientifiques et technologiques requièrent la participation publique dans le processus de prise de décision politique. Nous examinons ce point de vue, que nous appelons le paradigme de la participation publique [*Public Participation Paradigm*], en nous servant du cas des OGM. Nous suggérons qu'une raison centrale en faveur de l'appel à la participation publique se situe dans la croyance qu'une telle participation est requise par la légitimité démocratique. Nous montrons ensuite que la plupart des principales conceptions de la légitimité démocratique n'impliquent pas, en fait, que la participation publique puisse être requise pour les controverses scientifiques en général, et pour les OGM en particulier.

1. INTRODUCTION

In science policy and in science communication, it is widely assumed that controversial science requires public participation. In part, the background for this assumption is the experience of widespread public skepticism about the use of genetically modified organisms (GMOs). The initial assumption that public skepticism would be replaced by wide social acceptance once relevant information was provided to the public was quickly shown false. In response to the failings of this earlier approach, the view that extensive public participation both in the design of science policy and in policy-making processes in general was desirable gained popularity. The general idea expressed in writings on public participation is roughly that technology policy should be democratized— meaning that the skeptical public and the enthusiastic scientific community should debate their disagreements about controversial science and technology on an equal footing, and that policy makers should take the results of these joint deliberations into account when making decisions in the realm of controversial science. We call this the *Public Participation Paradigm* (PPP), and explain it more fully below. Legislation governing research and commercial use of GMOs is an area in which PPP would apply. Our aim in this paper is to offer a critical discussion of PPP, particularly of ways in which the paradigm might be defended, with the GMO case as an example. We will argue that, while widely endorsed, the Public Participation Paradigm is not well supported by currently influential theories of political legitimacy, and that this presents a challenge to the paradigm.

2. GMO SKEPTICISM

Public debate as well as empirical research establishes that there is considerable public skepticism (at least in the EU) about the use of genetically modified organisms (GMOs) in food production. This skepticism is mainly fuelled by concerns about unnaturalness (GMOs are considered to be unnatural, or to be produced by unnatural methods, or to be equal to playing God, and therefore to be morally objectionable), risk of adverse consequences (GMOs are perceived as involving unacceptable health risks or environmental risks), and a variety of socioeconomic consequences (risk of exploitation, monopolization, threats to traditional life forms) (see, e.g., Eurobarometer, 2010, pp. 18-32; Gaskell et al., 2010, pp. 36-39; Thompson, 2015, p. 201). For the sake of simplicity, let us refer to this broad class of views as GMO *skepticism*, and to the particular concerns that GMO skepticism is based upon as the *naturalness objection*, the *risk objection*, and the *social justice objection*, respectively.

As is familiar, scientists and some hard-nosed philosophers routinely reject GMO skepticism as irrational and ill informed (see, e.g., American Association for the Advancement of Science, 2012; Comstock, 2000, Ch. 5 & Ch. 6; Holtug, 2009; Thompson, 2011). For example, the naturalness objection is often dismissed as inconclusive and incoherent—it is difficult if not impossible to make satisfactory sense of the basic premise of the argument that something can

be morally bad merely in virtue of being unnatural. The risk objection would be cogent, were it not for the fact that no solid evidence have been forthcoming in the vast scientific literature that any such risk exists (National Academy of Sciences, 2016).[1] The social justice objection asserts that a range of unwanted economic consequences may follow from the use of GMOs in agriculture—for example, that large corporations may acquire monopoly power over central parts of the global food production system. These concerns are entirely reasonable. But, it may be replied, they are not specific to the use of GMOs as such—they apply to many technological innovations in agriculture and beyond (see Thompson 2007, Ch. 8). More importantly, whether unwanted socioeconomic consequences are likely to occur depends entirely on how the surrounding legislation (concerning, e.g., intellectual property rights, competition law, global trade rules, and broader food policy) is designed. As a result, the move from social justice concerns to an outright rejection of GMOs is tenuous. The use of GMOs is consistent with social justice, given the right legislation. And if one is really concerned about improving social justice in agriculture, focusing on GMOs is at best blinkered and at worst a red herring. So the distributive justice objection, while valid, misfires if it is used to support a general rejection of the use of GMOs.

This very brief presentation does not do justice to GMO skepticism or to the arguments against it made by scientists and philosophers. However, our aim is not to discuss GMO skepticism or the several different objections in any detail. Rather, we want to consider what should be done about the seeming conflict playing out between a skeptical public and at least parts of the science community. Given that sections of the public persistently disagree with the relevant experts about the cogency of the objections that give rise to GMO skepticism, how should democratic societies deal with such skepticism?

3. THE PUBLIC PARTICIPATION PARADIGM

A widely held view is that the conflict between scientists and skeptical citizens requires extensive public participation in the contentious parts of the policy-making process. Both public policy using contentious science and the science policy regulating that science can be covered by this demand. This view is what we call the *Public Participation Paradigm* (PPP). The PPP arose largely as a consequence of perceived weaknesses in earlier approaches to the relationship between science and the public. It will therefore be instructive to briefly recount the historical evolution of the PPP before we discuss the view systematically. We rely in the following on a number of accounts by Boerse & de Cock Buning (2012), Gregory & Lock (2008), Lock (2008), Bauer (2009), and Rowe & Frewer (2000).

3.1 The evolution of the Public Participation Paradigm

An initially dominant approach to public skepticism about scientific and technological developments assumed that the problem lay in a lack of public knowledge about the science. In the case of GMOs, this approach seemed

initially plausible: if only the public were better informed about the nature of GMOs, they would appreciate the benefits of GMOs, as well as the lack of genuine disadvantages. If the public were better able to reason about moral issues, then public skepticism would vanish or at least decrease. This position is what is sometimes referred to as the *deficit model*. The assumption in the deficit model is that skepticism towards technological developments is caused by a deficit in knowledge or rationality. However, sociologists have long pointed out that the deficit model does not provide an accurate picture of why public skepticism exists. An early empirical study revealed that increased knowledge of science was correlated with *less* favourable attitudes towards controversial science and technology, although also with *more* favourable attitudes towards noncontroversial technologies and towards science in general (Evans & Durant, 1995). Subsequent research has largely corroborated this conclusion (Allum et al., 2008; Sturgis & Allum, 2004). In the specific case of GMOs, the available empirical studies likewise do not uniformly find a correlation between higher levels of knowledge of GMOs and more positive attitudes towards them, and where studies have found correlations these have been relatively weak (Ahteensuu, 2012). Indeed, there is some evidence suggesting that more knowledge of GMOs is correlated with *more* skeptical attitudes (Eurobarometer, 2010, p. 18-32).[2] In general, the evidence suggests that the relationship between knowledge and attitudes towards controversial technologies is complex and is mediated by other factors, but it is fairly well established that simply improving people's knowledge will not alleviate skepticism towards GMOs or other controversial technologies.

Partly in reaction to the weakness of the deficit model, the view that increased public participation in the decision-making process regarding new science and technology is called for has gained popularity. If lay citizens' views about the desirability of scientific and technological developments are not simply expressions of ignorance or error, it has been argued, then democratic ideals seem to require that they be considered on an equal footing with the views of scientific experts (Durant, 1999; Fisher, 1999). This goes for both judgments about risks and likely consequences of using GMOs, on the one hand, and judgments about what moral views are rationally tenable, on the other. Public participation was suggested early on by prominent social scientists (e.g., Dryzek, 1989; Fischer, 1993; Wynne, 1991; Ziman 1991) and the general public-participation idea has since become widely accepted among scholars and policy makers, at least in Europe (e.g., Irwin & Wynne 1996; Winickoff et al. 2005; Jasanoff 2003, Nowotny 2003; see discussion and more references in Ahteensuu 2012). Summarizing this development, Gregory and Lock (2008, p. 1257) write:

> Ignorant or not, it was argued, the public should have opportunities to engage with the institutions of science in ways that took account of their views; and scientists should have opportunities to engage with the public to listen and learn as well as speak and teach. Where before the science–society issue had been conceptualised as a combative encounter between knowledgeable experts and ignorant lay masses, now, it became a collective exercise of citizenship in a participatory democracy.

The general theme expressed in writings on public participation is roughly that the skeptical public and the enthusiastic scientific community should debate their disagreement about GMOs on an equal footing, and that policy makers should take the results of these joint deliberations into account when making decisions in the realm of controversial science. As indicated, this is what we call the Public Participation Paradigm (PPP). Legislation governing research and commercial use of GM crops is a prime example of an area in which PPP would apply. Our aim in this paper is to offer a critical discussion of PPP, particularly of the ways in which the paradigm might be defended.

3.2 Justifications for The Public Participation Paradigm

The PPP is both widespread and highly influential, but it is not a view that one easily finds stated by proponents in a philosophically precise and detailed way, much less accompanied by a detailed philosophical justification. Most of the arguments for PPP are closely bound up with arguments *against* alternative models of policy-making—especially the deficit model and 'technocratic' approaches such as risk-cost-benefit analysis (Durant, 1999, p. 315; Stirling, 2008, p. 267).[3] In an influential paper, Fiorino (1990) identifies three types of argument against a "technocratic orientation" (arguments that are simultaneously arguments *for* increased inclusion of lay citizens in the policy-makin process): instrumental, substantive, and normative arguments (see also Stirling, 2008, pp. 268-273, and a similar, but slightly different classification in Durning, 1993). On an instrumental conception, public participation is desirable because it is thought to generally promote some independent aim. Commonly cited aims include securing public trust in the institutions that govern new technologies and ensuring public support (or at least acceptance) of a technology. But one might also speculate that public participation is sometimes promoted by technology-skeptic groups *because* it is likely to slow down or stop an unwanted technological development, such as the development and proliferation of GMOs in food production. On a substantive view, public involvement is desirable because it promotes an epistemically more qualified view of possible harms and benefits that may ensue from GMOs, and a better grasp of the normative issues at stake—for example, by including the experiences of those directly affected by the technology.

We will not discuss or question the instrumental or substantive rationales for PPP. Instead, we focus on normative justifications for PPP. In particular, we focus on the most important and widespread normative argument for PPP, namely that democratic legitimacy requires substantial public participation in the policy-making process—or at least that such public participation contributes significantly to the democratic legitimacy of the policies and governance that might eventually be enacted. More specifically, we assume that PPP is associated with two claims about democratic legitimacy of funding and governance of controversial science. First, processes recommended under PPP can contribute significantly to the democratic legitimacy of policies regarding funding and governance of controversial science. Second, ordinary democratic processes will

inevitably leave a significant legitimacy gap as regards the public policies governing controversial science—that is, in the absence of PPP-recommended processes, the policy-making process will be insufficiently legitimate.

The lack of a single precise and detailed defence of PPP on democratic grounds means that we must rely on what we take to be a reasonable rendering of PPP. The lack of precise definitions and justifications of PPP unavoidably gives our discussion a partially hypothetical character. We will systematically consider various ways in which PPP, as we conceive it, may be justified, and we hope that our systematic examination of a range of possibilities will be helpful. For the purposes of the following discussion, we will characterize PPP as follows:

> *The Public Participation Paradigm* (PPP): Democratically legitimate public policy and governance of controversial science requires special deliberative processes (in addition to ordinary democratic processes), featuring representatives from the public and from science, debating factual and normative issues on an equal footing, and issuing policy recommendations that should have significant normative weight for policy-makers.

A renowned, and in many ways paradigmatic, instance of a process licensed by PPP is the consensus conference (Joss & Durant, 1995). A consensus conference roughly proceeds as follows (our account here relies on Zurita, 2006; see also Andersen & Jæger, 1999; Nielsen, Lassen & Sandøe, 2007). A panel of experts and a panel of lay people get together over (typically) three or four days to debate some potentially controversial technology. The panel of lay people is chosen to be representative of the general population in terms age, gender, education, etc. Initially, the expert panel presents the scientific facts about the technology to the lay panel (which has been briefed so as to be able to ask experts for information they deem important). After this presentation and questioning, the lay panel debates, and drafts a recommendation for how the technology should be developed and regulated. This draft is reviewed by the experts to clear up factual errors and misunderstandings, and it is then presented to parliament in an official report. The aim of these conferences is thus to inform legislators of "the attitudes, hopes and concerns of the public" (Zurita, 2006, p. 20).

Another domain where PPP has been and is influential is science com-munication. Recent theorists propose that science communication should not be viewed as a one-way street, where the public receives information about the progress and potentials of science. Rather, science communication should essentially be conceived as a conversation between members of the public and members of the science community, where these are considered equals. So, according to this view, science communication is a vehicle for special participatory processes involving science when it is controversial, and it is reasonable to assume that this idea could also be motivated by PPP (see Bauer 2009 and Gregory & Lock 2008 for reviews of the various stages of thinking of the point and purpose of science communication).

As we prefer to think of it, PPP does not plausibly assume that designated public participatory processes such as consensus conferences or citizen hearings should *replace* ordinary democratic processes of public deliberation and democratic decision making. Rather, they should supplement those processes. Similarly, the idea is not, of course, that the outcome of designated public-participation processes should replace the formal authority of the parliamentary system to enact laws, etc.—consensus conferences should not supplant parliaments. Rather, consensus conferences and other ways of eliciting public opinion would represent a normatively significant input to policy makers in parliament and government. These decision makers would be morally at fault, or the decision procedures would be morally wanting, if they did not attach some significant weight to the outcome of such processes when formulating policy, though of course decision makers are not legally bound to respect the decisions of a consensus conference.

4. SHOULD WE ACCEPT THE PUBLIC PARTICIPATION PARADIGM?

While the Public Participation Paradigm has been widely accepted, our aim here is to offer a critical discussion of its justification—in particular, a discussion of whether PPP is normatively (that is, democratically) justified. We argue that, despite its popularity, it is actually difficult to mount a defence of PPP based on the most influential current views on democratic legitimacy. This conclusion at least presents a challenge to PPP and its proponents.

When we talk about democratic legitimacy here and below, we use a notion of legitimacy that is common in a broadly Rawlsian tradition (more about Rawls below). As Rawls observes, contemporary liberal societies inevitably feature a plurality of irreconcilable and comprehensive doctrines that underlie disagreements about which coercive policies are just, right, or best (Rawls, 1993). When we cannot reach agreement on what policies are just, right, or best, we should look for policies that are at least legitimate. Legitimate policies in a sense serve to reconcile our conflicting views about what would be right or just—the hope is that it is within our reach to come to agree that a particular policy is legitimate, even when we continue to disagree about whether it is just or right, or the best policy on the issue. Furthermore, acknowledging that a particular policy is legitimate should command some sort of respect, even among those who do not agree that it is the best or the right policy. It is this familiar concept of legitimacy that we assume here (see Estlund, 2008; Nagel, 1987; and Peter, 2009 for similar ways of conceiving of legitimacy).[4]

As we shall see in a moment, most theories of political legitimacy, including the theories that we shall be concerned with, assume that political legitimacy, in one way or another, requires that policies reflect the wills or preferences of those governed. Legitimate policies are those that win some form of qualified consent from those affected, even from those who think that the policies are not right, just, or best. This concept of legitimacy must be distinct from the concepts of rightness, justness, and optimality: since legitimacy is intended to serve what

Wall (2002, p. 387) calls a "reconciling function" in societies characterized by reasonable disagreement about what the right, just, or best policy is, it cannot be a requirement of legitimacy that policies be as a matter of fact right, just, or optimal. Note also that the fact that legitimacy requires *qualified* consent means that it is different from actual public support. A policy may have broad public support, and yet not be legitimate—say, if the public support has been garnered in deceptive and manipulative ways, or is based on false assumptions.

4.1 The Principle of Legitimacy

We will now consider various ways in which one might justify PPP in more detail. Our approach will be to review the most influential theories of democratic legitimacy, and then examine whether they support PPP. We argue that they do not. While PPP is quite popular, none of the currently influential theories of legitimacy supports PPP, at least as far as the case of the GMO controversy is concerned. If correct, this represents an important challenge to PPP.

The broad class of theories of democratic legitimacy that we will be concerned with endorse the following (Rawls 1993, p. 137, Waldron 1987, p. 128, Nagel 1991, p. 3, Gaus and Vallier 2009, p. 53).

> *The Principle of Legitimacy*. Policies are legitimate if and only if everyone affected could reasonably accept the policies (or could not reasonably reject them).

While PPP is not entailed by the Principle of Legitimacy, it may certainly seem that the Principle of Legitimacy supports PPP. However, closer inspection reveals that this is not so. The Principle of Legitimacy needs to be qualified in a number of ways to be plausible (or even interpretable). We argue that once we consider these qualifications, it turns out that the Principle of Legitimacy does not support PPP in the case of the GMO controversy.

Crucially, the qualifications needed for the Principle of Legitimacy concern the proper interpretation of the phrase 'reasonably accept.' Clearly, this clause cannot simply mean acceptance, as the principle would then say that policies are legitimate if and only if they are accepted by everyone involved. This would imply that policies that are rejected by some for reasons of blatant irrationality would count as not democratically legitimate. As one proponent puts it, "democratic governments should respond to people's values, not to their blunders" (Sunstein 2005, p. 126). Thus, one obvious, and familiar, way of interpreting the Principle of Legitimacy is in terms of hypothetical acceptance. On this view what matters for legitimacy is hypothetical acceptance, not actual acceptance. So, reasonable acceptance in the Principle of Legitimacy denotes just a specific form of hypothetical acceptance. Let us say that a subject S hypothetically accepts (in this, reasonable-acceptance sense) a policy if and only if S would have accepted it were S fully rational and fully informed about the relevant facts.

At first sight, it might seem that the hypothetical acceptance interpretation of the Principle of Legitimacy is congenial to PPP: the very idea is that lay citizens will be involved in a process of deliberation in which they will form views that are coherent and informed by the facts. The resulting views will thus be more rational and epistemically qualified. However, it is questionable whether PPP, when defended by the hypothetical acceptance interpretation of the Principle of Legitimacy, will do justice to the public skepticism we actually find: if the naturalness objections are hard to make sense of, and if worries about adverse effects of GMOs are misinformed, then we should infer that people's hypothetical (reasonable) judgments about policies should not be based on these stances. The hypothetical acceptance interpretation thus restricts the scope of PPP, since it does not justify the accommodation of citizens' actual views, "ignorant or not" (Gregory and Lock, 2008, p. 1257). The hypothetical acceptance interpretation of the Principle of Legitimacy would thus not support PPP as a vehicle of expression of at least some socially important forms of public skepticism (such as GMO skepticism). Consequently, the ideal that lay citizens' apprehensions and scientists' enthusiasm about GMOs should be treated as being of equal validity could not be upheld; and the very fact that GMOs are controversial would not in itself provide a reason for public participation. Furthermore, if PPP were to be a proxy of hypothetical acceptance, we should take steps to ensure that PPP approximates the ideals of full rationality and full informedness. This seems to reintroduce the view that the deficit model invited, namely that citizens need to be educated—only the means for education have changed. This is exactly the view that proponents of PPP do not want. It might be suggested that deliberation in diverse groups carries epistemic benefits that would make such education superfluous. The suggestion would be that when citizens deliberate with one another in diverse groups, they tend to give up unnaturalness objections and the risk argument, making PPP a good proxy for hypothetical acceptance. Again, this would turn the justification of PPP into a version of the deficit model.

None of this is to say that a version of PPP that self-consciously aims to determine the possible reasonable views on a controversial technology is not possible or desirable. But it seems to us, at least, that hypothetical acceptance interpretation of PPP would not do justice to the strong demands for equality between experts and lay citizens that is a frequent part of defences of PPP. Furthermore, if this is the aim, then that should be made explicit, and efforts should be made to design participatory institutions in a way that is conducive to it.

4.2 Rawls's political liberalism

Let us return now to the Principle of Legitimacy. A different interpretation of the Principle of Legitimacy is due to Rawls, and put forward in his influential book *Political Liberalism*. The basic idea is that democratically legitimate policies should reflect our shared commitment to the basic values of liberal democracy (values that are related to respecting one another as free and equal persons).

Rawls writes: "Our exercise of political power is proper only when we sincerely believe the reasons we offer for our political action may reasonably be accepted by other citizens as a justification of those actions" (Rawls 1993, p. xlvi). We interpret Rawls as proposing the following restraint on 'reasonable acceptance' in the interpretation of the Principle of Legitimacy: as citizens, we can offer only reasons that we reasonably and sincerely believe that other reasonable citizens can accept. In our justification of coercive policies, we cannot offer reasons that we know that other reasonable individuals do not accept. Public reasons are the set of reasons that meet this requirement—that is, the reasons that are shared by or acceptable to all reasonable comprehensive views. The set of public reasons is thus the intersection of the sets of reasons that are acceptable from each reasonable comprehensive view. A comprehensive view is a view that covers all or at least a wide range of metaphysical, religious, and moral questions. A comprehensive view is reasonable, according to Rawls, if it arises from the conscientious and correct application of theoretical and practical reason, and supports the liberal political conceptions. Since many difficult and subtle judgments go into formulating a comprehensive doctrine, we cannot expect all who use reason correctly and conscientiously to end up affirming the same comprehensive doctrines. Since we differ in our reasonable comprehensive views, not all parts of comprehensive views are part of the set of public reasons. So the restraint on public reason proposed by Rawls implies that some citizens will be prevented from expressing parts of their comprehensive views when seeking to justify coercive public policy—they are bound to restrain themselves to the parts of their comprehensive views that overlap with the comprehensive views of others.

Rawls also proposes a restriction in scope of the Principle of Legitimacy in that he suggests that it applies only to constitutional essentials and matters of basic justice. Arguably, legislation and governance of GM crops are not matters of constitutional essentials or basic justice. Some theorists have argued that it is difficult to justify restricting the application of the Principle of Legitimacy to cover only constitutional essentials—coercive legislation that does not belong to constitutional essentials may, it has been argued, be equally in need of public justification, and the Principle of Legitimacy should therefore be expanded to cover all legislative decisions (Quong 2011, p. 273-289). Others have pointed out that the distinction between constitutional essentials and ordinary legislation is impossible to maintain, since almost any piece of ordinary legislation could have some bearing on constitutional essentials (Habermas 2008, p. 123 fn.18; Greenawalt 1995, p. 1306-1308).[5] We will set this issue aside without taking any stance on it, but for the sake of argument we will simply ask what the implications would be if GMO legislation were required to be justifiable by Rawlsian public reasons. In particular, we ask whether PPP would be supported by such a requirement.

The implication of Rawls's view on public reason is that certain parts of our comprehensive views are excluded from public reason, no matter how sincerely held. Coercive policies must be justified by reasons shared by all. Our wider

normative, religious, and metaphysical commitments, which form our comprehensive views, contain elements that are not part of the set of public reasons, in so far as they are not shared by all reasonable citizens. These are sectarian or private views that are to be excluded from public reason. The view that GMOs are unnatural and therefore morally problematic would very likely not be part of public reason for Rawls, and so the naturalness objection cannot be part of what justifies public policies.[6]

Furthermore, Rawls assumes that public reason must include methods and results from science. As he writes, citizens "are to appeal only to presently accepted general beliefs and forms of reasoning found in common sense, and the methods and conclusions of science when these are not controversial" (Rawls 1993, p. 224). It is not clear from Rawls's writings how this criterion is to be interpreted. In particular, he does not say whether he has narrow intrascientific controversies in mind, or broader controversies involving established science on one hand, and a skeptical public on the other (Jønch-Clausen & Kappel, 2016). Clearly, according to the first narrow interpretation, only factual assumptions that are consistent with the methods and conclusions of science that are not controversial within the relevant scientific communities are part of public reason. This view would imply that views based on ignorance about the scientific consensus on the evidence of the safety of GMOs would be ruled out as parts of public reason, and on this interpretation PPP would be problematic. On the broader interpretation, methods and conclusions of science are excluded from public reason when they are subject to broader controversies, even if there is a consensus in the scientific community. This is an underdiscussed problem in the literature on public reason. The immediate implication of the broad interpretation would seem to be that both the view that GMOs are safe and the view that they are risky would be excluded from public reason.[7] So any policy or governance of GMOs would have to be based on other public reasons, or on shared reasons regarding decision making when a product's safety status is uncertain (if such are available).

So PPP is not warranted by appeals to a Rawlsian understanding of legitimacy as we have interpreted it. Quite the contrary, in fact: policy makers are obligated to disregard many of the views that are currently widely held by citizens, since these are not acceptable to some reasonable citizens.

Not everyone agrees that a plausible interpretation of Rawls on public reason implies that the naturalness objection should be excluded from public reason. In their instructive paper, Streiffer and Hedemann (2005) cite a number of influential authors (e.g., Comstock, 2000; Rollin, 1995; Thompson 2007 [1997]) who in various ways argue that what Streiffer and Hedemann refer to as intrinsic objections to GMOs have no role in public justification of policy since they are reasonably rejectable by reasonable comprehensive doctrines—that is, they are not part of the set of public reasons as we have defined it above. Intrinsic objections are objections pertaining to the very act of genetically modifying organisms, rather than to adverse effects. The most common intrinsic objections are

the naturalness objection and its theological cousin, the playing-God objection. But, Streiffer and Hedemann argue, since views about nature, naturalness, and even quasi-religious reasons for GMO skepticism are themselves part of some reasonable comprehensive doctrines, they cannot be excluded from public reason. Thus, excluding them as legitimate reasons for having certain regulatory frameworks and approval procedures would be illegitimate.

We will now consider this argument. In a crucial passage on the nature of public reason, Streiffer and Hedemann write:

> A reason for a political decision may reasonably be accepted by others as a justification for that decision only if it is consistent with those citizens' reasonable comprehensive doctrines. Thus, the principle of legitimacy says that legitimate political decisions must be justifiable in terms that are consistent with the reasonable comprehensive doctrines of the citizens governed by them. (2005, p. 196)

This sounds very much like our exposition of Rawls's theory of public reason above. However, this, and like passages, lead Streiffer and Hedemann to advance an argument which we suggest can be rendered as follows:[8]

(1) Public reason is the set of reasons that are consistent with all reasonable comprehensive views.

(2) Intrinsic objections (in particular the unnaturalness objection and the playing-God objection) are part of some reasonable comprehensive views.

(3) Therefore, the rejection of intrinsic objections is inconsistent with some reasonable comprehensive views.

(4) Therefore, the rejection of intrinsic objections is not part of public reason.

(5) So the naturalness objection and the playing-God objection cannot be legitimately excluded from public reason.[9]

So, Streiffer and Hedemann conclude, the intrinsic objections, or the views and values underlying them, should be "viewed as constraints on acceptable justifications for public policy about GE food" (2005, p. 206). And, they suggest, "the intrinsic objections will need to be weighed against other morally relevant factors, such as possible beneficial or harmful consequences" (p. 207). Moreover, rather than rejecting that intrinsic objections can be accorded any role in the justification of public policies, we should promote "a shift towards more majoritarian decision-making procedures, and, more broadly, a shift towards taking into account the level of public support for a given policy option" (*ibid.*), which would give more appropriate weight to intrinsic objections, and to citizens

holding these views. Finally, we should seek to accommodate intrinsic objections "without compromising other important political values." In the case of GMOs used in food, this points towards a compromise "allowing GE food, but with mandatory labeling" (p. 208).

We contend that this argument is based on an unusual and untenable interpretation of Rawls's thinking on public reason. Recall that, on the Rawlsian view, policies must be justified by reference only to reasons that are shared among all reasonable comprehensive views. The set of public reasons is, as we said above, the *intersection* of the sets of reasons that each reasonable comprehensive doctrine affirms. If—as we will assume, and as Streiffer & Hedemann assume— neither the intrinsic objections nor their rejection is affirmed by every reasonable comprehensive view, then neither the objections nor their rejection falls into this intersection. But the consequence of this, on a Rawlsian view, is that intrinsic objections must be set aside as not part of public reason, as must the assumption that intrinsic objections are misguided or mistaken. *Neither* can figure in the justification of public policies. Thus when Streiffer & Hedemann move from (4) to (5), they are misinterpreting (or they go beyond) the Rawlsian view. Consequently, it is not true that the intrinsic objections are constraints on acceptable justifications. What remains as reasons that may justify public policy is the secular (nonmetaphysical and nonreligious) concerns that we all agree about— e.g., concerns for liberty, equality, well-being, and risk and harms to humans and animals. As we pointed out above, this would exclude significant parts of what actually motivates public skepticism about GMOs in food production, and so would not support PPP in the case of public skepticism about GMOs.

4.3. Inclusive public-reason theories

Rawls's narrow conception of public reason has been criticized, in particular on the grounds that it prevents religious citizens from publicly offering the reasons that they find most important to themselves (Greenawalt 1995, Wolterstorff 1997, Eberle 2002, Perry 1988, Weithman 2002). Those attracted to the naturalness objection may voice the same complaint. Rawls's restrictive view of public reasons implies that the naturalness objection cannot justify public policies, and yet this very objection may be an important reason for some citizens to require restrictive legislation concerning GM crops.

It is worth considering public-reason theories that are more inclusive in that they reject Rawlsian restraints on public reason. An influential such theory is proposed by Gaus and Vallier (Gaus and Vallier 2009; see also Gaus 2011, Vallier 2011). According to Gaus and Vallier, coercive legislation is legitimate only if it can win the assent of everyone concerned, or be justified for everyone concerned. This is, of course, in broad agreement with the liberal tradition. However, what matters for assent and justification is convergence, not consensus. So citizens need not agree on the same reasons or rationales for coercive legislation. Instead, each of us must accept the legislation for his or her own reasons (i.e., we must converge on the policy outcome).[10] Hence, according to

Gaus and Vallier, we should reject the Rawlsian constraint that one can offer only reasons that one believes that other reasonable citizens also affirm. Any reason can be offered in the process of deliberation and justification—the only minimal requirement is that a reason being proposed be comprehensible as a reason to others. It is incompatible with respect for one's fellow citizens to offer reasons that are not comprehensible to them as reasons.

