

FORMULATION D'HYPOTHÈSES, EXPÉRIMENTATION, ÉVALUATION

Jan Carel ZADOKS

Section de phytopathologie, université agronomique de Wageningen,
B.P. 5025,
6700EE Wageningen, Pays-Bas

Résumé

Les rites de la communication scientifique, les conventions de la science, exigent une méthodologie stricte, afin que nos opinions personnelles puissent être intégrées dans le savoir commun qui est partagé entre scientifiques - dans notre cas, entre épidémiologistes. Le préalable à cette intégration au savoir commun est l'élimination d'idées subjectives, dont la valeur n'est pas fondée. Il existe, à cette fin, une trajectoire qui mène du préconçu à l'hypothèse, trajectoire qui passe par les phases de l'idée et de la prédiction. La trajectoire s'appelle « cycle empirique ». Ce cycle comporte une série d'étapes. La première est celle de l'observation. Le processus mental d'induction permet de parvenir à la seconde étape, celle de l'idée ou de la généralisation. Un autre processus mental, la déduction, mène à la troisième étape, la prédiction. La prédiction est mise à l'épreuve - en français moderne, elle est testée. L'épreuve, le test, confirme ou réfute la prédiction. En cas de confirmation, on déclare l'hypothèse « vraie ». En cas de réfutation, on la déclare « fausse ». Dans ce cas, le cycle empirique recommence à nouveau, avec les nouvelles informations acquises. La « vérité » obtenue n'est que temporaire : l'hypothèse est vraie aussi longtemps qu'il n'y a pas d'observation venant la réfuter. Le progrès de la science est comparable à celui d'un aveugle qui cherche son chemin en tâtonnant : ici on trébuche, là on se heurte, mais on progresse quand même. Nous ne connaissons jamais la vérité définitive, mais nous pouvons démontrer la fausseté d'une hypothèse. Heureusement, une hypothèse fautive nous permet de reprendre le cycle empirique à son départ. Cette description du progrès scientifique est un peu injuste. On acquiert, d'expérience en expérience, une certaine compréhension des phénomènes. L'image du cycle poursuivi et repris encore

à l'infini ne rend pas justice au travail scientifique. Une image plus juste est celle d'une spirale, et la trajectoire d'un cycle empirique constitue un tour de cette spirale. Une théorie n'est qu'une hypothèse, générale, de grande allure, qui est présentée comme « la vérité » pour des raisons didactiques ou politiques. Néanmoins, une théorie n'est qu'une hypothèse de durée de vie limitée, une hypothèse que l'on veut bien croire pour un temps. Dès qu'une nouvelle théorie apparaît, dont le domaine de validité est plus vaste que l'ancienne, celle-ci est mise à la poubelle. Le domaine de validité d'une théorie est une notion difficile à maîtriser. Une théorie peut être renforcée ici, alors qu'elle est réfutée ailleurs. Une théorie est inséparable de son domaine de validité. Ce constat trivial est important en biologie, parce que la notion juridique classique de tertium non datur (il n'y a pas de troisième solution) ne s'y applique pas. Très souvent, on trouve une nouvelle, une « troisième » possibilité.

Introduction

Les rites de la communication scientifique, les conventions de la science, exigent qu'une méthodologie stricte soit mise en œuvre afin que nos opinions personnelles puissent être intégrées dans le savoir que partagent tous les scientifiques, dans notre cas, tous les épidémiologistes (ZADOKS, 1972; 1978). Avant cette intégration, toute opinion subjective et sans valeur doit être éliminée. A cette fin, il existe une trajectoire menant du préconçu à la théorie, trajectoire qui traverse les étapes de l'hypothèse et de la prédiction. Cette trajectoire s'appelle « cycle empirique » (de GROOT, 1969).

Le cycle empirique

Le cycle empirique consiste en une série d'étapes (figure 1). Ce cycle commence avec l'observation. Le processus mental d'induction permet de passer à la seconde étape, c'est-à-dire l'idée, la généralisation, ou techniquement, l'hypothèse. Un autre processus mental; la déduction, permet de parvenir à la troisième étape, la prédiction. La prédiction est mise à l'épreuve, en français moderne, elle est testée. L'épreuve, le test, confirme ou réfute la prédiction. En cas de confirmation, on déclare que l'hypothèse est vraie. En cas de réfutation, elle est déclarée fausse. Alors le cycle empirique recommence, en utilisant de nouvelles informations.

Phase 1. **L'observation.** L'observation, qui inclue la collecte et la classification des données, est le premier acte intentionnel du chercheur. Un génie est un observateur plus précis et plus original qu'un chercheur moyen. Il « voit » plus dans et derrière les phénomènes.

