



HAL
open science

La psychologie doit la vérité scientifique pour la justice sociale : Une réponse à Canguilhem (1958)

Stéphane Vautier, Nadine Mazet

► To cite this version:

Stéphane Vautier, Nadine Mazet. La psychologie doit la vérité scientifique pour la justice sociale : Une réponse à Canguilhem (1958). 2020. halshs-02535106

HAL Id: halshs-02535106

<https://shs.hal.science/halshs-02535106>

Preprint submitted on 7 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La psychologie doit la vérité scientifique pour la justice sociale

Une réponse à Canguilhem (1958)

Stéphane Vautier et Nadine Mazet

Université de Toulouse¹

(Texte soumis pour publication à la Revue de Métaphysique et de Morale le 24 janvier 2020)

¹ Université de Toulouse Jean Jaurès, 5 allées Antonio Machado, 31058 Toulouse cedex 9. Courriel : vautier@univ-tlse2.fr

Résumé

Canguilhem (1958) s'en prenait à la bonne conscience méthodologique du psychologue-testeur : peux-tu me connaître si tes techniques d'observation (tests psychologiques) supposent que tu ne me reconnais pas comme ton égal ? Faut-il répondre nettement : l'objectivation impliquée par la curiosité scientifique en train de fabriquer des capteurs d'information et de les appliquer à des gens est possible. Mais elle doit une responsabilité scientifique, c'est-à-dire rationnelle et critique : ne pas tricher avec la nécessité logique et dire honnêtement ce qu'elle peut savoir et ce qu'elle ne peut pas savoir. Or, l'institution du testing psychologique s'est développée en fraudant avec cette exigence. Dès qu'elle confond mesurer et scorer, la psychologie porte atteinte à la vérité scientifique et trompe l'opinion publique en lui faisant prendre des conventions évaluatives pour des nécessités objectives. Cela est possible parce que les gens sont aussi des objets sociaux, circulant dans un réseau d'autorisations. Le zèle de la psychologie pour intervenir dans l'autorisation des gens dans ce réseau nécessite aussi qu'elle prenne ses responsabilités vis-à-vis d'une connaissance de l'utilité sociale, dans une époque où il faut reposer le problème de la nécessaire surveillance de l'Homme par l'Homme.

Abstract

Canguilhem (1958) criticized the methodological good consciousness of psychological testing: how can the psychologist know me if her/his observational techniques (psychological tests) presuppose that s/he does not recognize me as her/his alter ego? A neat answer is required: the objectification entailed by scientific curiosity which manufactures sensor systems and applies them to people is possible. But scientific curiosity works with scientific, i.e., rational and critical, responsibility: do not cheat the logical necessity and honestly say what can be known and what cannot be known. Yet psychological testing has grown by cheating such an obligation. As soon as "to measure" and "to score" are interchangeable in its discourse, psychology infringes the scientific truth and deceives public opinion, by presenting evaluative conventions as objective necessities. This is possible because people are social

objects that circulate into a flow of authorizations. The zeal of psychology to contribute to the authorization process requires also its responsiveness with respect to the knowledge of social utility, within a historical period that poses the problem of the necessary surveillance of human beings by human beings.

La psychologie doit la vérité scientifique pour la justice sociale

Une réponse à Canguilhem (1958)

Il y a plus d'un demi-siècle dans la Revue de métaphysique et de morale, méditant sur l'édification universitaire et professionnelle de la psychologie française (voir aussi Parot, 2017), Canguilhem (1958) exprimait « la répugnance [du testé] à se voir traiter comme un insecte, par un homme à qui il ne reconnaît aucune autorité pour lui dire ce qu'il est ou ce qu'il doit faire » (p. 24). L'examen terminait par un « conseil d'orientation » : en sortant de la Sorbonne par la rue Saint-Jacques, on monte vers le Panthéon, « Conservatoire de quelques grands hommes », ou on descend sûrement vers la Préfecture de Police. Ces phrases suffisent : le défi à répondre a été jeté et l'honneur académique est jeu. La provocation doit donc être honorée et la ligne de la défense qui s'impose est de ce tabac : lorsqu'il se prévaut de son diplôme universitaire, le psychologue peut aussi se tourner vers la cité tout entière pour dénoncer une erreur scientifique et, croit-il, morale de sa discipline puis, en gage de bonne foi, s'engager à la corriger.

La psychologie, discipline académique à laquelle l'institution du testing psychologique adosse son autorité scientifique vis-à-vis de l'opinion, se trouve dans la situation du fraudeur qui, sentant grandir la probabilité d'un contrôle, hésite : se dissimuler ou se livrer ? Mais la caricature est fautive parce qu'il n'existe pas de contrôleur de la psychologie. Sentir qu'on fraude suppose d'abord sentir un cadre tacite d'honorabilité. Il faut compliquer un peu : la psychologie se trouve dans la situation d'un notable qui ressent de plus en plus vivement la fragilité de son autorité ou de sa crédibilité ou de sa valeur sociale et qui, examinant sa pratique, pressent que les petits arrangements dont il a pris l'habitude, récompensés au sein même de son université et auxquels il se livre en toute conformité avec la pratique des confrères ou consœurs, sont « un peu » frauduleux. Il faut tout de suite ajouter que la mauvaise conscience scientifique touche les sciences humaines et sociales, probablement aussi les sciences biologiques et médicales. Il n'est donc pas question de faire de la psychologie une sorte de bouc émissaire et il faudrait, à tout prendre, saluer sa tolérance à la critique qu'elle permet ici ou là.

L'objet de la fraude est, aussi étrange que cela puisse paraître, une liberté épistémologique majeure : on s'autorise à tricher avec la nécessité logique (le moins possible, à chacun d'apprécier), ce qui constitue une fraude morale seulement si on accepte que le cadre d'honorabilité scientifique soit l'attitude rationaliste, dont on sait la contingence puisque il serait irrationnel de décréter que tous les phénomènes sont rationnels. Si on admet que science implique connaissance rationnelle, la fraude s'institutionnalise par une inconvenance académique (par ridiculisation collective) : la formulation de cette licence vis-à-vis de la nécessité logique ainsi que la formulation de la négation de cette licence sont forcloses de la méthodologie officielle² ; la nécessité morale (rationaliste) de ne pas tricher avec la nécessité logique est devenue une sorte de niaiserie, du même ordre mais pas du même âge, que la croyance enfantine selon laquelle les parents ne mentiraient pas : en parler serait inconvenant et la mise à jour (update) de la politique épistémologique de la discipline n'est pas encore à l'ordre du jour, pour la raison suffisante que la psychologie académique est née d'une différenciation d'avec la philosophie.