This should be understood on the basis of what Gaus and Vallier call the Liberty Principle, stating that "liberty should be the norm, [respect for persons as free and equal requires that] coercion always needs some special justification" (Gaus and Vallier, 2009, p. 53; Gaus and Vallier take this formulation from Joel Feinberg (1987, p. 9) and the square brackets are theirs). So the basic idea is that an absence of coercive legislation is legitimate by default, and that any deviation from this state in the form of coercive legislation requires the consent of everyone. There is, as one might say, a presumption in favour of liberty.

Clearly, this view implies that the naturalness objection is admissible as a part of public reason, as are religious objections asserting that genetic modification is objectionable because it amounts to playing God. Yet, the view does not immediately imply that a restrictive legislation and governance of GM is justified. In general, one can propose justifications for coercive legislation that are not shared by everyone. However, one should refrain from proposing coercive legislation that one is convinced cannot win the assent of everyone. Since some citizens accept only secular reasons, all coercive legislation must be justifiable by secular reasons, even if there are also citizens who accept this legislation purely for nonsecular reasons (2009, p. 63). It is clear how this carries over to the case of GMOs. Some citizens do not accept the naturalness objection (or the playing-God objection), which means that a restrictive governance and legislation of GMOs cannot be based solely on those objections; any restrictive legislation must be fully justifiable on nonmetaphysical and nonreligious grounds.

What then about the reverse situation? Suppose that a nonrestrictive legislation of GMOs is acceptable to those who are unmoved by the naturalness objection and the playing-God objection, but is rejected by those who find those objections convincing? Gaus and Vallier note that their view "implies that religious citizens must not have laws imposed upon them which they have no conclusive reason to accept. Even if a secular rationale is necessary in our society for a publicly justified law, it can be defeated by a reasonable religious conviction without any secular backing" (*Ibid*). Again, it is clear how it applies to the present context. Assume that a nonrestrictive governance and legislation of GMOs would be unacceptable to those who endorse the naturalness objection. This opens up the possibility that, on Gaus and Vallier's view, a permissive governance and legislation on GMOs could be rejected on the ground that proponents of the naturalness objection reject it. The basic idea proposed by Gaus and Vallier is that there is a fundamental asymmetry between the justifications needed for accepting a coercive legislation and those needed for rejecting it. One cannot

endorse a coercive legislation while admitting that it can be given only a religious justification. However, one can reject a coercive legislation, even if the rejection can be given only a religious reason, and no secular reason.

While Gaus and Vallier's view, as we have seen, is much more welcoming to reasons such as the naturalness objection and the playing-God objection than Rawls's view, it is still not clear how much support it lends to PPP. The crucial question for Gaus and Vallier will be whether a nonrestrictive governance of GMOs is *coercive* of those who endorse the naturalness objection or the playing-God objection (and perhaps the risk objection) in the sense that a nonrestrictive governance fails to respect their integrity and freedom of conscience. Consider an analogous case, a legislation that permits same-sex marriage by simply failing to prohibit it. This legislation is nonrestrictive in an obvious way. Consider Alf, who is not the least interested in same-sex marriage himself, but who objects to same-sex marriages on religious grounds, or because he views same-sex marriages as highly risky, though this is unsubstantiated by current evidence. Clearly, a nonrestrictive legislation regarding same-sex marriage would affect Alf's life—if the legislation is passed, he will be forced to live in a society where same-sex marriage is legally recognized, or emigrate to another country. However, can this legislation be considered coercive towards Alf in a way that defeats its justification as it fails to respect Alf's integrity and freedom of conscience?

Returning to the GMO case, the crucial question that we must answer in order to determine whether the naturalness objection is sufficient to block permissive GMO legislation is thus the following: Are permissive policies concerning GMOs coercive towards those who hold the naturalness objection? This question—of whether a nonrestrictive governance of GM is unduly disrespectful of some citizens—is orthogonal to the question of whether PPP is a good idea or not, and to any outcome that processes licensed under PPP may have. This renders special measures like PPP less relevant for determining the legitimacy (or contributing to the legitimacy) of policies.[11]

Note that among their many interesting observations, Gaus and Vallier point out that the justification of any policy one might propose depends on the reasons that other citizens have. However, "we do not know what reasons others have in large and complex societies. We have to discover what reasons people have" (p. 67). So to enable us to consider what coercive legislation is justified, we need institutions that broadcast the views of individuals for everyone to hear. Of course, central parts of public debate and the political system serve just that function. One can imagine that designated processes licensed under PPP could be seen as an important part of the set of institutions that serve to make publicly accessible what reasons for or against certain policy proposals individuals have. But note two things: First, this justification of PPP does not directly speak to the legitimizing function of processes licensed under PPP, but rather to their epistemic merits; basically, the processes serve to clarify what objections to GMOs exist in a polity. Second, given this rationale for PPP, such processes would make the most sense if we did not already know what objections there are, of if we had no other better way of finding this out.

4.4. Procedural views on democratic legitimacy

The views of legitimacy discussed so far have assumed that legitimacy can be defined in terms of what affected parties would accept, given certain levels of idealization, or given a more or less restricted set of public reasons they could appeal to. We have argued that none of these theories clearly supports PPP. However, an important strand of theories of democratic legitimacy is proceduralist, holding that democratically legitimate policies are those that are adopted by the right kind of decision procedures. This typically means some form of public deliberation followed by a majoritarian voting procedure, where this may include provisions about fair and free elections and reasonably equal access to the policy-making process by different interest groups, and so on (see Peter, 2008).

As may be familiar, some proceduralist theories are *purely* procedural, in that they deny that there is any procedure-independent criterion of legitimacy. The democratically legitimate (or correct) outcome is simply defined in terms of wherever the right procedure takes us. Other theories deny the assumption that there is no procedure-independent criterion of rightness, and advocate for proceduralism on partly epistemic grounds: impure procedural theories insist that there is a right outcome, and that deliberation and other democratic processes are our preferred means for tracking it (Estlund 2008). A combined theory argues that we don't avail ourselves of views about the right outcome, but we nonetheless impose epistemic constraints on what counts as the proper procedure (Peter, 2009).

Advocates of both pure and impure proceduralist views typically impose some requirement of epistemic competence of deliberators and impose epistemic constraints on what counts as the proper procedure. Deliberators need to take facts and reasons seriously (or act as if they do), and need to adjust their views accordingly. Views that blatantly ignore facts, are deeply incoherent, or are maintained in the face of contravening evidence and argument would, it is hoped, tend to be ignored by other deliberators and therefore not have much weight in the procedure. But such views should in any case in principle be disqualified from the deliberative procedure, according to most proceduralist views.

Would PPP be supported by proceduralist views of democratic legitimacy? On the one hand, it would seem that appropriately defined public participation could be part of the procedure that produces legitimacy, though it would of course be implausible to give too much weight to what goes on in such a participatory forum, and even less plausible to give such forums direct legislative power. On the other hand, proceduralism does not support the idea that public participation of the sort licensed under PPP—a procedure that goes beyond ordinary processes of democratic deliberation and legislation—would be a *necessary* additional legitimacy-conferring activity in the case of GMOs. There seems no reason why a proceduralist view could not say that legislation and governance of GMOs could be fully democratically legitimate merely by being the outcome

of the ordinary deliberative procedures. At any rate, a further argument is needed to show why legitimate decision making in the case of GMOs—and with regards to controversial science and technology more generally—requires direct public participation when (if) legitimate decision making concerning other controversial issues, such as tax or educational policy, do not.[12]

Moreover, certain features of procedural theories of legitimacy may be in tension with how PPP has often been implemented. Processes proposed under PPP carry the risk of employing an arbitrary or biased selection of a small number of participants, who are not elected in a fair process, and who do not represent interests or views in the political constituency in a systematic way. So, a PPP justified on proceduralist grounds must include a requirement that certain selection procedures be implemented—procedures that may be very different from what has characterized the kinds of mechanisms for participation that have been implemented.[13]

There is another proceduralist defense of PPP that appears much more plausible. This is the suggestion that PPP is warranted as a remedy for imperfections in the processes of public deliberation or democratic decision making. The idea is that what justifies PPP would be the fact that public deliberation or other parts of democratic decision making have somehow gone wrong it the case of GMOs, and that the remedy needed is the special processes devised under the PPP, such as consensus conferences or citizen hearings. Likewise, sociological surveys of the nature of skeptical public opinion may provide a necessary input to legitimate political decision making on the assumption that these views have somehow been suppressed, or insufficiently represented or efficacious in the democratic processes.

To assess this option, consider first the ways in which a deliberative process in a polity can be imperfect or distorted. There seem to be four (somewhat overlapping) general ways. First, certain individuals or groups may be excluded from participating—say, because they are denied the possibility of speaking, or because they have a more costly or difficult access to the venues where public deliberation occurs. Second, a deliberative process may be distorted if it excludes certain otherwise legitimate views from being expressed. This can happen, for example, if mainstream media are reluctant to report these views, or if powerful organizations and political parties are unwilling to represent them. A third way concerns the effects of admitting illegitimate views that deny basic liberal rights or equal standing of other citizens, such as racist or misogynist views, in the process of deliberation. Admitting such views in public deliberation may have chilling effects, making democratic participation more difficult for those who are targeted. Moreover, debating and contravening illegitimate views may divert attention from the real issues, and the representation of illegitimate views may have a distorting effect on what participants consider as reasonable political compromises (e.g., the presence of illegitimate hateful xenophobic rhetoric in the public debate may make us willing to accept compromises in our policies regarding refugees and migrants that we would otherwise not accept). Fourth,

democratic deliberation may be skewed or distorted when participants simply fail to respond rationally to evidence and arguments that are available or that are provided by other participants.

These all represent familiar and potentially serious ways in which public deliberation may be defective or may need amendment or corrective measures. As we said, one might try to defend PPP by appealing to one or more of these defects of deliberative processes, or other similar defects. It is far from obvious, however, that this can succeed in the case of GMOs. It is not credible that European debates over GMOs have been affected to a significant extent by any of the above. No one has been excluded from voicing his or her view about GMOs. No legitimate view about GMOs has been excluded, and this is true also of views critical of GMOs. It might certainly be true that mainstream scientific organizations of various sorts, including government organizations, have been unwilling to accept and propagate GMO-critical views. However, many other organizations, including the mainstream media, have provided ample space for these views.

Perhaps proponents of the naturalness objection (and its theological cousin, the playing-God argument) have felt that others were trying to exclude their arguments from serious consideration. Certainly, the naturalness argument has been met with resistance in academia and, one can imagine, in public debate as well. However, there is a charitable interpretation of such aversion to the naturalness objection. Those rejecting the naturalness objection can (and should) generally recognize that people are fully entitled to hold the view that GMOs are unnatural and therefore morally problematic, and to express that view in public. However, if one accepts a Rawlsian view on public reason, one can nonetheless insist that the naturalness objection is not suitable for justifying public policies, and that it should therefore be bracketed in those specific contexts. Moreover, even if one accepts that the naturalness objection is an eligible view, one can still object to it as incoherent and therefore as not rationally convincing, as we mentioned above. There is a difference between excluding a view or refusing to take it seriously, on the one hand, and not being rationally convinced by it, on the other. GMO skepticism and the grounds upon which it is based have been voiced and heard, but not everyone has been convinced. If anything, GMO skeptics could be blamed for having failed to respond rationally to evidence about the safety of GMOs. This could be considered a failure of the democratic deliberative process (and many do consider it as such), but it was not the one that proponents of PPP had in mind for correction. In conclusion, we suggest that PPP, as we perceive it, cannot be defended as a measure needed for correcting a failure of deliberative processes in the case of GMOs.

5. CONCLUDING REMARKS AND FURTHER REFLECTIONS

PPP endorses the uses of designated participatory procedures as a response to widespread GMO skepticism among citizens. However sympathetic this idea is, we have presented a challenge to it: it seems that the most influential and commonly discussed theories of democratic legitimacy do not support PPP in the

case of the GMO controversy. Designated public-participation processes about controversial issues might, in some instances, even be in tension with tenets of liberal democracy by violating restrictions on public reason, or by suggesting that coercive legislation limiting the use of GMOs can be publicly justified when it cannot. In some cases, such processes may effectively give certain views too much space and influence, thereby arguably detracting from the overall legitimacy of the democratic processes. We mentioned rationales for PPP that go beyond those having to do with democratic legitimacy. Two other general aims that could be furthered by designated public-participation processes are worth considering: first, the aim of increasing public support or acceptance of public policy and governance of controversial science and, second, the aim of providing an epistemically more qualified view of possible harms and benefits. The extent to which designated public-participation processes further these aims is, of course, variable. If what we have argued is correct, there might be a price to pay for invoking such processes, in the form of decreased democratic legitimacy, but there is no general answer to the question of whether these other benefits of such processes may be worth the price.

NOTES

[1] Inevitably, some disagree that the absence of evidence for adverse effects warrants the claim that GMOs do not have such effects. Critics typically argue that we cannot conclude that GMOs do not have harmful effects based on the fact that no evidence of such effects has been found, given the scope of and methods employed in the studies that have as yet been carried out (see, e.g., Hilbeck et al., 2015).

[2] In the study cited, the only proxy for knowledge is bare *awareness of* GMOs—i.e., an affirmative answer to the question "Have you ever heard of genetically modified (or GM) foods before?" (Eurobarometer, 2010, p. 13, n. 13). Such awareness is correlated with answers to questions concerning the benefits, risk, naturalness etc. of GMOs that are negative towards them (e.g., an affirmative answer to the question "GM food is not good for you and your familiy").

[3] There are likely other contributing causes of this lack of a clear and precise defence of PPP. First, many proponents of PPP are found among social scientists and administrative officials, and the cultures of these fields often include a substantial aversion to overtly normative argumentation. Second, the fact that that proponents typically conceive of themselves as doing empirical research gives rise to a tendency to frame PPP as a historical fact that is a consequence of how modern societies are structured. And third, it seems to be an unquestioned assumption among many writing on public participation (as it is in society generally) that democratic values require as much direct public involvement as possible—i.e., that PPP is 'more democratic' than alternative procedures.

[4] This section borrows from Kappel, 2017. See this paper for a further elaboration of the concept of legitimacy and how it relates to fact-dependent policy disagreements.

[5] Note that Rawls (2001, p. 91) acknowledges that many 'ordinary' political issues touch upon constitutional essentials and basic justice, including "policies to protect the environment and control pollution." However, he suggests that "the restrictions imposed by public reason do not apply to them, or if they do, at least not in the same way or so stringently." It is also worth noting that Rawls is in some places open to the possibility that issues not pertaining to constitutional essentials or basic justice should ideally be decided on the basis of public reasons. Thus he writes (1993, p. 215): "Some will ask: why not say that all questions in regard to which citizens exercise their final and coercive power over one another are subject to public reason? Why would it ever be admissible to go outside the range of public reasons? To answer: my aim is to consider first the strongest case where the political questions concern the most fundamental matters. If we should not honor the limits of public reason here, it would seem we need not honor them anywhere. Should they hold here, we can then proceed to other cases. Still, I grant that it is usually highly desirable to settle political questions by invoking the values of public reason. Yet this may not always be so."

[6] Connoisseurs of Rawls's writings may object that the requirement that policies be justified only by public reasons does not apply to ordinary citizens in ordinary public debate. However, the public-reason requirement *does* apply to ordinary citizens when they are engaged in political acts such as voting (Rawls 1997, p. 769). To the extent that consensus conferences are to have some weight in legislation, it is plausible to see citizens participating in them as engaged in such a political act, and thus under a duty to offer public reasons. Furthermore, while nonpublic reasons may be offered in ordinary public debate, this is only with the proviso that public reasons are presented "in due course"—at least before any legislation is enacted (Lafont 2007, p. 240; Rawls 1997, p. 784). So even if consensus conferences were seen as just another part of the ordinary public debate, public reasons for any policies recommended would still need to be given before they become law.

[7] One should be careful, however, about how one interprets the words 'safe' and 'risky.' As we understand these terms, they mean this: GMOs are safe iff it is known (or well established) that they are highly likely not to have any adverse effects on human health or the environment. They are risky iff it is known (or well established) that they are likely to have such adverse

effects. These two concepts thus do not exhaust the possible epistemic situations we can find ourselves in. In particular, we may face large *uncertainty* about the potential effects of GMOs—i.e., it might not be known (or well established) whether GMOs are likely to have adverse effects or not. This would seem to be the only legitimate factual basis for GMO policy on the broader interpretation of Rawls. But note that the view that we do not know that GMOs are unlikely to have adverse effects may be what some have in mind when they say that GM crops are risky (or at least that they are not safe). For a discussion of the relations between the concepts of safety, risk, and uncertainty, see Möller, Hansson & Peterson (2006).

[8] Streiffer and Hedemann advance their argument via an interpretation of the requirement that public justification must be neutral in certain respects. This detail does not affect the overall argument; the crucial question is what is, and what is not, included in public reason.

[9] Streiffer and Hedemann deny that the naturalness objection and the similar playing-God objection can be excluded from public reasons. They also argue that we should reject what they call the Requirement of Soundness (it is legitimate to dismiss an intrinsic objection on the grounds that it is unsound) (2005, p. 201), the Requirement of Secularity (religious views have no legitimate role in arguments about public policy) (p. 203), and the Requirement of Reason (it is legitimate to reject an intrinsic objection on the grounds that it is inchoate) (p. 204). Their basic argument for rejecting these further requirements is that reasonable comprehensive views can include views that do not meet the requirements. Consequently, premise (2) of their argument cannot be shown to be false with respect to the intrinsic objections by showing that those objections fail to meet one or more of the three requirements.

[10] Thus, an alternative way of framing the difference between Rawlsian theories and Gaus and Vallier's theory is in terms of what reasonable citizens must agree about. On Rawls's view, we must all agree on the *reasons* that are used to justify policies (but it is not a requirement that everyone, from within their comprehensive view, accept a policy if it is justified by public reasons only). On Gaus and Vallier's view, we must agree on the *coercive legislation* (but it is not a requirement that we all agree on what set of reasons justifies that piece of legislation). For an elaboration of the difference between the "reasons-for-decisions model" and the "coercion model" of public justification, see Lister (2013, esp. Ch. 1).

[11] It is worth noting that the asymmetry between accepting and rejecting coercive policies—which is an implication of the presumption of liberty—leads Gaus and Vallier's view towards fairly strong limits on state action: in Gaus's words, the theory has a "classical liberal tilt." Thus, a likely consequence of accepting this theory would be that development and marketing of GMOs by *private* actors would be allowed (at least insofar as such products do not present risks), while publicly funded research, relying on taxation and therefore coercion, would not be legitimate. Or in other words, Monsanto would be free to market "Round Up Ready" corn and soy, but governments would not be allowed to subsidize the development of crops such as vitamin-A-enriched "Golden Rice" that are intended to solve nutritional and other problems in the poorest parts of the world (and would perhaps not, therefore, be viable on market terms). This, we believe, is not the outcome most PPP advocates were expecting.

[12] Our argument against the view that PPP is justified by proceduralist theories of legitimacy is, in this sense, conditional: if the relevant proceduralist account requires direct public participation *generally* then of course such a conception of legitimacy supports direct public participation in the case of science and technology policy as well.

[13] This argument does not, of course, tell against participatory mechanisms that seek to mitigate these problems (see, e.g., Lafont, 2015).

REFERENCES

Ahteensuu, Marko, "Assumptions of the Deficit Model Type of Thinking: Ignorance, Attitudes, and Science Communication in the Debate on Genetic Engineering in Agriculture", *Journal of Agricultural and Environmental Ethics*, volume 25, 2012, pp. 295-313.

Allum, Nick, Sturgis, Patrick, Tabourazi, Dimitra & Brunton-Smith, Ian, "Science knowledge and attitudes across cultures: A meta-analysis", *Public Understanding of Science*, volume 18, issue 1, 2008, pp. 35-54.

American Association for the Advancement of Science, *Statement by the AAAS Board of Directors on Labelling of Genetically Modified Foods*, 2012, available at https://www.aaas.org/sites/default/files/AAAS_GM_statement.pdf

Andersen, Ida-Elisabeth & Jæger, Birgit, "Scenario workshops and consensus conferences: Towards more democratic decision-making", *Science and Public Policy,* volume 26, issue 5, 1999, pp. 331-340.

Audi, Robert & Wolterstorff, Nicholas, *Religion in the Public Square*, Lanham, MD, Rowman & Littlefield, 1997.

Boerse, Jacqueline E.W. & de Cock Buning, Tjard, "Public Engagement in Science and Technology", *in* Ruth Chadwick (ed.), *Encyclopedia of Applied Ethics*, 2nd ed., London, Academic Press, 2012.

Bauer, Martin W., "The evolution of public understanding of science—discourse and comparative evidence", *Science, Technology and Society*, volume 14, issue 2, 2009, pp. 221-240

Comstock, Gary, *Vexing Nature? On the Ethical Case against Agricultural Biotechnology*, Boston, Kluwer Academic Publishers, 2000.

Dryzek, John S., "Policy Sciences of Democracy", *Polity*, volume 22, issue 1, 1989, pp. 97-118.

Durant, John, "Participatory technology assessment and the democratic model of the public understanding of science", *Science and Public Policy,* volume 26, issue 5, 1999, pp. 313-319.

Durning, Dan, "Participatory Policy Analysis in a Social Service Agency: A Case Study", *Journal of Policy Analysis and Management*, volume 12, issue 2, 1993, pp. 297-322.

Estlund, David M., *Democratic authority: A philosophical framework,* Princeton, NJ, Princeton University Press, 2008.

Eurobarometer, *Eurobarometer 73.1: Biotechnology*, Brussels, 2010.

Feinberg, Joel, *Harm to Others: The Moral Limits of the Crimical Law, vol. 1*, Oxford, Oxford University Press, 1987.

Fiorino, Daniel J., "Citizen Participation and Environmental Risk: A Survey of Institutional Mechanisms", *Science, Technology and Human Values,* volume 15, issue 2, 1990, pp. 226-243.

Fischer, Frank, "Citizen participation and the democratization of policy expertise: From theoretical inquiry to practical cases", *Policy Sciences*, volume 26, issue 3, 1993, pp. 165-187.

—, "Technological deliberation in a democratic society: the case for participatory inquiry", *Science and Public Policy,* volume 26, issue 5, 1999, pp. 294-302.

Gaskell, George et al., *Europeans and biotechnology in 2010: Winds of change?* Report to the European Commission's Directorate-General for Research, 2010.

Gaus, Gerald F., *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*, Cambridge, Cambridge University Press, 2011.

Gaus, Gerald F. & Vallier, Kevin, "The roles of religious conviction in a publicly justified polity: The implications of convergence, asymmetry and political institutions", *Philosophy and Social Criticism,* volume 35, issues 1-2, 2009, pp. 51-76.

Greenawalt, Kent, *Private Consciences and Public Reasons*, New York, Oxford University Press, 1995.

Gregory, Jane & Lock, Simon Jay, "The Evolution of 'Public Understanding of Science': Public Engagement as a Tool of Science Policy in the UK", *Sociology Compass*, volume 2, issue 4, 2008, pp. 1252–1265.

Habermas, Jürgen, "Religion in the Public Sphere", *European Journal of Philosophy*, volume 14, issue 1, 2006, pp. 1-25.

Hilbeck, Angela et al., "No scientific consensus on GMO safety", *Environmental Sciences Europe*, volume 27, issue 4, 2015.

Holtug, Nils, "Creating and Patenting New Life Forms", *in* Helga Kuhse & Peter Singer (eds.), *A Companion to Bioethics, 2nd ed.*, Malden, Wiley-Blackwell, 2009.

Joss, Simon & Durant, John, *Public Participation in Science: The Role of Consensus Conferences in Europe*, London, Science Museum, 1995.

Jønch-Clausen, Karin & Kappel, Klemens, "Scientific Facts and Methods in Public Reason", *Res Publica,* volume 22, 2016, pp. 117-133.

Kappel, Klemens, "Fact-Dependent Policy Disagreements and Political Legitimacy", *Ethical Theory and Moral Practice*, volume 20, no. 2, 2017, pp. 313–331.

Lafont, Cristina, "Deliberation, Participation, and Democratic Legitimacy: Should Deliberative Mini-publics Shape Public Policy?", *Journal of Political Philosophy,* volume 23, issue 1, 2015, pp. 40-63.

Lister, Andrew, *Public Reason and Political Community*, London, Bloomsbury Publishing, 2013.

Möller, Niklas, Hansson, Sven Ove & Peterson, Martin, "Safety is more than the antonym of risk", *Journal of Applied Philosophy,* volume 23, issue 4, 2006, pp. 419-432.

Nagel, Thomas, "Moral Conflict and Political Legitimacy", *Philosophy & Public Affairs*, volume 16, issue 3, 1987, pp. 215–240.

National Academy of Sciences, *Genetically Engineered Crops: Experiences and Prospects,* Washington, DC, The National Academies Press, 2016.

Nielsen, Annika P., Lassen, Jesper & Sandøe, Peter, "Democracy at its Best? The Consensus Conference in a Cross-national Perspective", *Journal of Agricultural and Environmental Ethics*, volume 20, issue 1, 2007, pp. 13–35.

Nowotny, Helga, "Democratising expertise and socially robust knowledge", *Science and Public Policy*, volume 30, issue 3, 2003, pp. 151-156.

Perry, Michael, *Morality, Politics and Law*, New York, Oxford University Press, 1988.

Peter, Fabienne, *Democratic legitimacy*, New York, Routledge, 2009.

—, "Political Legitimacy", *in* Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2016 ed.), Available at: https://plato.stanford.edu/entries/legitimacy/

Rawls, John, *Political Liberalism*, New York, Columbia University Press, 1993.

—, "The Idea of Public Reason Revisited", *The University of Chicago Law Review*, volume 64, issue 3, 1997, pp. 765-807.

Rollin, Bernard E., *The Frankenstein Syndrome: Ethical and Social Issues in the Genetic Engineering of Animals*, Cambridge, Cambridge University Press, 1995.

Rowe, Gene & Frewer, Lynn J., "Public Participation Methods: A Framework for Evaluation", *Science, Technology and Human Values,* volume 25, issue 1, 2000, pp. 3-29.

Stirling, Andy, "'Opening Up' and 'Closing Down': Power, Participation, and Pluralism in the Social Appraisal of Technology", *Science, Technology and Human Values,* volume 33, issue 2, 2008, pp. 262-294.

Streiffer, Robert & Hedemann, Thomas, "The political import of intrinsic objections to genetically engineered food", *Journal of Agricultural and Environmental Ethics,* volume 18, issue 2, 2005, pp. 191-210.

Sturgis, Patrick & Allum, Nick, "Science in society: Re-evaluating the deficit model of public attitudes, *Public Understanding of Science*, volume 13, issue 1, 2004, pp. 55-74.

Sunstein, Cass R., *The Laws of Fear: Beyond the Precautionary Principle,* Cambridge, Cambridge University Press, 2005.

Thompson, Paul B., *Food Biotechnology in Ethical Perspective, 2nd ed.*, Dordrecht, Springer, 2007.

—, *From Field to Fork: Food Ethics for Everyone*, Oxford, Oxford University Press, 2015.

—, *Agro-Technology: A Philosophical Introduction,* Cambridge, Cambridge University Press, 2011.

Vallier, Kevin, "Convergence and Consensus in Public Reason", *Public Affairs Quarterly,* volume 25, issue 4, 2011, pp. 261-279.

Waldron, Jeremy, "Theoretical Foundations of Liberalism", *The Philosophical Quarterly*, volume 37, issue 147, 1987, pp. 127-150.

Wall, Steven, "Is public justification self-defeating?", *American Philosophical Quarterly,* volume 39, issue 4, 2002, pp. 385–394.

Weithman, Paul, *Religion and the Obligations of Citizenship*, Cambridge, Cambridge University Press, 2002.

Wynne, Brian, "Knowledges in Context", *Science, Technology & Human Values,* volume 16, issue 1, 1991, pp. 111-121.

Ziman, John, "Public Understanding of Science", *Science, Technology & Human Values,* volume 16, issue 1, 1991, pp. 99-105.

Zurita, Laura, "Consensus conference method in environmental issues: relevance and strength", *Land Use Policy,* volume 23, 2006, pp. 18-25.

DEMOCRATIC DECISION MAKING AND THE PSYCHOLOGY OF RISK

ANDREAS CHRISTIANSEN
POSTDOCTORAL FELLOW, DEPARTMENT OF MEDIA COGNITION AND COMMUNICATION,
UNIVERSITY OF COPENHAGEN

BJØRN GUNNAR HALLSSON
PH.D. STUDENT, DEPARTMENT OF MEDIA COGNITION AND COMMUNICATION, UNIVERSITY OF COPENHAGEN

ABSTRACT:

In many cases, the public (or large parts of it) want to restrict an activity or technology that they believe to be dangerous, but that scientific experts believe to be safe. There is thus a tension between respecting the preferences of the people and making policy based on our best scientific knowledge. Deciding how to make policy in the light of this tension requires an understanding of why citizens sometimes disagree with the experts on what is risky and what is safe. In this paper, we examine two highly influential theories of how people form beliefs about risks: the theory that risk beliefs are errors caused by bounded rationality and the theory that such beliefs are part and parcel of people's core value systems. We then discuss the implications of the psychological theories for questions regarding liberal-democratic decision making: (1) Should policy be responsive to the preferences of citizens in the domain of risk regulation? (2) What risk-regulation policies are legitimate? (3) How should liberal-democratic deliberation be structured?

