

Quand l'Ined rencontre Meetic



Marie BERGSTRÖM

Chargée de recherche à l'Institut national d'études démographiques

Certains thuriféraires des Big Data prédisent que ces données nouvelles vont supplanter les sources traditionnelles, telles que les enquêtes sur échantillon. Des chercheurs mettent au contraire en garde contre les données massives non structurées, en attirant l'attention sur leur manque de représentativité et sur leur instabilité. Rares sont les études qui utilisent les deux types de sources. C'est le cas de ce travail, qui conclut à leur complémentarité.

Le recours aux services de rencontres sur Internet est aujourd'hui une pratique fréquente¹. Contrairement aux annonces et aux agences matrimoniales dont l'usage est toujours resté marginal², les rencontres en ligne sont devenues courantes chez les jeunes comme les moins jeunes. C'est le cas en France où 12 % des femmes et 16 % des hommes, âgés de 26 à 65 ans, s'étaient déjà inscrits sur un tel site en 2013³. Parmi les personnes qui n'étaient pas en couple au moment de l'enquête, le taux d'usage était plus important : 25 % des femmes et 28 % des hommes⁴.

La diffusion de ces services leur vaut désormais une attention des chercheurs en sciences sociales. Ces sites enregistrent de nombreuses informations qui, lorsqu'extraites, codées et analysées par le chercheur, peuvent apporter des enseignements précieux sur la formation des couples. Les données issues des sites de rencontres ont ainsi été mobilisées dans plusieurs recherches récentes consacrées notamment à l'homogamie sociale et ethno- raciale⁵. Cette nouvelle source bouscule quelque peu la sociologie du couple qui repose traditionnellement sur l'exploitation d'enquêtes par questionnaire. En France, on pense notamment aux enquêtes de l'Ined dont *Le choix du conjoint*, réalisée par Alain Girard en 1959⁶ et *La formation des couples*, conduite par Michel Bozon et François Héran dans les années 1980⁷. Ces enquêtes ont prouvé leur capacité d'objectivation des rencontres amoureuses. Elles comportent cependant des limites face auxquelles les données issues des sites s'avèrent très complémentaires.

D'une part, ces sites permettent d'observer les *rencontres en train de se faire* plutôt que d'étudier les *couples déjà constitués* comme c'est le cas des enquêtes par questionnaire. Les sites

1. L'auteure remercie la société *Meetic* qui lui a fourni une partie des données mobilisées dans cet article. Elle tient aussi à remercier le programme OxPo (Oxford - Sciences Po) pour le soutien financier qu'il lui a accordé l'année où la recherche a été réalisée. Ce texte résume le résultat principal d'un article à paraître fin 2018 dans la *Revue française de sociologie*.
2. Bozon M. et Héran F. (1987), La découverte du conjoint. I. Évolution et morphologie des scènes de rencontre, *Population*, 42-6, 943-985.
3. Bergström M. (2016), Sites de rencontres : qui les utilise en France ? Qui y trouve son conjoint ?, *Population & Sociétés*, 530, 1-4.
4. Source : enquête *Épic*, Ined-Insee, 2013-2014.
5. Skopek J., Schulz F. et Blossfeld H-P. (2011), Who Contacts Whom? Educational Homophily in Online Mate Selection, *European Sociological Review*, 27-2, 180-195 ; Lin K-H. et Lundquist J. (2013), Mate selection in cyberspace: The intersection of race, gender, and education, *American Journal of Sociology*, 119-1, 183-215 ; Potârca G. et Mills M. (2015), Racial Preferences in Online Dating across European Countries, *European Sociological Review*, 31-3, 326-341 ; Schmitz A. (2016), *The Structure of Digital Partner Choice. A Bourdieusian perspective*, New-York, Springer.
6. Girard A. (2012 [1964]), *Le choix du conjoint. Une enquête psycho-sociologique en France*, Paris, Armand Colin.
7. Bozon M. et Héran F. (2006), *La formation du couple. Textes essentiels pour la sociologie de la famille*, Paris, La Découverte.

enregistrent en effet les comportements de contact des usagers et permettent ainsi d'observer le processus d'appariement des partenaires. Ce faisant, ils offrent aux sociologues du couple une opportunité méthodologique inédite : alors que les enquêtes ne captent que la *sélection* amoureuse, les sites enregistrent aussi les comportements de *refus* et d'*élimination* qui font partie intégrante de la séduction mais dont les enquêtes sont sans trace.

