

El estudio de viñeta para el análisis de las contrataciones laborales: una perspectiva crítica

Juan-Ignacio Martínez-Pastor
Irina Fernández-Lozano

UNED. Departamento de Sociología II
jimartinez@poli.uned.es; irina.fernandez@poli.uned.es



Recepción: 11-01-2022
Aceptación: 13-06-2022
Publicación: 20-10-2022

Cita recomendada: MARTÍNEZ-PASTOR, Juan-Ignacio y FERNÁNDEZ-LOZANO, Irina (2022). «El estudio de viñeta para el análisis de las contrataciones laborales: una perspectiva crítica». *Papers*, 107 (4), e3075. <<https://doi.org/10.5565/rev/papers.3075>>

Resumen

Un estudio de viñeta es, en esencia, una encuesta diseñada *ad hoc* consistente en evaluar escenarios ficticios por parte de las personas encuestadas, de tal manera que de dicha evaluación se puedan inferir opiniones o actitudes que en encuestas tradicionales no suelen salir a la luz. Sus ventajas principales son dos: en primer lugar, que, en virtud del diseño de la encuesta, se eliminan las correlaciones existentes entre variables que siempre se dan en poblaciones reales, lo que permite desentrañar mejor los efectos de esas variables; y, en segundo lugar, que las personas encuestadas responden indirectamente a las preguntas de investigación, por lo que esta técnica trata de eliminar sesgos de deseabilidad social. Tras revisar las características básicas de la técnica, el artículo plantea su utilidad desde un punto de vista crítico en relación con los siguientes aspectos: con los sesgos de deseabilidad social, con el hecho de plantear situaciones hipotéticas y no analizar comportamientos reales y, sobre todo, con su validez externa cuando se aplica a colectivos muy concretos (o hasta qué punto se pueden generalizar los resultados de este tipo de estudios).

Palabras clave: viñeta; estudios de viñeta; discriminación; género; validez interna; validez externa; encuestas factoriales; estudio factorial

Abstract. *The vignette study for the analysis of labour recruitment: a critical perspective*

A vignette study is, in essence, an *ad hoc* survey designed for respondents to evaluate fictitious scenarios, so that opinions or attitudes can be inferred that are not usually revealed in traditional surveys. Its main advantages are twofold: first, that, by virtue of the survey design, existing correlations between variables that always occur in real populations are eliminated and the effects of these variables can be better disentangled; and second, that respondents respond indirectly to the research questions, so this technique aims to eliminate social desirability biases. After reviewing the basic characteristics of the technique, the article considers its usefulness from a critical point of view in relation to the following aspects: social desirability biases; the fact of posing hypothetical situations and not analysing real behaviour; and, above all, external validity when applied to very specific groups (or the extent to which research results can be generalised).

Keywords: vignette; vignette studies; discrimination; gender; internal validity; external validity; factorial surveys; factorial study

Sumario

- | | |
|--|----------------------------|
| 1. Introducción | 5. Conclusiones |
| 2. ¿En qué consiste un estudio de viñeta? | Agradecimientos |
| 3. ¿Qué ventajas ofrecen los estudios de viñeta? | Financiación |
| 4. Los inconvenientes | Referencias bibliográficas |

1. Introducción

Los experimentos gozan de una tradición consolidada en algunas áreas de las ciencias sociales, como en la psicología o en la investigación de mercados. No sucede lo mismo en la sociología, aunque en los últimos años asistimos a una expansión de los diseños experimentales en nuestra disciplina. Dicha expansión se debe al grado de madurez alcanzado por ella, una madurez que ha conllevado una revisión crítica de la metodología aplicada hasta ahora y un afán por sobrepasar los límites marcados por los diseños, los métodos y las técnicas más habituales.

Tal revisión crítica se ha plasmado, por un lado, en que la frontera entre la investigación cualitativa y la cuantitativa se vaya difuminando en lo que respecta al diseño de la investigación y a la lógica inferencial (King et al., 2005); y, por otro, en el deseo de los investigadores de ir más allá del descubrimiento de correlaciones y de una descripción superficial de los hechos. En efecto, las personas que investigan en sociología, hartas de correlaciones y huérfanas de explicaciones, están poniendo el foco en el cómo y en el porqué suceden los hechos sociales, en establecer sus mecanismos causales. Es aquí donde los diseños experimentales, o cuasiexperimentales, adquieren relevancia (Wallander, 2009).