RÉSUMÉ :

Dans de nombreux cas, le public (ou une grande partie de celui-ci) veut restreindre une activité ou une technologie qu'il croit être dangereuse, mais que les experts scientifiques considèrent être sécuritaire. Il y a alors une tension entre le respect des préférences des gens et des politiques fondées sur nos meilleures connaissances scientifiques. Décider comment élaborer une politique à la lumière de cette tension nécessite de comprendre *pourquoi* les citoyens sont parfois en désaccord avec les experts à propos de ce qui est risqué et ce qui est sûr. Dans cet article, nous examinons deux théories très influentes sur la façon dont les gens forment des croyances sur les risques : la théorie selon laquelle les croyances liées au risque sont des erreurs causées par la rationalité limitée et la théorie selon laquelle ces croyances font partie intégrante des systèmes de valeurs fondamentales des personnes. Nous discutons ensuite les implications des théories psychologiques pour les questions touchant la prise de décision libérale-démocratique : (1) Les politiques devraient-elles être sensibles aux préférences des citoyens dans le domaine de la régulation des risques? (2) Quelles politiques de régulation des risques sont légitimes? (3) Comment la délibération libérale-démocratique devrait-elle être structurée?

1. INTRODUCTION

It is a common and immediately plausible thought that, in a liberal-democratic state worthy of the name, the public should play a substantial role in the policy-making process. It is an equally common and plausible thought that, in an enlightened state worthy of the name, policy making should be based on our best understanding of the relevant facts, which in many domains entails that policy making should be based on scientific knowledge. But now a puzzle presents itself: What to do in cases where the public (or large parts of it) want to restrict an activity or technology that they believe to be dangerous, but that scientific experts believe to be safe (or, conversely, where the public is sanguine about an activity or technology that experts believe to be highly risky)? How, if at all, can liberal-democratic and enlightenment values be reconciled? And if they cannot, how should the two conflicting sets of values be balanced?

In order to answer this question well, we need to understand *why* (parts of) the public sometimes disagree with the experts on matters of risk—we need a cognitive and social psychological understanding of public perceptions of risk. And once we have such knowledge, we need to reflect on what implications the psychological facts have for what role the public ought to play in liberal-democratic policy making. These are our two aims in this paper.

In the first part of the paper (§ 2), we will present and critically assess the evidence for two major and influential psychological theories of risk perception. One is the bounded rationality theory, according to which (nonexperts') thinking about risk is dominated by the use of fast heuristics that lead to predictable biases in risk perception. The other is the cultural cognition theory, which says that lay beliefs about many risks are a result of culturally (or ideologically) biased processing of evidence, and hence are strongly correlated with cultural (or ideological) worldviews. We will argue that, although both theories have their merits, cultural cognition seems to be at play in a majority of the cases where questions of risk regulation are salient politically.

In the second part of the paper (§ 3), we will examine the implications of the psychological theories for three influential liberal-democratic ideas: (3.1) that public policy should be *responsive* to the preferences of citizens; (3.2) that liberal-democratic *legitimacy* requires that policies are reasonably acceptable for all those subject to them; and (3.3) that the public should directly participate in policy making through public *deliberation*. We will focus on claims made by proponents of each of the psychological theories discussed concerning such implications. In particular, we will engage the views of Cass R. Sunstein, on the side of the bounded rationality theory (Sunstein, 2002; 2005; 2006), and of Dan M. Kahan, with a number of coauthors, on the side of the cultural cognition theory (Kahan, 2007; Kahan & Slovic, 2006; Kahan, Slovic, Braman & Gastil, 2006).

On Sunstein's view, the fact that public risk perceptions exhibit the biases characteristic of bounded rationality means that they should be disregarded, and that

policy should instead be determined by the experts using cost-benefit analysis. We will argue that, although Sunstein is right to point out that bounded rationality undermines the case for being responsive to public preferences for risk regulation, his alternative has its own problems.

According to Kahan and coauthors, the fact that risk perceptions are expressions of cultural or ideological worldviews means that they should be treated much as values are treated in liberal-democratic theory. We will argue that this is largely false. However, cultural cognition theory does contain important insights into how we can overcome the conflict between respecting people's values and respecting the truth when making policy concerning risk.

2. PSYCHOLOGICAL THEORIES OF RISK PERCEPTION

Risk perception research has made it clear that there are a number of domains where a substantial proportion of the public disagree with experts about risk-relevant facts. Genetically modified (GM) foods and global warming are two illustrative examples: according to a report by Pew (Pew Research Center 2015), 37% of US adults agree that it is safe to eat GM foods, while the corresponding number among AAAS scientists is 88%. 50% of US adults and 87% of AAAS scientists agree that global warming as a result of human activity is occurring, the latter number increasing to 97% among authors of peer-reviewed articles in climate science (Cook et al. 2013).

The psychology of risk perception aims at explaining such deviations by reference to features of human cognition. The field has been strongly influenced by seminal work by Amos Tversky and Daniel Kahneman on cognitive heuristics and their resulting biases on probability assessments and decision making, as well as their work on prospect theory (Tversky & Kahneman 1974; Tversky & Kahneman 1981; Kahneman 2011). A heuristic is a relatively simple cognitive mechanism that delivers a rapid answer to what may be a complex question, saving time and cognitive resources. While often accurate, the outputs of heuristics may systematically fail under some circumstances. It is these failures that are denoted as biases. So 'heuristic' refers to a cognitive mechanism while 'bias' expresses a normative assessment of the output of this mechanism, to the effect that something has gone wrong from the point of view of a certain normative theory of reasoning (usually probability theory or logic).

To provide an illustrative example: one of the most well-studied heuristics that is also highly relevant to risk perception is the availability heuristic. When using the availability heuristic to answer a question about the probability of an event, people rely on the ease with which they can recall or imagine instances of such events (Tversky & Kahneman 1974). While this may usually yield an acceptably accurate estimate, reliance on the availability heuristic leads to systematic biases in the assessment of probability. The probability of highly salient or widely publicized risks, such as tornadoes or homicides, tends to be overestimated, while the probability of less salient risks, such as heart disease or diabetes, tend to be underestimated (Folkes 1988; Lichtenstein et al. 1978).

Another heuristic whose more recent discovery had a profound impact on the psychology of risk perception is the affect heuristic (Slovic et al. 2004; Finucane et al. 2000). The affect heuristic denotes a tendency for people's judgments of risks and benefits to align along uniformly positive or negative affect towards the risk source. If someone believes that a technology or activity is high risk, she is also likely to believe that its benefits will be low, and vice versa, although there is little reason to suspect that such an inverse correlation usually obtains in reality. This goes beyond people starting with a positive or negative feeling toward a risk source and then generating beliefs about risk and benefits on that emotional background: simply providing people who are naïve with respect to some technology with information that it is high (or low) risk (or benefit) will tend by itself to generate affect, and therefore a belief about benefit (risk) that matches the valence of the initial information. So, if I inform you that a technology, which you currently have no opinion of, is highly risky, this alone will tend to cause you to form the belief that the technology carries little benefit, even in the absence of any direct information about its benefit. More generally, the affect heuristic is representative of an increased awareness within cognitive psychology of the important role emotion plays in risk perception (Roeser, 2010; Slovic et al. 2004).

Heuristic or emotional information processing is typically cast within a dual process framework where it is contrasted with more deliberate, analytical reasoning (Evans 2008; Reyna 2004). When someone is thinking about a technology or activity, a heuristic may yield an initial verdict about risk. Depending on motivation and ability, deliberate reasoning may then be used to scrutinize and possibly override this initial verdict with one that is the result of more deliberate processing (Evans & Stanovich 2013). Heuristics that yield strong intuitions or powerful emotional responses are naturally less likely to be overridden.

2.1. Bounded rationality theory

Psychologists are largely in agreement about the above core findings. Nevertheless, there is substantial disagreement about deeper theories of the psychology of risk perception. We first present bounded rationality theory. The term 'bounded rationality' is sometimes used simply to denote that we as humans are subject to limitations in our decision-making apparatus, compared to an ideally rational agent. This is not controversial. What we call *bounded rationality theory* is a more specific series of claims. It holds that our cognitive apparatus aims at providing accurate factual beliefs, but is fallible in achieving this aim because of overreliance on heuristics. When we form a belief about some risk-relevant fact, the function of that belief is to accurately represent some state of affairs to help us make better choices. However, beliefs may fail to fulfil this function because of cognitive limitations. Subjects may lack the time or processing capacity to engage in deliberate reasoning, and therefore rely on heuristics; and since heuristics are vulnerable to biases, our beliefs may be mistaken. These mistakes can be characterized as "blunders" (Sunstein, 2005): they stem from one's acceptance of the output of heuristic processing and failure to engage in

sufficient reasoning. When lay people disagree with experts about risk, the reason, according to bounded rationality theory, is that lay people often blunder.[1] They rely on heuristic processing, with their associated biases, in their assessment of risk, whereas experts tend to rely on deliberate reasoning including the scientific method and cost-benefit analysis.

Bounded rationality theory has a wealth of research to support it. It rests largely on the literature on core heuristics such as availability, the affect heuristic, framing, and anchoring—which is extensive and well replicated (Klein et al. 2014; Shafir & Leboeuf 2002; Kahneman 2013; Tversky & Kahneman 1981). Additionally, there is some support to the claim that many mistaken beliefs and bad decisions stem from heuristic processing and that increased deliberate processing tends to predict more accurate beliefs and better decisions. One line of research to provide this support is based on individual differences in rational thought (Stanovich & West 1998). People who score highly in one type of test of deliberate reasoning tend to score highly in others (Stanovich & West 2014), and often make better decisions. For example, they tend to make choices under uncertainty that are more utility maximizing compared to people who score low (Frederick 2005). Another approach is to experimentally impair deliberate reasoning through time pressure or a concurrent cognitive load task, or conversely to force a time delay or otherwise attempt to promote reasoning. Inhibiting reasoning consistently leads to errors and to more impulsive behaviour and risk aversion, while bolstering reasoning at least sometimes has the opposite effect (Benjamin et al. 2013).

An aspect of bounded rationality theory that will be important going forward is the implication that people would recognize many of their beliefs as erroneous if they were to engage in the deliberation required to correct their blunder. This hypothetical change of belief might then give rise to different assessments of risk, which would, by virtue of their increased accuracy, be better able to further people's own interests. Thus, adherents of bounded rationality theory can provide a justification for a policy that ignores people's actual beliefs by pointing out that, in addition to better serving their interests, the policy also respects the belief that people actually would have if they were to consider the issue more carefully.

Thus, if the bounded rationality explanation is correct, then we should expect that those parts of the population who disagree with expert judgment about risk-relevant facts do so in part because of a lack of cognitive resources. There are certainly cases where this is borne out. For example, people who tend to rely on intuitive processing profess greater belief in the efficacy of truly ineffectual treatments such as homeopathy to cure disease (Lindeman 2011). However, questioning the general truth of this prediction is at the heart of the cultural cognition critique of bounded rationality, to which we turn in the next section.

2.2. Cultural cognition theory

As mentioned, there is very little disagreement that humans do rely on heuristics and display biases in their thinking about risk.[2] However, the notion that mistaken factual beliefs as a rule are due to the operation of heuristics has come under strong empirical attack from cultural cognition theory. Cultural cognition theory has its roots in anthropological work that describes societal conflict over risk as structured along two cultural dimensions (Douglas & Wildavsky 1983). One dimension, individualism-communitarianism, classifies people according to the extent to which they prefer collective solutions to societal problems over individual and market-driven solutions. The other, egalitarianism-hierarchy, describes the extent to which one prefers firmly stratified social orderings in roles and authority. These two dimensions combine into cultural worldviews, which to a large extent determine people's perception of various risk factors depending on their congeniality or lack thereof to the worldview in question. For example, hierarchical individualists will tend to view regulation aimed at industry as questioning the competence of societal elites and the ability of market forces to solve problems, and therefore tend to view the activity of industry as low risk and not requiring such regulation.

This helps explain a feature of risk perception that is hard to make sense of from within a purely bounded-rationality framework: namely, that attitudes toward many risks form coherent clusters that are sharply divided along political and social fault lines. The above-mentioned figure of 50% of US adults affirming the reality of anthropogenic global warming hides a sharp division within the country: the number is only 15% among conservative republicans, but 79% among liberal democrats (Pew Research Center 2016). Likewise, if one denies the reality of global warming, one is also likely to profess the safety of nuclear power and to favour less gun control. One suggestion from bounded rationality theory might be that this shows one part of the population to be generally more disposed to rely on heuristics than the other. But one would then expect that this group would consistently hold beliefs that are contrary to scientific experts, which is not the case (e.g., as regards the safety of nuclear energy, Pew Research Center, 2015).

To the anthropological base, cultural-cognition theory adds work from psychology on confirmation bias, motivated reasoning, and identity-protective cognition, all of which describe how humans may be biased in their search for, and evaluation of, evidence (Nickerson 1998; Kunda 1990; Dawson et al. 2002). Humans tend to seek out and evaluate evidence in ways that are congenial to their believed or desired conclusions. We tend to accept evidence in favour of our favoured belief with little scrutiny. If the output of a heuristic bolsters a favoured position, then we are unlikely to engage deliberate reasoning to check and possibly overwrite this response. On the other hand, evidence against favoured beliefs is heavily scrutinized and subsequently tends to be deemed weak, while heuristic responses that run counter to a favoured belief will tend to activate deliberate reasoning in an attempt to find an alternative response (Taber & Lodge 2006;

Dawson et al. 2002; Kahan et al. 2017). In evidence-search situations, where people are given the choice between viewing evidence that supports or disconfirms their favoured view, subjects tend to select supporting evidence (Jones & Sugden, 2001; Taber & Lodge, 2006).

So, according to cultural cognition theory, cultural worldviews, not costs and benefits, to a large extent determine people's basic attitudes toward various risk sources. These worldviews furnish us with our basic values, which in turn cause us to engage in motivated reasoning in dealing with evidence, with the aim of reaching factual beliefs about these risk sources that protect and bolster the attitude in line with our values.

This suggests a flaw in the bounded-rationality picture. Mechanisms such as motivated reasoning and identity-protective cognition are not heuristics. They are instances of deliberate reasoning, but instances where the aim appears not to be merely a correct appreciation of the facts, but rather to provide support for a particular conclusion. When cultural worldviews are in play during evaluation of evidence regarding a risk source, we are likely to use our reasoning to assess the evidence such that it comes out supporting the position that confirms our worldview. This in turn predicts that widespread increased reliance on reasoning rather than heuristics will not necessarily bring about convergence towards a view closer to the truth. Rather, we should expect those with the greatest propensity and ability to engage in deliberate processing to be *best* at making the evidence yield their favoured conclusion (Kahan 2013).

In an illustrative study (Kahan et al. 2017), participants were asked to assess which of two conclusions the results of a (fictional) study supported. In the control version of the task, the study in question was on the efficacy of an experimental crème for the treatment of skin rash. The study's results were presented as a two-by-two matrix, with one dimension denoting whether study subjects' rash got better or worse, and the other denoting whether the subjects had received the treatment or the placebo. Each cell contained a number indicating how many people experienced a certain combination of these dimensions (e.g., people whose rash got better *and* who had received the treatment). Participants had to detect correlation between the variables in order to correctly solve the task. This was so difficult that less than half of participants provided the correct answer (i.e., the result was lower than chance), and performance increased with numeracy (a measure of deliberate processing ability) regardless of cultural background.

In the experimental version of the task, the study was on the effect of gun-control legislation on crime. Here, the cells corresponded to cities that had either implemented a gun-control law recently or not, and whether crime had increased or decreased (e.g., one cell contained the number of cities that had not implemented gun-control *and* had experienced a decrease in crime). Here, a sharp division along cultural lines was seen. If given a version where the correct answer was that crime had decreased as a result of gun control, then liberal participants were

likely to find the correct response, and this likelihood increased sharply with numeracy scores. However, conservative participants given this version were very unlikely to find the correct response, and increased numeracy had no effect on their likelihood to do so. The converse pattern was found for the version where the correct response was that crime had increased: conservatives were quite good at finding the correct response, and highly numerate conservatives much more so than less numerate ones, and liberals were bad at finding the correct response, with increased numeracy offering a very limited benefit. That is, increased capacity to engage in deliberate reasoning helped attaining true beliefs only when the evidence was supportive of one's worldview. This suggests that simply providing people with evidence or attempting to engage their deliberative faculty rather than heuristics will do little to correct false beliefs, when these false beliefs are congenial to their cultural worldview. It further suggests that, in general, one should not expect increased deliberative ability to lead to convergence on truth, but rather that one should find the greatest amount of cultural divergence among the most reflective, numerate, and educated.

Research from proponents of cultural cognition theory has borne this out. Across a great many culturally contested domains related to risk, such as global warming, gun control, the HPV vaccine, and fracking, cultural polarization is largest among those with the greatest reflective abilities (Kahan et al. 2012; Kahan et al. 2010; Kahan et al. 2013; Kahan 2015). It thus becomes highly problematic to refer to false beliefs that are the result of the mechanisms described by cultural cognition theory as blunders. In many cases, they may be the result of a large amount of deliberate reasoning, rather than an uncorrected heuristic. Likewise, the notion that policy-makers can assume that people's factual beliefs would align with those of scientific experts if only they were to reflect more becomes untenable. What one could expect is rather that increased reliance on deliberate reasoning would lead to attitude polarization: more extreme versions of current beliefs (Lord et al. 1979; Taber & Lodge 2006).

Naturally, far from all domains of risk are culturally contested. For example, there is no cultural conflict over artificial food colourings or sweeteners, cellphone radiation, the MMR vaccine, or genetically modified foods (in the US, but probably not in Europe), and in such domains one finds the expected pattern predicted by bounded rationality theory: that higher scientific literacy and reflective capacity increases the likelihood of agreeing with scientific experts, across cultural groups (Kahan 2015). Thus, one can view cultural cognition theory as describing an important class of exceptions to the general bounded rationality framework rather than as providing a full alternative.

It is an important and, to a large extent, unanswered question for cultural cognition theory why and how certain risks become culturally contested and whether this can be reversed: the HPV vaccine apparently became culturally salient only following a series of missteps on the part of its manufacturer (Kahan et al. 2010), and even global warming was not a particularly divisive issue in the early 1990s (McCright, Xiao, & Dunlap, 2014).

3. LIBERAL-DEMOCRATIC DECISION MAKING

We said at the outset that determining the appropriate balance between relying on experts and including lay citizens' views required understanding what causes citizens to sometimes disagree with experts about what things are risky and what things are safe. We have now seen that the answer is, it's complicated. With respect to some risks, the beliefs of (many) citizens are influenced by heuristics and, as a result, exhibit biases. In those cases, those who are the least scientifically literate and who rely the most on intuitive judgment tend to disagree most with the experts. However, for a substantial number of risks, lay opinion is divided along cultural lines. In these cases, agreement with experts is not correlated with scientific literacy or deliberate, careful reasoning—rather the opposite is true. Instead, an individual's beliefs about the riskiness of some phenomenon largely depends on whether that phenomenon is good or bad according to her basic cultural worldview—her basic values. Furthermore, cases where risk debates have become culturally charged are overrepresented among the risks that exhibit the conflict between experts and (some) citizens, which is our subject in this paper.

So what conclusion can we draw concerning risk management in a state that aims to respect liberal-democratic values and to be enlightened? As noted in the introduction, in assessing the political implications of risk psychology, we will focus on claims that proponents of the two theories we have presented have themselves made. We will structure our discussion according to three core ideas in liberal-democratic political theory. First, there is the idea that public policy should be *responsive* to the preferences of citizens—that is, that differences in public opinion should register as differences in the policies implemented. Second, there is the idea that policies should be such that they could enjoy the assent of all those subject to them. This is most famously engendered in liberal and 'public reason' accounts of political *legitimacy*. And third, there is the idea that the public should directly participate through some form of society-wide *deliberation* on policy issues. We will discuss the implications of the psychological theories for each of these ideas in turn. Before doing so, let us state a couple of clarifications and assumptions.

First, when we are talking about people's risk perceptions in a policy-making context, we are not typically talking about pure factual beliefs. Rather, we are typically talking about one of two things: (i) unprompted exclamations (letters to the editor, demonstrations, etc.) to the effect that a certain risk is *serious*, an activity is *dangerous*, or that *something must be done* about a risk, or (ii) support, in one form or another, for proposals to regulate the relevant risky activity (e.g., by expressing such support in surveys, by voting for such policies directly in referenda, or by basing one's vote for representative bodies on the risk-regulation platform of the relevant party or candidate). These are (more or less specific) opinions concerning *what policies should be enacted*—they are *policy preferences.*

Second, we will assume that there is in fact consensus among scientific experts concerning a given risk. Note here that *experts'* views of risk are typically not risk perceptions in the sense defined above (i.e., policy preferences). Rather, they are estimates of the probabilities of various (primarily negative) effects of a policy, such as deaths, other health effects, or environmental degradation. We will also assume that (parts of) the public express policy preferences that are at odds with this consensus, in the sense that the following three propositions are true: (a) the public want a technology or another potentially risky thing restricted, (b) this policy preference is based on a belief that the thing in question is risky, and (c) expert consensus is that the thing is not very risky.

3.1. Responsiveness

While it is fairly uncontroversial that it is an ideal of democratic systems that policies are responsive to the preferences of citizens, it is not clear what this ideal entails more precisely. In particular, it is not clear what 'public preferences' means—it might be public opinion as expressed in polls, the preferences expressed by those citizens who actively engage in political debate, or perhaps the preferences policy-makers perceive to be prevalent in the population (see Manza & Cook, 2002, pp. 631-632). Furthermore, it is not obvious what is required for policies to be responsive to such preferences. Typical explications merely hint at an answer, such as that politicians should *take preferences into account* or that policy should be *influenced* by public preferences (Brooks & Manza, 2006, pp. 474-475). How preferences should be taken into account or how much they should influence policy is left open—although most agree that "a perfect correspondence" is neither required nor desirable (Gilens, 2005, p. 778). We want here to set aside debates about what responsiveness is or should be. Instead, we focus on a more basic issue—namely, whether there is even a *prima facie* requirement that the policies of a democratic state should be responsive to citizens' risk perceptions when these are in apparent conflict with expert beliefs.

3.1.1. Sunstein

Sunstein can be seen as arguing that there is no such *prima facie* requirement. At least, he argues that citizens' policy preferences with respect to the regulation of risk-creating activities should play a relatively limited role in policy making. As an alternative, he argues that a major role should be given to cost-benefit analyses performed by experts in regulatory agencies. More precisely, he supports the current (as of 2016) United States system, in which a central agency of the federal government (OIRA, the Office for Informational and Regulatory Affairs) has a mandate to review and reject, on the basis of cost-benefit analyses, regulations suggested by the various technical agencies dealing with environmental, health, and safety policies (such as the Environmental Protection Agency or the Occupational Safety and Health Administration). A main reason for this is a belief that the technical agencies' regulatory priorities reflect public risk perceptions, rather than scientific estimates (Sunstein, 2002, p. 53, citing

Roberts, 1990). The details of Sunstein's proposals are complex, but the main underlying idea is that policy need not be responsive to public risk perceptions, since on his view these are largely (as we have seen above) the products of cognitive biases of various kinds. This conclusion he derives from a general principle: "democratic governments should respond to people's values, not to their blunders" (Sunstein, 2005, p. 126). Since risk perceptions are based on blunders, democratic governments are not required to be responsive to them.

Is he right about this? One possible reason to think that he is not arises if one thinks that the general principle—that democracies should respond only to values, not to blunders—is false. But it is an open question what it would mean for the principle to be false, since it is unclear what the principle says. The problem is that "values" and "blunders" are not exhaustive of the possible descriptions we may give of people's psychological attitudes. True factual beliefs, for example, are clearly neither values nor blunders. Sunstein's principle, then, says that policies *should* be responsive to people's normative beliefs, but *need not* be responsive to their false (or perhaps only obviously false) factual beliefs. This leaves entirely open what we should do when different people or groups hold divergent factual beliefs, none of which is clearly false. In other words, Sunstein's principle has nothing to say about the criteria for selecting which factual beliefs, beyond the clearly false ones, should be allowed to play a role in policy making.

A natural solution to this problem is to add in a principle for selecting respectable factual beliefs. One plausible such principle, congruent with the ideal of enlightened decision making we mentioned in the introduction, would be to use science as a standard-setter. On such a view, any belief conflicting with the scientifically established facts is not entitled to democratic responsiveness. There are ways of questioning this principle, and especially ways of questioning whether (and how) it could be justified given standard understandings of public reason and the nature of factual disagreements (see, e.g., Jønch-Clausen & Kappel, 2015; 2016). However, we believe the price of giving it up is exceedingly large; since the scientific method is the best known way of generating true factual beliefs, it seems that denying that science can act as gatekeeper for beliefs is tantamount to giving up on having any standards of right and wrong in the empirical domain. So we will accept that beliefs in conflict with established scientific fact are such that democratic governments need not respond to them.

An important caveat needs to be added. In a number of cases, among which are many that are policy relevant, scientific knowledge comes with sizeable uncertainties attached. This needs to be taken seriously by policy-makers. Uncertainty, in effect, means that a number of states of affairs are consistent with the available evidence. In the case of risk, a plausible (but perhaps too simple) way of fleshing this out is to assign only an *interval* of probabilities to a given event, rather than a precise probability (for instance, the probability per year of dying from exposure to pesticides may fall in the interval from one in one million to one in two million). In the case of discrete possibilities—for example, whether

gun control works to lower the number of gun-related deaths per year or not—uncertainty means that we cannot believe either discrete possibility very strongly (i.e., the maximum permissible credence for the proposition "gun control works" is relatively close to 0.5). Where uncertainty is involved, the scientific evidence thus does not permit us to give a unique answer to the policy-relevant question—e.g., what the probability per year of dying from pesticide exposure is, or whether gun control works to lower gun-related deaths. Instead, a number of unique answers are possible. It does not fall within the remit of scientific experts to select which of the set of scientifically permissible unique answers to use.

In many cases, however, policy choice depends on what unique answer is correct in the following sense: if p_1 is true, policy R_1 is required (or preferable), but if p_2 is true, R_2 is required. For example, if gun control works, then gun control is (arguably) required—but if gun control does not work, gun control is not required. In such cases, there is a gap between accepting Sunstein's values-not-blunders principle, and delegating decision-making authority to scientific experts, even granting that 'blunders' includes every belief that is contrary to what science says. Public risk perceptions may play some role in filling that gap.

A more important problem with the values-not-blunders principle is that the risk perceptions of ordinary people, being policy preferences, do not straightforwardly fall on either side of the normative-factual belief divide. Consider how an ideally rational person, of the kind one can meet in decision-theory textbooks, would form her policy preferences concerning a risky activity. Such a person would assign a probability and a value measure ("utility") to each possible outcome of each possible policy, multiply each probability by its utility and sum these products, and advocate the policy that has the highest expected utility. So, even for such a person, a call for a given policy is a consequence of a combination of factual and normative beliefs. Indeed, a policy preference can be made consistent with *any* factual belief, given that the appropriate adjustments are made to the person's normative beliefs. The mere fact that the person calls for a given policy does thus not in itself provide evidence that she has a factual belief that is in conflict with the scientific facts.

However, as we have seen above, the bounded rationality theory that Sunstein relies on provides positive reasons to think that people's factual beliefs concerning risk are often wrong. And (at least to a large extent) the basic fact that nonexperts' beliefs about the magnitude of risks often diverge from the best scientific estimates is not in dispute within psychology. So let us suppose that we can be fairly certain that at least some people have erroneous factual beliefs about the magnitudes of various risks. If it were possible to "implant" true beliefs into such people, then it seems plausible that their risk perceptions (i.e., their more or less precise beliefs about what policies should be enacted) would change.

A very plausible explication of the values-not-blunders principle is then this: what democracy requires is responsiveness to the preferences that people would have had if their factual beliefs were true (or at least not contrary to scientifically

established facts).[3] Call this their *counterfactual fact-based preferences*. In so far as policy preferences that ordinary people express currently—call this their *actual preferences*[4]—are different from their counterfactual fact-based preferences, actual preferences are not the kind of thing democracies need to be responsive to.

The normative appeal of this ideal of policy-responsiveness seems to us considerable (although one might want to consider some minimal criteria for what *normative* beliefs are above board as well). Its main problem is its hypothetical nature. We agree that the ideal form of democratic responsiveness is to the counterfactual fact-based policy preferences of citizens. But in order to implement responsiveness to counterfactual fact-based preferences, we must know (or have reasonably justified beliefs about) what specific preferences a citizen or group of citizens would have had, if they had believed the facts. Note that this is quite a lot harder than having a justified belief that citizens would not have had their actual preferences if they had believed the facts. The real challenge for those who wish to implement responsiveness to counterfactual fact-based preferences is to devise or point to some method for generating reasonably justified beliefs about the specific preferences citizens would have had if they had believed the facts. The only fail-safe way would be to make sure all citizens sincerely believe the facts, to have them determine their policy preferences given those beliefs, and then to make policy responsive to those preferences. But it is of course not possible to run a counterfactual fact-based version of the entire democratic process. So it seems that the best we can aim for is a method that we have reason to believe generates preferences that are reasonable approximations to people's counterfactual preferences.