D'autre part, les sites de rencontres captent des *pratiques* tandis que les enquêtes enregistrent des *déclarations de pratiques*. Le questionnaire suppose une réflexivité importante de la part des enquêtés et pose la question de la mise en cohérence des discours. Dans le cas des rencontres, on peut s'interroger tout à la fois sur la capacité, la volonté et les manières – socialement et sexuellement différenciées – de raconter « son histoire amoureuse ». Les données issues des sites de rencontres font, elles, l'économie de la déclaration pour autoriser, de façon originale, une *observation quantifiée* des pratiques affectives.

À ce titre, les sites de rencontres sont un bon exemple de l'intérêt que peuvent avoir les « données massives » pour la recherche en sciences sociales et leur complémentarité par rapport aux données d'enquête. On propose d'illustrer quelques-unes de ces opportunités méthodologiques et empiriques. Pour ce faire, on présente un exemple tiré d'une recherche sur l'écart d'âge entre partenaires hétérosexuels. L'étude confronte des données d'enquête (*Épic*) à des données d'un site (*Meetic*) qui donnent des réponses différentes à la question de comment l'écart d'âge advient.

Un nouveau site d'observation des rencontres

Dans les couples hétérosexuels, l'homme est souvent plus âgé que sa conjointe. En France en 2012, les unions cohabitantes étaient caractérisées par une différence d'âge moyenne de 2,5 en faveur de l'homme⁸. Cette asymétrie des âges s'observe – à des degrés variables – dans la quasi-totalité des pays et des époques connus⁹. Elle est un objet classique de la sociologie du couple, abordée principalement à partir d'enquêtes. Celles-ci informent avec précision sur l'ampleur et l'évolution de l'écart d'âge entre conjoints mais peinent à rendre compte de la manière dont cet écart se *produit*. Est-ce que ce sont les femmes, ou les hommes, ou les deux sexes qui désirent cette asymétrie ?

Pour tenter de répondre à cette question on mobilise des données issues du site *Meetic.fr*, obtenues grâce à un partenariat avec la société éditrice du site, *Meetic France*. Deux types d'informations ont plus précisément fait l'objet d'analyses. D'une part, on a travaillé sur des données anonymisées relatives aux *profils d'utilisateurs*. Il s'agit des informations que les usagers ont renseigné au sujet d'eux-mêmes et de leurs préférences (à l'exception des pseudonymes et des photographies), comme par exemple leur sexe, leur âge et leurs préférences d'âge. Une deuxième source d'information concerne les emails envoyés sur le site. En aucun cas on n'a eu accès au contenu des messages. Seulement les métadonnées relatives aux échanges ont été analysées : les identifiants de l'expéditeur et du destinataire ainsi que la date et l'heure de l'envoi. Recoupées avec les données de profil, ces informations permettent de savoir « qui contacte qui », notamment en fonction de l'âge. Pour cette recherche on se base, d'une part, sur un échantillon constitué de l'ensemble des profils d'utilisateurs enregistrés sur le site en 2014 et ayant envoyé au moins un mail dans l'année (environ 400 000 profils) et, d'autre part, un échantillon constitué de l'ensemble des emails échangés sur le site en 2014 (plus de 25 millions d'emails).

Si ces données « interactives » ont l'avantage de donner une image dynamique des rencontres, elles comportent de nombreuses limites. Une réserve importante concerne la véracité des informations. Sur Internet, on ne donne pas toujours son âge. On arrondit. C'est le cas sur *Meetic* où les utilisateurs renseignent leur date de naissance avec une inclination pour les nombres « cinquièmes » tels que 1975, 1980, 1985 etc. Cet ajustement de l'âge, qui consiste le plus souvent

8. Daguët F. (2016), De plus en plus de couples dans lesquels l'homme est plus jeune que la femme, *Insee première*, 1613, 1-4.