Este artículo se basa en la experiencia y en las reflexiones críticas de los autores tras haber aplicado una de las técnicas cuasiexperimentales que más expansión tienen en la última década: los estudios de viñeta (los llamados *vignette*).¹ El artículo se divide en cuatro partes. La primera ofrece una descripción sucinta de en qué consisten los estudios de viñeta; en la segunda se exponen las principales ventajas de dicha técnica; la tercera, más extensa, está dedicada a un análisis crítico de sus inconvenientes, con especial atención al problema de la validez externa debido a las muestras que suelen utilizarse; y, por último, incluimos un breve apartado de conclusiones. Es necesario subrayar que gran parte de la crítica aportada en este artículo —la validez externa de los estudios de viñeta— se centra en el análisis de las contrataciones laborales, por lo que no es una crítica generalizable a otros ámbitos temáticos en los que se ha aplicado esta técnica, como las actitudes con respecto al ingreso básico (Laenen et al., 2022), a las políticas redistributivas (Gniza et al., 2022) o a la prostitución (Escot et al., 2021), por citar solo tres ejemplos.

2. ¿En qué consiste un estudio de viñeta?

Un estudio de viñeta es, en esencia, una encuesta cuyo contenido nuclear es la evaluación de escenarios ficticios por parte de las personas encuestadas, de tal manera que de dicha evaluación se puedan inferir opiniones o actitudes. Así, la viñeta se distingue de una encuesta tradicional principalmente en su fase de diseño, y no tanto en la de análisis (aunque, como se verá más adelante, el propio diseño de la viñeta da pie a un análisis multinivel, lo que no necesariamente ocurre en otros tipos de encuesta). A menudo, dichos escenarios o situaciones ficticias presentan la particularidad de que se han construido usando todas las combinaciones posibles de una serie de dimensiones² (variables) cuyo impacto en opiniones, actitudes o incluso potenciales comportamientos se pretende analizar (Auspurg y Hinz, 2015). Dichas dimensiones son variables que los investigadores consideran socialmente sensibles con relación al tema de estudio y, por ello, sería muy problemático preguntar abiertamente por las mismas. Dos son las ventajas principales de un estudio de este tipo. En primer lugar, a través de una construcción ortogonal de los escenarios ficticios (a menudo, personas), se eliminan las correlaciones existentes entre variables en poblaciones reales, lo que permite aislar mejor sus efectos. En segundo lugar, las personas encuestadas, idealmente, responden de manera indirecta a las preguntas de investigación, lo que trata de eliminar los sesgos de deseabilidad social. Es decir, como en el resto de las técnicas experimentales o cuasiexperimentales, están

1. La investigación se publicó en una revista JCR del primer cuartil de sociología (Fernández-Lozano et al., 2020).
2. Estrictamente, en este caso hablaríamos de encuestas factoriales (*factorial surveys*), mientras que el término *viñeta* en sí hace referencia a la presentación de la situación ficticia. No obstante, en la literatura a menudo los términos *vignette* y *factorial surveys* se usan de manera intercambiable. En este artículo nos centramos en este tipo de estudios de viñeta basados en encuestas factoriales.

sometidas a un *tratamiento* del que no son conscientes. En el apartado 2.1. se muestran algunos ejemplos de preguntas indirectas. A continuación, se detallan e ilustran diferentes fases del diseño de un estudio de viñeta.

2.1. La selección del tema, hipótesis y dimensiones

Como en cualquier investigación social que se plantee, la selección del tema y las hipótesis ha de preceder a la selección de la técnica, por lo que antes de elegir la realización de un estudio de viñeta habría que plantearse su idoneidad para la pregunta de investigación frente a otras alternativas. La idoneidad de un diseño de viñeta es, al menos en la teoría, sencilla de evaluar: todas aquellas cuestiones sensibles que las personas no suelen responder abiertamente («¿contrataría a una mujer embarazada en su empresa?», «¿le importaría tener como vecino a una persona del colectivo X?», «¿a quién prefiere que vayan destinadas las ayudas sociales del municipio?») son propicias para este tipo de estudios. Evidentemente, muy pocas personas responderían abiertamente que *discriminan* a un determinado colectivo en una encuesta. Por el contrario, si no hay problemas de *deseabilidad social*, muy probablemente una viñeta (con la complejidad que lleva su diseño) no aportaría ningún valor frente a una pregunta directa acerca de la cuestión. Pero, además, la encuesta viñeta no es válida para indagar en hechos pasados específicos (aunque los mismos estén sujetos a deseabilidad social), sino en actitudes, opiniones y, en todo caso, comportamientos potenciales, ya que las viñetas representan escenarios ficticios creados específicamente para la investigación.