At least in some places, it seems that Sunstein believes that cost-benefit analysis is a procedure that realizes this. Cost-benefit analysis builds on the approach assumed in decision theory, where (as mentioned above) preferences are a function of separate factual beliefs and value judgments. With respect to factual beliefs, cost-benefit analysis uses the best scientific estimates of the magnitude of risks. As such, it clearly meets the criterion of nonresponsiveness to blunders (although doubts can be had as to whether cost-benefit analysts neglect scientific uncertainty (McGarity, 2002)). With respect to the value judgments, cost-benefit analysts assign a monetary value to a given risk (e.g., a one-in-one-hundred-thousand risk of death per year) based on studies of what people are willing to pay to avoid such a risk, or of what they demand to be paid in order to accept bearing such a risk. Typical ways of measuring willingness-to-pay are studies of wage differentials between risky and safe jobs, and surveys asking people directly for their valuations. Sunstein suggests that "the governing theory" behind this approach "follows [people's] own judgments about risk protection" (Sunstein, 2014, p. 86). Although he also stresses that the current practice does not fully realize the governing theory—in particular, it does not sufficiently take into account differences in risk valuations across individuals—he seems to believe that the general willingness-to-pay approach measures people's own valuations of a given risk (as he says, "the limitations [of current theory] are

practical ones" (Sunstein, 2014, p. 136)). By combining these valuations with the facts and assuming the framework of decision theory, cost-benefit analysis arrives at the preferences people would have had if they had believed the facts.

The idea that the methods of cost-benefit analysis tracks people's own valuations—their counterfactual fact-based preferences—is not universally accepted. It relies on extrapolation of behaviours in one context, in particular the labour market, to all other contexts, and on assumptions from economics and rational choice theory that are in many ways questionable (see, e.g., Anderson, 1993, ch. 9; Hausman, McPherson & Satz, 2017, ch. 9). Furthermore, the very same biases and heuristics that Sunstein is eager to expel from risk management through the use of scientific estimates are likely to influence people's valuations of risks in willingness-to-pay studies. Finally, survey studies frequently register a large number of so-called *protest valuations* (where people state a willingness to pay either nothing or an implausibly large amount, or perhaps decline to state a number at all), indicating a rejection of the very idea of using willingness to pay as a valuation measure for public goods (Kahneman, Ritov, Jacowitz & Grant, 1993). Such responses are typically disregarded, which suggests that cost-benefit analysis is ill equipped to deal with preferences that are not of the type typically relevant in markets. Thus it does not succeed in capturing the counterfactual fact-based preferences of those who reject treating a given policy domain as appropriately governed by the ideals of a market economy.

The conclusions that can be drawn from the above are limited. We have merely suggested that Sunstein's proposal of delegating much of the policy-making power to scientific experts doing cost-benefit analyses is not plausibly an ideal solution to risk regulation. So even if Sunstein is right that risk perceptions—of the unfiltered kind that are expressed in the various more or less precise calls for risk-regulating policies—are too tainted by their partial source in cognitive biases to be taken into account in policy making, his alternative may not be much better. At least, his alternative does not embody ideal responsiveness (i.e., responsiveness to counterfactual fact-based preferences). It is doubtful that ideal responsiveness *can* be fully realized in practice. It may be the case that the available realizable alternatives leave us with a dilemma: if we make policy responsive to expressed risk perceptions, we will be *overresponsive* to false or unscientific beliefs; but if we make policy unresponsive to these risk perceptions, we will be *underresponsive* to values. In other words, the seemingly simple ideal of responsiveness to values and nonresponsiveness to blunders may be an unattainable ideal. Call this the *responsiveness dilemma.*

3.1.2. Cultural cognition

Kahan and his coauthors argue that cultural cognition theory further undermines Sunstein's approach. Recall first what the cultural cognition theory says about how people form risk perceptions. On the cultural cognition model, risk perceptions are not formed in the way assumed by decision theorists (and by Sunstein)—that is, by combining pure factual beliefs about the numerical magni-

tude of risks (expected deaths, probabilities of ecosystem damage, and the like) with pure normative beliefs about how bad the various possible bad effects of a policy are. Instead, people assess (probably mostly unconsciously) the relationship between a possibly risky activity and their cultural worldview—and thus assess at the same time whether *restricting* the activity is justifiable, or perhaps required, according to their view of the ideal society. Thus, as we mentioned above, hierarchical individualists balk at regulation of industry because it questions the competence of elites (hierarchy) and assumes the inadequacy of market solutions (individualism). Conversely, egalitarians dislike the activity of capitalist industry generally, and thus welcome restrictions. Based on such general assessments of the value of activities and of restrictions on them, people form factual as well as normative beliefs about the risks and benefits of the activity, in a kind of post-rationalization procedure, in which motivated assessment of evidence concerning the effects of the activity and policy is central.[5] Consequently, "citizens invariably conclude that activities that affirm their preferred way of life are both beneficial and safe, and those that denigrate it are both worthless and dangerous," and even the factual aspect of risk perceptions (could they be isolated) "express [citizens'] worldviews" (Kahan et al. 2006, p. 1105).

Kahan et al. argue that cultural cognition theory undermines Sunstein's view in two related ways. First, they claim that Sunstein's strategy of using cost-benefit analysis to realize the values-not-blunders ideal "borders on incoherence" (Kahan et al., 2006, p. 1105). In other words, the fact that risk perceptions are due to cultural cognition means that the cost-benefit approach does not realize the ideal embodied in the values-not-blunders principle. On one reading, this would merely be the claim we have just made: that cost-benefit analysis fails to respect values. But of course this would be completely independent of the cultural cognition theory. The values we have argued are overridden in cost-benefit analysis are ordinary normative beliefs (about the value of a human life, say), not culturally influenced factual beliefs (about how many lives a certain activity will claim). Second, they suggest that "bringing the role of cultural cognition into view severely undermines the foundation for Sunstein's refusal to afford normative significance to public risk evaluations generally" (Kahan et al., 2006, p. 1004). That is, they suggest that acknowledging the role of cultural cognition undermines the case for nonresponsiveness to citizens' actual policy preferences.

How might the fact that people's risk perceptions are shaped by cultural cognition further undermine the cost-benefit analysts' approach and/or strengthen the case for responsiveness to actual preferences? We suggest that cultural cognition points to two different facts that may be important: (1) that the relationships between values (in the form of cultural worldviews), factual beliefs, and policy preferences are not as Sunstein and others assume, and (2) that risk perceptions are rooted in cultural worldviews, and therefore are expressions of citizens' values.

Let us first consider issue (1). Here, the claim would be that the fact that risk perceptions are due to cultural cognition means that they do not behave in ways that Sunstein and others assume—for example, that changes in factual beliefs do not change preferences in the way assumed—and that this undermines the strategy of cost-benefit analysis further and/or strengthens the case for responsiveness to actual preferences. Such a claim could be made in two ways:

(i) Since both factual beliefs and policy preferences are due to the same underlying cause, we should not expect changes in factual beliefs to change policy preferences. As Kahan et al. put it,

> risk perceptions originating in cultural evaluation are not ones individuals are likely to disown once their errors are revealed to them. Even if individuals could be made to see that their cultural commitments had biased their review of factual information … they would largely view those same commitments as justifying their policy preferences regardless of the facts. (Kahan et al. 2006, p. 1105)

On this reading, an individual's counterfactual fact-based preferences are likely to be the same as his actual preferences (i.e., the preference he would hold if he believed the facts is likely to be the same as the preference he currently holds). If that is the case, people's actual preferences are at least a good approximation of their counterfactual fact-based preferences. Thus we have a solution to the problem of how to achieve responsiveness to counterfactual fact-based preferences—namely, to use actual preferences. Or, to put the matter differently, it is not true that responsiveness to actual preferences is overresponsiveness to faulty factual beliefs, since actual preferences are not influenced by factual beliefs at all—faulty or not. Reading (i) would, then, give reason to be responsive to citizens' actual preferences.

Reading (i) faces two problems. The first problem is that the claim that changes in factual beliefs do not change policy preferences seems too strong, and it goes beyond what can be justified by the evidence that the cultural cognition theory relies on. Cultural cognition is primarily a thesis about how cultural commitments lead to biased assessment of evidence, such that one believes the evidence supports the factual beliefs that fits one's cultural commitments best. But it is possible to debias people at least to some degree, and to bring them towards mutual agreement on the facts. And furthermore, there is evidence that such debiasing alters people's policy preferences, bringing previously opposed parties closer together (Cohen et al., 2007). So it seems to us that the fact of cultural cognition does not justify ignoring the problem of overresponsiveness to false beliefs.

The second problem is that, at least in many policy domains, preferences may lose some of their claim to democratic responsiveness if they turn out to be too resistant to the facts. Resistance to changes in factual beliefs may reveal policy preferences to be based in kinds of value judgments that are unacceptable from

a liberal-democratic point of view—e.g., a desire to regulate purely private behaviour (such as sexual behaviour or harmless commercial activities) or worldviews that deny the fundamental equality of all citizens (such as racist or sexist views). If it were the case that citizens' policy preferences would not change regardless of what the facts are, we would at least need to examine the substantive content of those preferences in more detail—and to reserve judgment as to whether those preferences merit democratic responsiveness until we have a better understanding of what that substantive content is.

(ii) Since policy preferences and factual beliefs are both caused by people's cultural worldviews (i.e., their most basic values), any change in factual beliefs requires a change in basic values.

Suppose a given citizen actually has faulty factual beliefs, and that these beliefs are due to cultural cognition. According to reading (ii), the basic values this citizen actually holds are *not* the basic values she would hold in the counterfactual case where she came to believe the facts. The cost-benefit analysts' method is essentially an attempt to disentangle actual factual beliefs from actual value judgments. The analysis then recombines actual value judgments with the *true* facts, and thereby generates a policy preference. But on reading (ii), such an approach does not succeed in revealing citizens' counterfactual fact-based preferences. The cost-benefit method uses a citizen's *actual* values, but cultural cognition shows that these are likely to be different from her *counterfactual fact-based* values. In other words, a citizen's counterfactual fact-based preferences are not (as Sunstein believes) a function of her actual values and the facts, but a function of a *new* set of values and the facts.

Reading (ii) would show that the cost-benefit analysts' method does not successfully track people's counterfactual fact-based preferences. It also suggests that it is difficult to predict how people's preferences would change if they sincerely came to believe facts that are in conflict with their cultural worldviews. Thus it lends support to the use of more deliberative methods, wherein real flesh-and-blood people are allowed to undergo a change in their views in response to facts and arguments (unlike methods like cost-benefit analysis, which seeks to *infer* what people would prefer from data about what they actually believe, value, and prefer). Consequently, the "deliberative debiasing" methods Kahan et al. argue in favour of using are supported by this reading (Kahan et al., 2006, pp. 1100-1104).

Kahan et al.'s other claim—that cultural cognition supports responsiveness to actual preferences—is not supported by reading (ii). At best, reading (ii) shows cost-benefit analysis to be a worse approximation of the ideal of responsiveness to counterfactual fact-based preferences than we might otherwise have thought. However, this merely makes the responsiveness dilemma worse, by making one of the horns of that dilemma worse. It is not obvious, however, that reading (ii) is of much help in deciding how to choose when faced with a responsiveness dilemma—that is, if we have to choose between responsiveness to actually expressed preferences and (something like) cost-benefit analysis.

Let us now move to issue (2), the fact that cultural cognition theory shows risk perceptions to be expressions of values. Kahan et al. state that "when expert regulators reject as irrational public assessments of the risks associated with putatively dangerous activities … they are in fact overriding *values*" (Kahan et al. 2006, p. 1105). It is, unfortunately, not clear what is meant by "public assessments of … risks" in this quote. On the one hand, the phrase might refer to policy preferences, such as that a given activity A is dangerous and should be regulated. On the other hand, it might refer to people's purely factual beliefs about the magnitude of risks. Let us now consider each of these two readings of issue (2) in turn (we call them readings (iii) and (iv) to avoid confusion with (i) and (ii) above):

(iii) *Experts are overriding G_1's values because they implement a policy R_2 that is different from G_1's preferred policy R_1.*

Recall that the kind of case we are interested in has the following structure: (a) the public wants a technology or another potentially risky thing restricted, (b) this policy preference is based on a belief that the thing in question is risky, and (c) expert consensus is that the thing is not very risky. In the group-based framework of cultural cognition, 'the public' should be replaced with some cultural group. So we assume that a cultural group G_1 wants the activity A restricted through policy R_1, and that G_1 wants this because they believe p, that A carries certain risks. The experts, based on sound science, believe $Øp$ (i.e., that A does not carry those risks) and therefore implement a policy R_2 that does not restrict A appreciably.

In cases of this kind, it is hard to see why we should accept that implementing a policy other than R_1 overrides G_1's values. By assumption, G_1 prefers R_2 *because* they believe p—the implication being that they would not have preferred R_1 if they had believed $Øp$ (i.e., that R_1 is not their counterfactual fact-based policy preference). Once more, there are now two possibilities for what G_1's policy preference would then have been if they had believed $Øp$. First, G_1 might have preferred, or at least acquiesced to, R_2, the policy implemented by the experts. In that case, the expert decision procedure would have achieved its ideal aim. Thus there would be no reason to be responsive to G_1's actual preference, and we would have no reason to object to the experts' decision procedure either. Second, G_1 might have preferred some third possible policy R_3. In that case, we would still have no reason to demand that policy be responsive to G_1's actual preferences. However, there would be reason to complain that the experts' decision procedure has failed to be responsive to G_1's values. Insofar as we cannot tell *a priori* whether G_1 would have preferred (or acquiesced to) R_2 or not, the conclusion that follows is that we cannot be confident that the experts' decision procedure is responsive to G_1's values, in the absence of some effort to determine what G_1's counterfactual fact-based preferences are.

But perhaps the assumption that G_1 prefers R_1 *because* they believe p is not correct. That is, perhaps the case is one in which G_1 would prefer R_1 regardless

of the facts—G_1's factual belief that A is dangerous is merely a *post hoc* rationalization of the group's policy preference, which it holds for other reasons than that A is dangerous. Kahan and Braman (2008, pp. 51-54) suggest that it is only in cases of this kind—where people would not alter their policy preference even if they came to believe the facts—that there is a demand for policy responsiveness to preferences. At the same time, however, they speculate that people would *not* be inclined to hold on to their preferences if they were to realize that their factual beliefs are the product of cultural cognition, at least in the case they are discussing (cases of self-defence). The same might well be the case in typical instances of risk regulation. In the case where people *would* hold on to their policy preferences after coming to believe the facts, the problem we mentioned under reading (i) above recurs: G_1's preference for R_1 has some basis other than that A is in fact risky, and that basis may show the preference to be less reasonable than it initially seemed.

Consider, for example, the case of regulation of industry pollution. Recall that hierarchical individualists tend to be sceptical of such regulation because it casts doubt on the competence of societal elites and the ability of market forces to solve problems, and consequently tend to believe that the risks associated with industry pollution are low. But suppose hierarchical individualists were brought to sincerely believe that some industry's emission of a certain chemical C creates severe risks to the health of those exposed, but that they persisted in their policy preference (not to regulate). What could the basis of such that preference then be, other than a blatant disregard for the welfare of those who will likely suffer health problems? A similar problem arises for egalitarians, who are inclined to approve of restrictions of "commerce and industry, which they see as sources of unjust social disparities" (Kahan, 2012, p. 728), and who consequently tend to believe that the risks associated with industry pollution are high. Suppose egalitarians persisted in their desire to regulate emissions of C even after having sincerely accepted that C does not pose a serious risk to anyone. The only possible basis of such a preference is then a general anti-industry agenda. By persisting in their preferences, both the hierarchical individualists and the egalitarians would violate basic norms of risk regulation, such as that people have some right to be protected against serious risks and that harmless private behaviour cannot be restricted.

Thus it seems to us that in the case of risk regulation there is reason to be sceptical of policy preferences that would not change if people were to come to believe the facts. So, while the possibility that policy preferences would not change if people came to believe the facts does provide some reason to be responsive to those preferences, there will simultaneously be a reason not to be responsive. However, in cases where people merely *overestimate* risk (or underestimate, as the case may be), persisting in policy preference is less problematic. It may reflect, for example, a judgment that the aim of protecting people's health is very important relative to the aim of securing favourable conditions for business. But this is just the general problem with cost-benefit analysis we identified above. It is not obvious that the phenomenon of cultural cognition adds much to that problem.

(iv) Experts are overriding G_1's values by denying the pure factual beliefs of G_1 (i.e., p*), since those factual beliefs express values.*

Since believing p is an expression of G_1's values, the validity of G_1's values is denied when expert regulators implement a policy based on the fact that $Øp$ is true. We think the view that merely denying (a group of) citizens' factual views is to be underresponsive to their values has both strange and dangerous implications. Suppose, for example, that the experts in this case implement G_1's preferred policy R_1, but also believe (and state publicly) $Øp$. On the view considered, the implication would be that the experts' policy making is insufficiently responsive to the values of G_1 in this case, even though G_1 got its preferred policy implemented. That seems to us a strange implication, which requires an excessive demand for responsiveness.

Alternatively, consider a case like the one we mentioned above, where G_1 would at least acquiesce to the expert's implementation of R_2 if they were to come to believe the truth (i.e., $Øp$). One might think that, since the belief p is an expression of G_1's values, implementing R_2 exhibits a lack of responsiveness to G_1's values even though R_2 is G_1's counterfactual fact-based preference (or a least would be acceptable to G_1 in those counterfactual circumstances). In effect, this would amount to denying that policy preferences that unequivocally depend on factual beliefs that do not meet the required correctness criterion (i.e., beliefs that are blunders or contrary to scientifically established facts) do not merit democratic responsiveness. This seems to us a dangerous implication. In factual matters, priority must be given to the truth, and to our best methods for finding out the truth. And in fact, Kahan et al. seem to share our worry here. In a response to Sunstein's response to their original paper, Kahan and Slovic "admit to a fair measure of ambivalence about when beliefs formed as a result of cultural cognition merit normative respect within a democratic society," and concede that "if we came off sounding as if we think democracy entails respecting all culturally grounded risk perceptions, no matter how empirically misguided they might be, we overstated our position" (Kahan & Slovic, 2006, pp. 170-171).

In conclusion, Kahan et al.'s scepticism towards Sunstein's proposed use of expert cost-benefit analysis is largely warranted, but it is questionable if the fact of cultural cognition contributes much to the problems with cost-benefit analysis. To be sure, cultural cognition provides a *different* set of reasons for thinking that cost-benefit analysis does not succeed in tracking counterfactual fact-based preferences—but arguably that claim was already very well supported by other reasons. Furthermore, cultural cognition theory provides only very limited reason to be responsive to actual preferences in cases where these are in conflict with experts' scientific assessments of the riskiness of an activity. Cultural cognition theory therefore does not warrant solving the responsiveness dilemma in favour of responsiveness to actual preferences. It does, however, provide support for using deliberative debiasing techniques to solve that dilemma.

3.2. Liberal legitimacy

We now move from the democratic to the liberal aspect of the liberal-democratic ideal—more precisely, to the liberal conception of legitimacy. According to this conception, political power is legitimate only if could be reasonably accepted by all subject to it. While many philosophers are attracted to some version of the liberal legitimacy principle, there is no general agreement on what the principle precisely amounts to. It is controversial how demanding the requirement that political power be acceptable to all is—does it require that all can accept the basic procedure by which laws and policies are made (Rawls's view) or does it require that each law or policy be reasonably acceptable to all? The latter is obviously a much more demanding criterion. It is likewise controversial how demanding the reasonability clause is—should our conception of reasonability be such that the acceptance of most people as they really exist is required, or do we need to secure acceptance only from people whose views meet higher standards of justifiability? And there are more conflicts as well (for an overview, see Quong, 2013).

Kahan et al. suggest that the cultural cognition theory does have important implications for how policy may be made if it is to be legitimate on the liberal conception. On Kahan et al.'s explication of the liberal ideal, it consists in an "injunction that the law steer clear of endorsing a moral or cultural orthodoxy" (Kahan et al. 2006, p. 1106). They then go on to suggest that "it is questionable whether risk regulation should be responsive to public demands for regulation, since these express cultural worldviews"—that is, exactly the kind of views that it would be wrong for policy to endorse according to the liberal ideal. So even though Kahan et al. seem to believe that the dubious factual basis of risk-related policy preferences is not sufficient to strip them of their claim to democratic responsiveness, they suggest that there are *liberal* reasons for making policy nonresponsive to such preferences.

Kahan et al. do not elaborate what they mean by "endors[ing] a moral or cultural orthodoxy." But since they cite the writings of Bruce Ackerman and John Rawls in support of the principle, let us assume that the following, common liberal idea is what Kahan et al. have in mind: legitimacy requires policies to be justified only with reference to reasons that are public, in the sense that all reasonable citizens agree that these reasons count in favour of (or against, as the case may be) policies. Now suppose we have identified an exhaustive set of such reasons, and that these are the only ones actually given weight in the policy-making process. Obviously policies at the same time will reflect factual assumptions about how much various policies realize the values defined by public reasons. If the cultural cognition theory is correct, factual assumptions are not value neutral, since each set of factual assumptions expresses a cultural worldview.

What is the import of this for liberal legitimacy? The basic question is what it means that factual assumptions express worldviews and when that would be a problem. Suppose a policy is justified only on the basis of public reasons and the

facts. In that case, it seems to us strange to say that the policy in question is illiberal merely because the facts are (coincidentally) endorsed by adherents of one cultural worldview. 'Expressing a worldview' must refer to something more substantial than this kind of *correspondence to* a worldview if it is to be a liberal problem. This reflects the basic assumption we endorsed earlier—namely, that the facts, and scientific methods of establishing facts, ought to have priority in policy making.

Perhaps the problem arises only in cases where there is genuine uncertainty about what the facts are. Suppose that the scientific evidence concerning gun control allows for believing either that gun control does prevent deaths from firearm accidents and crimes (call this p) or that gun control does not prevent such deaths ($\varnothing p$).[6] And suppose further that the public reasons bearing on the case are such that if p is true, then gun control should be implemented, and if $\varnothing p$ is true, gun control should not be implemented (e.g., because there is a presumption of liberty). So policy *must* endorse either p or $\varnothing p$, in the sense that one policy follows from p and a different policy follows from $\varnothing p$. Supposing that p reflects the cultural worldview of one group G_1 and that $\varnothing p$ reflects the worldview of G_2, it seems that policy must endorse one group's worldview although the other group's view is not in conflict with science.

Suppose that one thinks that basing policy on either of p or $\varnothing p$ would be illiberal. Such a view would run into the following problem: it is a plausible requirement for any criterion of legitimacy that at least one policy is legitimate. But in the example given here, we must either say that *both* policies are legitimate or that *neither* policy is legitimate, since they are symmetrically situated with respect to their basis in both public reasons and factual assumptions. Since the view that neither policy is legitimate is not a viable option, we must say that both policies are legitimate. Consequently, G_1 does not have a viable complaint that a no-gun-control policy is illegitimate, although it does in one sense express the cultural worldview of G_2—and similarly G_2 has no legitimacy complaint against gun control.

Another possible interpretation of what it means that a policy preference expresses a worldview is that the worldview is the *real, causal explanation* for why a certain person or group has the preference. On this reading, calls for regulation of a given risk, although seemingly justified by reference to public reasons, are really caused by "an unjust desire to use the expressive capital of the law for culturally imperialist ends" (Kahan et al., 2006, p. 1107). Suppose the policy in question is above board in the sense that some combination of public reasons and scientifically acceptable factual assumptions would justify the policy. Would the fact that this legitimate rationale is not the real reason why the policy is implemented constitute a legitimacy problem? The assumption here is that the group implementing the policy sincerely (and correctly) believes that the policy has a legitimate rationale, a fact that they exploit in order to implement a policy that they desire in any case. Such a group could be accused of an unattractive opportunism. But this does not constitute a legitimacy problem on the orthodox inter-

pretations of the liberal legitimacy criterion.[7] The liberal criterion stresses the importance of all groups *being able to* reasonably accept the policy. Since the policy here is *ex hypothesi* justifiable based on a set of normative assumptions and a set of factual assumptions, both of which are reasonable (i.e., the set of public reasons and the set of scientifically accepted facts, respectively), all groups are able to reasonably accept the policy. It would be unreasonable for a group to demand that the factual assumptions best expressing *their* worldview be the basis of the law rather than another set of reasonable factual assumptions.

We conclude, then, that the fact that factual beliefs express cultural worldviews in the way the cultural cognition theory has revealed does not entail any problems from the point of view of the liberal conception of legitimacy in cases where policies are justifiable based on reasonable normative and factual beliefs.

3.3. Deliberation

In the previous section we discussed public reason as (a part of) a substantive account of policies' legitimacy. We were thus concerned with whether a certain class of reasons provide sufficient justification for a policy. But 'public reason' is also frequently used to refer to a certain norm of deliberation. Here, the concern is not so much whether a policy *could* be justified with reference to agreed-upon, public reasons, but what reasons we may make appeal to in the process of policy making—in public and parliamentary debate, in the civil service, and in courts. According to the deliberative norm of public reason, citizens, politicians, judges, and others may appeal only to reasons that are neutral between reasonable conceptions of the good. The idea, then, is to remove all appeals to contested worldviews from the public arena.

Kahan (2007) takes issue with this public-reason norm. On Kahan's reading, the public-reason norm has two rationales: First, it *disciplines* those in power by demanding that they pursue only policies that they sincerely believe are supported by public reasons. And second, it *protects* those out of power by ensuring that laws are such that they can accept them without thereby denouncing their vision of the good life (Kahan, 2007, p. 129). But, according to Kahan, the cultural cognition theory reveals that the public-reason norm fails to produce either of its promised effects. The demand for secular justifications does not prevent those in power from imposing their vision of the good on society, since even the sincerely held belief that a policy promotes the public good reflects a cultural worldview. And the demand does not ensure that political losers accept policies enacted by their opponents either. More likely, they will interpret opponents' arguments for those policies as disingenuous and reflecting a "smug insistence of their adversaries that such policies reflect a neutral and objective commitment to the good of all citizens" (Kahan, 2007, p. 131).

Kahan suggests that the public-reason norm be replaced with its polar opposite, which he calls the "expressive overdetermination" norm. According to this norm, justifications of policies in the public forum should not avoid references to

contested worldviews and conceptions of the good—they should instead attempt to show how the relevant policy promotes the substantial cultural commitments of all groups. Casting this in Rawlsian terms, we might say that the desire for overlapping consensus among adherents of rival comprehensive views should not lead us to ban reference to the content of these comprehensive views—say, to religious values, strongly egalitarian ideals, or free-market principles. Instead we should attempt to show that all of these values, in all their comprehensive thickness, support some policy (Kahan, 2007, pp. 131-132). The proposal builds on research from social psychology on self-affirmation. The kinds of biases in processing of evidence highlighted by cultural cognition theory stem from a motivation to defend one's identity by defending factual beliefs perceived to be important to the groups with which one identifies. Self-affirmation research has shown that these defensive motivations, and therefore the biases, are decreased when aspects of subjects' personal or social identities are affirmed—for example, by allowing them to write a brief essay outlining a value or group membership that is important to them. In effect, affirmation provides an identity buffer such that one can afford to lower one's cognitive defenses. People whose identities have been affirmed are thus more objective in assessing evidence and arguments, either written or during discussions (Sherman & Cohen 2002; Cohen et al. 2000; Correll et al. 2004; Cohen et al. 2007). Expressive overdetermination takes advantage of this: highlighting that a policy is in line with the values of one's group is taken to be one way of affirming the importance of that value. If so, one can expect people to be less biased in assessing the risks and benefits of the policy. Thus, expressive overdetermination is meant to achieve the goals of having public policy recognized by all groups as legitimate, and of diffusing the intensely conflictual nature of politics.

Kahan et al. (2015) provide direct evidence that expressive overdetermination may be effective. Hierarchical individualists were more likely to rate a study concluding that extant emission limits would be insufficient to avoid environmental catastrophe as valid and to express that climate change posed a high risk if they had previously been exposed to a study suggesting that geoengineering was a necessary element in combating climate change. Since geoengineering does not involve imposing restrictions on free enterprise or suggest that corporate elites are unable to solve collective problems, this framing highlighted that the reality of climate change need not threaten hierarchical individualist values. In fact, these values were affirmed insofar as a privately driven use of technology was cast as necessary to combat climate change. This allowed hierarchical individualists to assess the evidence more objectively without threat to their identity.

The realization that seemingly conflict-diffusing mechanisms, such as the public-reason norm, may in fact not work—or may even be counterproductive—seems to us to be the most directly useful insight for political philosophy that follows from the understanding of cultural cognition. Nevertheless, we do have some misgivings about the expressive-overdetermination norm and about Kahan's dismissal of the public-reason norm.