9. Mignot J-F. (2010) L'écart d'âge entre conjoints, *Revue française de sociologie*, 51-2, 281-320.

à se rajeunir, se pratique par les deux sexes au même titre : le degré de surreprésentation aux chiffres « ronds » est sensiblement le même pour les femmes et les hommes¹⁰. Il est aussi relativement limité. Contrairement à une idée reçue, et comme le montre plusieurs recherches¹¹, les écarts à la vérité sont certes *courants* sur Internet mais ils sont peu *importants*. Sur *Meetic* par exemple, les utilisateurs sont curieusement plus grands et plus sveltes que la population dans son ensemble, mais les écarts à la moyenne nationale sont assez faibles : environ 2 cm en plus pour les deux sexes, de même que 2 kg en moins pour les hommes et 5 kg en moins pour les femmes. Les chiffres « ronds » ont donc un effet centrifuge sur les nombres proches sans trop distordre la moyenne, ce qui veut dire que les ajustements se situent souvent dans les bornes de +/- 4. Il y a de bonnes raisons de penser qu'il en va de même pour l'âge. Cette tendance à arrondir l'âge n'implique pas moins d'interpréter les résultats avec prudence. Plutôt que de s'intéresser à un âge donné, il s'agit d'établir les tendances d'ensemble quant à la direction et à la variation des préférences d'âge telles que manifestées sur Internet.

Ces données « massives » issues du site *Meetic* sont pour partie comparées à la dernière enquête en date sur la formation des couples en France. Cette étude, intitulée *Étude des parcours individuels et conjugaux (Épic)* a été conduite par l'Ined et l'Insee en 2013-2014, sous la coordination de Wilfried Rault et Arnaud Régnier-Loilier¹². Au total, 7 809 personnes âgées de 26 à 65 ans ont répondu à cette enquête qui comportait notamment des questions sur les préférences d'âge des répondants. Comme nous le verrons, ces préférences déclarées lors d'une enquête ne sont pas les mêmes que celles déclarées *in situ* sur le site *Meetic*.

Préférence d'âge : ce que l'on dit et ce que l'on fait

Pour comprendre la manière dont l'écart d'âge se produit, l'enquête *Épic* a cherché à connaître les attitudes des femmes et des hommes envers cette caractéristique courante des couples. Il a été demandé aux répondants s'ils auraient « accepté facilement l'idée d'être avec quelqu'un qui aurait été plus jeune que vous, de 5 ans ou plus » ou « plus âgé que vous, de 5 ans ou plus ». Reprise de l'enquête sur la *Formation des couples* (1983-1984), la question permet de savoir si l'un des deux sexes tient plus que l'autre à ce que l'homme soit plus âgé dans les couples.

Le résultat est sans équivoque : en 2013 comme 30 ans auparavant, ce sont en premier lieu les femmes qui se disent attachées à un écart d'âge en faveur du partenaire masculin. Tandis que les hommes sont très nombreux à dire qu'ils accepteraient facilement une femme *plus âgée* – c'est le cas de quatre hommes sur cinq (79 %) – seulement un peu plus de la moitié des femmes s'imaginent avec un partenaire *plus jeune* (53 %). Cette réticence féminine envers un partenaire cadet est particulièrement forte chez les jeunes : parmi les 26-30 ans, seulement un tiers des femmes accepteraient facilement cette idée (33 %). Chez les hommes, les attitudes varient peu : à tous les âges, ils sont très majoritairement ouverts au scénario aussi bien d'une conjointe plus jeune que plus âgée.