Cependant, à chaque fois qu'elle est commise, la tricherie affaiblit la fiabilité de la vérité logique du discours. La validité déductive des énoncés qui se déduisent de ce discours n'est plus garantie, puisque l'implication $A \Rightarrow B$ ne transmet à la Vérité (logique) de B que la valeur « vrai » de la Vérité de A – si A est fausse, la valeur de Vérité de B est contingente, c'est-à-dire non-nécessaire³, connue en tant qu'inconnue dans le calcul logique. D'où l'atteinte à l'intérêt général, c'est-à-dire à la chose publique, pour autant que la vérité logique (avec une majuscule si on veut expliciter qu'on parle de l'ensemble {vrai, faux}) soit un bien public, comme Weil (1957) l'a réclamé. L'entorse à la nécessité logique dérègle la circulation de la vérité des propositions dans le discours, « infecte » le savoir à la manière d'un malware, par le bouche-à-oreille qui transporte le discours. L'absence de pointillisme en vigueur vis-à-vis de la vérité, qui revient à un détournement collectif de l'attention, n'efface pas la question sous-entendue par le philosophe lorsqu'il se méfie du psychologue-testeur : d'où vient, dans ces conditions, la vérité des descriptions que les psychologues font des gens testés ?

² Forclos, hors champ ; ce qui correspond à l'état « perdu dans l'automatisme de l'exécution » (p. 14) qu'utilise Canguilhem (1958) pour décrire une méthodologie qui a oublié son « projet », ainsi qu'au somnambulisme que Stengers (2013a p. 39) emploie pour décrire les « chercheurs » contemporains.

³ Non pas contingente « dans l'absolu », mais contingente dans l'algèbre descriptive (ou, en langage probabiliste, d'événements) générée par A et B.

Cinquante ans après, à notre connaissance, cette question reste ouverte et la psychologie le sent bien.

Mais nécessité logique oblige, faut-il aussi objecter rationnellement, la psychologie ne peut ni sentir, ni mentir, puisqu'elle n'a pas de pensée. Seuls les gens peuvent penser, c'est-à-dire des animaux dotés d'une psychologie et donc, admettons-le pour des raisons positivistes, d'une « neuro-cognition », en remarquant que les positivistes n'ont pas nécessairement tort sur tout même si la vérité complète qu'ils espéraient s'est avérée une impossibilité logique (Gödel, 1931). Or aucune institution n'est dotée d'une neuro-cognition. La psychologie n'est pas non plus une personne morale parce que cette dignité juridique suppose la possibilité d'une infraction pénale. Or même si elle atteint à la vérité scientifique, une fraude irrationaliste ne relève pas du pénal. En outre il ne s'agit pas d'accuser la psychologie d'irrationalité puisqu'elle est moralement irresponsable, la responsabilité morale supposant une conscience. Comment pouvons-nous attribuer une psychologie et donc une mentalité à l'institution « psychologie » en particulier et à une institution en général ?

Une institution est un fait social (Searle, 1995), c'est-à-dire une croyance commune, un « consentement » ou une convention. Par exemple, la vérité de l'énoncé A : « ce billet de cinq euros vaut cinq euros » ne tient qu'à la vérité de l'énoncé B_A : « suffisamment de monde croit ou admet qu'un billet de cinq euros vaut cinq euros ». On a $[(B_A \Rightarrow A) \text{ et } B_A] \Rightarrow A^4$. Si l'implication $B_A \Rightarrow A$ est fautive, il n'est logiquement pas nécessaire qu'un billet de cinq euros vaille cinq euros même si « tout le monde » le croit. Une institution est identifiable par quelques combinaisons de signifiants (la monnaie européenne par exemple). Elle est le produit culturel et historique d'un réseau social de psychologies individuelles qui se réclament de cette institution (via le signifiant). Ce réseau social évolue au gré de l'évolution continue des psychologies qui s'y rattachent. On approche par les termes flous de croyances et d'affects collectifs, ou encore d'opinion, les caractéristiques cognitives de ce réseau qui permettent de lui attribuer une *mentalité*. Si bien qu'il faut méthodologiquement trancher dans le vif puisqu'on ne sait pas décrire plus finement la chose : simplifier et modéliser par analogie en espérant que l'anthropomorphisme psychologisant de l'esquisse soit un procédé descriptif acceptable, au sens d'une description pauvre c'est-à-dire nécessairement réductrice, mais dont la généralité vaille.

⁴ « B_A » se lit « B qui suppose A ».

Grâce à une psychologisation simplificatrice des institutions (instances sociales impersonnelles et privées d'intention mais que l'on dote d'une personnalité pour résumer un état stable des croyances de leur affiliés), nous pourrions montrer comment la fraude épistémologique, que pressentait peut-être Canguilhem mais qu'il n'a pas diagnostiquée même s'il a créé la piste d'une réponse à apporter, se mue en un jeu de dupes qui s'adresse à l'opinion publique : il importe que l'opinion publique respecte l'autorité scientifique des méthodes de testing psychologique et donc des conclusions qui sont tirées de leurs applications. D'où la propagande : les tests sont des instruments de mesure scientifiquement validés (cf. l'article 24 du Code 2012 de déontologie des psychologues ; voir aussi la dénonciation du mythe du testing par Michell, 2009). Les psychologues peuvent tester les gens avec la scientificité des méthodes d'observation de leur discipline. Néanmoins certaines consciences au sein de l'institution, n'ignorant sans doute pas que la parole soit un pouvoir, notamment un pouvoir d'affecter les gens qu'elle décrit, ont des scrupules. Aussi le comité de rédaction du Code de déontologie des psychologues (ci-après le Code) est-il convenu de montrer de la mesure. Le comité du Code de 1996 écrit dans l'article 19 :

Le psychologue est averti du caractère relatif de ses évaluations et interprétations. Il ne tire pas de conclusions réductrices ou définitives sur les aptitudes ou la personnalité des individus, notamment lorsque ces conclusions peuvent avoir une influence directe sur leur existence.

Six ans plus tard le comité de rédaction du Code de 2012, remplaçant l'article 19 par l'article 25, ne reconnaît plus le pouvoir de la parole de l'évaluateur sur l'évalué et table sur l'évaluation des « ressources » humaines :

Le psychologue est averti du caractère relatif de ses évaluations et interprétations. Il prend en compte les processus évolutifs de la personne. Il ne tire pas de conclusions réductrices ou définitives concernant les ressources psychologiques et psychosociales des individus ou des groupes.

Mais il conserve la recommandation déontologique de *ne pas* tirer de conclusions « réductrices ou définitives » sur les gens, autrement dit de tirer des conclusions non-« réductrices » *et* « provisoires ». Toutefois ces précautions verbales ne règlent pas la question de savoir d'où viennent l'autorité, la crédibilité, la valeur des conclusions. En pratique, la vérité de la parole (ou du discours) des psychologues se passe de la nécessité logique

puisqu'elle tient à la crédulité de ceux qu'elle décrit ainsi qu'à celle, feinte ou pas, de ceux qui commandent les conclusions, c'est-à-dire qu'elle tient à la confiance... tant que la confiance est accordée. Or les temps changent, la confiance cède à la méfiance, l'opinion publique « émerge » d'un réseau de personnes qui peuvent penser et donc apprendre à douter, et il se pourrait que des comptes soient demandés puisque, transition écologique oblige, il faut reconstruire globalement la société et donc les règles qui gouvernent les pouvoirs et les libertés des personnes comme des institutions.