ANTICIPATION

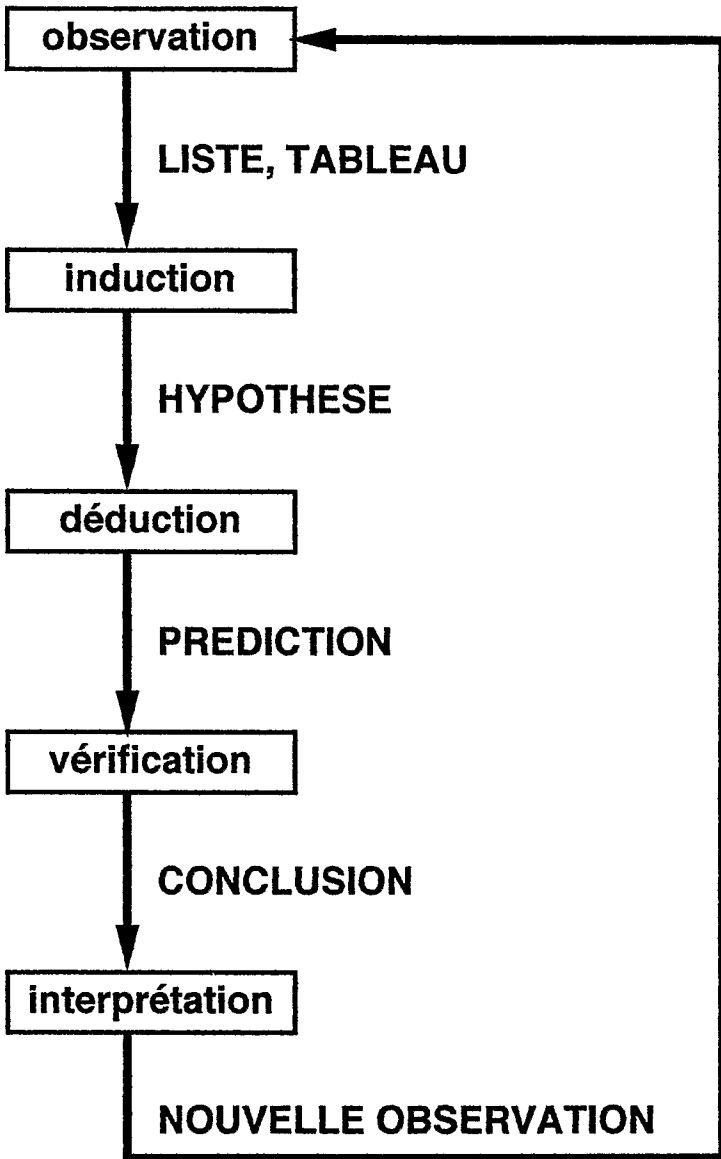


Figure 1
Le cycle empirique (d'après de GROOT, 1969).

Généralement, l'observation s'effectue avec une sorte d'**anticipation**. On observe rarement l'inattendu. Si l'anticipation est trop aiguë, on « voit » des choses qui n'existent pas. Si cette anticipation devient préoccupation, l'on ne voit que ce que l'on veut voir.

Phase 2. **L'induction**. L'induction est la généralisation d'observations déjà à « demi-digérées ». L'induction mène à l'hypothèse. Cette hypothèse n'est acceptable que si elle mène à une prédiction susceptible d'être évaluée dans une expérimentation.

Phase 3. **La déduction**. Si l'hypothèse a une valeur générale, alors elle doit aussi être valide dans un cas particulier, celui qui est choisi pour une prédiction. La déduction concerne des phénomènes nouveaux, qui n'ont pas été pris en considération dans le cadre de l'induction.

Voilà le raisonnement hypothético-déductif utilisé souvent en sciences naturelles. Il est indispensable de pouvoir vérifier la déduction par l'expérimentation.

Phase 4. **La vérification**. Une hypothèse est juste ou « vraie » si elle permet de prévoir précisément une situation nouvelle, fondée sur des données qui n'ont pas été utilisées lors de la construction de l'hypothèse. On dit que l'hypothèse est testée quant à son niveau de vérité. Il est typique des sciences naturelles que la vérification s'effectue grâce à une expérimentation bien définie et préparée à cette fin. Souvent, l'instrument utilisé est la statistique, avec ses deux catégories d'hypothèses : nulle ou alternative.

Selon les pensées de POPPER (1963) on ne connaîtra jamais la vérité définitive. Nous ne pouvons démontrer que la fausseté d'une hypothèse, pas sa justesse. Tant qu'une hypothèse n'a pas été prouvée fausse, on l'accepte comme vraie. Il existe, cependant, des vérifications incorrectes.

Phase 5. **L'évaluation**. La phase de vérification franchie, l'hypothèse est acceptée ou rejetée. Le résultat de la vérification constitue, en soi, une nouvelle observation soumise à l'interprétation scientifique. Cette interprétation pourra mener à un cycle nouveau, complémentaire ou modifié. C'est ainsi que le cycle empirique se clôt.