Como se verá más adelante, uno de los usos típicos del estudio de viñeta es la evaluación de candidaturas ficticias para un puesto de trabajo, aspecto en el que se centra gran parte de este artículo. A menudo, por tanto, esos escenarios ficticios toman la forma de currículums o perfiles más o menos complejos. Di Stasio analizó qué currículums eran considerados más fácilmente presentables a ojos de empleadores italianos (Di Stasio, 2014a). Karpinska et al. (2013), por su parte, evaluaron qué características deben tener los trabajadores de más edad para aumentar su probabilidad de seguir en la empresa en el contexto danés. Los sesgos de género o aquellos asociados particularmente a la maternidad en la contratación o la promoción laboral han sido objeto del estudio de varias investigaciones en Estados Unidos (Benard y Correll, 2010; Correll et al., 2007; Cuddy et al., 2004). No obstante, el objeto de estudio de una viñeta no se limita a la valoración de un CV por parte de empleadores potenciales o reales. Cualquier tema se presta al uso de una viñeta si lo que interesa es la identificación de prejuicios, estereotipos o discriminación. Kootstra (2016) analizó hasta qué punto los británicos y holandeses consideraban a individuos de diferente origen étnico y migrante «merecedores» de ayudas públicas. Jacobs y Gerson (2016), por su parte, estudiaron las opiniones de los estadounidenses en relación con el trabajo remunerado de las madres. La selección de las dimensiones o variables a estudiar será paralela al desarrollo de las hipótesis. Puesto que normalmente

los escenarios ficticios se refieren a personas, el origen étnico, el sexo, la orientación sexual, la condición de padre/madre o la edad son dimensiones típicas en estos estudios de viñeta.

2.2. *El diseño de las viñetas*

El diseño de las situaciones ficticias o viñetas conlleva en esencia tres fases —dos de carácter más técnico y una última más creativa. En primer lugar, se creará un universo de viñetas a partir de todas las combinaciones posibles de categorías (o «niveles») dentro de cada dimensión. Así, cuanto mayor sea el número de dimensiones a comprobar y el número de categorías dentro de estas dimensiones, mayor será el tamaño del universo de viñetas y, con ello, el tamaño de muestra necesario para poder analizar el efecto de dichas dimensiones evitando las correlaciones que se dan en la realidad.³ En suma, habremos creado una población ficticia donde no existe una correlación entre, por ejemplo, la condición de inmigrante y el nivel de ingresos, o el sexo y el nivel de responsabilidad dentro de una empresa, ya que todas las variables independientes principales responden a un diseño ortogonal o cartesiano. Así, en nuestro universo ficticio, mujeres y hombres tendrían la misma probabilidad de ser directivos o empleados subordinados, por lo que se podrá desentrañar el efecto de ambas variables sobre nuestra variable dependiente.

La segunda fase del diseño consiste en asignar dichas viñetas a diferentes modelos de cuestionario. Ello implica decidir a cuántas viñetas evaluará cada encuestado.⁴ Idealmente, si cada persona evalúa a un candidato, se evita el efecto fatiga y se esconden mejor las dimensiones a testar, ya que estas no se repiten. Sin embargo, este tipo de diseños es más costoso, ya que aumenta considerablemente el tamaño de la muestra necesaria para lograr la eficiencia estadística (quizás no por casualidad el número de viñetas evaluadas por cada encuestado es generalmente menor en los trabajos norteamericanos que en los europeos), y la mayoría de los estudios presentan entre 2 y 6 viñetas. Del universo completo de viñetas (que no es sino la combinación cartesiana de todas las categorías de todas las variables, como muestra la tabla 1), puede ser necesario extraer una muestra de perfiles a incluir en los cuestionarios. La gran mayoría de los estudios utilizan una muestra aleatoria, aunque algunos análisis han planteado la conveniencia de usar una muestra por cuotas (Dülmer, 2007). Distintos programas de análisis estadístico permiten obtener la muestra y distribuirla en modelos de cuestionario, de tal manera que se minimice la correlación entre variables y, con ello, se maximice la eficiencia estadística.⁵