Notre réponse à la question de Canguilhem : en sortant de l'université, la psychologie, c'est-à-dire chaque psychologue diplômé ou chaque aspirant au titre, peut explorer la seconde option qui s'offre à lui quand il hérite d'une fraude : (1) avouer l'erreur puis (2) prendre une résolution. (1) Dire la complaisance intéressée de sa discipline vis-à-vis de la vérité scientifique et corriger : on ne triche pas avec la nécessité logique. (2) Décider de prendre soin, dorénavant, de cette dernière, en prévision des besoins d'éclairage que le gros temps qui se prépare pour les institutions ne manquera pas de faire naître. Le programme d'exploration vise deux objectifs : (2.1) apprendre à connaître la nécessité logique, (2.2) prendre ses responsabilités morales et cognitives vis-à-vis des nécessités sociales qu'on accepte ou qu'on refuse. Dans un contexte socio-économico-politique troublé, une embellie en psychologie pourrait venir de l'officialisation de la différence profonde entre évaluer et mesurer (Vautier, 2015), ainsi que d'une déontologie de restriction de l'usage de l'évaluation psychologique officielle aux seules situations où celle-ci est une nécessité sociale (une utilité publique). L'évaluation de l'utilité sociale de l'évaluation psychologique est un immense chantier pluridisciplinaire mais, au sein de la psychologie et techniquement, il s'agit d'assumer, au titre de la responsabilité et de l'autonomie professionnelles revendiquées par le Code, que le scoring des comportements observés soit, par la nécessité logique qu'il faut saluer, une description scientifiquement fautive de la personne. En d'autres termes le scoring est une négligence scientifique donc lucide, parcimonieuse et susceptible de controverse rationnelle et critique. Rationnelle, c'est-à-dire telle que la vérité des énoncés qui s'échangent soit réglée par les règles de la logique ; critique, c'est-à-dire telle que la vérité des prémisses soit reconnue comme conventionnelle donc discutable.

1. Le jeu de dupes vis-à-vis de l'opinion publique

Pour situer le problème de l'autorité du testeur sur le testé, Canguilhem (1958) utilise l'analogie du « test à deux degrés qui [...] permet [à Gédéon, La Bible : Juges, Livre VII] de ne retenir d'abord que dix mille hommes sur trente deux mille, puis trois cents sur dix mille » (p. 24). Nous adopterons une analogie plus métabolique pour éviter le simplisme d'une société-troupeau gouvernée par un berger gouverné par l'Éternel. Admettons que tout membre de la société soit un objet dans un flux d'objets, dont la trajectoire dépend, entre autres choses, de ses *autorisations d'entrée* quand il se présente à telle ou telle « porte » de la société. Par exemple, après avoir passé le concours, Paul est autorisé à entrer à l'École Nationale de l'Aviation Civile pour devenir élève pilote de ligne. Après avoir passé les tests neuropsychologiques en vigueur, Gilberte n'est pas autorisée à entrer dans la salle d'IRM de tel service de santé auquel elle a adressé sa « plainte mnésique ».

L'évaluation psychologique met en scène un portier, un évaluateur, un testé, une clé et une porte à serrure sur fond d'opinion publique. Le portier est chargé du test de la serrure tandis que l'évaluateur est chargé du test de la clé. Le portier teste si la clé ouvre la serrure de la porte qu'il garde et si oui le testé est autorisé à entrer. L'évaluateur fabrique grâce à ses tests la clé qui *identifie* le testé face à la porte. Le testé se fait évaluer par le test de la clé et le résultat de l'évaluation est sa clé. Retenons quand même du test de Gédéon qu'il n'est pas toujours dit à l'avance quelles clés ouvrent la porte, la serrurerie imageant un montage institutionnel, dont les règles d'ouverture ou de fermeture sont contingentes⁵ – en principe, les procédés de fabrication des clés sont inconnus du portier comme du testé.

La contestation par les testés des résultats de leur test de la serrure est un risque de désordre social. Au pire, la contestation doit rester locale et embryonnaire. De plus le triage doit être suffisamment efficace à l'échelle de la porte sociale. Par exemple, il est indésirable qu'un élève pilote de ligne échoue sa formation (faux positif) ; il est indésirable qu'une personne non-dépistée par les tests de mémoire s'avère porteuse d'une tumeur cérébrale (faux négatif). Trop d'inefficacité institutionnelle menace l'institution et donc le portier qu'elle emploie. Dans le langage des « propriétés psychométriques » de la doctrine des tests psychologiques, la « validité prédictive » des clés psychotechniques relève d'une ingénierie sociale statistique et non pas d'une science psychologique (Lacot, Afzali, & Vautier, 2016;

⁵ Par exemple, Julie qui n'était pas autorisée à entrer à l'ENAC est maintenant autorisée suite au désistement d'un candidat autorisé.

Vautier, Veldhuis, Lacot, & Matton, 2012). La raison en est que le critère statistique d'une politique de décision décrit un rendement statistique... quand les données sont disponibles⁶. Si bien que, de manière « provisoirement cynique » pour prolonger l'attitude critique de Canguilhem, tout va bien pour le portier tant que son triage ne produit pas de conséquences socialement indésirables *visibles*.

D'où l'importance de la moralité du portier et du rôle d'expertise, sur fond de science psychologique, du psychologue : l'ouverture de la serrure dépend de la clé (et la question de la nécessité de l'existence de cette serrure est professionnellement forclosée) ; de manière analogue, l'autorisation du testé dépend du testé puisqu'elle dépend de *sa* clé, c'est-à-dire de ce qui le représente dans le langage des autorisations que parlent l'évaluateur et le portier.

L'autorité du couple que forment l'évaluateur et le portier vis-à-vis du testé repose sur l'opinion publique, état mental, conscient et inconscient, de quantités de personnes interconnectées, l'ordre de grandeur dépendant de la taille de la population démographique qui en est le substrat. Pour décrire ce que l'opinion sait des tests psychologiques, nous admettrons le « plus petit commun dénominateur » suivant. L'opinion concernée est telle que i à x sont vraies : (i) Elle a entendu que les tests « psy » sont des « instruments de mesure scientifiquement validés ». (ii) Elle ignore en quoi consiste exactement la « validation scientifique » de ces tests. Si elle s'informe, elle constate l'usage de statistiques, dont elle ne sait pas évaluer s'il triche avec la nécessité logique, et s'en tient aux signes : la modélisation statistique est signe de scientificité. (iii) Elle admet que les tests sont des instruments de mesure (sinon ils ne seraient pas « scientifiquement validés »). (iv) Elle admet aussi que les scientifiques, quand ils font leur métier de leur mieux, respectent les règles logiques, puisqu'ils sont « scientifiques ». (v) Elle ignore ce que sont ces règles dans le détail. (vi) Elle sait, si elle s'informe un peu, que des revues « scientifiques » filtrent les articles soumis à publication, même si elle ignore le fonctionnement du filtre⁷. (vii) Elle admet que la

⁶ Dans un concours par exemple, il n'existe pas de faux négatifs. Nécessité logique oblige, on ne peut pas ignorer que les statistiques dites « diagnostiques » d'une matrice de confusion (qui définit les faux positifs et les faux négatifs) sont des propriétés qui décrivent une matrice de confusion et non des propriétés qui décrivent les clés qui sont fabriquées ou les serrures que les clés sont supposées ouvrir.

