

Les thèses de complaisance : de l'acceptabilité d'un écart à l'intégrité académique

Magali Bringuier, Evelyne Decullier, David Malec et Jacques Py

Keywords: thesis of convenience, acceptability, ethics, scientific integrity, academic integrity

Mots clés: thèse de complaisance, acceptabilité, éthique, intégrité scientifique, intégrité académique

Résumé : Nous avons mené deux études afin d'identifier pourquoi les universitaires acceptent de décerner un titre prestigieux à des travaux non conformes aux attendus. À travers des entretiens semi-directifs conduits auprès de 19 universitaires, la première étude montre que les thèses de complaisance peuvent être soit envisagées comme telles par le directeur de thèse dès le début du travail doctoral, soit le devenir au fur et à mesure de l'avancée du doctorat. Dans la seconde étude, basée sur des scénarii ($N = 89$ universitaires), nous avons trouvé que l'acceptabilité de l'obtention d'un titre de complaisance dépend de la carrière envisagée par le doctorant et du type de faiblesses de la thèse. Ces recherches permettent, d'une part, de discuter comment l'éthique de la conviction des universitaires peut être supplantée par leur éthique de la responsabilité, et d'autre part, de soulever la question du référentiel, non partagé, qui de fait, permet une flexibilité quant aux attendus académiques.

Abstract: We conducted two studies to investigate why academics accept to award a prestigious title for unsatisfactory Ph.D. dissertations. Through semi-structured interviews with 19 academics, the first study showed that convenience theses could either be planned as such from the beginning of the doctoral work or become so as the doctorate develops. In the second, scenario-based study ($N = 89$ academics), we found that the acceptability of obtaining a title of convenience depends on the career envisaged by the doctoral student and the specific shortcomings of the dissertation. On the one hand, the present research allows to discuss how the academics' ethics of responsibility can supplant their ethics of conviction. On the other hand, it raises the question of the referential, which is not shared and allows for flexibility in requirements for delivering the PhD.

Publié dans *Actes du 2ème Colloque IRAFPA*, 2022, 243-254

<https://doi.org/10.56240/cmb9921>

Tout droit de reproduction réservé

Introduction

Le mouvement pour l'intégrité scientifique prend, depuis quelques années, une ampleur remarquable et se diffuse, certes lentement, dans le monde entier. En France, suite à la parution du rapport Corvol sur l'intégrité scientifique, l'Office Français d'Intégrité Scientifique (OFIS) a été créé et un réseau national des Référents à l'Intégrité Scientifique (RIS) s'est développé. En mai 2022, près de 200 universités et établissements d'enseignement supérieur et/ou de recherche ont nommé un RIS. Les organismes publics de recherche ont signé une charte de déontologie de la recherche, et les universités ont suivi ce mouvement. L'objectif est d'éviter les scandales de fraude scientifique, mais surtout les petits écarts aux bonnes pratiques scientifiques qui ont engendré une crise de la reproductibilité. Ainsi, la formation des jeunes chercheurs constitue une priorité. C'est pendant le doctorat que s'acquièrent les compétences du chercheur et l'assimilation des normes et valeurs morales propres à la science. Depuis mai 2016, un Arrêté français fixe le cadre national de la formation et les modalités de délivrance du Doctorat. Celui-ci impose une formation à l'éthique et à l'intégrité scientifique au cours de la thèse ; un serment du docteur qui s'engagera à respecter les principes et les exigences de l'intégrité scientifique dans sa future carrière sera très prochainement inclus dans la révision de cet Arrêté.

La recherche semble encadrée (Comité d'éthique, comité de protection des personnes, délégué à la protection des données), et les étudiants reçoivent une formation spécifique. Pourtant, le doctorat n'est pas exempt de problèmes d'intégrité. Les grands fraudeurs, comme le hollandais D. Stapel (Barthélémy, 2012), ont entrepris leur malversation scientifique dès la thèse. Parfois, les thèses témoignent de problèmes d'écriture, de manque de connaissance de la littérature pertinente, dans l'analyse des données ou dans la capacité à théoriser et interpréter. Ces insuffisances devraient se traduire par un échec à l'obtention du titre de docteur. Or cela n'arrive que très rarement (p. ex., en cas de fraude manifeste, comme un plagiat caractérisé), on évoque alors l'idée de thèses, ou de jurys de complaisance.