Sur *Meetic*, les choses sont assez différentes. Invités à renseigner leurs préférences d'âge sur le site (c'est une information obligatoire au vue de l'inscription), les utilisateurs indiquent une fourchette entre 18 et 99 ans. En comparant l'âge renseigné par les usagers avec leurs préférences quant aux âges *minimum* et *maximum*, il est possible de mesurer autrement l'attitude des femmes et des hommes envers l'âge du partenaire. Cette analyse tranche avec les résultats d'enquête. C'est ce que montre la figure 1 en comparant la part de répondants à l'enquête *Épic* ayant déclaré facilement accepter l'idée d'un écart d'âge inhabituel en faveur de la femme (de 5 ans ou plus) avec la part d'utilisateurs de *Meetic* ayant indiqué des préférences ouvertes à un tel écart.

10. Bergström M. (2015), L'âge et ses usages sexuels sur les sites de rencontres en France (années 2000), *Clio. Femmes, Genre, Histoire*, 2-42, 125-146.

11. Toma C., Hancock J. et Ellison N. (2008), Separating Fact From Fiction: An Examination of Deceptive Self-Presentation in Online Dating Profiles, *Personality and Social Psychology Bulletin*, 34-8, 1023-1036 ; Schmitz A., Sachse-Thürer S., Zillmann D. et Blossfeld H-P. (2011), Myths and facts about online mate choice. Contemporary beliefs and empirical findings, *Zeitschrift für Familienforschung*, 23-3, 358-381 ; Zillmann D., Schmitz A. et Blossfeld H-P. (2011), Lügner haben kurze Beine: zum Zusammenhang unwahrer Selbstdarstellung und partnerschaftlicher Chancen im Online-Dating, *Zeitschrift für Familienforschung*, 23-3, 291-318.

12. Rault W. et Régnier-Loilier A. (2015), La première vie en couple : évolutions récentes, *Population & Sociétés*, 521, 1-4.

Le décalage est apparent. Les utilisateurs de *Meetic* se montrent moins ouverts à un écart d'âge inhabituel que ne le font les répondants à l'enquête – c'est vrai pour les deux sexes. Or, ce décalage est bien plus important pour les hommes. Tandis que dans la situation d'enquête, les hommes se disaient – contrairement aux femmes – relativement indifférents à l'âge de leur partenaire, cette indifférence disparaît sur Internet. Une majorité d'utilisateurs sont certes ouverts aux femmes âgées jusqu'à 40 ans, mais après cet âge, c'est beaucoup moins le cas. Les déclarations des femmes sont, elles, plus cohérentes. Le taux d'acceptation d'un écart d'âge inhabituel est relativement similaire dans les deux situations : le coefficient de corrélation est de 0,73 contre seulement 0,24 pour les hommes.

Pourquoi ce décalage chez les hommes que l'on n'observe pas au même titre chez les femmes ? Une explication possible réside dans l'inégale réflexivité des deux sexes quant à leurs préférences amoureuses et sexuelles. Alors que les femmes sont socialisées et habituées depuis leur plus jeune âge à parler des relations affectives¹³, c'est moins le cas des hommes qui manifestent plus souvent une réticence à évoquer leur vie intime, notamment dans le cadre d'une enquête¹⁴. Peut-être est-ce d'abord cette socialisation différenciée aux discours sur l'intimité que captent les enquêtes. Cela voudrait dire que les réponses des femmes, plus tranchées, ne reflètent pas tant une attitude plus *intransigente* vis-à-vis de l'âge du partenaire mais avant tout des préférences plus *conscientes*. À l'inverse, l'indifférence manifestée par les hommes concernerait moins l'âge des partenaires que la *question* en tant que telle, à laquelle ils peinent peut-être à répondre. En tous les cas, *Meetic* indique que les hommes, plus que les femmes, « ne font pas toujours ce qu'ils disent qu'ils font » et interroge à ce titre les enquêtes sur la conjugalité et plus précisément la comparabilité des réponses données par les deux sexes.