⁷ « La communauté scientifique, à moins qu'elle ne soit déjà en coma dépassé, ne peut plus faire semblant de s'accommoder d'un fonctionnement interne dans lequel l'attribution des moyens repose sur la publication d'articles dans quelques journaux huppés dont le ticket d'entrée est – de l'aveu général – une loterie, où il est plus important pour faire carrière d'être un bon manager et un bon communicant que d'avoir de l'imagination et de la rigueur, et où l'évaluation anonyme par les pairs, clé de voûte de l'ensemble, s'apparente au lancer de dés » (Ségalat, 2009, p. 728).

publication atteste, par la renommée des « revues », la valeur scientifique des articles publiés. (viii) Elle raisonne que si les tests psychologiques sont des instruments de mesure scientifiquement validés ayant fait l'objet d'articles publiés, alors les mesures qu'ils permettent sont dignes de confiance (et d'ailleurs, l'opinion bien informée peut avoir entendu que les tests *possèdent* (faut-il souligner) des « propriétés psychométriques », et même, parfois et par erreur d'étourderie des communicants, des « propriétés métrologiques »). De plus (ix), l'opinion publique ignore comment sont prises les décisions concernant les gens qui sont testés en fonction des résultats (mesures, scores, clés, c'est tout un). (x) Elle ignore enfin comment les résultats eux-mêmes sont élaborés, ainsi que leurs interprétations quand elles sont explicitement différenciées des résultats – nous reviendrons sur ce point en concluant.

Le testé en tant que profane se trouve face à une puissance de véridiction, pour reprendre l'expression de Lordon (2019), qui institue l'évidence : la contester serait voué à l'échec. Ainsi l'emploi de l'évaluateur-fabriqueur de clés contribue-t-il à l'édification de l'autorité du portier : l'identification clé-testé est aussi inévitable que, par exemple, « je mesure un mètre soixante quinze » et il faut bien que le sort, le hasard ou la nécessité qu'on ignore décident. Un évaluateur-psychologue vaut mieux pour le public qu'un évaluateur-astrologue par exemple, l'astrologie n'étant pas un diplôme *académique*. L'évaluation psychologique est un métier, ce qui implique qu'elle soit effectuée par des professionnels diplômés (autorisés).

Mais il existe une contradiction, à mettre en perspective avec l'embarras du Code quant à la valeur des conclusions livrables (cf. les articles 19 et 25 cités ci-dessus) : les clés qui sont fabriquées au titre de l'autorité de la science psychométrique, laquelle n'est, en fait et comme on va le voir plus bas, qu'une docimologie probabiliste à la place d'une impossible métrologie des grandeurs psychologiques, les clés psychométriques disions-nous, sont des scores⁸. Or des scores ne sont pas des mesures mais des artefacts numérologiques indignes d'une science psychologique qui ne triche pas avec la nécessité logique (e. g., Johnson, 1943).

⁸ Nous simplifions pour clarifier. Il n'est pas nécessaire de supposer que les clés ne soient que des scores, ni même, puisque ce sont des messages, que ces messages contiennent au moins un score. Il suffit que les scores soient nécessaires à l'élaboration des clés.

D'où la question de la nécessité selon laquelle tel testé se trouve désigné par telle clé, si celle-ci ne lui était pas *naturelle*⁹.

D'où aussi le silence embarrassé, quand ce n'est pas la publicité mensongère à la « validation scientifique », c'est-à-dire la fraude scientifique, puisque l'autorité du psychologue évaluateur implique celle de sa science, et qu'une science qui publie à grand bruit et depuis ses débuts qu'elle sait mesurer ses grandeurs, alors que c'est faux (parce qu'elle a fondu les verbes « mesurer » et « évaluer » en un alliage sémantique qui, par l'interchangeabilité dans le discours qui crée la synonymie, escamote le problème du mesurage scientifique), se trouve en terrible posture vis-à-vis de l'opinion publique. Si l'opinion publique désavoue l'autorité scientifique de la psychologie, elle désavoue en cascade celle des professionnels de l'évaluation psychologique puis celle des portiers, c'est-à-dire des donneurs d'ordre. En effet, si la serrure s'ouvre en fonction de la clé, c'est par une nécessité naturelle, qui va de soi (même si clé et serrure sont des artefacts culturels) ; mais si l'autorisation d'entrer dépend de *scores* qui n'ont pas la nécessité empirique ou encore expérimentale d'un *mesurage*, quelle est la vérité psychologique des descriptions qui sont délivrées pour l'administration des testés ? D'où l'intérêt politique commun, au sens premier de lutte pour l'autorisation de circuler dans le flux social, pour le psychologue-évaluateur et son donneur d'ordre, de ne pas toucher à la respectabilité scientifique des méthodes de fabrication des scores qui, au demeurant, n'auraient qu'une portée technique, donc auxiliaire par rapport aux buts prioritaires. Mais ce serait oublier la priorité que l'on doit à la crédibilité scientifique si on veut en user : ne pas tricher avec la nécessité logique.

2. La patate chaude du scorage en psychologie

Il n'est pas nécessaire d'être grand métrologue pour savoir que mesurer nécessite que la variation du mesurant dépende de manière croissante (ou décroissance si nécessaire) de celle du mesuré (Vautier, 2016a). Ni d'être grand docimologue pour savoir que ce n'est pas le niveau scolaire qui fait la note, mais la note qui fait le niveau scolaire. Pourtant, il faut y

⁹ Cf. aussi Rosenvallon (2013) : « Le culte américain des tests, après celui de la phrénologie, a constitué un parfait révélateur de la grande contradiction qui a sous-tendu l'histoire des démocraties capitalistes libérales : la coexistence d'une philosophie égalitaire fondatrice et d'une réalité sociale marquée par de fortes inégalités. C'est en Amérique que cette tension a été la plus vive. Et c'est donc là que la tentation de naturaliser l'inégalité a été la plus manifeste. En étant de la sorte 'désocialisée', l'inégalité a pu prospérer dans un monde viscéralement attaché à l'idée que tous les hommes avaient été créés égaux. Elle n'a cessé d'être légitimée en étant massivement rapportée au seul mérite individuel et aux seuls talents hérités » (p. 145).

revenir, force est de constater la synonymie qu'ont acquis, en psychologie, les verbes « mesurer » et « évaluer », ainsi que la minuscule fréquence d'utilisation des verbes « noter » ou « scorer », dont l'usage attirerait l'attention sur le jeu avec les apparences : les mesures sont des nombres, les scores sont des nombres ; donc (illusion logique) les scores sont des mesures. Remarquons aussi que la modélisation statistique a besoin de nombres, tant mieux si ce sont des mesures, mais des mesures ne sont pas nécessaires : les statistiques en psychologie sont largement fondées sur des scores et leur utilité souvent obscure (Lakatos, 1978)¹⁰.