Processus et produit

Les cinq phases décrivent des processus, qui mènent à des produits. Le produit de l'observation peut être une liste ou un tableau, sur le papier, ou simplement présent à l'esprit. Le résultat de l'induction est une hypothèse; celui de la déduction, une prédiction. Le produit de la vérification est une conclusion; celui de l'interprétation, peut-être une nouvelle observation appartenant à un niveau hiérarchique supérieur. Dans la Figure 1, les processus sont indiqués en lettres minuscules, les résultats, en capitales.

Exemple trivial. Dans l'étang, on voit cinq cygnes blancs (observation). On admet que les cinq oiseaux sont probablement un échantillon aléatoire de la population de cygnes, et l'on tire la conclusion que, apparemment, tous les cygnes sont blancs

(induction). Puis on raisonne que, dans ce cas, les cygnes en Australie doivent aussi être blancs (déduction menant à une prédiction). On prend donc l'avion pour l'Australie. Arrivé on constate que, là, les cygnes sont noirs (vérification). La prédiction doit donc être réfutée, et, par conséquent, l'hypothèse est fausse (interprétation). L'interprétation est elle-même une nouvelle observation qui mène à une nouvelle hypothèse : les cygnes ne sont pas colorés, et ne peuvent être que blancs ou noirs. Et cetera.

Exemple épidémiologique. Comme VANDERPLANK (1963) lui-même l'a démontré, l'hypothèse de la croissance logistique (formule 1) décrivant la progression d'une maladie de plantes est fausse. Néanmoins, la formule logistique est très utile. Cela signifie qu'il existe un domaine de validité pour cette théorie. Ce domaine de validité est d'autant plus grand que les moyens (statistiques) dont nous disposons ne nous permettent pas de réfuter cette théorie dans de nombreux cas particuliers.

VANDERPLANK a amélioré sa théorie considérablement en posant une nouvelle hypothèse (formule 2) : celle d'une croissance para-logistique. KAMPMEIJER et ZADOKS ont démontré que cette nouvelle théorie est fausse :

(1) la théorie envisage une épidémie sans lui attribuer de dimension spatiale, tandis qu'en réalité, elle s'étend dans les trois dimensions de l'espace (les deux dimensions du champ, et celle, verticale, du couvert végétal);

(2) la théorie suppose implicitement une distribution régulière de l'inoculum, à chaque instant, ce qui est réfuté par le phénomène des foyers de maladie.

Il serait cependant injuste de ne pas insister sur la grande valeur éducative des deux formules de VANDERPLANK. Elles méritent d'être utilisées dans l'enseignement, en soulignant leur valeur en tant qu'hypothèse.

Interprétations

Le cycle empirique est une méthode commode pour décrire le processus mental qui se déroule dans le cerveau d'un chercheur. Même si, clairement, il s'agit d'une représentation stylisée d'une réalité toujours complexe et confuse, ce cycle empirique, utilisé comme canevas méthodologique, aide beaucoup à clarifier les pensées du chercheur individuel. Au début d'un projet de recherche, ce cycle est cependant difficile à mettre en œuvre.

La « vérité » obtenue est temporaire. L'hypothèse n'est vraie que tant qu'une observation vient la réfuter. Le progrès de la science est comparable à celui d'un aveugle qui cherche son chemin en tâtonnant : ici on trébuche, là on se heurte, mais on progresse quand même. Nous ne connaissons jamais la vérité définitive, mais il nous est possible de démontrer la fausseté d'une hypothèse. Heureusement, une hypothèse fausse nous permet de reprendre à nouveau le cycle empirique.

Cette description est un peu injuste. Au fur et à mesure des recherches, on accumule une certaine compréhension des phénomènes observés. L'image des cycles repris à nouveau et encore ne rend pas justice au travail scientifique. Une image plus juste est celle d'une spirale, et la trajectoire d'un cycle empirique constitue un tour de cette spirale.

Les phases de l'induction et de la déduction sont les plus difficiles à saisir. L'induction peut être le fruit d'années de travail du taxonomiste, ou apparaît au chercheur comme un éclair. Souvent, la déduction est le fruit d'un travail mental patient et systématique, tout spécialement lorsqu'elle aboutit à une expérience coûteuse, mais bien préparée du point de vue du (des) terrain(s) d'expérimentation, de la main-d'œuvre, et des techniques statistiques.

Le rôle de l'intuition pour l'anticipation, l'induction et la déduction ne doit pas être sous-estimé. L'intuition peut être renforcée par l'expérience. D'une part, l'intuition accélère l'induction comme la déduction, mais d'autre part, elle peut comporter des risques importants d'erreur. Certains chercheurs travaillent intuitivement, et leur progression ressemble à une succession de bonds et d'arrêts. D'autres scientifiques sont plus rationnels, progressent lentement, mais de manière continue. Généralement, ces différences correspondent à des différences de caractère, qui n'ont pas d'effet sérieux sur les résultats scientifiques obtenus.

