Crise de réputation : l’hypothèse de la cascade

Tyler Kohler . 23 octobre 2026 . note 03 . mis a jour le 10 novembre 2026

RESUME

Deux organisations touchées par un incident comparable peuvent connaître des trajectoires opposées. Cette note défend l’hypothèse qu’une crise de réputation se comporte comme une cascade : son ampleur se décide dans les conditions de reprise du fait initial, c’est-à-dire la distribution des dispositions à relayer, la présence d’amplificateurs et l’instant où la cible intervient. Le mécanisme est décrit en sept étapes, appuyé sur les modèles de seuil de Granovetter, sur les travaux de Goel et coauteurs sur la viralité structurelle et sur l’expérience de marché artificiel de Salganik et coauteurs. Je propose d’appeler point porte-voix l’instant où la réaction de la cible devient le premier amplificateur de sa propre crise. Les limites du modèle sont exposées.

Crise de réputation
Séquence au cours de laquelle un fait concernant une personne ou une organisation est repris par des tiers en nombre croissant, jusqu’à ce que l’information publiquement disponible sur elle soit dominée par ce fait pendant une durée qui excède celle des reprises.

Seuil de reprise
Nombre, ou proportion, de reprises déjà visibles à partir duquel un acteur donné relaie le fait à son tour : application aux reprises publiques du seuil que Granovetter (1978) définit pour le comportement collectif.

Amplificateur
Acteur ou dispositif dont le passage augmente l’audience d’un contenu sans modifier sa substance : compte à forte audience, reprise par un média, position obtenue dans un classement.

Viralité structurelle
Distance moyenne entre toutes les paires de nœuds d’une cascade de diffusion (Goel et coauteurs, 2016) : à nombre de personnes atteintes égal, elle sépare une transmission de proche en proche d’une diffusion émise en masse par une source unique.

Soit deux organisations traversant le même incident : un courriel interne qui circule, une phrase enregistrée sans précaution. L’exemple est construit pour les besoins de l’exposé et ne décrit aucun dossier. Six mois plus tard, la première pourrait n’en garder aucune trace publique, quand sur la seconde le fait serait devenu l’information dominante dès qu’on cherche son nom. On attribue d’ordinaire un tel écart à la gravité de ce qui s’est passé. Je propose de le chercher dans la mécanique de reprise.

L’hypothèse et ses conditions de validité

Hypothèse : une crise de réputation est un phénomène de cascade, et son ampleur se détermine pour l’essentiel dans les conditions de reprise du fait initial : la distribution des dispositions à relayer, la présence d’amplificateurs sur le trajet, et l’instant où la cible intervient.

Le fait initial délimite une plage de résultats possibles. À l’intérieur de cette plage, le résultat reste largement indéterminé.

L’hypothèse n’est pas universelle. Elle suppose d’abord que les conséquences passent par des reprises publiques entre acteurs qui s’observent les uns les autres. Une sanction prononcée par une autorité de surveillance, ou la rupture d’un contrat décidée par trois personnes dans une salle, suit d’autres règles, et un modèle de cascade n’y apporte rien. Elle suppose un plancher de plausibilité :

un fait qu'aucun lecteur ne tient pour possible ne franchit pas la première reprise. Elle suppose enfin que la cible dispose d'un canal d'expression visible. Sans canal, le troisième terme du modèle disparaît.

Quatre termes définis avant usage

Le vocabulaire courant mélange l'incident et sa trajectoire. Le terme de « bad buzz » désigne indifféremment une poussée de trois heures sans suite et une occupation durable de l'espace public. La définition que j'emploie porte sur la trajectoire.

Le terme suivant vient des modèles de comportement collectif. Granovetter (1978) appelle seuil d'un acteur le nombre, ou la proportion, d'autres acteurs qui doivent avoir agi avant qu'il agisse à son tour. Son article est un modèle mathématique, sans données de terrain, et l'auteur y signale lui-même la difficulté à mesurer les seuils.

Reste à nommer ce qui augmente l'audience d'un contenu sans en modifier la substance.

Le dernier vient de l'analyse des cascades : Goel et coauteurs (2016) mesurent la viralité structurelle d'une diffusion par la distance moyenne entre toutes les paires de nœuds de sa cascade.