Comment sont fabriqués les scores ? Au moins par *addition de points*. Comment sont fabriqués les points ? Par décret, convention, règle d'usage, habitude, imitation, formation universitaire, selon les mêmes nécessités consensuelles qui permettent les règles des jeux de société. Le format de réponse aux « items » du test est à la fois un dispositif de communication avec le testé et une échelle de points. Chaque réponse possible est associée à un nombre de points donné – *codage* (numérique) des réponses. Or les réponses aux items ne sont pas additives donc une somme de points n'a pas de valeur descriptive des événements ou faits dont le psychologue est le témoin professionnellement qualifié. Par exemple, l'addition de « Je suis anxieux : plutôt non » et de « Je me fais du souci : non » n'est pas définie, même si ces réponses s'écrivent (2, 1) dans le langage numérique des observations. Autre approche par le biais de la notation, il faudrait écrire plus précisément (2₁, 1₂) où les indices désignent les items du test. C'est parce qu'on efface les indices qu'il est possible de calculer 2 + 1 = 3. En effaçant les indices, on oublie facilement que les points codent des réponses structurées en n-uplets, avec n le nombre d'items du test, et on complète automatiquement le vide sémantique du résultat arithmétique par l'intuition d'une quantité de substance (qu'on peut nommer « niveau d'anxiété » par exemple). Si on calcule un score pour s'en servir d'observation, on triche avec la nécessité logique selon laquelle une observation est un événement possible. Or la somme des points codant les réponses à un test est une réponse impossible parce qu'elle n'existe pas dans le référentiel d'observation du test et donc elle n'est pas un événement possible (voir aussi Vautier, 2011).

¹⁰ "... one wonders whether the function of statistical techniques in the social sciences is not primarily to provide a machinery for producing phoney corroborations and thereby a semblance of 'scientific progress' where, in fact, there is nothing but an increase in pseudo-intellectual garbage. ... Thus the methodology of research programmes might help us in devising laws for stemming this intellectual pollution which may destroy our cultural environment even earlier than industrial and traffic pollution destroys our physical environment." (p. 88, footnote 4)

Pourtant, une telle erreur de manipulation symbolique des observations est détectable depuis longtemps en psychologie puisque Stevens (1946), pour défendre la science psychologique quantitative contre le scepticisme des physiciens, a fait publier qu'en psychologie, mesurer consiste à attribuer des nombres à des objets ou aspects selon des règles, et que la règle du codage nominal permettant de traiter les échelles de points comme des « échelles de mesure nominales » doit reposer sur un « isomorphisme » (« représentation » dans le langage de la théorie représentationnelle de la mesure, qui conserve les propriétés relationnelles de l'ensemble représenté). Stevens signale publiquement qu'on peut tricher avec cette exigence fondamentale lorsqu'il écrit que « L'intelligence, par exemple, est utilement évaluée [“assessed” et non pas “measured”] sur des échelles ordinales » (p. 679). S'il n'avait pas triché avec la nécessité logique, il aurait admis que les scores « d'intelligence », dont il parle évidemment, ne sont pas des mesures ordinales des réponses aux tests, parce qu'il n'existe pas d'isomorphisme entre l'ensemble des scores et celui des réponses. Il aurait aussi reconnu que sa typologie contourne le problème de savoir comment une réponse à un test de QI peut mesurer une quantité d'intelligence (Vautier, 2016b).

Encore fallait-il savoir, principe de non-contradiction, qu'un non-ordre ne peut pas équivaloir¹¹, être isomorphe, à un ordre ; que l'ensemble des n-uplets généré par le test, c'est-à-dire des réponses possibles au test, n'est pas un ordre au sens de l'usage courant du mot, qui entend qu'un ordre soit un ordre *complet*. Si on remplace non-ordre par ordre incomplet pour reconnaître les possibilités existantes, un ordre incomplet n'est pas équivalent à un ordre complet, donc une échelle de scores ne peut pas « mesurer ordinalement », c'est-à-dire de manière complètement ordonnée par une relation réflexive, des réponses au test, parce que les réponses au test sont incomplètement ordonnées (sauf exception empirique, point méthodologiquement crucial sur lequel nous reviendrons). Au lieu d'une avancée descriptive, qui aurait pu élever au niveau théorique (ou métathéorique) les propriétés mathématiques qui permettent de construire des ordres de différents types (e. g., Roberts, 1985) structurant les langages d'observation du psychologue, le plaidoyer de Stevens sert, intentionnellement ou pas, à détourner l'attention du public de la question scientifique de la mesurabilité de

¹¹ Ici est épinglé le problème, théorique, de la définition de départ, c'est-à-dire du concept, qu'on se donne de la valeur : par défaut c'est un segment de droite dont l'origine est zéro et la borne supérieure ce qu'on veut.

l'intelligence et autres « sensations », sachant que la question de la mesurabilité avait déjà été escamotée par les promoteurs de l'analyse dite factorielle des *scores* (Martin, 1997).

Du côté de la théorie classique des tests (Lord & Novick, 1968), l'opérationnalisme probabiliste a forgé le titre somptueux de « score vrai » (*true score*) pour nommer une chimère. Un score mesure un « score vrai », avec faut-il immédiatement concéder, une « erreur de mesure », ce qui frappe la théorie d'inutilité descriptive, pour deux raisons suffisantes, l'une empirique, l'autre théorique. Tout d'abord, le score vrai est défini comme l'espérance mathématique d'une variable aléatoire, dont l'expérience aléatoire réalise un événement empiriquement ou expérimentalement *unique*, tant qu'on n'a pas isolé, au sens du *contrôle expérimental* qu'implique le verbe « isoler », le système observé des « facteurs » qui « parasitent » le signal. Or, comme le remarquait Canguilhem (1958), « Nous ne pouvons ni sur nous-mêmes, ni sur autrui, nous livrer à des expériences » (p. 19, voir aussi Trendler, 2009), fussent-elles aléatoires, parce que nous butons sur la réplication à *l'identique* : identique à quoi ? Comment pouvons-nous contrôler une « variable dépendante » en jargon orthodoxe, c'est-à-dire obtenir ne serait-ce qu'une relative immobilité des états psychologiques d'une personne ? Tout testé qui recommence à répondre à un item de test n'est pas identique à lui-même d'un moment expérimental à l'autre, contrairement à un dé ou une pièce de monnaie. Ainsi, le score vrai s'avère une chimère statistique puisque son estimation, calcul d'une moyenne, nécessite un grand nombre de répétitions à l'identique d'une expérience aléatoire unique.

La seconde raison est que, comme les probabilités qui déterminent le score vrai en fonction des scores possibles sont théoriquement inconnues, l'erreur de mesure, définie comme différence entre le score réalisé (observé) et le score vrai, n'est pas utilement restreinte : quels que soient deux scores, on ignore complètement l'ordre des scores vrais correspondants (Vautier, 2019). L'opérationnalisme épistémologique de la psychométrie accouche d'un concept empiriquement inopérant, fondé, au demeurant, sur un angélisme probabiliste¹² contradictoire avec l'intention programmatique anti-réaliste ou anti-métaphysique affichée vis-à-vis des grandeurs psychologiques (voir aussi Vautier, 2014).