Il est vrai qu'anticipation et induction sont importantes. Selon PASTEUR : « l'aléa favorise l'esprit préparé ». D'après DARWIN : « il n'y pas de bon observateur qui ne soit pas un bon théoricien ». A cela s'ajoute le phénomène de « serendipity », mot qui n'existe pas dans le dictionnaire français. Ce néologisme anglais indique que l'on sait trouver des choses que l'on ne cherche pas, comme dans le conte du prince de SERENDIP, qui, d'ailleurs, n'a jamais existé. En science, on rencontre des cas de « serendipity », comme la découverte du rayonnement nucléaire.

Hypothèse et théorie

Une théorie n'est qu'une hypothèse, une hypothèse générale, de grande allure, présentée comme la vérité même, pour des raisons didactique ou politique. Néanmoins, une théorie n'est qu'une hypothèse de durée de vie limitée, une hypothèse que l'on veut croire temporairement. Aussitôt qu'une nouvelle théorie apparaît, dont le domaine de validité est plus vaste que celui de l'ancienne, celle-ci est mise à la poubelle.

Le domaine de validité d'une théorie est une notion difficile à maîtriser. Une théorie peut être renforcée ici, tandis qu'elle est réfutée ailleurs. Une théorie est inséparable de son domaine de validité. Ce constat trivial est important en biologie, parce que la notion juridique classique de *tertium non datur* (il n'y a pas de troisième solution) ne s'y applique pas. Très souvent, on trouve une nouvelle, une « troisième » possibilité. Un bon exemple est la détection de l'activité d'un stade sexué dans l'épidémie d'un champignon parasitaire, supposé jusque-là n'évoluer qu'au travers de cycles asexués (*Oidium tuckeri* de la vigne, *Septoria tritici* du blé).

Exemple trivial. Dans le cas des cygnes blancs, on pourrait esquisser la difficulté de l'existence de cygnes noirs en Australie, en avançant que l'hypothèse des cygnes blancs a un domaine de validité limité à l'Europe. Bien que cette limitation est justifiée pour les cygnes originaires d'Europe, elle montre la vulnérabilité de l'hypothèse, parce que dans un zoo européen, on peut voir des cygnes noirs importés d'Australie.

Exemple épidémiologique. La théorie des foyers, élaborée grâce aux mathématiques analytiques, aux mathématiques numériques (simulation dynamique) et à l'expérimentation en plein champ, nous a montré que le foyer d'une maladie de plantes progresse normalement avec un taux d'expansion constant (hypothèse). Pour tester cette théorie dans un cas nouveau non utilisé dans le cadre de son élaboration, on a formulé la prédiction qu'en culture mixte, le taux d'expansion serait proportionnel au logarithme de la proportion de plante sensible dans le mélange. Cette prédiction a été testée au champ, et vérifiée (VAN DEN BOSCH *et al.*, 1991).

FERRANDINO (1993) a proposé l'hypothèse des ondes dispersives, qui réfute l'hypothèse précédente, parce que le taux d'expansion des foyers augmente dans le temps. Ceci nous oblige à déclarer que le domaine de validité de la théorie des foyers à taux constant d'expansion a un domaine de validité particulier, domaine d'ailleurs qui n'est pas encore clairement défini.

Références

- de GROOT A.D., 1969. *Methodology, Foundations of Inference and Research in Behavioral Sciences*. La Haye, Mouton. 490 p.
- FERRANDINO F.J., 1993. Dispersive epidemic waves: I. Focus expansion within a linear planting. *Phytopathology* 83: 795-802.
- KAMPMEIJER P. & ZADOKS J.C., 1977. *EPIMUL, a simulator of foci and epidemics in mixtures, multilines, and mosaics of resistant and susceptible plants*. Simulation Monographs, Wageningen, Pudoc. 50 p.
- VAN DEN BOSCH F, VERHAAR M.A. & BUIEL A.A.M., HOOGKAMER W. & ZADOKS J.C., 1990. Focus expansion in plant disease: IV. Expansion rates in mixtures of resistant and susceptible hosts. *Phytopathology* 80: 598-602.
- VAN DER PLANK J.E., 1963. *Plant Diseases: Epidemics and Control*. New York, Academic Press. 349 p.
- ZADOKS J.C., 1972. Methodology of epidemiological research. *Annual Review of Phytopathology* 10: 253-276.
- ZADOKS J.C., 1972. *Methodology of epidemiological research*. pp. 63-96, in: *Plant Disease, an Advanced Treatise*, vol. II. (HORSFALL J.G. & COWLING E.B., Eds.), Academic Press, New York, 436 p.