3. Más allá de que, como se verá más adelante, un tamaño de muestra suficiente también es importante para lograr la validez externa de la investigación.
4. En ocasiones, se añaden viñetas «falsas» (comunes a todos los cuestionarios y que no formarán parte del análisis central) para, por ejemplo, evitar el efecto del orden u otros sesgos (como en Castilla y Bernard, 2010, que al añadir una tercera viñeta trataban de evitar que se descubriese que el género era una de las dimensiones analizadas).
5. Los investigadores suelen utilizar la macro de SAS '%mktblock'.

Por último, se diseñará el cuestionario, incluyendo una *cover story* (narrativa que sirve de «pretexto») que tratará de esconder el objetivo último de la investigación para no poner «en alerta» a la persona entrevistada. Cualquier elemento que se añada a la narrativa ha de ser neutral con relación a nuestro objetivo de investigación, bien común a todas las viñetas, bien controlado como una dimensión más por el investigador.⁶ Es necesaria una gran minuciosidad en el diseño de una viñeta, porque cualquier elemento de la narrativa o presentación de las viñetas puede introducir potencialmente una variable no controlada que correlacione con nuestra variable dependiente.⁷ Más allá de eso, hay múltiples formas de presentar la situación a evaluar.⁸ A menudo las dimensiones se presentan de manera indirecta. Por ejemplo, el origen étnico o el sexo se pueden presentar a través del nombre de la persona, como en Correll et al. (2007) o en Fernández-Lozano et al. (2020). El cuestionario incluirá, necesariamente, una o varias preguntas cerradas de evaluación de las viñetas, que serán nuestras variables dependientes.⁹ A menudo estas preguntas serán una manera indirecta de aproximarnos a nuestro objeto de estudio, de un modo análogo a como abordamos en nuestra vida diaria cuestiones delicadas (si preguntamos «¿qué tal con Pedro?» en lugar de «¿qué te parece la manera en que se dirige a nosotros?»). En ocasiones hay más de una variable dependiente, por ejemplo, al simular las etapas de un proceso de selección real (Wolf y Velden, 2001) o, más habitualmente, porque algunas de estas variables se consideran variables mediadoras que permiten entender los mecanismos mentales que llevan a una determinada actitud (Cuddy et al., 2004; Fernández-Lozano et al., 2020).

2.3. El análisis de los resultados

Una vez administrado el cuestionario,¹⁰ obtendremos una base de datos donde las unidades últimas de análisis serán las viñetas (los candidatos ficticios, a menudo, como se ha visto) agrupadas en *personas encuestadas*. Esta estructura multinivel permite distinguir los distintos niveles de varianza que habrá que tener en cuenta (entre encuestados y para un mismo encuestado). Como se verá más adelante, es importante contar con una muestra suficiente de personas encuestadas para garantizar la validez externa del experimento.¹¹

6. A menudo se introducen dimensiones de control no relacionadas directamente con las hipótesis de investigación con el objetivo de añadir complejidad a las viñetas presentadas o descartar correlaciones espurias.
7. Como, por ejemplo, el orden de presentación de las viñetas (véase el análisis de Auspurg y Jäckle, 2015).
8. Siempre con el doble objetivo, en ocasiones difícil de compatibilizar, de esconder el tratamiento, para evitar los problemas de deseabilidad social y la fatiga de la persona encuestada.
9. Además de variables de clasificación de la persona entrevistada u otras.
10. La administración de la viñeta no difiere de cualquier otro tipo de encuesta, si bien no hemos encontrado ninguna evidencia de administración de una viñeta mediante encuesta telefónica.
11. Para mayor detalle sobre el análisis multinivel de un estudio viñeta, véase el apéndice metodológico en Fernández-Lozano et al. (2020).