¹² La valeur numérique de la *probabilité* d'un événement, comme la valeur « féminin » ou « masculin » du *sexe* d'un ange, sont des êtres imaginaires. La probabilité mathématique ne pose pas de problème scientifique tant qu'elle n'est pas utilisée comme propriété d'un événement empirique unique. Si c'est le cas, il faut reconnaître qu'il existe des possibilités quantitatives dans l'ontologie, l'objet en construction ou encore le projet, dans le

Mais la stérilité opératoire et descriptive des « variables latentes » dont se nourrit la psychométrie classique (et contemporaine) n'empêche pas que ces « variables » constituent une mine de « labels qualité » pour la propagande de la validité scientifique des tests : la propriété psychométrique nommée « fidélité » (reliability), qui présente à la limite un intérêt esthétique (e. g., Vautier & Jmel, 2003; Vautier & Pohl, 2009; Vautier, Steyer, & Boomsma, 2008) joue le rôle, dans le discours promotionnel des tests, d'une mesure de la « précision de la mesure ». Ce qui revient à confondre d'une part « approximation ou imprécision de lecture d'un point sur un cadran », le point étant un concept non-empirique s'il en est, d'autre part « propriétés statistiques d'un échantillon de *points* », non pas cercles de diamètre nul mais *nombres ordinaux* (e. g., Essex & Smythe, 1999) codant les réponses aux items de test. Si, en psychologie comme ailleurs, tel nombre de points était une mesure au sens scientifique et pragmatique du marchand de légumes, ce serait le nombre qui code dans une échelle ordinale un intervalle d'étendue inconnue de la grandeur mesurée, ce qui constitue une hypothèse testable dès qu'on reconnaît que la description est multivariée (Vautier, 2016a).

Malgré la connaissance pratique des ordres partiels en psychométrie (Coombs, 1964; Guttman, 1944), on n'a pas voulu considérer l'immensité de l'évidence falsifiant qu'un n-uplet de réponses à un test soit une mesure. Personne ne nie que la réponse à un item de test résulte d'un processus neuro-cognitif très largement inconnu, dont l'approche par les processus de « traitement de l'information » interdit qu'il soit réductible à l'image d'une grandeur par une fonction croissante, donc qu'il soit mesurable. Du point de vue théorique, il n'existe pas de grandeur psychologique mais des « processus psychologiques » qu'on essaie de rapprocher des processus cérébraux – ce qui n'empêche pas que ces processus, c'est-à-dire ces transitions d'états, puissent avoir des propriétés quantitatives mesurables au même titre qu'un métabolisme. Du point de vue empirique, lorsqu'on répète l'expérience et même en ignorant nécessairement la variation théorique qu'on veut mesurer, les n-uplets qu'on peut observer peuvent être incomparables, c'est-à-dire non-ordonnables. Si c'est effectivement le cas, l'hypothèse de dépendance croissante du mesurant est falsifiée puisqu'elle implique que

langage de Canguilhem, de cette science, et poser le problème préalable de leur mesurage si l'on ne veut pas d'une science fondée sur une hypothèse auxiliaire dogmatique, c'est-à-dire taboue – l'interdit consistant à admettre la possibilité qu'elle soit fautive, ce qui entraînerait que les conséquences empiriques de cette théorie ne soient pas nécessairement vraies, autrement dit que la théorie pourrait entrer en contradiction avec d'autres connaissances sans que cela affecte sa *valeur* ni la valeur des autres connaissances (ne pas tricher avec la nécessité logique).

le mesurant soit complètement ordonné. D'où le « point crucial » signalé plus haut : si les n-uplets observés étaient complètement ordonnés, alors on aurait fait une découverte empirique étonnante, autorisant peut-être l'espoir qu'on ait saisi quelque chose de mesurable au sens d'abord ordinal du terme, métrique plus tard peut-être, et ce serait une révolution métrologique. Pour autant qu'on sache, il n'existe pas de fonction croissante dont le domaine de définition serait la grandeur à mesurer et le domaine de valeur l'ensemble complètement ordonné des n-uplets observables avec quelque test psychologique que ce soit. Donc les réponses aux tests ne sont pas des mesures, semble-t-il et si on ne triche avec la nécessité logique. La signification des scores que fabrique le psychologue évaluateur pour décrire le testé n'est pas logiquement déductible d'un mesurage miraculeux. Par conséquent, si elle existe, la nécessité des clés qui identifient le testé vis-à-vis des portiers auxquels il est présenté ou se présente, ne relève pas de la science psychologique, qui ne sait pas expliquer ce qu'elle observe et qui ne sait pas bien ce qu'elle observe, mais d'une administration des testés dans la société, aidée d'une docimologie, c'est-à-dire d'un art du scorage.

3. Conclusion

À Canguilhem (1958) s'arrêtant sur l'ambiguïté de la psychologie institutionnelle vis-à-vis de son objet, nous avons répondu que « lorsqu'il se prévaut de son diplôme universitaire, le psychologue peut aussi se tourner vers la cité tout entière pour dénoncer une erreur scientifique et, croit-il, morale de sa science puis, en gage de bonne foi, s'engager à la corriger ». Non pas qu'il existe une troisième voie. Le psychologue monte en défendant, pour l'honneur de sa science, qu'un score ne soit pas une observation comportementale et que, par conséquent, il n'y a pas à ignorer les conséquences du fait que ses grilles d'observation soient multivariées, même s'il ou on veut les utiliser pour parler de choses qu'il ou on voudrait qu'elles mesurent au sens que la science fait du mesurage une question ouverte à l'expérience et, donc, il n'y a pas à ignorer que ses techniques d'observation soient *par défaut* irréductibles à une quelconque « échelle de valeur » – l'évaluation (univariée) n'est pas de mise, mais la description logiquement cohérente et la moins réductrice possible. Nous tombons d'accord avec le diagnostic épistémologique de Canguilhem : « La psychologie ne peut donc être que descriptive » (p. 19) et il s'agit de prendre au sérieux ce qu'est la description scientifique : l'invention d'algèbres descriptives ou encore de langages d'observation permettant des

énoncés descriptifs valables, c'est-à-dire logiquement vrais, permettant la recherche de lois expérimentales soigneusement formulées (Vautier, 2011).

Si le psychologue descend, c'est peut-être mais pas nécessairement vers la « Préfecture de Police ». En fait il se trouve déjà « en bas », dans l'univers des portes qui s'ouvrent ou se ferment et des personnes qui s'y meuvent et s'y trient, c'est-à-dire dans la société. À une époque qui a viré, croyons-nous, à l'eschatologie tant que le mythe de la nécessaire croissance économique a été insuffisamment heurté par les effets des limites écologiques de l'organisme/corps physique dont dépend son existence (e. g., Bouleau, 2018; Giraud & Renouard, 2012; Jacquard, 1991; Stengers, 2013b), il est temps de croître en maturité, c'est-à-dire en sagesse comme *réunion* de vérités logiques ou morales : l'anti-surveillance libertaire adolescente doit être abandonnée pour une responsabilité de surveillance, tant il est vrai que, selon la formule que nous croyons devoir à Lévinas (1971), *l'autre me regarde*, c'est-à-dire, en ce début de siècle, une responsabilité civique ou, pour profiter du sens historique de Parot (2017), républicaine : pas de société sans portes, donc la surveillance est nécessaire ou encore inévitable. La surveillance est impliquée par l'information qui circule concernant les autorisations de quiconque, qui regardent quiconque parce qu'elles sont, peu ou prou, constitutives du métabolisme de son milieu social ou naturel. Si le psychologue contribue à tel ou tel triage, il fait son métier de psychologue-évaluateur, technicien ou ingénieur selon la nature des problèmes dont il s'occupe, à condition de ne pas ignorer la nécessité de l'évaluation de l'utilité sociale de ce tri (sinon c'est un agent à la Milgram, 1984) – ou bien, si l'on rejette tout triage social, il faut aussi rejeter les scorages qui ont autorisé Canguilhem par exemple à franchir la porte de la Sorbonne entre autres portes prestigieuses. Si le psychologue-évaluateur contribue à l'architecture (en train de se faire) du système de portes, il monte vers le Panthéon ou disparaît en Enfer, selon la manière dont on juge le rouage institutionnel ainsi fabriqué.