Il n'existe aucune littérature scientifique à propos des thèses de complaisance. D'ailleurs, la notion n'existe pas en anglais, ce qui rend impossible une recherche bibliographique informatisée. Les termes les plus proches pour évoquer une thèse de complaisance seraient « fake Ph.D. thesis » qui pourraient concerner aussi bien des thèses reposant sur un large plagiat que l'invention d'un diplôme. Pourtant, tous les universitaires connaissent ces termes, voire les emploient à l'occasion. La notion de « thèse de complaisance » est évoquée lorsque le jury de thèse est adapté à l'examen du travail d'un doctorant jugé trop faible pour se soumettre à une analyse rigoureuse de celui-ci. Est ainsi décerné un titre prestigieux à des travaux dont on pourrait estimer qu'ils ne le méritent pas. De leurs côtés, les autorités académiques,

de par le monde, développent des stratégies modifiées régulièrement, pour mieux cadrer l'attribution du titre de docteur. Le seul fait que ces stratégies varient régulièrement indique qu'aucune solution satisfaisante n'a été trouvée. Toute thèse a vocation à tendre vers l'excellence. Il est, en revanche, beaucoup plus complexe de déterminer actuellement des seuils. De plus, lorsque les évaluations intermédiaires sont passées avec succès : bilan/avis de fin de première année (lorsqu'il existe), avis du Comité de Suivi Individuel (CSI) pour l'inscription en troisième année, avis favorable des rapporteurs, la soutenance apparaît comme une formalité pour la délivrance du diplôme. À partir de quel seuil peut-on considérer qu'une thèse est trop éloignée de l'objectif de l'excellence, et donc a fait l'objet de complaisance de la part du jury si elle a été soutenue avec succès ?

En l'absence de littérature scientifique, notre projet est donc exploratoire et a pour objectifs : 1) d'identifier les raisons (relationnelles et institutionnelles) pour lesquelles des universitaires acceptent de décerner un titre prestigieux à des travaux non conformes aux attendus et les stratégies déployées par les directeurs de thèse pour rendre acceptables ces travaux doctoraux; et 2) de mesurer les facteurs d'acceptabilité des thèses de complaisance.

Néanmoins, nous formulons quelques hypothèses qui ont permis d'envisager les thèmes abordés dans les entretiens menés et les facteurs manipulés dans les scénarii. Nous postulons que les directeurs de thèse et membres de jury seront sensibles à plusieurs éléments contextuels : le type de financement de la thèse, le fait que la thèse soit réalisée avec un partenariat structurant pour le laboratoire, le type de carrière envisagée par le candidat et le type de "faiblesses du candidat" (conceptualisation, écriture, traitement des données), cette dernière pouvant avoir une incidence probablement différente en fonction de la discipline (littéraire, sciences humaines et sociales, ou sciences de la terre ou de la vie).

Les raisons de l'acceptabilité : Entretiens semi-directifs

La première étude repose sur une série d'entretiens semi-directifs menés auprès de directeurs de thèse et de membres de jury de thèse suffisamment expérimentés provenant de disciplines diverses. Les participants ont été sollicités directement dans les sphères de connaissances des auteurs, en veillant à obtenir une certaine diversité de profils, notamment disciplinaire.

Le guide d'entretien comprenait une question de départ sur les thèses et jurys de complaisance : « si je vous dis « thèse de complaisance ou jury de complaisance », qu'est-ce que cela vous évoque ? ». Ensuite, différentes thématiques ont été abordées dont : les thèses peuvent ou doivent-elles tendre vers l'excellence ? Comment définir un seuil d'acceptabilité ? Les entretiens se sont déroulés via

l'application Zoom ou en présentiel avec enregistrement dont seule la trace audio a été conservée pour la durée de l'étude. Les entretiens, d'une durée d'une à deux heures, se sont déroulés de manière très fluide, les interviewés se prêtant très facilement à la narration orale de leurs expériences ainsi qu'à l'analyse des situations décrites en qualité de participant ou de témoin.

Participants

Au 3 juin 2022, 19 membres de jurys de thèse ont participé à un entretien d'une durée d'une à deux heures. Les caractéristiques des participants sont les suivantes :

- ❖ champs disciplinaires: Sciences de la vie et de la Terre, Chimie, Physique, Génie électrique, Électronique, Mécanique, Santé, Psychologie sociale, Géographie, Droit ;
- ❖ jurys de thèse : de 20 à 150 participations à des jurys de thèse dirigées par des collègues ;
- ❖ encadrement de thèse : de 8 à 20 thèses amenées en soutenance.