Tabla 1. Ejemplos de distinta forma de presentación de las viñetas en estudios de promoción / contratación / evaluación de personal

	Viñetas por cuestionario (reales + de «relleno»)	N.º total de diferentes perfiles*	Ratio encuestados / perfiles	Presentación de los perfiles
Wolf y Velden (2001)	20	$2 \times 3 \times 6 \times 2 \times 2 \times 3 \times 2 \times 4 \times 6 \times 2 \times 2 \times 2 \times 2 = 331.776$ (se seleccionó un universo de 138.240 perfiles realistas)	< 1	CV y cartas de presentación que variaban en <u>sexo</u> , <u>edad</u> , <u>tipo de estudios (2 dimensiones)</u> , <u>resultados académicos (2 dimensiones)</u> , <u>experiencia profesional</u> , <u>actividades extra-académicas (2 dimensiones)</u> , <u>4 rasgos de personalidad y apariencia (a través de una fotografía previamente testada)</u> .
Cuddy, Fiske y Glick (2004)	3 (1 + 2)	$2 \times 2 = 4$	30	« <u>Kate (Dan)</u> es una consultora asociada de 32 años que se graduó con un MBA. Lleva seis años trabajando en su campo actual. Cuando trabaja con un cliente, sus funciones incluyen la identificación de problemas, la planificación y realización de entrevistas y análisis, la síntesis de conclusiones en forma de recomendaciones y la ayuda a la implantación de cambios en las organizaciones de sus clientes. Sus aficiones incluyen la natación y el tenis. (<u>Kate y su marido [Dan y su mujer] han tenido recientemente su primer bebé.</u>) Ella (Él) vive en el centro de Nueva Jersey, se desplaza al trabajo dos días a la semana y trabaja a distancia tres días a la semana».
Castilla y Bernard (2010)	3 (2 + 1)	$2 \times 2 = 4$	29/ 57 /25 (3 muestras)	Formulario de evaluación del rendimiento y del desarrollo del personal, que incluía el <u>género</u> de la persona y autodescripción de la empresa como <u>meritocrática o no</u> . Se incluía también una evaluación cuantitativa y cualitativa del desempeño del empleado/a, equivalente en todos los perfiles.
Rudman y Mescher (2013)	1	$2 \times 4 \times 2 = 16$	23	Transcripción de una entrevista entre un empleado/a y un directivo/a de RRHH que variaba en el <u>tipo de permiso solicitado (para cuidado de menor / progenitor / ninguno)</u> , <u>origen racial (negro / blanco) y sexo</u> .
Di Stasio (2014a)	18	$2 \times 2 \times 2 \times 3 \times 3 \times 2 \times 3 \times 2 = 864$	< 1	Curriculum vitae que variaba en 8 dimensiones (<u>género</u> , <u>experiencia profesional previa</u> , <u>realización de beca en empresa</u> , <u>nivel educativo</u> , <u>campo de estudio</u> , <u>expediente académico</u> , <u>duración de los estudios</u> , <u>actividades extracurriculares</u>).

Notas: se subrayan las dimensiones.

* a partir de la combinación cartesiana de todas las categorías de todas las dimensiones.

Fuente: elaboración propia.

3. ¿Qué ventajas ofrecen los estudios de viñeta?

La ventaja más importante, compartida con otras técnicas experimentales, es la de manipular las condiciones a las que están expuestos los participantes en el experimento para comprobar las hipótesis planteadas. Además, al ser estudios diseñados *ad hoc*, se mide realmente lo que se pretende, sin utilizar aproximaciones ni remedios. Estas ventajas atañen a la validez de constructo y a la validez interna de una investigación. La validez de constructo se refiere «a la correspondencia entre una teoría y un diseño de la investigación que se propone comprobar esa teoría», mientras que la validez interna nos remite «a la verdad de una proposición respecto de la muestra elegida» (Gerring, 2014: 457).

La alta validez interna y de constructo posibilita establecer relaciones causales con mucha más seguridad que en los estudios observacionales. Así pues, a diferencia de esos estudios —más habituales en las ciencias sociales—, en los de viñetas el investigador manipula el conjunto de factores (equivalente a las variables en los análisis cuantitativos) y niveles (equivalentes a las categorías de esas variables) presentados a la persona que responde, de tal modo que las correlaciones entre los factores se minimizan a un valor cercano a cero. Tal y como se ha señalado, es posible así separar los efectos de las variables que en el mundo real están correlacionadas, al evitar el problema de la multicolinealidad y permitir una interpretación causal de las relaciones entre las variables (Di Stasio, 2014b: capítulo 4).