Le mécanisme proposé

Première étape, la mise en circulation. Un fait devient disponible pour des tiers. Sa forme compte à ce stade, et je le pose comme hypothèse : un document consultable, un enregistrement ou une capture d'écran se reprendraient plus facilement qu'une affirmation sans support. Aucun des travaux cités ne mesure cet écart.

Deuxième étape, la première couche de reprise. Elle vient des acteurs dont le seuil est nul ou proche de zéro, ceux qui relaient sans attendre que d'autres l'aient fait : parties déjà engagées, concurrents, comptes spécialisés, anciens collaborateurs. Je fais l'hypothèse que cette couche est petite et stable dans sa composition. Elle n'est pas mesurée ici. Les dispositifs qui fabriquent artificiellement une première couche relèvent d'un autre sujet, traité dans la note sur ce qui distingue une campagne d'une colère.

Troisième étape, l'asymétrie de traitement. Baumeister et coauteurs (2001) recensent, dans un grand nombre de domaines, une asymétrie récurrente au bénéfice du négatif : les événements, les émotions et les retours défavorables auraient des effets plus forts que leurs équivalents favorables. Leur article est une synthèse narrative de littérature, sans taille d'effet agrégée, et les auteurs signalent eux-mêmes des exceptions à l'asymétrie qu'ils décrivent. Je formule donc l'étape comme une hypothèse : le seuil de reprise d'un fait défavorable serait en moyenne plus bas que celui d'un fait favorable comparable. Les ratios chiffrés que l'on attribue parfois à ce travail n'y figurent pas.

Quatrième étape, le franchissement d'un amplificateur. Goel et coauteurs (2016), sur environ un milliard d'événements de diffusion, établissent deux choses qui importent ici. La très grande majorité des cascades reste minuscule. Et la popularité d'un contenu dépend surtout de la taille de sa plus grande diffusion directe, donc d'un mécanisme de diffusion de masse ; la transmission de proche en proche y joue un rôle secondaire. J'en tire une conséquence que l'étude n'établit pas : le passage par un seul amplificateur modifierait l'échelle davantage que des centaines de reprises latérales. Goel et coauteurs ne mesurent aucune capacité de prédiction au niveau d'un contenu particulier.

Cinquième étape, le seuil de bascule. C'est le point du modèle où la prévision devient impossible. Granovetter démontre formellement que le résultat collectif se déduit de la distribution des seuils, et

que des variations infimes de cette distribution produisent des équilibres radicalement différents : avec deux populations presque identiques, émeute générale ou rien du tout. Il en tire une conséquence plus importante encore : on ne peut pas inférer les dispositions individuelles à partir du résultat collectif observé.

Une expérience éclaire ce point. Salganik, Dodds et Watts (2006) ont construit un marché musical artificiel : 14'341 participants, huit univers parallèles exposés aux mêmes morceaux de groupes inconnus. L'affichage du nombre de téléchargements augmente simultanément l'inégalité des succès et leur imprévisibilité d'un univers à l'autre. La qualité mesurée sans influence sociale ne prédit que faiblement le classement final : les meilleurs morceaux échouent rarement, les pires réussissent rarement, et entre ces deux bornes le résultat est indéterminé. Le dispositif est un laboratoire en ligne, avec des artistes inconnus et un horizon de quelques jours ; les auteurs notent que l'influence sociale hors laboratoire est plus forte et de nature différente.

Sixième étape, la réaction de la cible. Dans l'hypothèse proposée, elle ajouterait au dossier une reprise émise par la source la plus crédible sur le sujet, et elle fournirait aux amplificateurs un motif de reprise indépendant du fait initial. Aucun travail cité ne mesure cet effet.

Septième étape, la persistance. Chaque reprise laisse une trace consultable, et la répétition agit sur la validité attribuée à un énoncé. L'effet de vérité illusoire, décrit par Hasher, Goldstein et Toppino (1977), s'observe même chez des participants qui possèdent la connaissance contredisant l'énoncé (Fazio et coauteurs, 2015). Les énoncés de 1977 étaient des énoncés de culture générale plausibles, de vérité incertaine pour les participants. Fazio, Rand et Pennycook (2019) observent l'effet à tous les niveaux de plausibilité, y compris pour les énoncés implausibles, le plus marqué portant sur les énoncés ambigus : la plausibilité d'un énoncé ne met pas à l'abri de l'effet. Ces travaux mesurent des jugements de vérité déclarés en laboratoire, sur des énoncés courts et sans source identifiée. Les transposer à des contenus d'actualité demeure une extrapolation.