L'ambiguïté fondamentale de la psychologie qui gêne Canguilhem est multiforme, bien que son étiologie remonte à l'inoculation d'un seul germe : on peut tricher avec la nécessité logique. L'évaluation psychologique en est une forme particulièrement saillante. De deux choses l'une : soit le psychologue sait au moins quelque chose sur le testé, soit il ne sait rien sur le testé, à l'exception nécessaire qu'il sait qu'il ne sait rien sur le testé (récursivité du langage). Or pas d'évaluation psychologique sans savoir sur le testé. Faut-il alors contrôler,

surveillance sociale oblige ou, si on préfère l'euphémisme, interdépendance systémique oblige, la puissance de ce savoir. Qui dit que le sort qu'on jette au testé, concernant son autorisation de franchir la porte à laquelle il s'est présenté dans la société, est moins puissant lorsqu'il est chuchoté ? L'histoire montre que des sorts d'immense puissance ont été jetés à voix basse ; auraient-ils été encore plus puissants s'ils avaient été hurlés ? Le Code invite au chuchotement en déclarant que les conclusions (clés) livrées doivent être non-réductrices et provisoires. En effet seule l'intention d'atténuation du volume du son du message vaut, pour deux raisons dont chacune est suffisante.

D'abord, une conclusion *non-réductrice* est une impossibilité logique : l'information suppose une réduction, c'est-à-dire la suppression d'au moins une possibilité. Donc une conclusion est nécessairement réductrice sinon elle ne vaut que comme paquet vide d'information, ce qui supposerait que le psychologue soit employé par le portier pour faire semblant de donner de l'information psychologique sur le testé et il est bien évident qu'il n'est pas question d'envisager de telles extrémités¹³. Parce qu'ils l'apprennent à l'université, les psychologues sont supposés savoir que décrire scientifiquement implique réduire (nécessité logique).

Ensuite, il est évident que la vérité qui est accordée aux phrases ou aux discours *change* au gré des époques, le point est *hors sujet*. Il est aussi évident que la vérité de la description que le psychologue effectue pour fabriquer la clé qu'il va attribuer au testé est supposée connue, c'est ici qu'intervient la crédibilité de sa science comme science d'observation objective, c'est-à-dire capable de fournir des observations *vérifiables*. Le caractère *provisoire* de la conclusion n'est pas à gagner parce que si la clé était provisoire, alors il faudrait y adjoindre une date de péremption. Or, à notre connaissance, aucun psychologue-évaluateur ne conclut en précisant la date de validité/péremption de sa conclusion. Il suffit de laisser entendre, par exemple, que l'enfant Paul rencontrerait des difficultés s'il était autorisé à entrer, pour que la possibilité *agisse* dans les esprits, formant la connaissance que « tout le monde » sait qu'elle existe et que « tout le monde » le sait et ainsi de suite à l'infini, pour que la possibilité devienne une nécessité collectivement *rêvée*, pour utiliser le pouvoir littéraire du langage chamanique (pouvoir magique de la parole, dont la

¹³ Le score est un paquet vide d'information si on veut le score vrai de la théorie classique des tests ou la valeur numérique prise par une quelconque « variable latente ».

pratique, par conséquent, doit être impeccable, voir Ruiz, 2016) et, finalement et définitivement, qu'une décision s'exécute.

Réglons rapidement la question de savoir comment différencier les observations des interprétations. Il y a deux types d'interprétation : celles qui sont nécessaires compte tenu (i) des observations et (ii) de la théorie qu'on admet, celles qui ne le sont pas. Dans le premier cas, les interprétations sont des déductions, vraies pour autant que leurs prémisses le soient aussi. Dans le second cas, la vérité des interprétations est contingente. Par exemple, « Vous êtes 20,33 % plus stressé qu'une personne normale »¹⁴ est une description dont la vérité est contingente. Lorsqu'on calcule la différence entre un score et un score moyen puis, par une règle de trois, une proportion, que l'on exprime ensuite en pourcentage, on peut écrire « si votre score vaut 100 %, il se trouve 22 % au-dessus du score moyen de tel échantillon de référence », mais cela n'implique pas qu'on sache mesurer l'anxiété d'une « personne normale » ni votre anxiété avec ces scores. On sait même de source scientifique qu'on ne sait pas le faire et que, de surcroît, l'anxiété n'est pas sérieusement définie comme une grandeur. Il faut reconnaître froidement que la vérité de ce type d'interprétation est irrationnelle. Faut-il s'apprêter à défendre le droit à l'irrationnel ? Chacun fait ce qu'il doit faire mais alors, pas en combattant un soi-disant totalitarisme rationaliste ou logique en s'animant de l'énergie affective d'un libertarisme adolescent qui refuse la nécessité, y compris la nécessité logique ou rationnelle. D'abord, en notant A les observations et la théorie, B les interprétations, $A \Rightarrow B$ ne transmet pas nécessairement la valeur « vrai » à B lorsque A est fausse. Ensuite rien ne dit qu'il soit nécessaire que $A \Rightarrow B$ soit vraie. Enfin s'il est vrai que $A \Rightarrow B$ soit fausse, alors il est vrai que la valeur « vrai » de A n'implique pas la valeur « vrai » de B. Donc le rationalisme scientifique ne contrôle pas toute la vérité et il sait même que c'est impossible : il admet que la vérité est nécessairement contingente dans des conditions réelles. Il faut ici rendre hommage à Popper (2013), dont l'œuvre devrait être citée en extension mais, dans le contexte de cette réponse à la difficile question que Canguilhem adressait à la psychologie, que le lecteur pardonne la sélection d'un plaidoyer pour une « société ouverte », la liberté y étant équivalente à la responsabilité parce qu'elles s'impliquent réciproquement (voir aussi Artigas, 1999).

¹⁴ Voir par exemple le site <https://www.idrlabs.com/fr/anxiete-stress-depression/test.php>

Une autre approche déontologique que le chuchotement de l'information existe. Défendre la valeur de ce que le psychologue sait sur le testé, peu importe que ce savoir soit positif ou négatif (Vautier, Lacot, & Veldhuis, 2014), lucratif ou non-lucratif. S'il utilise un savoir, nécessairement positif, pour fabriquer des clés de tri social, défendre la valeur de la clé, ainsi que la valeur de la serrure qu'elle est supposée ouvrir à la faveur des dieux. Les dieux existent, il suffit de localiser puis d'illuminer les postes de triage dans la société pour trouver ceux, celles, ou les automates qui s'en occupent et savoir comment ils font (e. g., de Terssac, 1992; Dejours, 2003).