Let us start with the latter. Is it really true that the public-reason norm fails to deliver on both of its promises? First, consider whether the norm disciplines those in power. The cultural cognition theory shows that the mere fact that those in power *sincerely believe* policies to be supported by public reasons does not ensure that policies *are in fact* so supported. However, it remains plausible that the public-reason norm *contributes* to the aim of liberally legitimate policies. The mere demand that evidence that a certain policy promotes publically recognized goods must be produced will likely provide some constraints on what policies will be implemented by conscientious adherents to the public-reason norm. Although processing of evidence is culturally biased as described above, there are limits to the degree to which people can pick the evidence that suits them (Kunda 1990). Furthermore, there is evidence (Vinokur & Burnstein 1978; Luskin et al. 2012; Cohen et al. 2007) that deliberations between adherents of conflicting worldviews or ideologies brings these people closer together with respect to their factual beliefs. Insofar as the willingness (and perhaps even active desire) to engage with the arguments of political opponents is also a part of the public-reason norm, it has resources to diffuse the kind of conflicts that arise from cultural cognition as well.

Second, consider the protective aim of public reason. A corollary of the above is that the public-reason norm does not plausibly increase the likelihood that liberally illegitimate policies will be enacted (rather, it plausibly lowers that likelihood). So there is no reason to think that losers are less well protected under the public-reason norm than in the case where appeals to "thick" values can freely be made. What the cultural cognition theory shows with respect to losers is that they are likely to feel aggrieved even when they have no right to do so (since policies are legitimate). So only if the goal is to ensure *actual* acceptance on the part of losers does the public-reason norm fail. This is a worthy goal, but less important than protecting them from illiberal cultural imperialism.

Now what about the expressive-overdetermination norm as an alternative? Supposing that Kahan accepts the standard public-reason account of *legitimacy*, expressive overdetermination does not contribute to the legitimacy of policies. On that account, a policy that is *in fact* justifiable by reference to public reasons only is legitimate. The fact that a group falsely believes that a policy is not so justifiable does not alter the fact that it is. Furthermore, expressive overdetermination does not contain any resources that increase the likelihood of policies that are in fact legitimate, or any resources that lower the likelihood of policies that are not legitimate.

There are nonstandard accounts of public reason that may be more conducive to seeing expressive overdetermination as having a legitimacy-creating role. On the convergence view of Gerald Gaus, for example, legitimacy requires that each citizen be able to support the policy from within her own total view (Gaus, 2011; Gaus and Vallier, 2009). Gaus's main argument for viewing legitimacy in this way is that reasons that people hold as part of their comprehensive view, but which are not public reasons, may *defeat* the justification of a policy based on

public reasons. Consequently, people would not be able to sincerely accept the imposition of that policy. This line of argument meshes well with the protection function of deliberative norms as Kahan describes it. However, the convergence view faces the potential problem that there will often not be a policy that can gain support from all comprehensive points of view. Additional principles for determining what policies are legitimate in such cases are then needed. Gaus has developed an elaborate theory for this purpose, but nothing Kahan has written suggests that he would go along with Gaus in this regard. If a legitimacy-incurring role for expressive overdetermination is to be grounded in an account like Gaus's, much work remains to be done to flesh out the theory.

Return now to more standard accounts of public reason. Since expressive overdetermination does not contribute to policies' legitimacy, it seems that the expressive-overdetermination norm can be justified only instrumentally, as a means to an end. The most immediately obvious end that the norm serves is to ensure actual acceptance of policies by all groups. And actual acceptance is presumably valuable because it realizes the aims of disciplining the powerful and protecting the powerless. But there is some reason to be sceptical that actual acceptance will realize those goals. Expressive overdetermination can be used to secure acceptance from groups without substantially respecting their values. Consider an example that Kahan points to—namely, French abortion law. This law gives women access to abortions, but in order to secure acceptance from conservatives, this access is available only in an "emergency" (Kahan, 2007, p. 132). However, no criteria for what constitutes an emergency were included, and no questioning of a woman's own declaration that an emergency exists is allowed. In effect, then, the emergency clause is substantively empty, and was included only for its symbolic meaning. While this construction did succeed in creating a consensus on the policy, it is hard to see why those who believe in any serious way that non-emergency abortions is a problem *should* have been satisfied with this law.[8]

On the other hand, expressive overdetermination might be used for another end—namely, to enable people holding conflicting views to converge on the facts (cf. the climate change study described above), and hence to diffuse or avoid cultural conflict over factual questions. Kahan et al. have provided strong evidence that the public-reason norm does not realize this goal particularly well, and that a norm of expressive overdetermination can (perhaps somewhat counterintuitively) realize the goal better. However, and as Kahan himself recognizes, expressive overdetermination is merely one tool for achieving fact convergence.[9]

4. CONCLUSION

We have argued above that the psychological facts of risk perception are complex. Divergences between experts and lay citizens are sometimes at least partly a reflection of lack of scientific literacy and overreliance on heuristics on the part of some citizens. But in other cases, cultural worldviews seem to be

behind differences of opinion over what is risky and what is not. And in fact those seem to be the cases that are most interesting politically, such as global warming, environmental issues, or GM foods (in Europe).

However, we have also argued that the fact that faulty beliefs express people's basic values has few implications for how liberal-democratic states should go about formulating policy with respect to putatively risky activities and technologies. Contrary to what proponents of cultural cognition argue, the fact that risk perceptions express cultural worldviews does not give us stronger reasons than we would otherwise have for making policy responsive to such perceptions. Similarly, the fact that factual beliefs about risks express visions of the ideal society does not undermine the legitimacy of using scientifically accepted facts as the basis for policy making.

This largely means that we are stuck with the responsiveness dilemma that we identified in our discussion of Sunstein's view: if policy is insulated from the people, we risk being underresponsive to citizens' values, and if policy is made in a more populist manner, we risk overresponsiveness to false beliefs. However, the cultural cognition theory does provide some important insights into how this dilemma can be resolved. It supports the case for using structured deliberation methods to determine what citizens' preferences would be if they were to come to accept scientific facts. And it provides significant guidance for those of us who want to reform political discourse in a way that enables reasonable discussion of policies based on common acceptance of the relevant facts.

NOTES

1 This is a bit of a simplification. Bounded rationality is also consistent with mistakes being due to a lack of information or to social processes such as information cascades or group polarization (Anderson & Holt 1997; Moscovici & Zavalloni 1969; Sunstein 2002).

2 However, the ecological rationality programme of Gerd Gigerenzer and colleagues points out that, far from being a source of ubiquitous bias, heuristics can often be beneficial, providing "fast and frugal" decision procedures that can rival or even beat analytical approaches (Gigerenzer & Goldstein 1996; Czerlinski et al. 1999).

3 Discussing the phenomenon of "nudging," where policy proposals have similarly been justified with reference the psychology of heuristics and biases, one observer suggests that the people arguing for such policies "generally believe that social policy should aim to satisfy purified preferences" (Hausman, 2016). "Purified" preferences are preferences people would have had, if they had not been the victims of biases.

4 Here, and generally in the paper, we use the word 'actually' to indicate what is the case in the *actual* world, as that concept is typically used in possible-worlds ways of speaking of counterfactuals and alethic concepts such as necessity and possibility. That is, we use 'actual' to indicate what is currently the case in the world in which we find ourselves.

5 There are two likely mechanisms at play: First, people form beliefs about whether a given type of risk regulation is desirable, based directly on their cultural worldview. That is, there is a direct causal link from worldviews to policy preferences. Second, people form *factual* beliefs—about the numerical magnitude of risks—through motivated cognition, wherein cultural worldviews affect people's assessment of the evidence concerning the riskiness (or safety) of an activity. Here, the causal link goes from worldviews to assessment of evidence, and thus to pure factual beliefs, and then in a second step from those factual beliefs to policy preferences. Since pure factual beliefs are not easily disentangled from policy preferences (see, e.g., Kahan & Slovic, 2006, pp. 166-168), it is difficult to test which of these mechanisms is the dominant one. However, in a study of self-defence cases, Kahan and Braman found support for the view that "the influence that values exert over outcome judgments is mediated by the impact of the commitments on individuals' perceptions of the facts" (Kahan and Braman, 2008, p. 45)—i.e., for the second mechanism.

6 Kahan's own treatment of this case (2007, pp. 120-122) seems to imply that this is the case. However, more recent evidence suggests that gun control does, in fact, lower gun-related injuries and fatalities. See Santaella-Tenorio, Cerdá, Villaveces and Galea, 2016.

7 There is some debate among theorists of public reason regarding the appropriate role of *sincerity*. Some views within this debate hold it to be a requirement for legitimacy that public reasons are the actual motivation for people's advocacy of a given policy (see Schwartzmann, 2011, pp. 387-390). Kahan et al. may of course defend their position by endorsing such a view, but in doing so they would no longer be able to claim the support of the liberal principle of legitimacy *tout court*.

8 Of course, one might not think that the anti-abortion party's views were such that they ought to be respected on a liberal view of legitimacy—but the example is illustrative of the risk that expressive overdetermination can be used to manipulate groups to accept policies that illegitimately trample their values.

9 http://www.culturalcognition.net/blog/2014/4/5/cognitive-illiberalism-expressive-overdetermination-a-fragme.html (comment by "dmk38").

REFERENCES

Anderson, Elizabeth, *Value in Ethics and Economics,* Cambridge, MA, Harvard University Press, 1993.

Anderson, Lisa R. & Holt, Charles A., "Information Cascades in the Laboratory", *American Economic Review*, volume 87, issue 5, 1997, pp. 847-862.

Benjamin, Daniel J., Brown, Sebastian A. & Shapiro, Jesse M., "Who is 'behavioral'? Cognitive ability and anomalous preferences", *Journal of the European Economic Association*, volume 11, issue 6, 2013, pp. 1231–1255.

Blackburn, Simon, *Practical Tortoise Raising: and Other Philosophical Essays,* Oxford, Oxford University Press, 2010.

Brooks, Clem & Manza, Jeff, "Social Policy Responsiveness in the Developed Democracies", *American Sociological Review,* volume 71, issue 3, 2006, pp. 474-494.

Cohen, Geoffrey L., Sherman, David K., Bastardi, Anthony, Hsu, Lillian, McGoey, Michelle & Ross, Lee, "Bridging the Partisan Divide: Self-Affirmation Reduces Ideological Close-Mindedness and Inflexibility in Negotiation", *Journal of Personality and Social Psychology*, volume 93, issue 3, 2007, pp. 415-430.

Cohen, Geoffrey L., Aronson, Joshua and Steele, Claude M. "When Beliefs Yield to Evidence: Reducing Biased Evaluation by Affirming the Self", *Personality and Social Psychology Bulletin*, volume 26, issue 9, 2000, pp. 1151-1164.

Cook, John et al., "Quantifying the consensus on anthropogenic global warming in the scientific literature", *Environmental Research Letters*, volume 8, issue 2, 2013, pp. 1-7.

Correll, Joshua, Spencer, Steven J. and Zanna, Mark P. "An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength", *Journal of Experimental Social Psychology*, volume 40, issue 3, 2004, pp. 350-356.

Czerlinski, Jean, Gigerenzer, Gerd & Goldstein, Daniel G., "How good are simple heuristics?", *in* Todd, Peter M., Gigerenzer, Gerd & ABC Research Group (eds.), *Simple Heuristics that make us smart*, Oxford, Oxford University Press, 1999.

Dawson, Erica, Gilovich, Thomas & Regan, Dennis T., "Motivated Reasoning and Performance on the Wason Selection Task", *Personality and Social Psychology Bulletin*, volume 28, issue 10, 2002, pp. 1379-1387.

Douglas, Mary. & Wildavsky, Aaron B., *Risk and culture: an essay on the selection of technological and environmental dangers*, Berkeley, University of California Press, 1983

Evans, Jonathan St. B.T., "Dual-processing accounts of reasoning, judgment, and social cognition", *Annual Review of Psychology*, volume 59, issue 1, 2008, pp. 255-278.

Evans, Jonathan St. B.T. & Stanovich, Keith E., "Dual-Process Theories of Higher Cognition: Advancing the Debate", *Perspectives on Psychological Science*, volume 8, issue 3, 2013, pp. 223-241.

Finucane, Melissa L., Alhakami, Ali, Slovic, Paul & Johnson, Stephen M., "The affect heuristic in judgments of risks and benefits", *Journal of Behavioral Decision Making*, volume 13, issue 1, 2000, pp. 1-17.

Folkes, Valerie S., "The Availability Heuristic and Perceived Risk", *Journal of Consumer Research*, volume 15, issue 1, 1988, pp. 13-23.

Frederick, Shane, "Cognitive Reflection and Decision Making", *Journal of Economic Perspectives*, volume 19, issue 4, 2005, pp.25-42.

Gaus, Gerald, *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World,* Cambridge, Cambridge University Press, 2011.

Gaus, Gerald & Vallier, Kevin, "The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Assymmetry, and Political Institutions", *Philosophy and Social Criticism,* volume 35, issue 1, 2009, pp. 51-76.

Gigerenzer, Gerd & Goldstein, Daniel G., "Reasoning the fast and frugal way: models of bounded rationality", *Psychological review*, volume 103, issue 4, 2005, pp. 650-669.

Gilens, Martin, "Inequality and Democratic Responsiveness", *Public Opinion Quarterly*, volume 69, issue 5, 2005, pp. 778-796.

Jones, Martin & Sugden, Robert, "Positive confirmation bias in the acquisition of information", *Theory and Decision*, volume 50, issue 1, 2001, pp. 59-99.

Jønch-Clausen, Karin & Kappel, Klemens, "Social Epistemic Liberalism and the Problem of Deep Epistemic Disagreements", *Ethical Theory and Moral Practice*, volume 18, issue 2, 2015, pp. 371-387.

—, "Scientific Facts and Methods in Public Reason", *Res Publica*, volume 22, 2016, pp. 117-133

Kahan, Dan M., "The Cognitively Illiberal State", *Stanford Law Review*, volume 60, issue 1, 2007, pp. 115-140.

—, "Cultural Cognition as a Conception of the Cultural Theory of Risk", *in* Roeser, Sabine, Hillerbrand, Rafaela, Sandin, Per & Peterson, Martin (eds.), *Handbook of Risk Research: Epistemology, Decision Theory, Ethics and Social Implications of Risk,* Dordrecht, Springer, 2012.

—, "Ideology, motivated reasoning, and cognitive reflection", *Judgment and Decision Making*, volume 8, issue 4, 2013, pp. 407-424.

—, "Climate-science communication and the measurement problem", *Political Psychology*, volume 36, issue S1, 2015, pp. 1-43.

Kahan, Dan M., Slovic, Paul, Braman, Donald & Gastil, John, "Fear of Democracy: A Cultural Evaluation of Sunstein on Risk", *Harvard Law Review*, volume 119, 2006, pp. 1071-1109.

Kahan, Dan M. & Slovic, Paul, "Cultural Evaluations of Risk: 'Values' or 'Blunders'?", *Harvard Law Review Forum,* volume 119, 2006, pp. 166-172.

Kahan, Dan M. & Braman, Donald, "The Self-Defensive Cognition of Self-Defense", *American Criminal Law Review,* volume 45, issue 1, 2008, pp. 1-65.

Kahan, Dan M. et al., "Who fears the HPV vaccine, who doesn't, and why? an experimental study of the mechanisms of cultural cognition", *Law and human behavior*, volume 34, issue 6, 2010, pp. 501–516.

Kahan, Dan M., Jenkins-Smith, Hank & Braman, Donald, "Cultural Cognition of Scientific Consensus", *Journal of Risk Research,* volume 14, number 2, 2011, pp. 147-174.

Kahan, Dan M. et al., "The polarizing impact of science literacy and numeracy on perceived climate change risks", *Nature Climate Change*, volume 2, issue 10, 2012, pp. 732-735.

Kahan, Dan M, Peters, Ellen, Dawson, Erica C. & Slovic, Paul, "Motivated Numeracy and Enlightened Self-Government", *Behavioral Public Policy*, volume 1, issue 1, 2017, pp. 54-86.

Kahan, Dan M., Silva, Carol L., Tarantola, Tor, Jenkins-Smith, Hank & Braman, Donald, "Geoengineering and Climate Change Polarization: Testing a Two-Channel Model of Science Communication", *Annals of the American Academy of Political and Social Science,* volume 658, issue 92, 2015, pp. 192-222.

Kahneman, Daniel, *Thinking, Fast and Slow*, New York, Farrar, Strauss & Giroux, 2011.

Kahneman, Daniel, Ritov, Ilana, Jacowitz, Karen E. & Grant, Paul, "Stated Willingess to Pay for Public Goods: A Psychological Perspective", *Psychological Science,* volume 4, issue 5, 1993, pp. 310-315.

Klein, Richard A. et al., "Investigating Variation in Replicability", *Social Psychology*, volume 45, issue 3, 2014, pp. 142-152.

Kunda, Ziva, "The case for motivated reasoning", *Psychological Bulletin*, volume 108, issue 3, 1990, pp.480-498.

Lichtenstein, Sarah, Slovic, Paul, Fischhoff, Baruch, Layman, Mark & Combs, Barbara, "Judged Frequency of Lethal Events", *Journal of Experimental Psychology: Human Learning and Memory,* volume 4, number 6, 1978, pp. 551-578.

Lindeman, Marjaana, "Biases in intuitive reasoning and belief in complementary and alternative medicine", *Psychology and Health*, volume 26, issue 3, 2011, pp. 371-382.

Lord, Charles G., Ross, Lee & Lepper, Mark R., "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence", *Journal of Personality and Social Psychology*, volume 37, issue 11, 1979, pp. 2098-2109.

Luskin, Rober C., O'Flynn, Ian, Fishkin, James S, & Russell, David, "Deliberating across Deep Divides", *Political Studies*, volume 62, issue 1, 2012, pp. 116–135.

Manza, Jeff & Cook, Fay Lomax, "A Democratic Polity? Three Views of Policy Responsiveness to Public Opinion in the United States", *American Politics Research*, volume 30, issue 6, 2002, pp. 630-667.

McCright, Aaron M., Xiao, Chenyang & Dunlap, Riley E., "Political polarization on support for government spending on environmental protection in the USA, 1974-2012", *Social Science Research*, volume 48, 2014, pp. 251-260.

McGarity, Thomas O., "Professor Sunstein's Fuzzy Math", *The Georgetown Law Journal,* volume 90, 2002, pp. 2341-2377.

Hausman, Daniel, McPherson, Michael & Satz, Debra, *Economic Analysis, Moral Philosophy, and Public Policy (3rd ed.)*, Cambridge, Cambridge University Press, 2017.

Moscovici, Serge & Zavalloni, Marisa, "The group as a polarizer of attitudes", *Journal of Personality and Social Psychology*, volume 12, issue 2, 1969, pp. 125-135.

Nickerson, Raymond S., "Confirmation bias: A ubiquitous phenomenon in many guises", *Review of general psychology*, volume 2, issue 2, 1998, pp. 175-220.

Pew Research Center, *Public and Scientists' Views on Science and Society*, January 29, 2015.

—, *The Politics of Climate*, October 4, 2016.

Quong, Jonathan, "Public Reason", *in* Zalta, Edward N. (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2013 edition)*. URL: https://plato.stanford.edu/archives/sum2013/entries/public-reason/

Rawls, John, *Political Liberalism*, New York, Columbia University Press, 2006.

Reyna, Valerie F., "How people make decisions that involve risk: A dual-process approach", *Current Directions in Psychological Science*, volume 13, issue 2, 2004, pp. 60-66.

Roberts, Leslie, "Counting on Science at EPA", *Science*, volume 249, issue 4969, 1990, pp. 616-618.

Roeser, Sabine (Ed.), *Emotions and Risky Technologies*, Dordrecht, Springer, 2010.

Santaella-Tenorio, Julian, Cerdá, Magdalena, Villaveces, Andrés & Galea, Sandro, "What Do We Know About the Association Between Firearm Legislation and Firearm-Related Injuries?", *Epidemologic Reviews*, volume 38, issue 1, 2016, pp. 140-157.

Shafir, Eldar & Leboeuf, Robyn A., "Rationality", *Annual Review of Psychology*, volume 53, 2002, pp. 491–517.

Sherman, David K. and Cohen, Geoffrey L. "Accepting Threatening Information: Self-Affirmation and the Reduction of Defensive Biases", *Current Directions in Psychological Science*, volume 11, issue 4, 2002, pp. 119-123.

Slovic, Paul, Finucane, Melissa L., Peters, Ellen & MacGregor, Donald G. "Risk as analysis and risk as feelings: Some Thoughs about Affect, Reason, Risk and Rationality", *Risk Analysis*, volume 24, issue 2, 2004, pp. 311-322.

Stanovich, Keith E. & West, Richard F., "Individual differences in rational thought", *Journal of Experimental Psychology: General*, volume 127, issue 2, 1998, pp. 161-188.

—, "The Assessment of Rational Thinking: IQ ≠ RQ", *Teaching of Psychology*, volume 41, issue 3, 2014, pp. 265-271.

Sunstein, Cass R., "The Law of Group Polarization", *Journal of Political Philosophy*, volume 10, issue 2, 2002, pp. 175-195.

—, *Risk and Reason*, Cambridge, Cambridge University Press, 2002.

—, *Laws of Fear: Beyond the Precautionary Principle,* Cambridge, Cambridge University Press, 2005.

—, "Misfearing: A Reply", *Harvard Law Review,* volume 119, no. 4, 2006, pp. 1110-1125.

—, *Valuing Life: Humanizing the Regulatory State,* Chicago, The University of Chicago Press, 2014.

Taber, Charles S. & Lodge, Milton, "Motivated skepticism in the evaluation of political beliefs", *American Journal of Political Science*, volume 50, issue 3, 2006, pp. 755-769.

Todd, Peter M., Gigerenzer, Gerd & the ABC Research Group, *Ecological Rationality: Intelligence in the World*, Oxford, Oxford University Press, 2012.

CONSENSUS AND LIBERAL LEGITIMACY: FROM FIRST TO SECOND BEST?

XAVIER LANDES
STOCKHOLM SCHOOL OF ECONOMICS IN RIGA

ABSTRACT:

In this article, consensus, defined as the consent of all citizens, is argued to be the first best for part of the liberal tradition on political legitimacy. Consensus would be the foundation of the liberal society that, when out of reach, needs to be approximated through, for instance, voting (majority rule). I build on the timid attempts in political theory at using the theorem of the second best as a tool to settle difficult decision making in applied political theory. More precisely, I defend the view that consensus would be the first best for part of the liberal tradition on political legitimacy. Furthermore, I illustrate how moral, factual, (and, incidentally, epistemological) disagreements may create second-best problems, especially in terms of stability. Finally, I spell out some reasons why such problems directly affect a liberal order, on pragmatic grounds. The final purpose is to contribute to the literatures on legitimacy and on the application of the theorem of the second best outside economics. The key idea is to stress that much more work has to be done for "importing" the theorem of the second best into political theory.

RÉSUMÉ :

Dans cet article, le consensus, défini comme le consentement de tous les citoyens, est présenté comme remplissant la fonction d'optimum de premier rang [*first best*] pour une partie de la tradition libérale quant à la légitimité politique. Le consensus fonderait la société libérale qui, lorsque hors d'atteinte, nécessiterait d'être approximer au travers, par exemple, du vote (*règle majoritaire*). Je m'appuie sur les timides tentatives en théorie politique d'utiliser le théorème de l'optimum de second rang [*second best*] comme outil pour résoudre des prises de décision difficiles en théorie politique appliquée. Plus précisément, je défends l'idée que le consensus serait l'optimum de premier rang pour une partie de la tradition libérale quant à la légitimité politique. De plus, j'illustre comment les désaccords moraux, factuels (et, incidemment, épistémologiques) peuvent créer des problèmes d'optimum de second rang, en particulier en termes de stabilité. Enfin, j'expose quelques raisons pour lesquelles de tels problèmes affectent directement un ordre libéral, d'un point de vue pragmatique. Le but final est de contribuer aux littératures sur la légitimité et sur l'application du théorème d'optimum de second rang à l'extérieur des sciences économiques. L'idée principale est de souligner qu'il reste beaucoup de travail afin d'« importer » le théorème d'optimum de second rang en théorie politique.

Consensus holds a preeminent position in philosophy (Rescher, 1993), especially within liberal theory of political legitimacy (for a critical discussion of the importance of consensus in political theory and a defence of compromise, please refer to Bellamy, 1999; Thrasher & Vallier, 2013; Vallier, 2011). Rescher characterizes consensus as a "widespread and pervasive agreement" (Rescher, 1993, p. 45). *Merriam-Webster* defines it as "a general agreement about something," "an idea or an opinion that is shared by all the people in a group" ("unanimity" is offered as a synonym).[1] Consensus-oriented theories claim that political decisions and institutions are legitimate insofar as all citizens (would at least hypothetically) consent to them in one way or another.[2]

The underlying idea is that the legitimacy of political decisions and institutions depends on reasons and justifications that should be understandable and hypothetically agreed on by all citizens (Vallier 2011, p. 262). The influence of consensus is visible in concrete decision-making procedures too, such as United Nations General Assembly resolutions, Occupy Wall Street decision-making procedure, and Quakers' decision-making procedure. The Nice Treaty (2001) set consensus as the decision rule for the EU Council before the Lisbon Treaty (2007) replaced it by qualified majority.

Such influence is visible not only in consensus-based mechanisms, but also in the positively valued proximity in political theory between consensus and other principles or procedures, most notably majority rule and its variants. In liberal-democratic theory, part of the justification for majority rule is that, when consensus is impossible, the best alternative is to approximate it by getting the consent of as many individuals as possible or of a majority. Qualified majority is a closer approximation than simple majority and, thus, strengthens the legitimacy of ensuing decisions.

In economic parlance, consensus appears as the first best for public decision making for part of the liberal tradition (equivalent to the Pareto optimum). According to mainstream economics, even if the first best is unattainable because one of its conditions is constrained, achieving other conditions would still be desirable. Thus, if we accept the relevance of economic concepts for political theory, and the possibility that consensus could be presented as a first best, then a manner to look at situations where consensus is not reachable is to call to the second-best theorem (TSB hereafter).

> The general theorem for the second best optimum states that if there is introduced into a general equilibrium system a constraint which prevents the attainment of one of the Paretian conditions, the other Paretian conditions, although still attainable are, in general, no longer desirable. (Lipsey and Lancaster, 1956-1957, p. 11)

The TSB captures the idea that there are situations where it is preferable not to try getting as close as possible to the first best or to try fulfilling as many first-best conditions as possible (Gaus, 2016, pp.14-15). As later noted by

Richard Lipsey, "a 'second best situation' referred to any situation in which the first best was unachievable" (Lipsey, 2007, p.352). Another point needs to be stressed: the departure should be general (i.e., all first-best conditions dropped).

Authors have underlined the fecundity of the TSB for political, moral, or legal theory (Coram, 1996; Estlund, 2014; Gaus, 2016; Lipsey, 2006; Margalit, 1983; Räikkä, 2000; Rebello, 2002; Rodrik, 2008; Sher, 1997), but often in brief mentions without elaborated discussion of how to interpret Lipsey and Lancaster's theorem. Attempts at deciphering TSB's meaning and implications for political theory have been rare [Juha Räikkä (2000) is one of the very few exceptions].[3]

This article is an exercise in applied political theory that partially addresses such a lack by starting from Lipsey and Lancaster's formulation and tackling an original issue: consensus as a liberal first best. It initiates an evaluation of TSB's usefulness for liberal legitimacy, without providing a full-fledged analysis. The goal is to contribute to the few initiatives that aim at evaluating the promises and challenges of importing the TSB in political theory. Thus, I hope to contribute to discussions about the process of importing the TSB itself, the relevance of the theorem for political theory (which question remains under-studied) while applying the theorem to a specific field: political legitimacy.

In the first section, I argue that consensus stands as the *first best* for part of liberal theory. In welfare economics, the first best is the Pareto optimum[4] resulting from the fulfilment of "first-best conditions." Microeconomics textbooks introduce perfect competition as the result of conditions (homogeneity of goods, atomicity of actors, perfect information, free entrance/exit from the market, and perfect mobility of productive factors). First-best theories stipulate that if one condition is unattainable, the others should be fulfilled anyway. For example, if homogeneity of goods cannot be achieved, fulfilling the other conditions will still enhance social welfare. Structural reforms imposed by the International Monetary Fund on developing countries follow this premise: if all first-best conditions for perfect competition cannot be met, it is still preferable to implement some rather than none (Rodrik 2008).

The TSB is useful because it provides the vocabulary for identifying consensus as the first best for part of the liberal tradition of political legitimacy. According to this tradition, if the consent of all individuals on a complete set of propositions (which is the only condition retained here for the discussion) cannot be obtained, it is still desirable to approximate it. Two versions have been popular among liberal theorists: tacit and hypothetical consent. John Rawls's overlapping consensus is another illustration of this strategy of approximating a full-blooded consensus in the original position through an agreement on a deflated set of conditions.

The second section builds on Richard Bellamy's criticism of the central role of consensus in political philosophy, including liberal theory. His criticism tackles

the assumption of and the search for consensus on moral matters in pluralistic societies. My aim is to show how consensus raises similar issues regarding factual and epistemological disagreements.[5] A liberal society depends on consensus not only for political matters, but also for facts and epistemological rules.