Analyse du contenu des entretiens

Cette série d'entretiens, permet de dégager quelques grands enseignements. D'abord, la notion de complaisance est d'emblée associée à des fraudes, en particulier à des plagiat manifestes. Concernant les thèses qui ne relèvent pas d'une fraude, mais dont la qualité n'apparaît pas à la hauteur des attentes des membres du jury, de nombreux participants semblent mal à l'aise par l'absence d'un référentiel commun. Chacun est bien conscient qu'il existe des thèses de qualité variable, certaines s'avérant médiocres, soit parce qu'il apparaît un manque de conceptualisation ou de maîtrise de la littérature, ce que les participants traduisent souvent par le fait qu'ils n'apprennent rien à la lecture de l'introduction générale, soit parce que l'apport scientifique est faible, parfois assorti d'un traitement des données approximatif.

Chaque témoignage est unique, car chaque situation est différente, mais il ressort globalement une classification en deux grandes catégories :

- ❖ les thèses qui démarrent a priori de manière classique ;
- ❖ les thèses qui démarrent sciemment sur un format de thèse de complaisance et qui se terminent naturellement par un jury de complaisance.

Dans la première catégorie, il n'y a pas, au moment du recrutement, d'a priori négatif sur les capacités du candidat à mener un travail de thèse en pleine autonomie et maîtrise du sujet, celui-ci ayant été recruté après un processus de sélection exigeant. La complaisance s'esquisse en fin de thèse car le travail n'est pas à la hauteur, mais la thèse doit absolument être soutenue. Quelles sont les motivations de la direction de thèse (parfois en binôme) pour amener en soutenance un travail

insuffisant ? Plusieurs raisons sont invoquées. Certaines se regroupent sous la formule : 'le système ne nous incite pas à arrêter une thèse avant la diplomation'. Derrière cette phrase se classent les arguments suivants :

- ❖ financement des thèses : il faudra argumenter les raisons de l'échec et justifier le mauvais recrutement. Le directeur de thèse risque alors de ne pas être classé comme prioritaire à la prochaine attribution de financement de thèses (par l'équipe de recherche, le laboratoire, l'école doctorale, etc.) ;
- ❖ évaluation des chercheurs et enseignants-chercheurs : un abandon de thèse pourra être perçu comme un échec, et se remarquera par la baisse des publications ;
- ❖ réciprocité : comme il nous arrivera peut-être un jour de forcer un peu le passage d'une thèse avec un doctorant de faible niveau, il faut également répondre à la sollicitation d'un collègue nous demandant d'être magnanime avec le sien ;
- ❖ maintien des relations internationales : un arrêt de thèse pourrait ternir ou faire périlcliter la relation entre les universités partenaires.

Dans la seconde catégorie de thèses, il y a une volonté délibérée de pousser le candidat jusqu'au bout, quel que soit son niveau. Ces thèses peuvent être liées ;

- ❖ soit à des enjeux professionnels : il s'agit de faire travailler dans les laboratoires ou en entreprise des doctorants sur des thématiques de recherche nécessitant beaucoup de travail expérimental sur paillasse ou du développement de programmes informatiques. En effet, leur coût est moindre que celui d'un ingénieur de recherche. Le titre de docteur est alors considéré comme une compensation du travail effectué à bas coût ;
- ❖ soit, plus souvent, à des enjeux personnels : la course aux publications et aux thèses soutenues, point abordé plus haut, peut pousser certains encadrants en manque de reconnaissance et avides de promotion à fabriquer des doctorats, au détriment de la qualité de la recherche.

En sciences humaines et sociales, on relève aussi de nombreuses thèses réalisées par des professionnels, dont l'immense majorité ne se destine pas à une carrière académique. En droit, en particulier, un titre de docteur permet d'éviter l'examen d'entrée – très sélectif – à l'École d'avocat. Des thèses sont aussi réalisées par des ressortissants de pays du Sud qui leur permettront d'obtenir un statut de haut fonctionnaire de retour au pays d'origine. On n'aura pas le même niveau d'exigences pour tous ces types de candidats, mais, parfois, le directeur de thèse écrit lui-même une bonne partie de la thèse pour qu'elle soit acceptable.

Ces éléments nous ont amenés à construire une étude quantitative dans laquelle nous allons évaluer le niveau d'acceptabilité d'une thèse de complaisance en fonction de quelques facteurs qui sont ressortis des entretiens.