Le schéma représente la trajectoire d'un fait en quatre phases : déclencheur, amplificateurs, seuil de bascule où le résultat devient indéterminé, puis persistance, le fait restant consultable après la fin des reprises.

Ce que montre la littérature, et à quelles conditions

Le corpus de référence sur la diffusion des rumeurs est celui de Vosoughi, Roy et Aral (2018) : environ 126'000 cascades, diffusées par environ trois millions de personnes sur Twitter entre 2006 et 2017, leur véracité ayant été établie par six organismes de vérification indépendants. Les cascades fausses y atteignent plus de personnes, plus vite et plus loin que les vraies, l'écart étant maximal pour l'information politique. Ce travail porte sur des rumeurs controversées. Il ne porte pas sur des contenus visant une personne ou une organisation, et les informations non contestées, soit l'essentiel du flux, restent hors de son échantillon.

La réanalyse de Juul et Ugander (2021) apparie les cascades sur leur taille : les différences de profondeur, de largeur, de viralité structurelle et de vitesse entre vrai et faux disparaissent alors. Elle porte sur l'opposition entre vrai et faux. Elle invalide l'idée d'une forme de propagation propre au faux. L'étendre aux contenus défavorables serait une extrapolation de ma part. La différence de taille, donc de portée, subsiste. J'ai discuté ces deux travaux et ce que leur confrontation laisse debout dans la note sur l'influence de masse ; ce que la cascade exige ici se limite à cette différence de portée.

Sur la structure du voisinage, deux lignes de travaux s'opposent. Centola et Macy (2007)

démontrent, par simulation sur des réseaux de type Watts-Strogatz et sans aucune donnée empirique, que pour une contagion à seuil les liens longs et isolés cessent d'accélérer la diffusion : la propagation dépend alors de la largeur des ponts entre groupes, soit du nombre de liens qui les relient. Centola (2018) rapporte des expériences allant dans le même sens sur des réseaux construits en laboratoire. Ni l'un ni l'autre ne mesure la propagation d'un comportement réel. Eckles et coauteurs (2024) renversent ce résultat, par une démonstration de modèle elle aussi : il suffit d'une faible probabilité d'adoption en dessous du seuil pour que les liens longs accélèrent la propagation. Leur critique porte sur la robustesse du résultat relatif à la structure des liens ; l'existence du renforcement social reste hors de cause. Les deux lignes conduisent à des recommandations opposées sur la structure d'un réseau. Aucune des deux ne permet de prescrire quoi que ce soit à une organisation exposée.

Sur la réponse elle-même, la référence est la théorie situationnelle de la communication de crise (Coombs, 2007). Elle tient que l'efficacité d'une réponse dépend de la responsabilité attribuée à l'organisation, laquelle dépend du type de crise, de l'historique de crises et de la réputation antérieure. C'est un cadre théorique, sans données nouvelles. La méta-analyse de Ma et Zhan (2016) confirme la relation centrale entre responsabilité attribuée et dégradation de la réputation, et rapporte un soutien plus faible et plus hétérogène aux prescriptions d'appariement entre type de crise et stratégie de réponse. Le cadre de 2007 ignore par ailleurs entièrement l'environnement numérique : rien sur les plateformes, rien sur la visibilité en ligne. Ce qui borne l'effet d'une réponse, avant toute question de registre, fait l'objet de la note sur la communication de crise.

Un concept proposé : le point porte-voix

Jansen et Martin (2015) proposent un cadre d'analyse de l'effet Streisand. La tentative de suppression se retourne contre son auteur lorsque les tactiques habituelles de réduction de l'indignation échouent toutes : dissimuler l'acte, dévaloriser la cible, réinterpréter les faits, recourir à des procédures officielles d'apparence légitime, intimider. Leur article est conceptuel. Les cas y sont retenus parce que la censure y a échoué, et les suppressions réussies sont par nature invisibles, donc absentes de l'échantillon : ce cadre ne fournit aucune fréquence et aucune probabilité de retour de flamme. Ce défaut d'échantillon est examiné dans la note sur l'effet Streisand.

Il y manque un repère temporel. Je propose d'appeler point porte-voix l'instant d'une cascade à partir duquel l'audience des interventions de la cible excède celle du fait initial.