The third section is about reasons for taking seriously factual and epistemological disagreements. The territorial claims of Canadian First Nations illustrate how such disagreements *may* challenge consensus-based theories of legitimacy by creating second-best issues. Such disagreements are endemic in liberal democracies and, more importantly, they offer reasons for modifying epistemological standards, especially for stability reasons, which raises second-best issues.

The fourth section presents conditions under which factual and epistemological disagreements undermine a liberal society based on consensus as the first best. The debate on GMOs illustrates such a challenge. I claim that considerations of trust and stability offer a solid ground for considering that factual and epistemological disagreements *may* create second-best issues for consensus-based approaches to legitimacy. Then, I present two issues that political theorists interested in "importing" the TSB into their field need to address. The first has to do with identifying a general departure from first-best conditions that is liberal. The second is to take seriously stability as a constraint on potential second bests. I end the article by noting that the TSB intersects with debates on the value of contextualism and pragmatism in political theory.

1. CONSENSUS AS LIBERAL FIRST BEST

As shown by Bellamy (1999), some liberals (and other political theorists) value consensus as the first best for the legitimacy of political decisions, institutions, or regimes. This value implies that when the consent of all is out of reach, the inclination is to approximate it. This also implies that political decisions or institutions become more legitimate the closer they get to consensus (which explains, for example, that a qualified majority appears to offer a stronger legitimacy to a political decision than a simple majority).

Few words about this process of approximation are necessary here. On matters of political legitimacy, the first best could be understood as encompassing consensus (i.e., everyone truly agrees/consents to all features of a decision, policy, institution, or regime). Hence, the first best actually has two dimensions: extension (everyone truly agrees/consents to a decision) and depth (everyone agrees upon all the features of the decision). Then, if consensus is out of reach, approximation could be to obtain the consent of either as many people as possible on a fixed set of propositions (e.g., majority rule) or all people on an altered set of propositions (e.g., compromise or Rawls's overlapping consensus).[6] My discussion will mostly focus on (without being restricted to) the extension dimension. Consensus could also be used hypothetically—that is, as a justification for political principles (for instance, under idealized and counterfactual conditions).

This does not mean that consensus is the core of all liberal conceptions of legitimacy. It just means that consensus, like freedom or equality, is a central concept for many liberals. The liberal attachment to consensus flows from a commitment to individual autonomy—more precisely, to a regime of equal liberty, where every citizen has access to the same maximum level of fundamental liberties, including political ones (e.g., expression, autonomy).

Consensus is about individuals' consent—that is, a situation where all individuals agree to a decision or a proposition. In a political regime, it applies to the justifications of public decisions and institutions. Indubitably, some conceptions of political legitimacy are independent of individual consent or agreement, such as conservative conceptions that are based on tradition and religion, or technocratic conceptions that are based on expertise or bureaucracy. But the specificity of liberal legitimacy is to be rooted in the fundamental idea that, ultimately, political decisions and institutions are legitimate insofar as individuals who are affected by them accept them in a less or more formal manner (i.e., consent or agree to them).

Then, *consensus characterizes situations where all the parties to the decision-making process agree on or consent to decisions, institutions, or political regimes.* Additionally, consensus could be interpreted as the agreement or consent of all *bearing on all the features of the decisions, institutions, or political regimes.* (This paper mostly focuses on the first dimension.)

Jeremy Waldron clearly expresses this importance of the consent of all for liberal theory.

> Liberals are committed to a conception of freedom and of respect for the capacities and the agency of individual men and women, and that these commitments generate a requirement that *all aspects of the society* should either be made acceptable or be capable of being made acceptable to *every last individual*. (Waldron, 1987, p. 128, emphasis added)

One of the reasons why, within a liberal society, social rules may not be seen to be as restricting of individual freedom as in other political regimes lies in consensus *qua* consent of all. The fundamental liberal thesis is that "a social and political order is illegitimate unless it is rooted in the *consent of all those who have to live under it*" (Waldron, 1987, p. 140, emphasis added). In other words, consensus grounds the legitimacy of political decisions and institutions in a liberal *order* (i.e., regime or society).

The idea could be found in Locke's *Second Treatise* (II, section 95), where individual consent is placed at the heart of liberal societies: "Men being, as has been said, by nature all free, equal, and independent, no one can be put out of this estate, and subjected to the political power of another, without his own consent" (Locke, 2003, p. 141). Even if Locke immediately introduces the majority rule for enacting the consent of all after the initial establishment of the political community, which confirms the status of the majority rule as a substitute for

consensus, the foundation of the political community and the ultimate justification for the political regime yet lie in the consent of all.[7]

The idea of political legitimacy as founded in the consent of *all* is the backbone of contractarianism (Vallier, 2011). It expresses itself through the principle of popular sovereignty (e.g., as originated in Hobbes, Locke, and Rousseau).[8] Indeed, consensus can be interpreted as being *the* first best—*the* optimum that serves as a benchmark for suboptimal alternatives—of legitimacy for part of the liberal tradition.

As a practical matter, consensus is rarely reachable and few theorists seriously consider that legitimacy requires the actual, explicit consent of all to any democratic decision or institution. They have instead taken one of these two paths: *tacit* or *hypothetical consent.* Both are altered versions of consensus, retaining some of its features. Their value in terms of legitimacy largely stems from the extent to which they approximate the first best.

Still in his *Second Treatise* (II, 119), Locke (2003, pp. 152-153) contrasts express and tacit consent. If the former is the unmistakable expression of one's liberty, the second is simply assumed when individuals benefit from the security offered by the government. The rationale behind tacit consent is quite clear: when actual consent of all is impossible to obtain, the strategy is to retreat to a solution that approximates the first best.

The problem with *tacit consent,* in the absence of further qualifications, is that it may express less consent than imposed choice or restricted option, especially in non-ideal settings. Moreover, tacit consent alone (i.e., unqualified) is too loose a criterion for political legitimacy. It is too encompassing. If individuals can be shown to benefit from the security offered by an, even totalitarian, government, they could be assumed to be tacitly consenting.

The second path is *hypothetical consent*. In that case, there is no need to identify some eventual benefit. Consent is assumed under some hypothetical conditions and settings: political decisions and institutions are legitimate if it can be demonstrated that individuals would consent to them under specific conditions [e.g., of rationality (Kant, 1996)].

John Rawls's original position (Rawls, 1971, p. 13; Rawls, 1980) illustrates hypothetical consent. The original position is a device that hypothetically generates the consent of all (or shows that such consent is a reasonable assumption) on the principles of justice that regulate the basic structure of the society. From *A Theory of Justice* to *Political Liberalism,* the role of consensus is not restricted to the original position and the "values of political justice"; it includes the "values of public reason" too (Rawls, 1993, p. 224). The "liberal principle of legitimacy" stipulates that

> our exercise of political power is fully proper when it is exercised in accordance with a constitution the essentials of which *all citizens* as free

and equal may reasonably be expected to endorse in the light of princi-
ples and ideals acceptable to their common human reason… Only a polit-
ical conception of justice that *all citizens* might be reasonably expected
to endorse can serve as a basis of public reason and justification. (Rawls,
1993, p. 137, emphasis added)

However, securing the consent of all hypothetical agents and securing the
consent of all actual individuals are two different things. The latter individuals
know their position in the society, hold religious or moral views, belong to
socioeconomic categories, and so forth. They might be tempted to adopt self-
serving principles. Moral pluralism is also a serious obstacle for consensus-based
legitimacy. It strips consensus of part of its appeal.

The issue is not limited to pursuing an ideal in a constrained situation. The exis-
tence (or reasonable assumption) of pervasive pluralism undermines the value
and traction exerted by consensus. Rawls's answer is to propose to ground legit-
imacy on "overlapping consensus" (Rawls, 1993, pp. 133-172), which is a light
version of full-blooded consensus. Despite the "fact of pluralism," Rawls consid-
ers that consensus is still possible among individuals holding "not unreasonable
comprehensive doctrines" (Rawls, 1993, p. 140). The consensus is limited to
the values of public reason, while "it is left to citizens individually…to settle
how they think the values of the political domain are related to other values in
their comprehensive doctrine" (Rawls, 1993, p. 140).[9]

In short, consensus plays a central role in liberal legitimacy for some authors
(e.g., for Rawlsians), at least when it concerns the basic structure of the society
and public reason. In other words, it is *the* first best: the legitimacy of political
decisions or institutions in a liberal regime depends on the actual or hypotheti-
cal consent of all individuals. When consensus is not attainable under reasonable
assumptions, the solution is to approximate it, by retaining some first-best condi-
tions: the consent of all is kept, but, for instance with Rawls, restricted to a
limited set of principles or values.[10] This restriction stems from a deflationist
strategy that distinguishes metaphysical and political principles, insulates the
latter from moral controversies, and searches for a consensus on these principles
(and not on a full-blooded conception of the society).

Consensus is more than a theoretical concern for liberals. It also pervades
concrete decision-making procedures. The United Nations General Assembly
uses consensus and unanimity rules. The Council of the European Union used
it until the Lisbon Treaty of 2007. Representative democracy is another exam-
ple. Given the impossibility of consensus-based political decision making, liberal
states rely on the representation principle, the representatives being chosen by
voting. Representative liberal democracy approximates consensus in two
manners: the consent of all is replaced by the consent of elected representatives
and the consent of all elected representatives is usually replaced by the consent
of the majority. The value of the majority rule may be interpreted as deriving
from its approximation of consensus. Finally, qualified majority is a closer
approximation of consensus than simple majority.

In sum, consensus is not only the theoretical first best for part of the liberal tradition of political legitimacy. Its value is not restricted to the choice of principles that regulate the basic structure of the society. It also inspires concrete practices and decisions ranging from representation to voting.

2. CONSENSUS, FROM MORALITY TO EPISTEMOLOGY

According to Bellamy (1999), consensus is at the core of liberalism and large segments of political theory (e.g., Rawls, Walzer, and Hayek). This pre-eminence explains why compromises are sometimes regarded by theorists as being politically, if not morally, inferior. Bellamy stresses that the liberal quest for a *minimal* consensus works best "when principles are largely agreed or else kept off the agenda" (Bellamy, 1999, p. 98). Problems emerge, however, when political values clash. In a liberal society, consensus is reachable only because, by definition, disputable matters have been put off the agenda (e.g., Rawls posits that individuals are capable of distinguishing political from moral principles, and of finding arguments in the latter for supporting the former). Ultimately, Bellamy challenges the assumption that the "underlying consensus is a liberal one" (Bellamy, 1999, p. 99).

While Bellamy underlines the shortcomings of consensus (as the first best) in cases of moral disagreements in a liberal society, my aim is to extend the discussion to factual disagreements (Kappel, 2017). First there is one objection to consider. One may argue that Rawls, and other liberals, place factual and epistemological issues outside the scope of their consensus-based approach. If right, that would entail that discussing consensus for factual issues would be irrelevant.

The objection calls for two remarks. Firstly, Rawls himself is explicit on consensus playing a role outside moral/political disagreements, at least for the foundations of the political community. He indicates that justifications "on matters of constitutional essentials and basic structure" should "appeal only to presently accepted general beliefs and forms of reasoning found in common sense, and the methods and conclusions of science when these are not controversial" (Rawls, 1993, p. 224). Thus, consensus applies to factual and epistemological components as a central part of the justification of liberal institutions and policies (their legitimacy) or, at least, consensus is the condition for allowing factual components to enter the justifications for political decisions and institutions (Jønch-Clausen and Kappel, 2015).

Furthermore, many liberal theorists consider factual and epistemological material as more hospitable to consensus than axiological material since the validity of epistemological rules and facts may be assessed by scientific enquiry. Thus, factual and epistemological disagreements do not benefit from the same attention from liberal theory. If moral controversies on issues such as the purpose of life or religious education might appear so unsolvable and prone to endless disagreements, even among reasonable agents, factual disagreements have been assumed to belong to the "solvable" category, viz. given proper enquiry, they

could (at least potentially) be answered by scientific observation. In other words, if political liberalism conceives moral disagreements as being permanent (and, furthermore, being the unescapable consequence of the freedom of moral agents), in contrast, epistemological and factual disagreements are seen as transient.

Thus, facts and epistemological material are part of the decision-making process and consensus *a priori* applies to this material—that is, to the part that has not been falsified yet. They are filtered out in ways similar to moral material: by skimming off the controversial (i.e., scientifically challenged) part. This process matches Rawls's "method of avoidance" (Rawls, 1971, p. 231), but applies to facts, reinforcing Bellamy's point that liberal legitimacy would hold only because the most contentious (moral or epistemological) issues are pruned off during public decision making.

Secondly, if we move away from the legitimacy of the basic structure and other "essentials" to the actual legitimacy of public policies and institutions, legitimacy often depends on a mix of moral, political, and factual assumptions. Moreover, the various types of assumptions may be difficult to disentangle from one another or to divide into controversial and noncontroversial elements. Alternatively, reasons for policies may rely on justificatory blocs where moral and factual elements are tied together.

For instance, public-health policies rely on moral judgments (e.g., it is a good thing to be healthy) and factual claims about the effects of lifestyles (e.g., smoking causes cancer), some being controversial [the net effect of physical activities on health in urban environment (e.g., biking, running in car fumes) or on obesity (Malhotra et al., 2015)]. The legitimacy of public-health policies is conditional on factual assumptions that are probable in a statistical sense. Decision makers can resent or object to this lack of certainty, which could fuel discontent. Furthermore, assumptions' being true and factual does not mean that they are uncontroversial.[11]

Consensus plays a prominent role for political legitimacy on moral and epistemological grounds. The liberal society relies more or less explicitly on facts and epistemological elements that are assumed to be uncontroversial—for example, assumptions about human nature (instrumental rationality, autonomy), epistemology (the existence of truth, the hypothetic-deductive method as a firm basis for science), historic events (presence on a delimited territory, unity of a nation), and so forth. Liberal policies assume causation in the domains of health (e.g., between lifestyles and pathologies), the economy (among macroeconomic aggregates), politics, and so forth. Once that point is acknowledged, the next step is to determine how factual and epistemological disagreements may challenge consensus-based conceptions of legitimacy. Moreover, how do they affect decision rules that approximate consensus (e.g., majority rule, representation)?

3. EPISTEMOLOGICAL AND FACTUAL CHALLENGES TO CONSENSUS

In some circumstances, factual and epistemological disagreements may jeopardize consensus-based conceptions of political legitimacy. Numerous decisions or institutions in a liberal democracy need to be grounded on facts or epistemological rules, which can be controversial due to lack of conclusive evidence (e.g., health effects of nanoparticles, macroeconomic relations between inflation and unemployment). In such cases, there is no possibility of setting aside controversial facts or epistemological procedures to reach *a minima* consensus: the decision process cannot abstain from controversial facts and procedures. In other words, the first best cannot be reached by political institutions.

In addition, the possibility of producing scientific evidence may not settle controversies among citizens. Epistemological or factual disagreements may not be suppressed, even in the presence of scientific evidence (e.g., disagreements about genetically modified organisms, GMOs hereafter), because citizens who are part of the public debates or decision-making process are too committed to their epistemic beliefs. As a result, decision makers and citizens could remain deeply divided on policies without any hope of consensus, which again presents a challenge for consensus as a liberal first best.

Situations where factual or epistemological disagreements impair the legitimacy of political decisions and institutions by fragmenting the population include debates on the recognition of past wrongs (e.g., the Armenian genocide, the Soviet occupation of the Baltic States, the positive role of French colonization or its qualification as a crime against humanity). A contemporary example is Indigenous land claims.

In Canada, proving an ante-European, and at least since colonization, continuous presence on a territory is a prerequisite for the First Nations in order to obtain the recognition of an "aboriginal title" or "ancestral rights" as protected by the 1982 Constitution Act (section 35).[12] Some claims can be settled by historians and archaeologists producing evidence before governments and courts whereas, for others, evidence is controversial, subject to conflicting interpretations. Then nothing is to be expected from scientific enquiry, at least not in the current state of the art of science. A third category covers cases where available evidence will not lessen factual disagreements. For instance, historical presence on a given territory could be so central to Indigenous or non-Indigenous identity that evidence cannot win the agreement of all or contribute to attenuating controversies. A last category includes cases of true epistemological disagreements (i.e., disagreements on epistemological rules and not only on the facts). The society at large (Indigenous and non-Indigenous citizens) may be deeply divided on public policies because of controversial epistemological aspects.

The Canadian jurisprudence illustrates the depth that factual and epistemological disagreements can reach in liberal democracies. For asserting their pre-European and continuous presence on the claimed territory, Indigenous peoples often

use oral histories—for example, stories, legends, or oral accounts transmitted from one generation to another (Etinson, 2008). Courts have traditionally refused to consider such accounts as proofs equivalent to written documents or archaeological artefacts. Under the general regime of proof, oral accounts, when not stemming from direct witnesses, amount to no more than hearsays. However, the jurisprudence has been evolving. Some courts have placed oral accounts seemingly on par with other kinds of evidence (e.g., written historical documents) to prove Indigenous presence on a given territory.[13]

Why consider individuals' epistemological commitments in examining claims such as Indigenous peoples' ancestral rights? In situations characterized by epistemological uncertainty, the absence of indisputable evidence may advocate for giving more weight to such commitments. However, the argumentative force is not in the commitment, but in the absence of indisputable evidence. What about situations where evidence exists or could be produced? Why consider controversial factual or epistemological beliefs?

Three reasons may support considering factually ungrounded or controversial epistemological beliefs and, thus, diverging from the first best. The first is that *redressing past injustices and inequalities* justifies modifying epistemological standards. However, such exemption is constrained: it cannot be used in matters unrelated to the past wrongs that justified modifying standards in the first place.

The second reason is that *respect is due to the decision makers as participants in public reason or debates*. The justification is that "relaxing" standards *might* be necessary for including decision makers in the democratic process. However, such a justification is too broad: standards cannot be "relaxed" for all decision makers in any situation, except at the price of a collapse of rational deliberation. Such a justification needs to be circumscribed by identifying a conception of respect that could justify "relaxing" standards only for some decision makers (e.g., for Indigenous peoples and not for other citizens who inhabit the disputed territory).

In liberal theory, respect is usually grounded on the (posited) possession of given faculties (viz. reason, rationality, or autonomy) by all individuals. However, "relaxing" epistemological standards only for some decision makers does not do justice to decision makers in terms of reason, rationality, or autonomy. A possibility for qualifying the second reason is to elaborate on the respect due to specific decision makers. Liberal theory already justifies differential treatment. Take the example of impaired capacity. Because of their limited abilities could justify a differential treatment. Because of their impaired capacity to distinguish true from false statements or to follow sound epistemological rules, respect imposes to relax epistemological standards for some decision makers. Hence all decision makers will be included in the public exchange of reasons. It is then possible to present second-best issues in relation to consensus as expressing a tension between an epistemic defence of democracy, according to which democracy and truth enquiry are mutually supportive, and democracy as equal stand-

ing and inclusion (MacGilvray, 2014, p.117). Simply put, democracy as a process to hold true beliefs might threaten democracy as equality in the public exchange of reasons.

While justified for individuals with limited abilities (e.g., children, people with dementia), such reason is abhorrent when used to justify "lowering" epistemological standards for Indigenous peoples or any mobilizing group acting upon different standards. The differential treatment is rooted in a depreciative view of the individuals who hold different epistemological views. Such a ground for lowering standards adds insult to the injury of injustice. Furthermore, if respect implies treating people equally and treating them as equals (Dworkin, 1985, p. 190), this is also disrespectful of decision makers to whom original standards continue applying.

A third reason is to allow controversial facts or epistemological considerations to enter the decision-making process for pragmatic reasons, such as political stability. Because the exclusion of such considerations may nurture political instability, it may be better, everything considered, to give weight to them during decision making. For instance, proofs may exist that a genocide took place one century ago. Part of the population may ask the government to release an official recognition of the tragedy or to compensate for the victims. Another part of the population may disagree so fiercely on the reality of the tragedy or on its consequences that the government may finally decide to give way to their opposition by not enacting any kind of recognition or compensation.

While the first argument makes sense when dealing with justice issues such as Canadian First Nations and the second argument is controversial for the reasons I already mentioned, the third one is interesting due to the general framework it offers for thinking about second-best departures from the first best for reasons of stability, which occupies a central place in the liberal thought. Many developments of Rawls's *Political Liberalism* are devoted to it and to the means for securing individuals' support to democratic institutions (Barry, 1995; Klosko, 1994). Stability is so central for Rawls that it is one pillar of political liberalism, alongside justice (Klosko, 1994, p. 1183). (I will come back to stability in the next section.)

In any case, this does not imply that the three reasons are convincing. They illustrate that reasonable grounds exist for considering epistemologically controversial arguments during public decision making, even where evidence could be available. Then, the following question emerges: in situations where only controversial facts or epistemological elements are available or where decision makers are too committed to their epistemological beliefs, why not call for traditional decisional procedures that approximate consensus (e.g., voting)? In other words, why do such factual and epistemological disagreements undermine consensus as the first best?

4. SECOND BEST AND POLITICAL LEGITIMACY

The previous sections offer two arguments: (a) consensus is the first best for part of liberal legitimacy, and (b) factual and epistemological disagreements *may* challenge the consensus-based approach, especially for pragmatic reasons like political stability. There are two manners in which to understand this challenge: one radical and another moderate. The former is that disagreements *necessarily* incur a second-best issue and, consequently, a general departure from first-best conditions (as suggested by Lipsey and Lancaster). The latter is that disagreements *under specific conditions* impose a general departure from first-best conditions (i.e., from consensus approximations). In other words, only *some* disagreements would constitute or raise second-best issues. Only the latter is plausible, though, since it is difficult to see how *all* instances of factual and epistemological disagreements could raise second-best issues.

Before proceeding, I need to establish a preliminary condition. As seen, a second-best problem emerges when the first best is not attainable due to factual or epistemological disagreements. There are cases where a full consensus on facts or epistemological rules is not attainable because of uncertainty[14] [e.g., disagreements over risk evaluation, such as for the "Black Swans" (Taleb, 2008)] or decision makers' epistemological commitments. In addition, the disputed facts or epistemological rules should be necessary for decision making—that is, for justifying a political decision, policy, or regime.

For instance, when regulating chemical pollution, public institutions cannot avoid using contamination thresholds. They need measurement and impact studies to determine which chemical to regulate and how. These elements are necessary for policy making and often cannot be set aside, even in cases of deep disagreement. If the disputed material is not necessary (e.g., if the disagreement is about the kind of cancer a chemical favours, while there is no disagreement about the chemical creating a hazard), it could be ignored during the decision-making process (e.g., when setting a release threshold for the industry) and a consensus could be hammered out from the undisputed material (the hazardous aspect of the chemical and the necessity to regulate).

The preliminary condition for a second-best problem is that the first best cannot be reached due to factual or epistemological disagreements bearing on element(s) that are necessary for decision making. Although necessary, this condition is not sufficient. If the first best—consensus—is not attainable, a second-best problem does not necessarily arise. Approximating the first best could still be possible or desirable. In general, a second-best issue arises only if the remaining first-best conditions are impossible or undesirable (Räikkä, 2000).

Concerning the first best for liberal legitimacy *qua* consensus, the preliminary condition requires that the consensus not be able to be reached and that approximating the consent of all is either impossible or undesirable. Only then does a second-best issue arise. Theoretically, it means that second-best issues charac-

terize situations where any mild or altered form of consensus, such as the overlapping consensus, is either impossible or undesirable. Practically, it implies that aiming at getting a majority through voting and representation may be neither possible nor desirable.

The preliminary condition raises a question: under which conditions is pursuing the remaining first-best conditions either impossible or undesirable, and, in particular, what kind of disagreement could bar approximating the consent of all regarding the legitimacy of institutions, policies, or public actions? As previously evoked, the main risk posed to legitimacy by factual and epistemological disagreements occurs when they deeply fragment the political community or accentuates prior profound division. For a second-best problem to arise, citizens would have to be so divided that trying to get the consent of as many citizens as possible on the contested facts or epistemological rules would undermine the legitimacy of political decisions and institutions. This condition has two parts: citizens would have to be divided on the factual and epistemological issues at stake and the divide would have to cut so deep it threatens legitimacy *when institutions try approximating consensus*.[15]

The kind of social division threatening a liberal regime is one that endangers stability. In addition, stability represents a key liberal concern. Rawls proposes an overlapping consensus because it increases the probability of individuals adhering to the principles of a liberal society and, therefore, enhances political stability.[16] Then, various situations may present a threat for liberal legitimacy.

An important category embraces situations where the combination of the pursuit of the first best and factual/epistemological disagreements undermines *mutual* or *institutional trust* (i.e., trust among citizens or citizens' trust towards their institutions) and where it is very likely that consensus approximation will undermine trust. The second-best issue could worsen if, due to damaged trust, citizen participation in liberal institutions declines to the point that the perpetuation of such institutions might be endangered.

An illustration is GMOs. European citizens do not trust national and European institutions on genetically modified food (van Kleef et al., 2006, p. 58). In these cases, vote or representation may generate or accentuate distrust, whereas metastudies have underlined the absence of proved harms to human health by GMOs (Barrows et al., 2014; Nicolai et al., 2014), while no comparable metastudy has proved the contrary.[17] Such situations are risky for liberal democracies in several respects.

First, the risk is that consensus approximation (mostly voting) may worsen tensions between the majority, opposed to GMOs (Gaskell et al., 2010), and the minority, including experts who do not evaluate GMOs as more harmful than their non-GM equivalents. Another risk is the possibility that consensus approximations (viz. voting) support suboptimal public policies in a Paretian sense—

that is, policies that would make no one better off without degrading the situation of someone else when at least one alternative exists where an agent could be better off with no one else being worse off.

In relation to suboptimal outcomes, a stronger claim is that opposing GMOs actually harms people (Potrykus, 2010). The future collapse of traditional crop yields, overpopulation, deforestation, and climatically driven changes (e.g., water shortages, irregular rainfalls, higher drought frequency, increase of extreme climatic events, salt saturation in coastal lands, mineral depletion in soils) carry significant risks that might be (partially) addressed by broad cultivation of GM crops (under certain conditions).[18] Therefore, the rejection of GMOs based on consensus approximations might turn out to be detrimental to most of humanity within a few decades.[19]

If we accept scientific evidence that GMOs are not more harmful to human health than their non-GM equivalents, and if we accept GMOs' gains in yield and in pest resistance (e.g., Klümper and Qaim, 2014), public opposition to GMOs creates social gridlocks. These gridlocks happen when the conditions of agents (e.g., consumers, farmers) could be improved without worsening the condition of anyone else (or, at least, without worsening the condition of another agent in proportions larger than would have been the case with non-GMOs).

A third risk is the rejection of scientific evidence as a legitimate ground for public policy (when scientific evidence and popular sentiments conflict and consensus is approximated), which may lead to grounding public policies on unproven or false premises, a risk comparable to the one posed by populism. A further challenge is the danger of increasing the intensity of factual/epistemological disagreements in the future, because of science's loss of authority (due to its declining perceived legitimacy within a population).

Hence, the question is the following: *Which resources could the TSB offer to public decision making embroiled in factual and epistemological disagreements?* This question captures the challenge of mobilizing the second best: to offer specific enough theoretical and practical resources. The issue is that the literature on the TSB in political theory is underdeveloped. Moreover, this literature is crippled by misunderstandings and approximations (Räikkä, 2000). Most of political theorists who appealed to the TSB did so in a very brief manner (in no more than a couple of paragraphs or sometimes a few sentences).[20] Therefore, there is still a long way for the TSB to help deal with the political issues of legitimacy, education, institutions, justice, and so forth.

To make the TSB relevant for liberal legitimacy (and political theory), two issues ought to be addressed head on. The liberal theorists who want to mobilize the TSB need to specify how to understand the general departure form first-best conditions in a liberal sense and the constraints stability imposes on second bests. However, this task should be carried out while keeping in mind Lipsey's reservations and warnings. First, finding a second best seems more difficult than

achieving the first best (Lipsey, 2007, p. 356). Second, finding a second best is about "piecemeal improvements in welfare" (i.e., about adopting a pragmatic approach).

4.1. Liberal departure

According to Lipsey and Lancaster's theorem, a second best implies a "general departure" from first-best conditions (i.e., a change of all conditions). Therefore, for the TSB to be meaningful for liberal theory, it is important to constrain the principles or rules used for identifying a second best. In the context of the present discussion, these principles should be compatible with liberalism, but alien to a consensus-based approach. The principles supporting a second best should constitute a general departure from first-best conditions (viz. no voting or other approximation procedure) *and* be liberal. They need to be liberal because liberalism is, by definition, the encompassing justificatory architecture for the conception of political legitimacy discussed in this article.

Therefore, the issue of what could constitute liberal second bests for political legitimacy (among other issues) emerges. Without discussing the matter in detail, let me consider just a few possibilities to give the reader a rough idea of the kind of conceptual work to be undertaken by anyone who wants to apply the TSB to political theory.

One possibility is to adopt a rights-based approach. When factual and epistemological disagreements jeopardize the realization of the first best, political decisions and policies could be compared and adopted based on their impact on individual rights. This presupposes mobilizing a background theory of the nature of rights and rules of adjudication between competing claims, which excludes many group-rights approaches (which are anti-individualistic, limit individual autonomy, etc.). Another option is to adopt a welfarist approach. The legitimacy of decisions and policies may be assessed by assessing their impact on welfare. There again, further discussions are necessary on the nature of welfare, the possibility of making welfare comparisons across individuals, groups, situations, and so forth. As for the rights-based approach, the justificatory scheme ought to be compatible with liberalism.