Acceptabilité : étude par questionnaire

Pour évaluer la mesure avec laquelle l'obtention d'un titre de docteur délivré par un jury de complaisance est acceptable, nous avons utilisé la théorie fonctionnelle de la cognition d'Anderson (1981) qui permet une mesure indirecte de l'acceptabilité, relativement peu sensible à la désirabilité sociale.

Les participants suivaient un lien les menant sur la plateforme Qualtrics. La première page de l'étude permettait de recueillir leur consentement à la participation. Ensuite, la passation se déroulait en deux phases : une phase de familiarisation, une phase d'expérimentation. Durant la première phase était présenté 11 % du matériel final, soit 4 scénarii sélectionnés aléatoirement parmi les 36 créés pour l'étude. Cette phase sert d'entraînement, et permet aux participants de comprendre la construction des scénarii, la question posée et l'échelle de réponse proposée. Pendant cette phase, le participant peut, s'il le souhaite, revenir en arrière et modifier ses réponses, ce qui ne sera plus le cas par la suite.

Matériel

Les scénarii décrivent M. Dupond, convié à participer à un jury de thèse, et qui, lors de l'évaluation, comprend que la thèse n'est pas conforme aux attendus, mais qui décide, comme le reste des membres du jury, d'accorder le titre de docteur au candidat. Nous avons croisé 4 facteurs (36 scénarii): le financement de la thèse (CDU; CIFRE), le partenaire avec lequel a été menée la thèse (importance majeure; mineure), les faiblesses du candidat (Conceptualisation ; Écriture ; Traitement des données) et la carrière envisagée par le candidat (Carrière universitaire en France ; à l'étranger ; non académique). Chaque scénario est suivi de la question « À quel point pensez-vous que le comportement de M. Dupond est acceptable ? ». Les participants jugeaient de l'acceptabilité de la situation sur une échelle en 11 points allant de « pas du tout acceptable » à « totalement acceptable ».

Des analyses de la variance ont été menées, ainsi que des analyses en clusters (Hoffmans et Mullet, 2013) pour mettre en évidence l'existence de différents profils de réponses.

Résultats

Un total de 89 participants a complété entièrement l'étude. Les caractéristiques des participants sont présentées dans le Tableau 1.

		Total	Cluster1	Cluster2
Sexe	Homme	50	21	29
	Femme	39	18	21
Âge	45 ans ou moins	28	14	14
	46 ans ou plus	61	25	36
Avec ou Sans habilitation à diriger les recherches (HDR)	Sans précision	1	0	1
	Avec HDR	73	31	42
	Sans HDR	15	8	7
Champ disciplinaire	LSHS	29	13	16
	Science Techno Santé	60	26	34
Expérience des jurys de thèse	Moins de 5	25	12	13
	5 fois et plus	64	27	37

Nb de thèses dirigées soutenues	Moins de 5	37	20	17
	5 et plus	46	16	30

Tableau 1. Description de l'échantillon

Les résultats de l'ANOVA indiquent que seuls 2 des 4 facteurs manipulés ont un effet significatif sur le niveau d'acceptabilité accordé à la validation d'une thèse qui n'est pas strictement conforme aux attendus : les faiblesses du candidat ($p < .001$) et la carrière envisagée par le candidat ($p < .001$). Les faiblesses d'écriture sont significativement plus acceptables que les faiblesses de conceptualisation et que les faiblesses de traitement des données. Il est également plus acceptable de donner le titre de docteur à une personne souhaitant ne pas continuer dans le milieu académique qu'à une personne souhaitant poursuivre une carrière académique à l'étranger ($p < .001$). Enfin, vouloir poursuivre une carrière académique en France est la modalité avec laquelle l'obtention d'un titre de complaisance est la moins acceptée.

Deux clusters ($n = 39$ et $n = 50$) ont été mis en évidence. Le Tableau 1 détaille la composition de chaque cluster. Le Cluster 1 est associé à une moyenne d'acceptabilité significativement plus élevée que le Cluster 2 ($p < .001$). Comme le suggère la Figure 1, l'effet d'interaction entre les facteurs « cluster » et « carrière envisagée » est significatif. Dans le cluster 1, la modalité la moins acceptée est le fait d'envisager une carrière académique en France. Dans le cluster 2, les modalités les moins acceptées sont lorsque le candidat envisage une carrière académique (sans distinction selon que la carrière se déroulerait en France ou à l'étranger).

