

SEMINAR



Social Sciences in question: The leading epistemological and methodological controversies of our time

Compte-rendu de la 20^e séance du séminaire CEE-CERI

Les sciences sociales en question :

grandes controverses épistémologiques et méthodologiques

La prévision en relations internationales est-elle possible ?

16 octobre 2014

Samy Cohen (Sciences Po CERI) ouvre le séminaire en présentant Gerald Schneider, professeur à l'Université de Constance et président de l'European Political Science Association, auteur de *Forecasting in International Relations, One Guest, Three Approaches* (2011) et de *Forecasting the Risk of Extreme Massacres in Syria* (2014). Son intervention est discutée par Gilles Andréani, Conseiller maître à la Cour des comptes, ancien directeur du Centre d'analyse et de prévision (CAP) du ministère des Affaires étrangères entre 1995 et 1999 et entre 2001 et 2004 et professeur associé à l'Université Panthéon-Assas (Paris 2).

Pour Gerald Schneider la prévision en relations internationales est possible et même nécessaire. Le chercheur commence par un rappel des échecs en ce domaine, avant de présenter les différents types d'approches. Personne n'avait prévu le déclenchement de la Première Guerre mondiale. Ironiquement, *The Great Illusion*, le livre qui valut le prix Nobel de la paix à Norman Angell, écrit cinq ans avant 1914, pointait au contraire les facteurs économiques qui devaient dissuader les grandes puissances de se faire la guerre. De même, aucun chercheur en relations internationales, comme le souligne l'historien John Lewis Gaddis dans son article *International Relations Theory and the End of the Cold War* (1992/1993), n'avait anticipé la fin de la guerre froide¹. Et ni la CIA ni le FBI n'avaient prévu les attentats du 11 septembre 2001. Certains en viennent à mettre en doute la possibilité même de prédictions d'experts scientifiquement fondées ; à partir de l'évaluation des prédictions de centaines d'experts. Philippe E. Tetlock montre que ceux-ci ne font guère mieux que les amateurs (*Expert Political Judgment : How good is it ? How can we know ?* Princeton University Press, 2005)².

Gerald Schneider estime pourtant que certains événements dramatiques auraient pu, et auraient dû, être prédits. L'étude de Scott R. Feil (*Preventing Genocide : How the Early Use of Force Might Have Succeeded in Rwanda*, Carnegie Corporation, 1998) montre que le génocide rwandais, prévisible, aurait pu être évité par l'envoi préventif de troupes en avril 1994. Il en va de même pour le tremblement de terre d'Aquila, où les mises en garde répétées d'un technicien de laboratoire, Gianpaolo Giuliani, plusieurs semaines avant le drame n'ont pas été prises au sérieux par les autorités locales ou les scientifiques (Richard A. Kerr, Earthquake Prediction, *Science magazine*, 324, avril 2009, p. 322).

¹ *International Security*, (17 (3), Winter 1992/1993 :5-58).

² Voir aussi leurs recommandations aux services de renseignements américains : Michael C. Horowitz, Philip E. Tetlock, « National Security.Trending Upward: How the intelligence community can better see into the future », *Foreign Policy*, Septembre 2012.

Avec l'avènement des *big data*, le chercheur souligne le fait que de nouvelles possibilités de prédiction ont émergé, à partir de la veille systématique sur les médias et les réseaux sociaux. De nouveaux outils sont apparus comme GDELT (Global Database of Events, Language, and Tone)³ Analysis Service ou encore HealthMap⁴. GDELT a détecté un jour avant HealthMap des indices de la flambée du virus Ebola. Ces outils ne sont pas parfaits, ils ont leurs détracteurs⁵. Il n'en demeure pas moins que les données de Twitter et de Facebook sont aujourd'hui fréquemment utilisées pour la prédiction en relations internationales. Ainsi, Thomas Zeitzoff a utilisé Twitter pour suivre l'évolution du conflit entre le Hamas et les Israéliens⁶. Dans un article à paraître, co-écrit avec Sebastian Wörle, « Costly Tweets: The Impact of Negative Twitter Messages on Sovereign Spreads during the Eurozone crisis » (2014), Gerald Schneider s'intéresse à la crise de la zone euro et à l'influence des tweets négatifs sur les obligations souveraines des Etats en crise par rapport à leur relation avec l'Allemagne.

Pour le chercheur, il est particulièrement important de chercher à prédire certains événements, ce qui nécessite des théories innovantes. La qualité d'une théorie prédictive se juge à sa capacité à prévoir des événements extrêmes (massacres, attentats, épidémies) situés dans un futur plus ou moins proche. Il estime également que les techniques de prédiction se rôdent en équipe, et qu'il serait bon d'encourager la compétition entre les équipes, de généraliser les concours de prédictions pour sélectionner les meilleures d'entre elles.