These options do not come without conceptual and practical difficulties. Moreover, they do not cover all possibilities. Nonetheless the liberal-departure issue shows that one might still endorse consensus as a first best, but, due to specific constraints, adopt a backup justificatory scheme. Thus, justifications might come into play at different levels (which is a different point from the distinction between ideal and non-ideal theories). In other words, the TSB may advocate for a pragmatic decompartmentalization between competing conceptions of political legitimacy. It argues for tuning down some differences among theories of political legitimacy and adopts some sort of justificatory pluralism based on pragmatic considerations.

4.2. Stability

This liberal concern, present in Rawls, requires that the application of second-best principles not undermine political stability (e.g., by fuelling distrust toward democratic institutions). This condition could be made stronger by requiring that the application of these principles strengthen stability. In other words, on top of being compatible with liberalism, the principles that provide guidance for identifying a second best ought to also guarantee social stability. This condition is practical and, as such, open to debate. However, the important point to retain is that, again, the TSB seems to point to highly contextual reflections. In that sense, further investigations of the TSB will show a strong potential overlap with methodological discussions on contextual approaches in political theory (e.g., Carens, 2004; Kukathas, 2004).

5. CONCLUSION

Looking back at the TSB fifty years on, Lipsey (2007, p. 356) indicates that second-best issues call for pragmatism, especially because no global second-best solutions can be tailored for economics (while the case is still open for politics). Therefore, only piecemeal changes are left to policy makers. Coupled with the underqualification of the TSB, this notice of caution could fuel the view that the TSB's fecundity for political theory remains limited. Moreover, one could nurture the view that whether factual and epistemological disagreements undermine political legitimacy is a purely contingent matter. Then, the conclusion could be that the TSB offers no compelling or workable challenge to consensus-based conceptions of legitimacy.

Certainly, the existence of factual and epistemological disagreements does not necessarily undermine political legitimacy, and, as indicated, the existence of such disagreements does not necessarily create a second-best problem—that is, a general departure from first-best conditions (consensus). In this article, I recognize these points and discuss a more modest thesis: factual and epistemological disagreements *could* create second-best problems for the liberal tradition of political legitimacy based on consensus. I tried to clarify two points: (1) what consensus as the first best means for part of the liberal tradition as concerns moral *and* factual/epistemological disagreements and (2) how such disagreements *could* create second-best problems for consensus-based approaches. I ended up by underlining, on the one hand, the risks associated with first-best approximations in situations shaped by a deeply divided society and, on the other hand, the necessity to evaluate more closely the resources the TSB could offer to political theory.

That being said, it is true to some extent that the claim that factual/epistemological disagreements create second-best issues for consensus-based legitimacy depends on contingent factors (e.g., the degree of trust, the intensity of disagreements, the nature of disagreeing parties, etc.). Nevertheless, the challenge for liberal legitimacy founded on consensus is not contingent (in the sense of *anec-*

dotal). For any political theory that relies on the proximity to consensus for assuring the legitimacy of its foundational principles and institutions, any reasonable expectation of dissensus is a direct challenge to that theory's plausibility and expected outcomes, like stability, (and therefore its traction) (Thrasher and Vallier, 2013). The existence of enduring factual and epistemological dissensus forces consensus-based approaches to legitimate the rules and outcomes of decision making with other principles, which should be compatible with liberalism. This dimension is largely (if not totally) absent from political theory.

However, for the political theorists interested in the task, a suspicion looms over the whole project. To the question "Are there general policy rules for piecemeal improvements?," Lipsey (2007, p. 358) answers with a clear "no," reducing the entirety of the TSB to a piecemeal approach. Thus, in the absence of serious discussions of the TSB by political theorists, it seems that appeals to the second best are doomed to either remain rhetorical (as false equivalents for non-ideal approaches, for instance) or advocate for political casuistry.

NOTES

[1] See *Merriam-Webster Learners Dictionary*, s.v. "consensus," accessed October 25, 2017, http://www.learnersdictionary.com/definition/consensus. I do not claim that consensus is *the* ultimate principle of all liberal conceptions of political legitimacy. I discuss consensus as the ultimate principle for *part of* the liberal tradition.

[2] Consensus and consent share the same Latin origin: *consentire* (*com-*, together, *-sentire,* to feel), from which is derived *consens* ("agreed").

[3] Moreover, the TSB is sometimes interpreted as making the case for non-ideal theories. According to Räikkä (2000, p. 213), "John Rawls defined non-ideal theory as a theory that answers the question of what should be done in circumstances where social arrangements are unjust or individual conduct morally blameworthy", whereas "a problem of second best, however, may arise even in perfect just and morally acceptable circumstances". I would broaden this point. If ideal is understood in a more colloquial sense, it remains true that the second best could be as idealistic than the first best, or the first best conditions could potentially be implemented, but with undesirable effects (meaning that a the divide is not between idealistic, first-best, and non-idealistic, second-best, approaches). In short, the distinction idealism/non-idealism fails to capture what is at work here.

[4] A Pareto optimum is reached when it is impossible to improve the situation of one economic agent without worsening the situation of another agent.

[5] I mostly focus on factual disagreements even if part of the content applies to epistemological disagreements too—i.e., disagreements on the rules regarding the production and validation of knowledge.

[6] In the case of Rawls, propositions are deflated of their metaphysical or philosophical content when it comes, for instance, to agreeing to principles of justice.

[7] "When any number of men have so consented to make one community or government, they are thereby presently incorporated, and make one body politic, wherein the majority have a right to act and conclude the rest" (Locke, 2003, p. 142).

[8] For Hobbes, the Sovereign body's legitimacy is grounded on the "consent" of all.

[9] Rawls is not alone in trying to identify the minimal basis for a consensus in a morally diverse society. According to Larmore (1990), liberal institutions are based on a double commitment from all citizens to the norms of "rational dialogue" and "equal respect for persons."

[10] Rawls privileges the *extension* dimension of consensus (the consent or agreement of all) to the detriment of the *depth* dimension (the consent to all aspects of a political decision or institution).

[11] Legitimacy works at two levels here: the plausibility of the set of assumptions that justifies a given policy and the accessibility of such plausibility to individuals who are affected by the policy.

[12] *R v. Sparrow [1990] 1 S.C.R. 1075.*

[13] The Supreme Court decision *Delgamuukw* initiated the change at the federal level. The jurisprudence was applied by lower jurisdictions—e.g., in the *Tsilhqot'in* case (*Delgamuukw v. British Columbia [1997] 3 S.C.R. 1010, Tsilhqot'in Nation v. British Columbia, [2007] BCSC 1700*).

[14] Uncertainty does not necessarily create disagreements. The claim is that uncertainty *may* create disagreements. As a matter of fact, uncertainty is one of the main sources of such disagreements on political matters (e.g., GMOs, financial crisis, and regulation).

[15] The social divide could result from widespread distrust (among citizens or towards institutions), and not from controversial (epistemological) features of a political decision or justifications for a political institution.

[16] Rawls's concern is not isolated within the liberal tradition. Political liberalism is not only a theory of individual rights and of the protection of the private sphere against external intrusion (from the state, community, etc.). It is also a theory elaborated during troubled times

(European wars of religion) and marked by the necessity of guaranteeing individuals' security (see, e.g., Hobbes) through stable political institutions. Also, nothing guarantees that individual rights are forcibly conducive, under any circumstance, to stability.

[17] This absence of large-scale evidence does not, of course, prove that the consumption of GMOs is not harmful to health. It just shows that, according to the state-to-the-art of scientific investigation, there is no proof that *GMOs would be relatively more harmful than non-GMOs.*

[18] GM technology offers the advantage of adapting plants at a much quicker pace than conventional breeding would ever allow. This pace of adaptation is particularly relevant when one considers the fact that climate change carries environmental alterations that are too fast for most of animal and vegetal species.

[19] This does not mean that it must be the case, only that GMOs represent a reasonable case where trying to get the consent of as many individuals as possible *might* turn to be harmful for liberal societies. Most of the discussion implies factual assumptions (e.g., about population growth, the comparative lower yields of organic or traditional agriculture by comparison with GM agriculture) and scientific evidence that definitely exceed the scope of this article.

[20] Among political theorists, only Räikkä, to my knowledge, has conducted a consistent evaluation of the TSB.

REFERENCES

Barrows, Geoffrey, Sexton, Steven and Zilberman, David, "Agricultural Biotechnology: The Promise and Prospects of Genetically Modified Crops", *The Journal of Economic Perspectives*, vol. 28, no. 1, 2014, p. 99-119.

Barry, Brian, "John Rawls and the Search for Stability", *Ethics*, vol. 105, no. 4, 1995, p. 874-915.

Bellamy, Richard, *Liberalism and Pluralism*, London and New York, Routledge, 1999.

Carens, Joseph, "A Contextual Approach to Political Theory", *Ethical Theory and Moral Practice*, vol. 7, no. 2, 2004, p. 117-132.

Coram, Bruce Talbot, "Second Best Theories and the Implications for Institutional Design", *in* Richard E. Goodin (ed.), *The Theory of Institutional Design*, Cambridge, Cambridge University Press, 1996, p. 90-102.

Dworkin, Ronald, *A Matter of Principle*, New York, Oxford University Press, 1985.

Estlund, David, "Utopophobia", *Philosophy & Public Affairs*, vol. 42, no. 2, 2014, p. 113-134.

Etinson, Adam Daniel, "Aboriginal Oral History Evidence and Canadian Law", *The Central European Journal of Canadian Studies*, vol. 6, no. 1, 2008, p. 97-104.

Gaskell, George, Stares, Sally, Allansdottir, Agnes, Allum, Nick, Castro, Paula and Jackson, Jonathan, *Europeans and biotechnology in 2010: winds of change?* Luxembourg, Publications Office of the European Union, 2010.

Gaus, Gerald, *The Tyranny of the Ideal*, Princeton and Oxford, Princeton University Press, 2016.

Jønch-Clausen, Karin and Kappel, Klemens, "Social Epistemic Liberalism and the Problem of Deep Epistemic Disagreements", *Ethical Theory and Moral Practice*, vol. 18, no. 2, 2015, p. 371-384.

Kant, Immanuel, "On the Common saying: That May Be Correct in Theory, But It Is of No Use in Practice" (1793), *in* M. J. Gregor (ed.), *Practical Philosophy,* Cambridge, Cambridge University Press, 1996, p. 273-310.

Kappel, Klemens. "Fact-Dependent Policy Disagreements and Political Legitimacy", *Ethical Theory and Moral Practice*, vol. 20, no. 2, 2017, p. 313-331.

van Kleef, Ellen, Frewer, Lynn J., Chryssochoidis, George M., Houghton, Julie R., Korzen-Bohr, Sara, Krystallis, Thanassis, Lassen, Jesper, Pfenning, Uwe and Rowe, Gene, "Perceptions of Food Risk Management Among Key Stakeholders: Results from a Cross-European Study", *Appetite*, vol. 47, no. 1, 2006, p. 46-63.

Klosko, George, "Rawls' Argument from Political Stability", *Columbia Law Review*, vol. 94, no. 6, p. 1882-1897.

Klümper, Wilhelm and Qaim, Matim, "A Meta-Analysis of the Impacts of Genetically Modified Crops", *PLoS ONE*, vol. 9, no. 11, 2014, p. 1-7.

Kukathas, Chandran, "Contextualism Reconsidered: Some Skeptical Reflections", *Ethical Theory and Moral Practice*, vol.7, no. 2, 2004, p. 215-225.

Larmore, Charles, "Political Liberalism", *Political Theory*, vol. 18, no. 3, 1990, p. 339-360.

Lipsey, Richard G., "Reflections on the General Theory of Second Best at its Golden Jubilee", *International Tax and Public Finance*, vol. 14, no. 4, 2007.

Lipsey, Richard G., and Lancaster, Kelvin, "The General Theory of Second Best", *The Review of Economic Studies*, vol. 24, no. 1, 1956-1957, p. 11-32.

Locke, John, *Two Treatises of Government and Letter Concerning Toleration,* New Haven and London, Yale University Press, 2003.

MacGilvray, Eric, "Democratic Doubts: Pragmatism and the Epistemic Defense of Democracy", *The Journal of Political Philosophy*, vol. 22, no. 1, 2014, p. 105-123.

Malhotra, Aseem, Noakes, Timothy, and Phinney, S., "It is time to bust the myth of physical inactivity and obesity: you cannot outrun a bad diet", *British Journal of Sports Medicine*, 2015, p. 1-2.

Margalit, Avishai, "Ideals and Second Best", *in* Seymour Fox (ed.), *Philosophy for Education*, Jerusalem, Van Leer Jerusalem Foundation, 1983.

Moss, David A., *When All Else Fails*, Cambridge, Harvard University Press, 2002.

Nicolai, Alessandro, Manzo, Alberto, Veronesi, Fabio and Rosellini, Daniele, "An Overview of the Last 10 Years of Genetically engineered crop safety research", *Critical Review in Biotechnology*, vol. 34, no. 1, 2014, p. 77-88.

Potrykus, Ingo, "Regulation Must Be Revolutionized", *Science*, vol. 466, no. 7306, p. 561.

Räikkä, Juha, "The Problem of the Second Best: Conceptual Issues", *Utilitas*, vol. 12, no. 2, 2000, p. 204-218.

Rawls, John, *A Theory of Justice*, Cambridge, Belknap Press, 1971.

—, "Kantian Constructivism in Moral Theory", *The Journal of Philosophy*, vol. 77, no. 9, 1980, p. 515-572.

—, *Political Liberalism*, New York, Columbia University Press, 1993.

Rebello, Joseph, "The Problem of the Second Best: Are Partial Equilibrium and Third-Best Analyses Solutions?", Chicago, University of Chicago, 2002.

Rodrik, Dani, "Second-Best Institutions", *American Economic Review*, vol. 98, no. 2, 2008, p. 100-104.

Sher, George, *Approximate Justice*, Lanham, Rowman & Littlefield Publishers, 1997.

Taleb, Nassim Nicholas, *The Black Swan*, London, Penguin, 2008.

Thrasher, John, and Vallier, Kevin, "The Fragility of Consensus: Public Reason, Diversity and Stability", *European Journal of Philosophy*, 2013, vol. 23, no. 4, p. 933-954.

Vallier, Kevin, "Consensus and Convergence in Public Reason", *Public Affairs Quarterly*, vol. 25, no. 4, 2011, p. 261-279.

Waldron, Jeremy, "Theoretical Foundations of Liberalism", *The Philosophical Quarterly*, vol. 37, no. 147, 1987, p. 127-150.

NEW TROUBLE FOR DELIBERATIVE DEMOCRACY

ROBERT TALISSE
DEPARTMENT OF PHILOSOPHY, VANDERBILT UNIVERSITY

ABSTRACT:

In the past two decades, democratic political practice has taken a deliberative turn. That is, contemporary democratic politics has become increasingly focused on facilitating citizen participation in the public exchange of reasons. Although the deliberative turn in democratic practice is in several respects welcome, the technological and communicative advances that have facilitated it also make possible new kinds of deliberative democratic pathology. This essay calls attention to and examines new epistemological troubles for public deliberation enacted under contemporary conditions. Drawing from a lesson offered by Lyn Sanders two decades ago, the paper raises the concern that the deliberative turn in democratic practice has counter-democratic effects.

RÉSUMÉ :

Au cours des deux décennies passées, la pratique politique démocratique a pris un tournant délibératif. Plus précisément, la politique démocratique contemporaine s'est de plus en plus concentrée sur la manière de faciliter la participation citoyenne dans l'échange public de raisons. Si ce tournant est le bienvenu pour plusieurs raisons, les avancées technologiques et communicationnelles qui l'ont facilité ont également rendu possibles de nouvelles pathologies démocratiques et délibératives. Cet essai examine les nouveaux problèmes épistémologiques pour la délibération publique contemporaine. Tirant la leçon des travaux menés par Lyn Sanders il y a deux décennies, l'article s'interroge sur les effets antidémocratiques du tournant délibératif.

I. INTRODUCTION

Deliberative democracy is an appealing yet elusive ideal. In its canonical versions, it is the thesis that democracy's value—its legitimacy, authority, authenticity—lies in its ability to base political decisions on citizens' reasons rather than simply on their preferences or votes.[1] To be sure, there is a great variety of views in currency claiming the name, and so the foregoing summary statement requires a good deal of clarification. Yet any version of the canonical view gives rise to a series of questions that help to explain deliberative democracy's elusiveness. What does it mean to *base* political decisions on citizens' reasons? How are reasons to be distinguished from preferences? Why do reasons confer legitimacy (or authority, or authenticity) on collective decisions? Does deliberative democracy rest upon the assumption that citizens already share a view about what reasons *are*? Hence extensive and rapidly growing literatures have emerged around these (and other) questions.

It is somewhat surprising, then, to find in a review essay published nearly twenty years ago James Bohman (1998) declaring that deliberative democracy has "come of age." Noting that the core idea of deliberative democracy had been a central and explicit theme in much democratic theory since at least the 1980s, Bohman was canvassing the then-recent theoretical developments aimed at showing that the deliberative ideal could be regarded as feasible. That is, by 1998, the central challenges to deliberative democracy concerned not its theoretical architecture, but rather its practical implementation. According to Bohman, deliberative democracy's principal theoretical commitments—including its normative superiority to aggregative, pluralist, and elitist models of democracy—had been widely accepted. By 1998, deliberative democracy had "come of age" in that it had established itself as one of the central frameworks for normative democratic theory. As Bohman presents it, the remaining task is that of *enacting* the theory.

Writing two years later, Samuel Freeman sounds a similar tone. Freeman observes that "deliberative democracy" had become "more than just another popular label"; it is, he contends, a distinctive family of views united by more than simply a common rejection of an entrenched opposing view (2000, p. 371). Freeman notes that, as it is a positive program within normative democratic theory, deliberative democracy is the site of several internal disputes. After carefully working through many of these, Freeman concludes that deliberative democracy provides a vision of the democratic ideal that is indeed superior to that offered by well-established non-deliberative theories. However, his enthusiasm is somewhat measured. In the end Freeman remains "sympathetic" to deliberative democracy, but he is not fully an advocate; he expresses the apprehension that deliberative democrats have yet to demonstrate that the deliberative ideal is practically feasible (2000, p. 418).

It is safe to say that deliberative democracy's popularity has only grown in the intervening decades. Today, deliberative democracy is arguably the predominant framework in normative democratic theory; hence, it is difficult to find a normative democratic theorist who does not embrace *some* version of the core deliberativist thesis that democracy's value (authority, legitimacy, justice) is owing to democracy's ability to shape political decision in response to the open exchange of ideas, reasons, and arguments of citizens.[2] Moreover, the loudest opposition to deliberativism tends to originate from theorists who are

suspicious of normative accounts of democracy as such.[3] Still, the concerns over deliberative democracy's feasibility remain, and these challenges have led deliberative theorists to adopt increasingly stylized models of public deliberation. These models vary significantly over fundamental issues: *Who* deliberates? *When, where, among whom, and for how long* should deliberation occur? *What* questions are suitable for public deliberation? Is *deliberativeness* primarily a feature of interpersonal communicative interactions, or is it rather a property of certain *systems* of collective decision? Each of these questions is the focus of intense and ongoing debate. Hence, even though deliberative democracy dominates normative democratic theory, it remains a highly troubled framework.

I will not canvass these longstanding and intricate debates here.[4] Instead, I will revisit a kind of critique of deliberative democracy that targets its *desirability* under social conditions that currently prevail and should be expected to persist. To explain: in 1997, Lyn Sanders argued that the social dynamics of race, gender, and class significantly impact deliberative encounters in ways that replicate the patterns of exclusion, disadvantage, and marginalization that obtain in the society at large. Drawing on robust studies of jury behavior, Sanders noted that women, non-whites, and the economically disadvantaged speak less frequently, are almost never selected for the role of foreman, and are more likely to be interrupted and dismissed while speaking, as compared with fellow jurors who are white economically privileged males. Sanders argued that the jury studies suggest that, even under institutionally favourable conditions, deliberative democracy is likely to further entrench existing patterns of social inequality.[5] She thus urged caution in calling for the deliberativization of existing democratic practice. Sanders argued that, in order to play their intended role in enriching democracy, deliberative institutions and practices must operate against the background of broader egalitarian social commitments, and that these are commitments that deliberation alone cannot foster; furthermore, she showed that they are commitments that deliberation, when enacted under conditions in which they are insufficiently entrenched, can undermine. In a nutshell, then, Sanders's lesson is that we do not necessarily make progress towards the deliberative democrats' political ideal by adjusting existing democratic institutions so that they more closely approximate the ones prescribed in deliberative democratic theory.[6]

The trouble is that in the past two decades democratic political practice has taken a deliberative turn. That is, contemporary democratic politics has become increasingly focused on facilitating citizen participation in public argumentation. Owing largely to advances in communications technology and social media developments, the public sphere is saturated with outlets, sites, and forums for public political discourse, from comments sections on news websites to feeds on Facebook and Twitter. Even nightly news programming is presented in a pro-and-con debate format, where viewers observe an exchange of competing reasons and then are primed to draw their own conclusions. Just as Sanders would predict, our politics has become increasingly divisive and uncivil; more importantly, the prevailing divides have become less a matter of *disagreement* among democratic citizens and more a power struggle among conflicting visions of what it means to be a democratic citizen, with each side *condemning* the others as fundamentally opposed to a proper political order. Under such conditions, civil disagreement is hardly possible, as the contending parties are apt to regard each other as peddling a distorted or perverted conception of democracy itself. In the US in particular, the increased emphasis on public argument has helped

to bring about conditions under which no real political debate is possible; adherents of contending views do not debate, but merely challenge each other's competence, sanity, and fitness for citizenship. As Sanders warned, increased deliberativeness under existing conditions seems to have exacerbated underlying social divisions.

In this essay, I raise the worry that deliberative democracy in practice is unavoidably vulnerable to the kind of pathology to which Sanders called attention. However, the core of the specific concern I will raise differs importantly from what Sanders described. Sanders argued that, in order to be democratically enriching, deliberation needed to operate against a background of the kind of egalitarianism that could dismantle entrenched hierarchies of race, gender, and class. To be sure, Sanders's argument identifies a demanding precondition for deliberative democracy. The worry I will raise points to *epistemological* prerequisites for democratic deliberation that arguably are even more demanding.

In the next section (II), I will show that, although deliberative democracy is most commonly presented as a moral ideal, it nonetheless has a decidedly epistemological dimension, and thus places on democratic citizens distinctively epistemological requirements. This means that there are specifically epistemological ways in which deliberative democracy can falter. In the third section, I will review a familiar way in which public deliberation can fail epistemologically. Then I will argue, in the fourth section, that there are unique epistemological problems that arise from the fact that democratic deliberation is conducted not merely between contending parties, but among contending parties *arguing in front of an onlooking audience*. Once it is noticed that public deliberation is frequently conducted *for the sake of the onlookers*, new occasions for epistemological pathology arise. These third-party epistemological pathologies are difficult to counteract within a democratic framework—hence the "new trouble" announced in my title. The concluding section (V) will draw some admittedly bleak upshots of the foregoing analysis.

II. THE EPISTEMOLOGICAL DIMENSION OF DELIBERATIVE DEMOCRACY

Deliberative democracy is most frequently proposed as a centrally *moral* ideal. The idea is that, in collectively deciding how the coercive power of the democratic state is to be exercised, citizens owe to each other reasons for favouring a given policy over its alternatives.[7] The deliberativist claims that, when collective political decision is driven by activities of public deliberation, political policy emerges less as an imposition upon the democratic citizenry, and more as an expression of the popular will; political decisions preceded by public deliberation are thus said to realize the traditional ideal of collective self-government. Moreover, the deliberativist contends that public deliberation helps to *legitimize* collective decision by giving citizens access to the reasons behind public policy, reasons which can subsequently be challenged, revised, or overturned in ongoing public discourse. In this way, again, public deliberation is proposed as a means for making collective decisions that each citizen can regard as something more than a raw exercise of power; deliberativism regards the state and its policies as vulnerable to the reasoned contestation of democratic citizens, and, in this way, power is rendered *accountable* to the citizens. Finally, the deliberativist holds that processes of public deliberation manifest an attractive conception of citizenship in that when citizens deliberate, they must *civilly* give and

receive reasons, acknowledge each other's points of view, and respectfully argue in ways that provide others not only with a chance to speak, but also with an opportunity to be heard. Hence deliberative democracy invokes a particular conception of the traditional democratic idea of an active and engaged citizenry.[8] To put these points together: the deliberative democrat's central contention is that public deliberation is necessary in order to realize the democratic ideal of *collective self-government among morally equal and active citizens*.

Articulated as such, deliberative democracy is clearly a moral ideal. But it also makes indispensable use of several epistemic concepts.[9] *Deliberation* itself is unavoidably epistemic; it is a process of discerning and evaluating *reasons*. And *reasons* are most certainly epistemic items, as they are (on anyone's view) considerations that count *in favour* of some conclusion. Note, moreover, that deliberativists hold that democracy calls for public deliberation precisely because there is *disagreement* among citizens over how the state's coercive power should be exercised. Disagreement is not deployed here as a merely descriptive term, indicating the fact that unanimity does not prevail. Rather, the deliberativist acknowledges that there is disagreement because there is a *clash* among reasons, and different reasons favour different policies. The task of deliberation is hence that of attempting to *consider* the full range of reasons and discerning their respective weight so that one could decide which policy outcome is best supported by the reasons. It is difficult, to say the least, for an individual citizen to survey the full range of reasons in play with respect to any given public policy, so citizens must deliberate *together*; they must share, exchange, and scrutinize each other's reasons. Accordingly, public deliberation is partly—perhaps largely—a process of *public argumentation* where citizens *make the case* for their favoured public policy to each other, consider cases made by others for alternative policies, and all stand ready to be challenged.

The ideal of a deliberatively engaged and arguing citizenry is undeniably demanding, and, again, some have criticized it on that ground.[10] My present point, however, is that, although deliberative democrats most frequently offer *moral* reasons to hold that public deliberation is necessary for proper democracy, the processes of democratic deliberation themselves cannot be identified except by reference to epistemic concepts. We might say, then, that the deliberative democrat proposes that citizens morally owe each other civil participation in a collective epistemic activity. Hence deliberative democracy involves a moral requirement and an epistemic requirement; citizens must interact civilly, and their interactions must rise to the epistemic level of deliberation.

Once we see that deliberative democracy is a both moral and epistemic proposal, we also see that the normative core of deliberative democracy is partly epistemological. It would be hard to imagine any democratic theorist endorsing deliberativism in the light of a demonstration that public deliberation, even when conducted civilly, always produces epistemically disastrous results. That is, part of the normative appeal of deliberative democracy lies in the presumed potential for civil public deliberation to yield epistemological benefits of some kind. Hence deliberativists often claim that public deliberation produces epistemically better collective decisions, more rational policies, better informed voters, more intelligent citizens, and the like.

I will not examine here the question of whether public deliberation actually yields epistemological benefits.[11] My point is that deliberative democracy can

falter normatively in at least two ways: first, it falters *morally* when citizens are not able or are not inclined to engage each other civilly on political questions; second, it falters *epistemologically* when citizens indeed engage each other, and may even do so civilly, but engage in ways that severely fall short of the epistemic ideals of public deliberation (better-informed judgments and voters, better-reasoned policy decisions, greater public understanding of public policy, increased accountability, and so on).

To be sure, although these two kinds of failure are conceptually distinct, in practice they comingle. For example, it is common for incivility among deliberating parties to have its root in accusations of epistemic incompetence. And one especially potent form of incivility consists in the systematic impugning of others' epistemic capacities or credentials. Indeed, Sanders notes that many of the varieties of deliberative incivility and exclusion that she discusses have their root in a prior judgment that non-white non-males lack "epistemological authority" (1997, p. 349). Now, it would be optimistic to claim that deliberative democratic incivility always has its source in an unjustified negative assessment of the epistemic condition of one's opposition. Surely a considerable portion of democratic incivility is due to unadulterated bigotry and garden-variety intolerance. But distinctively epistemological failings of deliberative democracy are prevalent and have accordingly attracted a good deal of attention. Reviewing a familiar kind of epistemic pathology of deliberation will set the stage for a new kind of difficulty.

III. FAMILIAR TROUBLE: THE POLARIZATION DYNAMIC

Deliberative democrats hold that citizens should engage in public deliberation. In public deliberation citizens do not merely *announce* the reasons driving their political advocacy; rather, they participate in a collective epistemological activity that involves public political *argument*. Deliberative democratic citizens *reason together*; they present their arguments to each other for the sake of advancing the rational collective investigation into some public and political issue. This collective aspect of public deliberation provides occasion for a range of epistemological pathologies. I here focus on a common dynamic among at least three such pathologies, and the dynamic begins with a well-studied and common phenomenon known as *group polarization*.

Group polarization is the phenomenon where members of a doxastically homogeneous deliberative group predictably move, imperceptibly to themselves, towards a more extreme version of the view they held prior to deliberating. It is important to notice that the trouble with group polarization is that the doxastic shift is driven by group dynamics rather than by reason. When groups polarize, it is not due to the introduction of new information or better arguments favouring a more extreme position; polarization occurs simply as the psychological consequence of immersing oneself in what Cass Sunstein has described as an epistemic enclave, a cognitive environment of relative unanimity where one hears "louder echoes" of one's own voice (2007, p. 13). And as Sunstein has noted repeatedly (2003; 2009), the technology that structures most of our political communication enhances individuals' ability to preselect the political valence of their interlocutors and even their news and information. As a result, discussion within epistemic enclaves is rampant, and group polarization prevails.