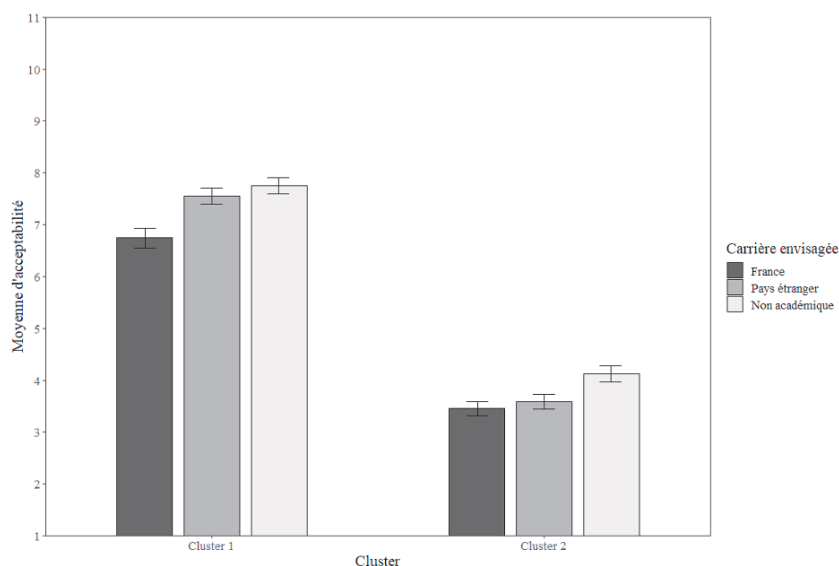


Figure 1. Acceptabilité moyenne en fonction du cluster et de la carrière envisagée

Il y a une différence d'acceptabilité générale selon le cluster. Elle est de 3,62 (sur une échelle variant de 1 à 11) pour les participants du cluster 2, les plus exigeants (et les plus nombreux) ; elle est de 7,17, soit le double, pour les participants du cluster 1. À noter que le seuil de basculement entre acceptabilité et non-acceptabilité est situé à 6. Les membres du cluster 2 expriment donc clairement une faible acceptabilité à propos d'écart aux attendus, alors que les membres du cluster 1 expriment une certaine acceptabilité. Aucune variable sociodémographique ne rend compte de la répartition des participants entre les deux clusters. Ce n'est donc ni l'âge, ni le sexe, ni le statut universitaire, ni le nombre de participations à des jurys de thèse, ni le nombre de thèses dirigées amenées en soutenance, ni même le secteur disciplinaire (LSHS vs sciences) du répondant qui permet de comprendre la grande différence d'acceptabilité entre les deux groupes. Des débriefings semblent nécessaires pour explorer les caractéristiques qui pourraient différencier ces deux groupes de sujets. A ce stade, on doit simplement conclure qu'il existe deux populations de membres de jurys de thèse qui se prononcent très différemment sur la question de l'acceptabilité d'écart aux normes académiques, l'une se montrant plus sensible aux éléments de contexte que l'autre.

Discussion

Éthique de la conviction et éthique de la responsabilité

Notre étude quantitative, et plus encore les entretiens que nous avons menés indiquent que le contexte institutionnel incite les membres d'un jury de thèse à se détacher d'une éthique de la conviction pour assumer une éthique de la responsabilité. Le devoir du chercheur est la recherche de la vérité. Cette exigence de vérité, que Weber désignait comme la valeur-finalité de la science, s'accompagne d'une exigence de probité, d'honnêteté intellectuelle (Hottois, 1996). Cette recherche de la vérité a trouvé son expression dans le principe d'évaluation par les pairs, que nous avons qualifié de « pire système d'évaluation de la production scientifique, à l'exception de tous les autres » (Py, 2021). Pour un travail doctoral, le système d'évaluation est assez proche d'un système d'évaluation par les pairs. Le travail du jury consiste, dans son principe, à évaluer honnêtement et rigoureusement une production scientifique. Or, les contraintes imposées par le système de l'université du 21^e siècle obligent les membres du jury à pondérer une éthique de la conviction par une éthique de la responsabilité, celle que Weber attribuait à l'homme d'action et à l'homme politique quand il considérait que la première devait caractériser l'homme de science. En effet, la production scientifique réalisée par le doctorant s'inscrit dans un contexte de production qui comporte de fortes incitations (financement, contraintes temporelles, production d'articles, etc.) nécessaire à une insertion universitaire. Il ne s'agit pas de faire avancer la science ; il s'agit de produire des « bouts de science », un assemblage de recherches, pas nécessairement bien articulées ni de réel intérêt scientifique, permettant chacune la production d'un article scientifique. Pour le directeur de thèse, la direction d'une thèse arrivée en soutenance est un élément d'appréciation important de son curriculum vitae. Si la thèse n'arrive pas à la soutenance, la réputation du directeur de recherche s'en trouvera affectée, du point de vue de l'École Doctorale, de celui du financeur, voire des pairs qui auront à évaluer sa carrière. Ainsi, face au produit à évaluer, le jury (soumis aux mêmes contraintes) va mettre les choses en perspective, se sentir responsable du devenir du doctorant comme du directeur de thèse, et teinter son éthique de la conviction par une éthique de la responsabilité. Nous ne discutons pas ici de la nécessaire hauteur de vue lorsqu'on est en position d'évaluer autrui. Nous discutons d'une forme de confusion dans la prise en compte de valeurs pertinentes et de valeurs qui ne devraient pas être prises en considération.