Dans cette hypothèse, une prise de parole antérieure au point resterait contenue dans l'audience du fait et s'adresserait à des personnes déjà informées. Une prise de parole postérieure atteindrait des personnes qui ignoraient le fait et le leur apprendrait. Le déplacement serait double : la cible deviendrait le premier amplificateur de sa propre crise, et sa réaction constituerait un fait nouveau, susceptible de sa propre cascade.

Ce point est une hypothèse. Aucun travail publié ne le mesure, et je ne connais pas d'indicateur permettant de le situer à l'avance. Formulé ainsi, il n'est pas réfutable. Il le deviendrait si l'on spécifiait une grandeur observable, par exemple l'audience constatée du fait comparée à celle du canal par lequel la cible répond. Son intérêt est un intérêt de décision : il oblige à estimer une audience avant de choisir le volume d'une réponse.

Limites et zones d'incertitude

Le modèle repose sur une grandeur rarement mesurable. Granovetter est explicite sur la difficulté à mesurer les seuils. J'en conclus qu'il ne permet aucune prévision, et cette conclusion m'appartient.

Un usage courant du modèle, expliquer après coup pourquoi telle affaire a pris, est celui que son résultat central interdit.

L'indétermination a des bornes. Salganik et Watts (2008) ont inversé artificiellement les classements affichés dans une expérience de même type, menée sur plus de 12'000 participants : les faux classements s'auto-réalisent partiellement, puis la qualité intrinsèque des morceaux se réaffirme progressivement. Le récit du succès ou de l'échec purement arbitraire ne tient pas.

Le modèle n'explique pas pourquoi certaines crises produisent des conséquences matérielles et d'autres aucune. La littérature mobilisée porte sur des diffusions et sur des jugements déclarés. Aucun de ces travaux ne mesure une perte de clientèle ni un retrait de mandat.

Le poids que j'attribue aux amplificateurs de visibilité repose sur une base étroite. Pan et coauteurs (2007) montrent, par oculométrie, que la position d'un résultat agit comme un signal de crédibilité autonome : les participants cliquent majoritairement le premier résultat même lorsque son résumé est jugé moins pertinent par des juges indépendants. L'effectif est très faible, seize participants retenus sur vingt-deux recrutés, et l'étude n'établit aucun taux de clic par rang généralisable. Les parts de clic par position que l'on cite couramment ne viennent pas de ce travail. Elle ne dit rien non plus du comportement sur mobile, des pages de résultats actuelles ni des autres moteurs. Les dispositifs de recommandation automatique sortent de ce que les travaux cités permettent d'examiner, et je les laisse hors du modèle.

Enfin, l'essentiel des données de diffusion mobilisées ici vient d'une seule plateforme, sur une période close, et de corpus sélectionnés par ce qui avait attiré la vérification. La transposition à d'autres espaces reste une hypothèse de travail.

Implications

Ce qui suit hérite des limites exposées plus haut.

Compte tenu de la distribution des tailles de cascades observée par Goel et coauteurs, une poussée hostile est souvent plus petite qu'elle ne paraît à la personne qui la subit, seule à la lire en entier. Je n'ai pas de mesure de cet écart.

Le volume d'une réponse se dimensionne sur l'audience constatée du fait. Estimer cette audience avant de parler suppose un chiffre, rarement disponible au moment où la décision se prend.

La seule relation que la méta-analyse de Ma et Zhan confirme solidement lie la responsabilité attribuée à la dégradation de la réputation. Elle ne montre pas qu'une réponse déplace cette attribution. Je fais l'hypothèse que l'attribution est le paramètre le plus accessible à une réponse, et cette hypothèse n'est pas testée. Les prescriptions plus fines d'appariement entre type de crise et registre de réponse reposent sur une base plus faible et plus hétérogène, ce qui invite à les traiter comme une grille de lecture.

La vitesse n'a pas de valeur propre. Son signe dépend du moment où l'on intervient par rapport au point porte-voix, et ce moment n'est pas observable à l'avance.

L'objectif atteignable se mesure en durée d'occupation de l'espace public. La persistance des traces rend l'effacement improbable. L'effet de la répétition sur la validité attribuée à un énoncé, lui, a été mesuré en laboratoire, sur des énoncés courts et sans source identifiée : il éclaire le problème sans le trancher.