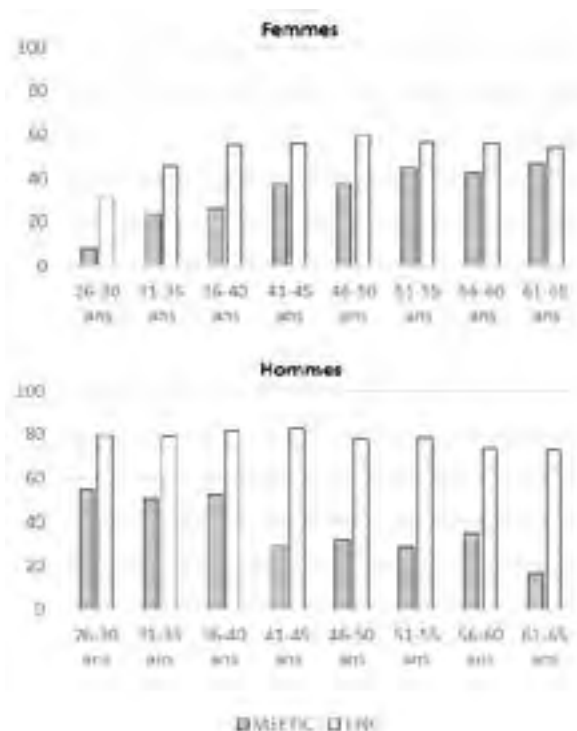


Figure 1. Part de personnes qui acceptent un écart d'âge (de 5 ans ou plus) en faveur de la femme – comparaison de réponses à l'enquête *Épic* et sur le site *Meetic* (%)¹⁵

13. Monnot C. (2009), *Petites filles d'aujourd'hui, l'apprentissage de la féminité*, Paris, Autrement.

14. Blin O. (1997), *Paroles d'amour, mots du cœur ou vues de l'esprit ?*, Mana. Revue de sociologie et d'anthropologie, 3, 57-72 ; Duret P. (1999), *Les jeunes et l'identité masculine*, Paris, Presses universitaires de France

15. **Lecture** : lors de l'enquête *Épic*, 83 % des hommes âgés de 41 à 45 ont affirmé qu'ils auraient « facilement accepté l'idée d'être avec quelqu'un qui aurait été plus âgé que [eux], de 5 ans ou plus ». Sur *Meetic*, 29 % des profils présentant des hommes de même âge ont affiché des préférences ouvertes à une partenaire plus âgée, de 5 ans ou plus.

Champ : compte d'utilisateurs actifs enregistrés sur Meetic en 2014 ; personnes vivant en France métropolitaine.

Source : base d'utilisateurs de *Meetic.fr*, Meetic Group, 2014 ; enquête *Épic*, Ined-Insee, 2013-2014.

Une histoire d'amour est une histoire de compromis

L'analyse des comportements de contact sur *Meetic* permet d'aller plus loin. Elle fait apparaître un inversement des attitudes envers l'écart d'âge avec l'âge propre des utilisateurs. Les jeunes femmes se montrent – sur *Meetic* comme ailleurs – très réticentes envers un partenaire plus jeune, mais avec l'âge les utilisatrices s'ouvrent manifestement à cette idée. Les femmes seniors se montrent même plus intéressées par des hommes plus jeunes que par des hommes plus âgés. Ainsi, l'écart d'âge moyen entre les utilisatrices et leurs interlocuteurs passe de 6,6 ans en moyenne en faveur de l'homme chez les 18-24 ans, à 3,9 ans en faveur de la femme chez les 60-70 ans.

Chez les hommes, les tendances suivent une même tendance : l'intérêt va croissant pour un partenaire plus jeune avec l'âge. C'est ainsi que les utilisateurs de 18-24 ans échangent souvent avec des femmes âgées – l'écart d'âge avec leurs interlocutrices est de 2,2 ans en moyenne en faveur de la femme – alors qu'à l'autre bout de l'échelle des âges, les hommes sollicitent des femmes bien plus jeunes – 7 ans en moyenne.

Ces observations indiquent que l'écart d'âge est le résultat commun de processus différents. En début de parcours affectif, ce sont surtout les femmes qui réclament cet écart, alors qu'aux âges qui correspondent aux deuxièmes unions, ce sont plutôt les hommes qui revendiquent une différence d'âge en leur faveur. Alors que la force statistique de l'écart d'âge conduit souvent à y voir une *préférence forte et symétrique des deux sexes*, il est sollicité tantôt par les femmes, tantôt par les hommes.