Dans un second temps Gerald Schneider dresse une typologie des types d'approche : démarche idéal-typique, prédiction basée sur les jugements d'experts, sur la théorie des jeux, modèles « structurels », modèles

³ Pour une présentation détaillée de cette monumentale banque de données voir le site <http://www.gdeltproject.org/about.html>

⁴ Surveillance mondiale des épidémies et maladies contagieuses : <http://healthmap.org/fr/>

⁵ Voir notamment David Lazer et al. (2014) « The Parable of Google Flu: Traps in Big Data Analysis », *Science*, mars 2014, 343 (6176): 1203-1205.

⁶ « Using social media to measure conflict dynamics: An application to the 2008-2009 Gaza conflict », *Journal of Conflict Resolution*, December 2011, 55 (6): 938-969).

« dynamiques », etc. La plupart de ces approches s'appuient sur des sources d'origine humaine, ce qui les différencie est l'usage fait des informations.

Les jugements d'experts sont utiles dans la mesure où ces derniers sont familiers avec le contexte de l'événement à prévoir. Cependant, Philip Tetlock dans son article de 2005 a démontré que leurs prédictions ne sont pas nécessairement plus fiables que celles des non académiques. Des travaux plus récents aboutissent à des conclusions inverses. Comparant les prévisions des médias et des *traders* à Tel-Aviv quant à la durabilité de vingt et un cessez-le-feu dans différents conflits du Proche-Orient, Naomi Bosler et Gerald Schneider, dans « The Oracle or the Crowd? Experts Versus the Stock Market in Forecasting Ceasefire Success in the Levant » (2013), montrent que les *traders* ont été plus performants que les journalistes. En panachant trois critères, la précision, la mémoire et l'exactitude, ils infirment l'hypothèse de Tetlock.

La théorie des jeux est à la base des modèles élaborés par Bruce Bueno de Mesquita à l'Université de New York. En consultant des experts, il détermine les questions sensibles, les acteurs clés, la capacité et le pouvoir de ces derniers, informations qu'il combine à l'aide d'un algorithme. Ses prédictions ont un taux élevé d'exactitude, qu'il s'agisse de la seconde Intifada ou du successeur de Leonid Brejnev. Mais il est impossible de les répliquer puisque Bruce Bueno de Mesquita n'a jamais rendu publique la formule de cet algorithme, ce qui jette d'ailleurs un doute sur la démarche. Divers modèles de théorie des jeux ont également été testés par Gerald Schneider et ses collègues pour prévoir le résultat des négociations au sein européennes sur le e-commerce (Gerald Schneider, Daniel Finke, Bailer, « Bargaining Power in the EU : An Evaluation of Competing Game-Theoretic Models », *Political Studies*, 58 (1), February 2010 :85-103).

La prévision dite structurelle est comparable à celle que font les sismologues quand ils prédisent un futur tremblement de terre, en se fondant sur les

caractéristiques structurelles d'une région ou d'un pays et sur les épisodes antérieurs. Elle est utilisée pour prédire l'instabilité politique, les guerres civiles, les conflits armés. La Political Instability Task Force (auparavant State Failure Task force), subventionnée par le gouvernement américain, utilisent couramment ce type de modèle. Ces modèles prévoient toutefois mieux les cas de coopération que les conflits et souvent dans un avenir trop lointain pour que les conclusions soient très intéressantes⁷.

La prévision « dynamique » emprunte aux modèles probabilistes pour lesquels l'état présent d'un système est conditionné par ses états antérieurs et conditionne ses états futurs. Elle est fréquemment utilisée pour prédire des événements de même nature que les précédents (guerres civiles, massacres). Gerald Schneider a utilisé des modèles économétriques pour expliquer le massacre de Srebrenica et plus récemment, ceux qui ont eu lieu en Syrie après 2012⁸. Des informations sur le nombre de pertes civiles par province et par jour ont été récoltées sur un site web en 2013 puis traitées pour prédire l'occurrence d'événements similaires dans l'avenir. En observant si la distribution des données suivait une « loi de puissance », à l'opposé d'une distribution « normale »⁹ l'objectif était de prédire le risque d'un massacre similaire à celui de Homs dans les trois mois suivants sur la base des données disponibles pour les trois mois précédents. Selon Gerald Schneider, les probabilités les plus fortes vont dans le sens d'une stabilité des pertes aux alentours de 500 hommes par jour, un chiffre très élevé mais qu'on ne peut assimiler à un génocide. Son article,

⁷ Voir par exemple Jack A. Goldstone and al., « A Global Model for Forecasting Political Instability », *American Journal of Political Science*, 50 (1), January 2010 :190-208 et Håvard Hegre, Joakim Karlsen, Håvard Møkleiv Nygård, Håvard Strand, Henrik Urdal, « Predicting Armed Conflict, 2010–2050 », *International Studies Quarterly*, 57 (2), June 2013: 257-270.