Now, the antidote to group polarization is doxastic heterogeneity amidst social norms that invite disputation and welcome dissent. In short, deliberating groups need to take steps to ensure that critical voices are encouraged and heard; they need to inoculate themselves against their tendency to construct echo chambers (Sunstein, 2003). It may seem an easy fix, yet such countermeasures are more difficult to implement than one might suppose. Consider: group polarization tends to encourage a closely related phenomenon, *epistemic closure*.[12] As groups polarize, they become less able to countenance the *possibility* of reasoned and sincere disagreement; opposing views come to sound like confused noise, critics begin to look craven and ignorant, and the view favoured by the group comes to be regarded as the only rational view there could be. As there's obviously no point in trying to argue with craven noisemakers, members of polarized groups become less able to deliberate with anyone who is not already within their fold. The strong sense of an epistemic in-group (and out-group) encourages yet another closely related pathology—namely, the *epistemic marginalization* of dissenting voices; this involves not only the tendency to decline to engage in deliberation with dissenting others, but also the denial of their epistemic capacities as such. The epistemically marginalized are not merely ignored; they are overtly regarded as incapable of knowing, or even of serving as sources of information. Such marginalization is obviously correlated with other forms of social disadvantage, including violations of democratic equality.[13]

It's not difficult to see, then, that in real-world political deliberation group polarization, epistemic closure, and epistemic marginalization operate in a dynamic of mutual reinforcement.[14] The degree to which a group is polarized tracks the degree to which members fail to recognize their critics as even rational, let alone as possibly correct or even as sources of valuable information. Call this the *polarization dynamic*. It goes without saying that the polarization dynamic is poisonous from the perspective of deliberative democracy. Recall that deliberative democracy is premised on the idea that stark disagreement over public policy is possible among well-intentioned, sincere, and duly informed democratic citizens. The polarization dynamic not only dissolves civility, but also disables public deliberation by encouraging the idea among the citizenry that ultimately there is nothing to deliberate about because reasonable disagreement is in fact not possible. Accordingly, the directive to group members to welcome dissent and invite criticism might be useful for *preventing* polarization, but is of limited help in *counteracting* polarization once it has emerged within a deliberating group.

Sunstein's own prescription hence is to introduce legal measures that could limit a doxastic group's capacity to enclave. These proposals rely less on group members' inclination to welcome dissenting voices and more on institutional design aimed at making political echo chambers more difficult for groups to construct. Among his more notorious suggestions is that politically extremist websites should be legally required to carry links to opposing websites (Sunstein 2007, p. 204). Of course, the efficacy of this policy still depends largely on individual visitors' willingness to actually follow the opposing links and investigate the opposing viewpoints open-mindedly. And it is not difficult to imagine ways in which Sunstein's envisioned "must carry" laws could be subverted so that they *contribute* to group polarization. To see this, consider a politically progressive site that features dozens of opposing links, but only to the most unhinged and irresponsible conservative sites. This would serve to confirm the progressive

group's favoured image of their opposition, and thus would contribute to their polarization. So maybe there is no failsafe against the polarization dynamic, but there still could be legal interventions to combat it.

In fact, one could argue that our current media environment is well suited to the task of combating group polarization. Since actual democracy has taken a deliberative turn, it is difficult in our day-to-day lives to escape the clash of political viewpoints, much less deny that there are clashes of this kind. News outlets, televised and online, are now almost entirely devoted to report-and-discuss formatting, where a host first presents a story, and then moderates a brief panel discussion among proponents of differing political perspectives. Viewers are explicitly tasked with weighing reasons and evaluating the arguments presented by the panelists. Online media allow for ongoing exchanges of views and arguments among citizens who otherwise would not interact. And, judging from the popularity of politically oriented news programming, online sites, and social media, citizens are largely interested in participating in public deliberation. From the perspective of our communications and media technologies, we should be living in a deliberative democrat's paradise. Yet, as we all recognize, public political discussion is horrendous, both morally and epistemically. What's going wrong?

IV. NEW EPISTEMOLOGICAL TROUBLE

A lot of work on deliberative democracy intentionally employs an avowedly simplified model of deliberation. Often, it is presumed that there is a single question under consideration, which admits only of a binary, yes-or-no response, and the deliberation is conducted by only two parties. The deliberating parties are taken to be addressing only each other, each evaluating the other's reasons while also proposing arguments of their own that are designed to move the interlocutor. Of course, no deliberative democrat is really committed to the idea that real world deliberative encounters are so simple. The typical models are intentionally simplified for purposes of theoretical manageability; everyone acknowledges that actual political deliberations will be far more complicated. But if public deliberation can't be made to look theoretically appealing under highly idealized conditions, there's no reason to think it worthwhile under more complex circumstances. Nonetheless, simplifying measures can sometimes omit too much, rendering a model unduly simplistic and hence unable to capture relevant phenomena.

What standing models of public deliberation seem to omit is that even when deliberation is indeed conducted between only two parties who are explicitly addressing only each other, deliberative exchanges are frequently nonetheless *public performances* enacted in the presence of an onlooking audience. In fact, in our current communications environment, public deliberation is most frequently conducted *for the sake of* the onlookers. That is, although the participants in the deliberative exchange might explicitly address each other, they are often implicitly addressing the audience as well, and it is the latter that is their central, yet only implicit, focus; the reasons entered into the deliberative exchange by the deliberators are commonly designed to *move the audience* rather than convince the interlocutor. Importantly, the audience typically comes to the exchange for the sake of gaining information about the issue under debate. Perhaps more commonly, the audience views the debate for the sake of seeing how their favoured view stacks up against its competition. They look on

precisely because they want to see how their favoured view overcomes or prevails against the opposition; they watch the debate unfold as a means of learning about the dialectical situation that obtains among the positions in play. Even though the onlookers may have adopted a position with respect to the issue under discussion, they are, so to speak, as yet *uninformed* about the *relative* strengths and weaknesses of the options in play. They seek information about the relative merits of the competing views by watching those who have this knowledge engage in public argumentation.

Public political argument among interested and purportedly informed deliberators before an audience who may have formed opinions but do not yet know the comparative strengths and weaknesses of the competing views creates opportunities for a new kind of epistemological failure in public deliberation. In our collaborative work, Scott Aikin and I have developed the idea of a *dialectical fallacy* (Aikin and Talisse, 2014b). Dialectical fallacies are ways in which argumentation fails *specifically* with respect to the onlooking audience. They are *distinctive* failures in that a dialectical fallacy can be committed by an arguer who nonetheless does *not* commit a formal or informal fallacy against his or her interlocutor. Further, the deployment of a dialectical fallacy need not involve any incivility towards one's interlocutor. Thus, when an arguer commits a dialectical fallacy, he or she need not have thereby violated any of the standing moral or epistemological requirements identified by extant conceptions of deliberative democracy. And yet that arguer will have acted in a way that is objectionable from the point of view of deliberative democratic citizenship.

To get a better sense of what dialectical fallacies are, consider the contrast between the informal *Straw Man Fallacy* and the dialectical *Weak Man Fallacy*.[15] In the textbook version of the Straw Man, an arguer misrepresents his or her interlocutor's view so that it is easier to refute; the arguer then validly refutes the more flimsy version of this interlocutor's view, but presents himself or herself as having refuted this interlocutor. Note that, to describe the Straw Man, it is necessary to refer to an audience to whom the misrepresentation is projected.[16] The Weak Man also involves a misrepresentation projected to an audience, but is importantly distinct. An arguer who commits the Weak Man is one who seeks to discredit a view by engaging with an especially inept proponent of it; the arguer then validly refutes this proponent's actual argument, but presents himself or herself as having refuted *the best the opposition has to offer*. Unlike the Straw Man, the Weak Man need involve no mistreatment of the perpetrator's interlocutor; indeed, the Weak Man can be deployed with the utmost respect, fairness, civility, and epistemological integrity towards one's discursive partner.

Crucially, successful deployments of the Weak Man are formally sound and informally cogent; proper Weak Man arguments indeed refute one's interlocutor. The fallaciousness of the Weak Man occurs entirely at the level of the onlooking audience; the perpetrator has misrepresented not his or her specific interlocutor's argument, but rather has projected a distorted view of the *dialectical situation* that obtains between that interlocutor's view and its opposition. The perpetrator has presented the state-of-play in the dialectic as one in which his or her own view obviously prevails against its opponents, all of whom are at least as feeble as the one he or she has just refuted decidedly. The onlookers are thus left with the impression that there is but a single viable view in play, and the best opposition to it is easily shut down. When fully successful, the Weak Man

creates the impression among the audience that there is no use in seeking out or listening to further opponents of the prevailing view, since the best of the opposition has been decisively rebuffed.

In the context of political argumentation, the Weak Man serves to shut down public deliberation as such; it overtly promotes the idea that there is ultimately nothing to deliberate about since there is but one responsible view in play and all other perspectives are weak and confused. The attempt to deliberate with those who disagree comes to be seen as a waste of time; what one's opponents need is liberation from their ignorance, not an exchange of reasons among equals. The Weak Man hence produces conditions ripe for the polarization dynamic.

Consider as a second example what Scott Aikin and I have playfully called *Modus Tonens* (Aikin and Talisse, 2008). Modus Tonens is the tactic of restating an interlocutor's claim in an incredulous tone of voice. Now, assuming that proper argumentative exchange allows for some degree of biting and snarky engagement, incredulously restating what an interlocutor has said is not necessarily out of bounds; one is surely permitted to express exasperation and surprise when signaling to an interlocutor that one finds what he or she has said in need of clarification or more deliberate affirmation. But like the Weak Man, Modus Tonens is deployed for the sake of projecting to one's audience a particular conception of the dialectical situation between the views. More specifically, Modus Tonens is the attempt to project to the onlookers that one's interlocutor is dialectically subordinate, someone who needs additional prompting and special assistance in articulating his or her own views. In a successful deployment of Modus Tonens, the perpetrator presents himself or herself as the *teacher* of the interlocutor, the more *intellectually mature* party to the discussion who must hence *enact* and *enforce* proper norms of serious intellectual discussion. Again, the tactic does not necessarily involve any mistreatment of one's interlocutor, but it serves to project to the relatively uninformed onlookers the view that, of the positions in play, only one is worthy of serious discussion.

The Weak Man and Modus Tonens are but two kinds of dialectical fallacy. There are many others. But I will not labour the point by cataloguing them here.[17] The important thing to note is that when public argumentation occurs in the presence of an onlooking and as-yet uninformed audience, interlocutors are incentivized to implicitly address the onlookers in a strategic way. In such cases, arguers may address their reasons directly to their interlocutor, and in this they might violate no standard norm of civility or principle of proper epistemic conduct owed among deliberative partners. That is, interlocutors might trade only in mutually acceptable reasons, sustain a respectful and unaggressive tone, listen sincerely to each other, invite objections and questions, and so on, while nonetheless arguing with a view towards projecting to the onlookers a particular conception of the dialectical situation that obtains among the interlocutors and their respective positions. These projections can serve to *miseducate* the onlookers in ways that serve to disable deliberation among them. To return to the case of a successful deployment of the Weak Man: the onlookers will be convinced that there really isn't anything to deliberate about. They will hold that there is but one position that is well-informed and defensible, and all of its critics have been handily repudiated; and they will conclude that anyone who sees fit to engage the question any further must be badly misinformed and thus not worth arguing with. Thus, enclave deliberation is encouraged, and the polarization dynamic is set in motion.

The crucial point bears repeating once again: when the polarization dynamic is initiated by the deployment of dialectical fallacies, no public deliberators need to have behaved uncivilly or in an epistemically improper manner toward their interlocutors. The perpetrators of dialectical fallacies can have clean hands from the perspective of their actual deliberative encounters. What renders them criticizable is something about their deployment of *legitimate* forms of dialectical exchange with their interlocutors. Perpetrators of dialectical fallacies opportunistically exploit the fact that onlooking audiences are often relatively uninformed about the issue under debate, and indeed are watching the debate for the purpose of learning about the respective merits of the disputing sides. The engagement with the actual interlocutor hence becomes incidental with respect to the actual aim of moving the audience by constructing for them a projection of the state-of-play among the disputants and the opposing views. In the cases most worth the attention of deliberative democrats, dialectical fallacies are deployed for the sake of projecting to an audience the view that there is but one intellectually responsible and defensible position to take on a given question, and thus all opposition is misguided, ignorant, or vicious. This kind of conception of the dialectical situation among the going positions in a dispute sets the polarization dynamic in motion, and thus undermines deliberative democratic practice by attacking the very presumption upon which it relies—namely, that there could be reasoned and sincere but stark disagreement among duly informed democratic citizens over important matters of public policy.

V. BLEAK UPSHOTS

Now, it seems to me that something like the account presented in the previous section is at work in a great deal of our current politics. We confront conditions where democratic citizens are not only increasingly enclaved, but also increasingly of the view that there could be no reasonable opposition to their political perspectives. As one would expect under such conditions, our public discourse is saturated with *medicalized* accounts of political disagreement; for many citizens, those with whom they disagree are not to be *reasoned with* but simply *diagnosed* as cognitively or morally impaired. The seemingly endless parade of political panels and debates operate as public spectacles that pantomime deliberation when actually serving only to confirm audience biases. In short, although our political practice has taken a deliberative turn, we are now living within a *simulation* of deliberative democracy, a context where our aspirations and attempts to realize them are systematically turned against themselves, resulting in a *distortion* of our democratic ideals and their further dissolution. However, I won't continue lamenting in this way; I don't think there's anyone who would vigorously dispute the claim that the contemporary state of democracy is worthy of serious concern. I will conclude instead by identifying a few dispiriting upshots of the foregoing analysis for democratic theory.

First, the account offered presents a supplement to Sanders's initial concerns. Recall that she cautioned against the deliberativization of democracy under political conditions where certain forms of inequality prevail; she argued that, when conducted amidst social inequality, deliberative democracy will simply reinforce those inequalities. The argument above suggests a similar lesson, but with an epistemological bent. Roughly: when conducted amidst a population that is epistemologically unequally situated, in that some are informed about the political issues under discussion whereas others are not, deliberative democracy creates distinctive opportunities for strategic arguers to create epistemic enclaves

among the less-informed, thereby initiating the polarization dynamic and disabling public deliberation. To put this point in a different way, deliberative forums and episodes of public deliberation may be helpful when it comes time for antecedently well-informed citizens to *evaluate* the reasons in play with regard to a given policy issue. But public deliberation is not the right format for finding out what reasons there are; it is not the right way to set about *informing* oneself about an issue. It seems then that deliberative democracy requires robust nondeliberative institutions by which citizens can prepare to deliberate; it can function only against a social epistemic backdrop of nondeliberative but shared sources of reliable information.

The trouble, of course, is that this very suggestion seems to contravene much of the spirit of deliberativism. After all, it is a call for what looks like a return to old-fashioned media and news formatting, where a newscaster presents the day's stories of note, and those stories become the basis for subsequent deliberations. Familiar hazards abound with this model. And one suspects that when fully elaborated, the suggestion involves a subtle form of epistemic paternalism, where citizens' everyday political talk must be facilitated and curated by experts who supply the framework and parameters within which citizen deliberation is to occur. This always involves the risk of degrading into something decidedly nondemocratic.[18]

Still, the reality remains that our media and communications technologies have already made the deliberative turn. Our politics is now increasingly conducted by means of pro-and-con argumentation and discussion performed purportedly for the sake of helping citizens to become informed and make up their minds about the pressing political issues of the day. The actual result of all the talking is that our politics have become increasingly *argumentative* and *disagreeable*, but far less reasoned and almost entirely devoid of actual disagreement. In fact, we now confront a media landscape of opposed "political realities," each with its own unique markers of epistemic reliability and norms of civility, sharing so little intellectual and moral ground that no discourse across such "realities" is possible. Amidst a permeation of pantomimed public deliberation, deliberative democracy is undermined.

The second upshot is that these conditions might not be remediable. Under existing conditions, democratic citizens need to be able to counteract the polarization dynamic. And this requires in many instances the ability to discern and diagnose dialectically fallacious performances of public argumentation. Yet, as we have seen above, the very idea of a dialectical fallacy is complicated. In fact, the entire enterprise of analyzing and assessing instances of *argumentation* calls for a family of robust and sometime unwieldy concepts, arguably an entire meta-language not unlike the kind deployed in formal logic. For example, a few moments' reflection on the concept of *hypocrisy* demonstrates the need for a fairly robust menu of epistemological concepts: in order to say what *hypocrisy is*, one must introduce, at the very least, distinctions between saying and doing, and intending and not-intending that are philosophically far more slippery than is often noticed. Or, to consider the matter from a slightly different direction, it looks as if any account of hypocrisy will have to investigate a possible conceptual connection between instances of *hypocrisy* and instances of *lying*; but we know that it is surprisingly difficult to state clearly what it is to *lie*.[19] Matters get only *more complex* when one attempts to devise the requisite theoretical tools for assessing and evaluating speakers' performances in argumentative encounters.

Yet deliberative democracy *requires* such tools to be both forged and mastered. *Prima facie,* the requirement looks doomed; one should not build into one's conception of democratic citizenship the requirement to master a substantive theory of argumentation. It may be retorted that what is needed is not that citizens learn argumentation *theory*, but only that they gain competence in the metalanguage by which argumentative performances can be assessed.

This retort is correct as far as it goes. However, the current state of argumentation theory does not allow for a clean distinction between the concepts employed in the metalanguage and the theoretical apparatus designed to elucidate them. To put the matter starkly: it is difficult, even among the professionals, to keep the theory and the theorized phenomena distinct. What looks to one argumentation theorist as a garden-variety case of some specific fallacy will look to the proponent of a different theory of argument as no fallacy at all. One worries that in order to get deliberative democracy right, one must first complete the task of argumentation theory as an academic discipline. Such completion is a long way off. And democracy can't wait.

I conclude with a third and related upshot. No matter how things stand with respect to the demandingness of the theoretical apparatus needed to detect and diagnose fallacious argumentative performances, deliberative democracy remains a demanding proposition. As argued above, the deliberative ideal involves not only a set of moral requirements; there are epistemological requirements, too. In light of the discussion of dialectical fallacies above, we might say that deliberative democracy calls for a substantive epistemological ethic, a set of norms delineating what one owes, epistemically, to one's interlocutors and to one's audience. These norms will include prescriptions outlining when one must concede a critic's point, revise one's view, change sides, be silent, admit one's error, suspend judgment for the purpose of gathering more information, and much else. We all know how difficult it is to abide by such norms in the context of relatively low-stakes exchanges at academic conferences and departmental meetings. To expect citizens as such to adopt them, and enact them reasonably successfully, in the relatively high-stakes contexts of politics seems naïve.

Of course, deliberative democrats will concede that the deliberativist epistemological ethic is demanding. They will agree that existing citizens are unlikely to embrace the requisite norms. However, they will next add that deliberative democracy must be learned, cultivated, and practiced. Again, this is correct as far as it goes. The trouble is that, if the arguments above are roughly correct, it seems that we cannot learn good deliberative democratic epistemological habits by engaging in public deliberation. To repeat my variation on Sanders's lesson: deliberative encounters under existing conditions should be expected to initiate the polarization dynamic. And it is not yet clear to anyone how demanding epistemological norms can be reliably cultivated.

We seem to have reached an impasse. Or, perhaps more accurately, we seem to be caught between two commitments that don't ultimately sit well together. On the one hand, we tend to take our own political views to be competently reasoned and well-informed. Additionally, we tend to take ourselves to be politically fair-minded, duly responsive to countervailing considerations, and welcoming of good criticism from formidable critics. Yet, on the other hand, we tend to see the vast majority of our political opponents to be short-sighted, less than rational, ignorant, unwittingly in the grip of various biases, ideologies, and illusions,

unwilling to fairly engage with their critics, unable to respond to objections, and worse. We take ourselves to be fully invested in the democratic project, and often see that project as involving public discourse among citizens who staunchly disagree; yet we often find ourselves unable to formulate the most powerful objections to our own most cherished political commitments. In fact, in unguarded moments, we are inclined to flatly deny that there are any such objections to the view we most deeply hold. Accordingly, we tend to see democracy's present ills wholly as the result of *others'* failings. We think democracy would get back on track if only everyone else would see the light. Too often, we tacitly think that when others "see the light," they will adopt our own most cherished political beliefs; we attribute the facts that our preferred candidate lost the election and our preferred policy lost at the polls to the ignorance, gullibility, immorality, and selfishness of others. Crucially, we never attribute our political wins to those same forces; when our side wins in democracy, it is always due to a triumph of public virtue and good sense. Consequently, we tend to see democratic progress as requiring deeper and expanding levels of unanimity, and an increasingly diminishing field of matters about which there could be reasoned disputation. Perhaps the deliberative turn has failed us all.

ACKNOWLEDGEMENTS

NOTES

[1] I take Habermas (1996), Benhabib (1996), Cohen (1997), Rawls (1997), and Gutmann and Thompson (2004) to be offering canonical versions of deliberative democracy.

[2] One telling example here is Ronald Dworkin. Dworkin never explicitly embraced deliberative democracy, but his vision of democracy grows increasingly deliberative. By 2006 (Dworkin, 2006), the view is thoroughly deliberativist.

[3] See, for example, Posner (2003), Somin (2016), and Achen and Bartels (2016). Other theorists are suspicious of deliberative democracy's demandingness; they hold that in general democratic citizens are not cognitively capable of public deliberation in the deliberative democrat's sense; see especially Brennan (2016), Kelly (2012), and Ahlström-Vij (2013).

[4] To get a flavour of these debates, one may consult three somewhat dated but still representative collections: Bohman and Rehg (1997), Elster (1998), and Fishkin and Laslett (2003). For a collection of more current work, see Steiner (2012).

[5] See also Young (1996; 2003). See Dryzek (2000) and Talisse (2005) for responses.

[6] It is worth noting that Gaus (2016) argues that *all* ideals contain this kind of danger: moves on the ground in the direction of realizing the idea involve unanticipated violations of the ideal.

[7] See Gutmann and Thompson (2004, p. 3-7) for a canonical articulation of deliberative democracy as a moral ideal. See also Habermas (1996), Cohen (1997), and Benhabib (1996) for alternative formulations of the idea that deliberative democracy is fundamentally a moral ideal.

[8] The connections between deliberative democracy and participatory models of democracy are worth exploring in their own right, though I cannot discuss them here. The association of the deliberative ideal with that of active participation is made explicit in Mansbridge (1983), Barber (2004), Ackerman and Fishkin (2004), and Pettit (2012).

[9] See Cohen (2008) and Estlund (2008).

[10] For example, see Achen and Bartels (2016), Brennan (2016), Somin (2016), and Posner (2003).

[11] For skepticism, see Ahlström-Vij (2013) and Kelly (2012). For a defense of the epistemic value of democracy, see Landemore (2013).

[12] Hardin (2002) provides an early analysis of the phenomenon, calling it "crippled epistemology."

[13] See Fricker (2007) and the materials collected in Kidd and Medina, eds. (2017).

[14] See Sunstein (2003) and Sunstein (2017) for reviews of the relevant empirical materials.

[15] See Aikin and Talisse (2006).

[16] That is, in the absence of an onlooking audience to whom the Straw Man is projected, there is simply a *mischaracterization* by one interlocutor of the other's view. The Straw Man involves a mischaracterization that is projected to an audience that is *not* one's interlocutor, plus the spectacle of knocking down an opponent.

[17] See Aikin and Talisse (2014a) for a fuller taxonomy.

[18] The real bite of the arguments presented by Ahlström-Vij (2013) comes from the evidence he provides that suggests that we tend to be unable to correct ourselves epistemically; we *need* paternalistic intervention to improve epistemically.

[19] On this, Saul (2012) is exemplary. Saul demonstrates that the seemingly simple task of identifying *what a lie is* in fact requires a remarkably subtle and intricate architecture of philosophical concepts drawn from epistemology, philosophy of language, philosophy of mind, and ethics.

REFERENCES

Achen, Christopher and Larry Bartels, *Democracy for Realists*, Princeton, Princeton University Press, 2016.

Ackerman, Bruce and James Fishkin, *Deliberation Day*, New Haven, Yale University Press, 2004.

Ahlström-Vij, Kristoffer, *Epistemic Paternalism: A Defence*, New York, Palgrave Macmillan, 2013.

Aikin, Scott and Robert Talisse, *Why We Argue (And How We Should*), New York, Routledge, 2014a.

—, "Why We Argue: A Sketch of an Epistemic-Democratic Research Program," *Inquiry*, vol. 29, no. 2, 2014b, p. 60-66.

—, "Modus Tonens", *Argumentation,* vol. 22, no. 8, 2008, p 521-529.

—, "Two Forms of the Straw Man", *Argumentation*, vol. 20, no. 3, 2006, p. 345-352.

Barber, Benjamin, *Strong Democracy*, Berkeley, University of California Press, 2003.

Benhabib, Seyla, "Towards a Deliberative Model of Democratic Legitimacy", *in* Benhabib, Seyla (ed.), *Democracy and Difference*, Princeton, Princeton University Press, 1996.

Brennan, Jason, *Against Democracy*, Princeton, Princeton University Press, 2016.

Bohman, James, "The Coming of Age of Deliberative Democracy", *Journal of Political Philosophy*, vol. 6, no. 4, 1998, p. 400-425.

Bohman, James and William Rehg (eds.), *Deliberative Democracy*, Cambridge, MIT Press, 1997.

Cohen, Joshua, "Procedure and Substance in Deliberative Democracy", *in* Bohman, James and William Rehg (eds.), *Deliberative Democracy*, Cambridge, MIT Press, 1997.

Cohen, Joshua, "Truth and Public Reason", *Philosophy & Public Affairs*, vol. 37, no. 1, 2009, p. 2-42.

Dryzek, John, *Deliberative Democracy and Beyond*, New York, Oxford University Press, 2000.

Dworkin, Ronald, *Is Democracy Possible Here?*, Princeton, Princeton University Press, 2006.

Elster, Jon (ed.), *Deliberative Democracy*, Cambridge, Cambridge University Press, 1998.

Estlund, David, *Democratic Authority*, Princeton, Princeton University Press, 2008.

Fishkin, James and Peter Laslett (eds.), *Debating Deliberative Democracy*, Oxford, Blackwell, 2003.

Freeman, Samuel, "Deliberative Democracy: A Sympathetic Comment", *Philosophy & Public Affairs*, vol. 29, no. 4, 2000, p. 371-418.

Fricker, Miranda, *Epistemic Injustice*, New York, Oxford University Press, 2007.

Gaus, Gerald, *The Tyranny of the Ideal*, Princeton, Princeton University Press, 2016.

Gutmann, Amy and Dennis Thompson, *Why Deliberative Democracy*, Princeton, Princeton University Press, 2004.

Habermas, Jürgen, *Between Facts and Norms*, Cambridge, MIT Press, 1997.

Hardin, Russell, "The Crippled Epistemology of Extremism", *in* Albert Breton (ed.), *Political Extremism and Rationality*, Cambridge, Cambridge University Press, 2002, p. 3-22.

Kelly, Jamie, *Framing Democracy*, Princeton, Princeton University Press, 2012.

Kidd, Ian James, José Medina, and Gaile Pohlhaus, Jr. (eds.), *The Routledge Handbook of Epistemic Injustice*, New York, Routledge, 2017.

Landemore, Hélène, *Democratic Reason*, Princeton, Princeton University Press, 2013.

Mansbridge, Jane, *Beyond Adversary Democracy*, Chicago, University of Chicago Press, 1983.

Pettit, Philip, *On the People's Terms*, Cambridge: Cambridge University Press, 2012.

Posner, Richard, *Law, Pragmatism, and Democracy*, Cambridge, Harvard University Press, 2003.

Rawls, John, "The Idea of Public Reason Revisited," *University of Chicago Law Review*, vol. 64, no. 3, 1997, p. 765-807.

Sanders, Lyn, "Against Deliberation", *Political Theory*, vol. 25, no. 3, 1997, p. 347-376.

Saul, Jennifer Mather, *Lying, Misleading, and What is Said*, New York, Oxford University Press, 2012.

Somin, Ilya, *Democracy and Political Ignorance*, second edition, Stanford, Stanford University Press, 2016.

Steiner, Jürg (ed.) *The Foundations of Deliberative Democracy*, Cambridge, Cambridge University Press, 2012.

Sunstein, Cass, *Why Societies Need Dissent*, Cambridge, Harvard University Press, 2003.

—, *Republic.Com 2.0*, Princeton, Princeton University Press, 2007.

—, *#Republic: Divided Democracy in an Age of Social Media*, Princeton, Princeton University Press, 2017.

Talisse, Robert, "Deliberativist Responses to Activist Challenges", *Philosophy and Social Criticism*, vol. 31, no. 4, 2005, p. 423-444.

Young, Iris, "Communication with the Other: Beyond Deliberative Democracy", *in* Seyla Benhabib (ed.), *Democracy and Difference*, Princeton, Princeton University Press, 1996, p. 120-135.

Young, Iris, "Activist Challenges to Deliberative Democracy", *in* James Fishkin and Peter Laslett (eds.), *Debating Deliberative Democracy*, Oxford, Blackwell Publishing, 2003, p. 102-120.