La question du référentiel

Lorsqu'on évoque la notion de complaisance, on se réfère à un ensemble de normes qui ne seraient pas respectées. Or, il n'y a pas réellement d'étalon de mesure en matière de travail scientifique. Lorsqu'on évalue une thèse, on peut trouver le travail médiocre tout en ayant conscience que cette appréciation est relative ; elle est référée aux productions réalisées au sein d'une communauté relativement circonscrite, et ne sera pas partagée par des collègues se référant à d'autres communautés de production. La question du référentiel a été évoquée par les participants de nos entretiens. Il est relativement aisé de s'opposer à la soutenance d'une thèse frauduleuse, car elle représente un écart majeur au devoir de vérité inhérent au chercheur. Il apparaît difficile de s'opposer à la soutenance d'une thèse médiocre, car les écarts par rapport aux attendus sont difficiles à cerner, ces attendus n'étant pas définis et ne faisant pas l'objet d'un référentiel universellement, ni même localement, reconnu.

Quel que soit le type de thèse de complaisance, la disparition des mentions et félicitations du jury permet de masquer la médiocrité du travail puisque le curseur qui permettait jadis de quantifier la qualité du travail de recherche a disparu. Pour certaines disciplines, l'absence d'un véritable rapport de thèse contribue également à rendre invisibles les faiblesses du travail produit. Cet état de fait cause du tort à l'image du doctorat qui continue de « souffrir de complexes » par rapport aux diplômes délivrés par les grandes écoles. Lutter efficacement contre les thèses de complaisance nécessite de remettre une éthique de la conviction au centre du travail du chercheur, y compris dans ses tâches d'évaluation des pairs qui sont constituantes de son travail d'Homme de science. Pour cela, il convient que la distribution des renforcements valorise les conduites utiles d'évaluation au détriment des conduites perverses (Beauvois, 1976).

Bibliographie

Anderson, N. H. (1981). *Foundations of information integration theory* (Vol. 578). Academic Press New York.

Barthélémy, P. (2012, December 09). Le scandale Stapel, ou comment un homme seul a dupé le système scientifique. *Le Monde*.
https://www.lemonde.fr/passeurdsciences/article/2012/12/09/le-scandale-stapel-ou-comment-un-homme-seul-a-dupe-le-systeme-scientifique_5986311_5470970.html

Beauvois, J.-L. (1976). Problématique des conduites sociales d'évaluation. *Connexions*, 19, 7-30.

Ben-Shachar, M., Lüdtke, D., & Makowski, D. (2020). Effect size: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, 5(56), 2815. <https://doi.org/10.21105/joss.02815>

Hofmans, J., & Mullet, E. (2013). Towards unveiling individual differences in different stages of information processing: A clustering-based approach. *Quality & Quantity*, 47(1), 455–464. <https://doi.org/10.1007/s11135-011-9529-7>

Hottois, G. (1996). « Éthique de la responsabilité et éthique de la conviction ». *Laval Théologique et Philosophique*, 52(2), 489-498.

Py, J. (2021). Déontologie et processus de publications scientifiques. In M. Bergadaà & P. Peixoto (Eds.), *L'urgence de l'intégrité académique*. EMS - Editions Management & Société ; 131-148.

Weber, M. (1919, 1959). *Le savant et le politique* (trad. par J. Freund). Plon.