⁸ Adam Scharpf, Gerald Schneider, Anna Nöh und Aaron Clauset, « Forecasting the risk of extreme massacres in Syria », *European Review of International Studies*, 1 (2), 2014:51-68. Consultable sur le site :

<http://www.polver.uni-konstanz.de/gschneider/arbeitspapiere/replikationsdaten/>

⁹ La distribution normale d'un phénomène suit une courbe en cloche ou courbe de Gauss. Par exemple, sur l'échelle gauche-droite ; on trouve peu d'individus aux deux extrêmes de l'échelle, la majorité se regroupe au centre (centre droit, centre et centre gauche). En revanche, certains phénomènes, comme le cas étudié, ont une distribution asymétrique. Il y a une loi de puissance si les fréquences (des massacres) diminuent fortement quand leur taille (nombre de morts) augmente.

il le déplore, n'a pas suscité l'intérêt des médias ni de la classe politique, sans doute précisément parce que il ne prévoyait pas une intensification des massacres. Il n'a été l'objet de débat à la radio et à la télévision qu'après l'attaque chimique d'août 2013.

Tout en reconnaissant les limites de certaines prévisions, Gerald Schneider insiste d'abord sur l'obligation morale qu'il y a à prédire dans les relations internationales, dans les cas de violences de masse, comme l'a fait le journaliste allemand Kurt Tocholsky pour les nombreux meurtres politiques qui ont secoué les premières années de la République de Weimar. Il avance ensuite des raisons scientifiques : la prédiction est un outil de recherche permettant de tester la pertinence des théories explicatives existantes. Il ne suffit pas de disposer des données, l'important est d'avoir de bonnes théories. Il lui semble enfin nécessaire d'entraîner les étudiants et les professionnels à la prédiction, de promouvoir une véritable culture de la prédiction dans l'administration publique.

Gilles Andréani

En tant que témoin et acteur, ce dernier revient sur l'intérêt et les utilisations de ces travaux de prédiction. Il ouvre le débat sur l'apport de ces méthodes, à aborder selon lui avec modestie. S'il reconnaît que la prédiction est possible en science politique et en relations internationales, il plaide néanmoins pour une approche empirique. Pour lui, une bonne information, une réflexion collective et une certaine culture historique permettent d'aboutir à des conclusions tout aussi pertinentes. Alors que les données sont fiables dans certains domaines, par exemple l'économétrie, il met en garde contre le manque de données en relations internationales et la mauvaise qualité de celles qui existent. Selon lui, l'analyse quantitative doit aller de pair avec l'analyse qualitative.

Tout comme Gerald Schneider, Gilles Andréani trouve fondée l'analogie avec la sismologie, mais il ajoute deux dimensions, « tectonique » et « éruptive » à la

prédiction. Pour lui, le « tectonique », c'est-à-dire le déplacement de grands champs de force et les tendances lourdes, est prévisible tandis que c'est plus difficilement le cas pour « l'éruptif » qui contient une part d'émotion, d'irrationnel, d'accident et de chance, comme tous les domaines de la politique.

Finalement, il revient sur les raisons de l'audience limitée des chercheurs. Le fait que, souvent, les hommes politiques ont leur propre vision des choses rend selon lui difficile l'intégration de la prédiction, souvent exprimée en termes probabilistes, dans leur schéma de décision.

Gerald Schneider

Répondant à Gilles Andréani, Gerald Schneider explique que ce sont moins les données qui manquent en relations internationales que les théories pour appréhender ces dernières. Il reconnaît que les hommes politiques sont très peu familiers des probabilités et les grands chiffres, mais également que certains professionnels ont tendance à être trop confiants dans leurs prédictions. Et il rappelle que l'agrégation des opinions d'individus – la sagesse des foules – est parfois plus pertinente que l'opinion d'un expert.

Débat avec la salle

Plusieurs questions portent sur la capacité à prédire des événements qui n'ont pas d'antécédents, qui n'ont pas encore eu lieu ou dont on n'a pas encore idée. Vincent Pouliot (CERI) souligne que Gerald Schneider tend à prédire la continuité et l'extrapolation d'une tendance plutôt que les déviations, les *black swans* (événements imprévisibles qui ont une faible probabilité de se dérouler).

Gerald Schneider répond qu'il existe des théories pour appréhender des cas spécifiques comme les attentats du 11 septembre 2001 ou la flambée d'Ebola, mais que prédire quelque chose dont on n'a pas encore idée est un défi réel, tant

d'un point de vue théorique que de celui de données adéquates, dont la collecte n'est pas systématique. Dans la mesure où il est nécessaire de connaître un processus pour faire des prédictions, il reconnaît qu'extrapoler à partir de ce qu'on sait est plus facile que de prédire les cas déviants ou extrêmes.