Plus précisément, cette caractéristique courante des couples apparaît comme un « arbitrage » entre les préférences féminines et masculines qui divergent souvent par ailleurs. La figure 2 confronte les critères d'âge déclarés par les utilisateurs des deux sexes dans leur profil avec la différence d'âge effectivement observée dans les contacts établis. Elle montre que les rencontres se font à mi-chemin des préférences affichées par les femmes et les hommes. Plutôt qu'une *préférence partagée*, l'écart d'âge apparaît comme le compromis entre deux aspirations accomodées.

Ces observations remettent en cause la notion « choix » du conjoint, chère aux sociologues du couple. La rencontre amoureuse est une interaction complexe où les deux sexes s'affrontent et se frottent aux désirs de l'autre. Les données de *Meetic* révèlent les concessions et les renoncements qui font aussi les rencontres mais au sujet desquelles les enquêtes restent silencieuses.

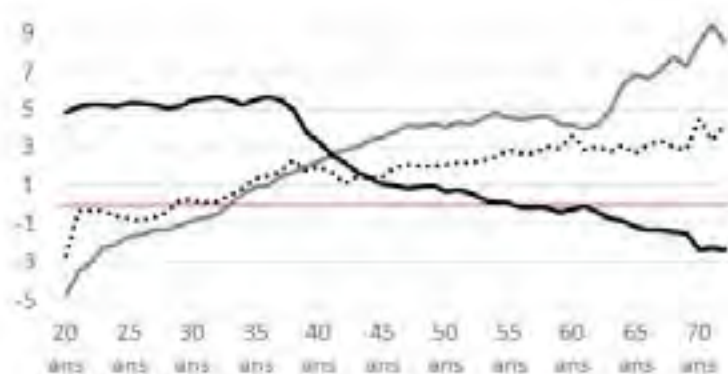


Figure 2. Écart d'âge moyen en faveur de l'homme tel que demandé dans les profils et observé dans les contacts établis¹⁶

16. **Lecture** : les profils présentant une femme de 55 ans affichent une préférence pour un écart d'âge moyen de 0,1 ans en faveur de la femme tandis que les profils présentant un homme de même âge affichent une préférence pour un écart moyen de 4,9 en faveur de l'homme. L'écart d'âge moyen des contacts établis par un expéditeur de 55 ans est de 2,7 ans.

Champ : compte d'utilisateurs actifs enregistrés sur *Meetic* en 2014 ; l'ensemble des premiers contacts ayant reçu une réponse sur *Meetic* en 2014.
Source : base d'utilisateurs de *Meetic.fr*, Meetic Group, 2014.

In fine, les sites de rencontres illustrent deux intérêts majeurs des données massives. D'une part, elles captent de façon avantageuses les *processus* qui, mal saisis par les enquêtes, sont pourtant centraux dans la réflexion en sciences sociales. D'autre part, elles offrent aux quantitativistes des possibilités nouvelles *d'observation* – une méthodologie jusque-là associée davantage au travail qualitatif. Ainsi, ces données ne sont pas seulement une source alternative mais proposent des *perspectives d'analyse statistique* nouvelles qui se soldent aussi par des résultats différents. Si les enquêtes conçues par les chercheurs restent inégalées de par leur richesse d'information, les données massives peuvent éclairer leurs zones d'ombres. Aujourd'hui, lorsque confrontés aux limites de leurs enquêtes, les chercheurs recourent aux hypothèses et aux « modèles d'acteurs ». La diffusion de nouvelles données nous incite à aborder autrement, et empiriquement, les problématiques à laquelle les questionnaires peinent à répondre. Cela implique de bricoler et d'accepter de travailler avec des données non représentatives. Les *Big Data* nous invitent en effet à élargir les critères de qualité des données quantitatives pour les juger autant au regard de leur pertinence et leur contenu informatif que de leur seul caractère représentatif.