Une fois qu'on connaît les règles du jeu d'autorisation, on peut viser l'exigence d'un triage juste, ce qui suppose une théorie de la justice. Mais comment évaluer la justice de modes de triage furtifs, dont la surveillance est *réservée*, c'est-à-dire conditionnée à des autorisations de savoir ? S'il prétend être reconnu comme spécialiste d'une certaine information, s'il se résout alors à assumer son rôle d'informateur, le psychologue-évaluateur peut apporter beaucoup à l'édification d'une science sociale à visée post-eschatologique, au titre de sa pratique des jeux de clé et de serrures sociales. Si bien que, tout conte devant terminer par une morale, il peut se poser une question morale et, puisque la psychologie implique la possibilité d'avoir peur en pensant, essayer de contrôler sa peur si elle survient tandis que la réponse ne viendrait pas : ici et maintenant, dans la solitude qui est la mienne, si j'admets que je dois quelque chose à ma science, comment puis-je y contribuer par les effets de mon pouvoir dans la société-milieu qui l'abrite ?

Références

Artigas, M. (1999). *The ethical nature of Karl Popper's theory of knowledge*. Bern: Peter Lang.

Bouleau, N. (2018). *Le mensonge de la finance : Les mathématiques, le signal-prix et la planète*. Paris: Les Editions de l'Atelier.

Canguilhem, G. (1958). Qu'est-ce que la psychologie ? *Revue de Métaphysique et de Morale*, 1, 12-25.

Coombs, C. H. (1964). *A theory of data*. New York: Wiley.

- de Terssac, G. (1992). *Autonomie dans le travail*. Paris: Presses Universitaires de France.
- Dejours, C. (2003). *L'évaluation du travail à l'épreuve du réel : Critique des fondements de l'évaluation*. Paris: INRA Editions.
- Essex, C., & Smythe, W. E. (1999). Between numbers and notions : A critique of psychological measurement. *Theory & Psychology, 9*, 739-767.
- Giraud, G., & Renouard, C. (2012). *Le facteur 12 : Pourquoi il faut plafonner les revenus*. Paris: Carnets Nord.
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. *Monatshefte für Mathematik und Physik, 8*, 173-198.
- Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*, 139-150.
- Jacquard, A. (1991). *Voici venu le temps du monde fini*. Paris: Seuil.
- Johnson, H. M. (1943). Index-numerology and measures of impairment. *American Journal of Psychology, 56*, 551-558.
- Lacot, E., Afzali, M. H., & Vautier, S. (2016). Test validation without measurement : Disentangling scientific explanation of item responses and justification of focused assessment policies based on test data. *European Journal of Psychological Assessment, 32*, 204-214.
- Lakatos, I. (1978). Falsification and the methodology of scientific research programmes. In J. Worrall & G. Currie (Éd.), *Imre Lakatos : Philosophical papers. Vol. 1 : The methodology of scientific research programmes* (p. 8-101). Cambridge: Cambridge University Press.

- Lévinas, E. (1971). *Totalité et infini*. Paris: Le Livre de Poche.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA.: Addison-Wesley.
- Lordon, F. (2019). *La condition anarchique : Affects et institution de la valeur*. Paris: Seuil.
- Martin, O. (1997). *La mesure de l'esprit : Origines et développements de la psychométrie, 1900-1950*. Paris: L'Harmattan.
- Michell, J. (2009). Invalidity in validity. In R. W. Lissitz (Éd.), *The concept of validity : Revisions, new directions, and applications* (p. 111-133). Charlotte, NC: Information Age Publishing.
- Milgram, S. (1984). *Soumission à l'autorité : Un point de vue expérimental* (2e éd.). Paris: Calmann-Lévy.
- Parot, F. (2017). *La psychologie française dans l'impasse : Du positivisme de Piéron au personnalisme de Fraisse*. Paris: Editions Matériologiques.
- Popper, K. R. (2013). *The open society and its enemies*. Princeton, NJ: Princeton University Press (Original work published 1945).
- Roberts, F. S. (1985). *Measurement theory, with applications to decision-making, utility, and the social sciences*. Reading, MA: Addison-Wesley.
- Rosanvallon, P. (2013). *La société des égaux*. Paris: Seuil.
- Ruiz, D. M. (2016). *Les quatre accords toltèques : La voie de la liberté personnelle*. Genève: Editions Jouvence.
- Searle, J. R. (1995). *The construction of social reality*. New York: The Free Press.
- Ségalat, L. (2009). *La science à bout de souffle ?* Paris: Seuil.

- Stengers, I. (2013a). Avoir l'étoffe du chercheur. In *Une autre science est possible ! Manifeste pour un ralentissement des sciences* (p. 27-50). Paris: La Découverte.
- Stengers, I. (2013b). Cosmopolitiques : Civiliser les pratiques modernes. In *Une autre science est possible ! Manifeste pour un ralentissement des sciences* (p. 113-141). Paris: La Découverte.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 667-680.
- Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory & Psychology*, 19, 579-599.
- Vautier, S. (2011). The operationalization of general hypotheses versus the discovery of empirical laws in Psychology. *Philosophia Scientiae*, 15, 105-122.
- Vautier, S. (2014, mai 28). La théorie classique des tests : Un triple désastre scientifique. Consulté à l'adresse Epistémologie de la psychologie website:
<http://epistemo.hypotheses.org/1054>
- Vautier, S. (2015). La psychotechnique des aptitudes : Pour différencier une sociotechnique de l'évaluation sans mesurage et une psychologie balbutiante de la compréhension de la performance. *Pratiques Psychologiques*, 21, 1-18.
- Vautier, S. (2016a). *Mesurer en psychologie : Un état des connaissances [cours vidéo]*. Consulté à l'adresse <http://epistemo.hypotheses.org/cours-video>
- Vautier, S. (2016b, juillet 12). 40a. Les échelles de mesure : Lire Stevens (1946). Consulté à l'adresse Epistémologie de la psychologie website:
<http://epistemo.hypotheses.org/3696>

Vautier, S. (2019, octobre 7). 54. Zéro information des scores en théorie classique des tests.

Consulté à l'adresse Epistémologie de la psychologie website:

<https://epistemo.hypotheses.org/6580>

Vautier, S., & Jmel, S. (2003). Transient error or specificity ? An alternative to the staggered equivalent split-half procedure. *Psychological Methods, 8*, 225-238.

Vautier, S., Lacot, E., & Veldhuis, M. (2014). Puzzle-solving in psychology : The neo-Galtonian vs. Nomothetic research focuses. *New Ideas in Psychology, 33*, 46-53.

Vautier, S., & Pohl, S. (2009). Do balanced scales assess bipolar constructs ? The case of the STAI scales. *Psychological Assessment, 21*, 187-193.

Vautier, S., Steyer, R., & Boomsma, A. (2008). The true-change model with individual method effects : Reliability issues. *British Journal of Mathematical and Statistical Psychology, 61*, 369-399.

Vautier, S., Veldhuis, M., Lacot, E., & Matton, N. (2012). The ambiguous utility of psychometrics for the interpretative founding of socially relevant avatars. *Theory & Psychology, 22*, 810-822.

Weil, S. (1957). *Ecrits de Londres et dernières lettres*. Paris: Gallimard.