M. Guilhem (Ministère de la Défense) demande à l'auteur s'il connaît le projet de prédiction en ligne Hypermind¹⁰ composé de « pronostiqueurs d'élite » amateurs. Gerald Schneider ne connaît pas ce projet mais souligne que Tetlock coordonne un projet similaire.

A propos des échelles de prédiction, Philippe Bonditti (CERI) demande à l'auteur s'il fait la différence entre prédiction de phénomènes de masse et prédiction d'actes individuels. Jean Leca (Sciences Po) s'interroge sur la portée de la prédiction au sens où l'entend Gerald Schneider : s'agit-il de prédictions globales recouvrant une multiplicité de phénomènes (par exemple, l'éclatement d'une guerre) ou de mettre au jour des liens de causalité dans un domaine spécifique (par exemple le *Process tracing* en relations internationales qui essaie de faire de l'analyse causale sans pour autant faire des prédictions).

Gerald Schneider déclare vouloir faire des prédictions autant au niveau macro que micro. Il admet qu'il existe davantage de données et de théories sur les processus d'escalade au niveau international qu'au niveau national.

Plusieurs questions portent ensuite sur la réception controversée des prédictions et des probabilités par les décideurs politiques, la pertinence de ces travaux, nonobstant la réticence manifestée par ces derniers à leur égard, le manque d'études comparatives pour attirer leur attention sur le sujet. Cet aspect est également abordé par Samy Cohen (Sciences Po, CERI), qui souligne les limites aux prédictions face aux préjugés des dirigeants et les relations parfois difficiles existant entre experts et hommes politiques. Il donne l'exemple de la guerre de

10 <https://hypermind.lumenogic.com/hypermind/app.html#welcome>

Kippour et de la mobilisation tardive de l'armée israélienne face à l'attaque égyptienne du 6 octobre 1973. Les dirigeants israéliens avaient bien été informés de l'imminence d'une attaque, mais ils étaient convaincus que l'armée égyptienne n'avait aucune chance de réussir. Dans certains cas, les hommes politiques ne font simplement pas confiance aux experts. C'était le cas de Roland Dumas qui travaillait en comité très restreint et excluait le Centre d'analyse et de prévision du processus de décision.

Pour Gerald Schneider, l'entraînement à la compréhension des modèles probabilistes est très important pour améliorer le jugement des décideurs. Il ajoute néanmoins que tous les phénomènes ne présentent pas un niveau de complexité élevé et que certains nécessitent des réponses simples, binaires, qui peuvent facilement être identifiées par les décideurs. Toute prédiction devrait d'ailleurs à ses yeux être donnée avec un niveau de fiabilité, un intervalle de confiance. Selon lui, les décideurs ne sont pas intéressés par la comparaison avec d'autres mais par leur seul cas individuel.

Nonna Mayer (Sciences Po, CEE) s'interroge sur la fiabilité des *big data* par rapport aux autres types de données (sondages d'opinions, évolution des marchés, presse) et sur l'utilisation des modèles prédictifs développés en physique (contingence, chaos et entropie).

Pour Gerald Schneider les modèles développés en physique ne sont pas capables de saisir les interactions sociales dans toute leur complexité.

Un autre intervenant demande comment dans son article à paraître sur le rôle des tweets durant la crise de l'Eurozone crisis » l'auteur différencie l'impact des commentaires de Twitter de celui des agences de presse et des médias. Gerald Schneider explique que son analyse prend également en compte l'impact du développement des marchés et des agences de notation, ainsi que l'interaction entre ces indicateurs économiques et les commentaires de Twitter.

Nonna Mayer (Sciences Po, CEE) pose une dernière question sur la position des ministères concernés en France par l'utilisation de ce type de données.

Gerald Schneider estime l'usage des *big data* peu répandu en France, mais reconnaît ne pas bien connaître le contexte. Gilles Andréani précise que l'intérêt premier des réseaux du CAP (Centre d'analyse et de prévision du Ministère des Affaires étrangères) était d'amener des chercheurs et des experts susceptibles d'apporter de nouvelles idées et de nouvelles données pour éclairer le processus de décision en relations internationales. Il ajoute que le centre n'était pas intéressé par la prédiction des accidents ou des événements ponctuels. Gilles Andréani était particulièrement attentif à révéler (et moins à prédire) les tendances sous-jacentes, le « tectonique », et les moments de rupture. Un raisonnement probabiliste n'aurait pas intéressé les décideurs et aurait présenté un risque pour l'image du centre en cas d'erreur d'appréciation.

Il conclut sur le fait que la prédiction est possible, mais qu'il n'existe pas de déterminisme et que l'horizon et le contenu de la prédiction sont souvent très limités.