Le travail qui précède la crise est l'objet de ce que mesure un classement quand la requête est un

nom : il porte sur ce qui existe déjà quand le fait sort. Ce que la première page de résultats laisse voir une fois la cascade retombée relève d'une autre question, traitée dans la note sur la gestion de la perception.

Une crise traversée sans dommage ne prouve pas la qualité de la réaction. Du résultat observé, on n'infère ni la gravité du fait initial, ni la justesse de ce qui a été fait.

REFERENCES

GRANOVETTER, Mark (1978). Threshold Models of Collective Behavior. American Journal of Sociology, vol. 83, no 6, p. 1420-1443.

SALGANIK, Matthew J., DODDS, Peter Sheridan et WATTS, Duncan J (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. Science, vol. 311, no 5762, p. 854-856.

SALGANIK, Matthew J. et WATTS, Duncan J (2008). Leading the Herd Astray: An Experimental Study of Self-Fulfilling Prophecies in an Artificial Cultural Market. Social Psychology Quarterly, vol. 71, no 4, p. 338-355.

GOEL, Sharad, ANDERSON, Ashton, HOFMAN, Jake et WATTS, Duncan J (2016). The Structural Virality of Online Diffusion. Management Science, vol. 62, no 1, p. 180-196.

VOSOUGHI, Soroush, ROY, Deb et ARAL, Sinan (2018). The spread of true and false news online. Science, vol. 359, no 6380, p. 1146-1151.

JUUL, Jonas L. et UGANDER, Johan (2021). Comparing information diffusion mechanisms by matching on cascade size. Proceedings of the National Academy of Sciences, vol. 118, no 46, article e2100786118.

CENTOLA, Damon et MACY, Michael W (2007). Complex Contagions and the Weakness of Long Ties. American Journal of Sociology, vol. 113, no 3, p. 702-734.

CENTOLA, Damon (2018). How Behavior Spreads: The Science of Complex Contagions. Princeton, Princeton University Press, 312 p.

ECKLES, Dean, MOSSEL, Elchanan, RAHIMIAN, M. Amin et SEN, Subhabrata (2024). Long ties accelerate noisy threshold-based contagions. Nature Human Behaviour, vol. 8, no 6, p. 1057-1064.

BAUMEISTER, Roy F., BRATSLAVSKY, Ellen, FINKENAUER, Catrin et VOHS, Kathleen D (2001). Bad is Stronger than Good. Review of General Psychology, vol. 5, no 4, p. 323-370.

JANSEN, Sue Curry et MARTIN, Brian (2015). The Streisand Effect and Censorship Backfire. International Journal of Communication, vol. 9, p. 656-671 (acces libre).

COOMBS, W. Timothy (2007). Protecting Organization Reputations During a Crisis: The Development and Application of Situational Crisis Communication Theory. Corporate Reputation Review, vol. 10, no 3, p. 163-176.

MA, L. et ZHAN, M (2016). Effects of attributed responsibility and response strategies on organizational reputation: A meta-analysis of situational crisis communication theory research. Journal of Public Relations Research, vol. 28, no 2, p. 102-119.

HASHER, Lynn, GOLDSTEIN, David et TOPPINO, Thomas (1977). Frequency and the conference of referential validity. Journal of Verbal Learning and Verbal Behavior, vol. 16, no 1, p. 107-112 (titre reproduit tel que publie).

FAZIO, Lisa K., BRASHIER, Nadia M., PAYNE, B. Keith et MARSH, Elizabeth J (2015). Knowledge does not protect against illusory truth. Journal of Experimental Psychology: General, vol. 144, no 5, p. 993-1002.

FAZIO, Lisa K., RAND, David G. et PENNYCOOK, Gordon (2019). Repetition increases perceived truth equally for plausible and implausible statements. Psychonomic Bulletin & Review, vol. 26, no 5, p. 1705-1710.

PAN, Bing, HEMBROOKE, Helene, JOACHIMS, Thorsten, LORIGO, Lori, GAY, Geri et GRANKA, Laura (2007). In Google We Trust: Users' Decisions on Rank, Position, and Relevance. Journal of Computer-Mediated Communication, vol. 12, no 3, p. 801-823.

Kohler, T. (2026). Crise de réputation : l'hypothèse de la cascade. Notes, tylerkohler.ch. https://tylerkohler.ch/notes/crise-de-reputation/