Los estudios de viñeta presentan también ventajas en comparación con otros experimentales más *arriesgados* que simulan situaciones reales (escondiendo de manera total la realización de una investigación), como los estudios *audit*,¹² que conllevan más implicaciones éticas. Las personas que participan en un estudio de viñeta saben en todo momento que están participando en una investigación social, aunque no se les suele revelar el verdadero propósito de esta. Por otro lado, el coste de un estudio de viñeta sería el mismo que el de cualquier encuesta *ad hoc*, porque, en esencia, se trata de una encuesta.

Otra de las ventajas del análisis de viñetas es que, en algunos temas para los que la sociología suele usar este método, como el análisis de la contratación o de la promoción de personal, esta técnica evita los problemas de autoselección de los candidatos. Al combinar aleatoriamente los niveles de las viñetas, cualquier candidato puede optar a un puesto, en cambio, en el mundo real, muchos candidatos directamente no se presentan a una oferta de empleo porque creen que no tienen ninguna posibilidad de ser elegidos para el puesto. La técnica de viñetas permite comprobar hasta qué punto pueden serlo o no (Di Stasio, 2014b: capítulo 4).

12. Los estudios *audit*, también llamados estudios de correspondencia, implican la simulación de candidaturas en un contexto real (normalmente mediante el envío de CV ficticios a empleadores). No se trata, por tanto, de un experimento basado en una encuesta. Véase, por ejemplo, Booth y Leigh (2010), Cortina et al. (2021) y Gaddis (2015).

Tabla 2. Principales ventajas e inconvenientes de los estudios de viñeta

Ventajas	Inconvenientes
Alta validez interna. Sirve para descubrir relaciones causales entre hechos sociales.	Se analizan situaciones hipotéticas y con candidatos ficticios.
Alta validez de constructo.	Eliminar todos los sesgos de deseabilidad social («esconder» el tratamiento) es mucho más sencillo que en una encuesta normal, aunque no existen garantías completas de que se logre.
Evita el problema de la autoselección o la autocensura de algunos candidatos para el tema que se analice (por ejemplo, la contratación de personal), que se da en el mundo real. Con las viñetas, puede saberse hasta qué punto las características autocensuradas tienen peso en los procesos de contratación o promoción.	Al igual que en otras encuestas <i>ad hoc</i> , lograr una muestra representativa suficiente que garantice la validez externa es muy costoso.
Plantean menos problemas éticos que otros diseños experimentales, como el <i>audit</i> .	No se está evaluando una situación real.

Fuente: elaboración propia.

4. Los inconvenientes

4.1. Situaciones hipotéticas y deseabilidad social

Las ventajas no esconden los inconvenientes de esta técnica. Aunque nos detendremos en los que se derivan de las muestras de estos estudios, cuestión que suele ser tratada muy de soslayo en la literatura, es preciso mencionar antes otros inconvenientes no menores. El primero es que se evalúan situaciones hipotéticas, no reales. Tomando como ejemplo los estudios sobre la contratación o la promoción de personal, los trabajadores propuestos son ficticios, como también lo es la propia contratación o la promoción. Esto supone una desventaja en comparación con otra técnica experimental como el *audit*, cuyos candidatos a un empleo también son ficticios, pero la persona encuestada lo desconoce y los empleos ofertados son reales. Al ser una técnica relativamente reciente en la sociología, existen pocos estudios que hayan analizado los efectos de presentar situaciones hipotéticas en lugar de reales. A este respecto, algunas investigaciones han demostrado que no hay diferencias significativas entre evaluar una situación u otra. Eso concluyeron Hainmueller et al. (2015) en un experimento sobre actitudes hacia la inmigración en Suiza; y Gutfleisch et al. (2021), en otro sobre contrataciones laborales en Luxemburgo.

Relacionado en parte con el anterior, otro de los mayores inconvenientes atañe a los posibles sesgos de deseabilidad social que pueden conllevar los estudios de viñeta (McDonald, 2019). Se ha argumentado que esta técnica es más apropiada que las encuestas precisamente para evitar dichos sesgos, al no preguntar directamente sobre un tema sensible y presentar una situación hipotética en la que se suele enmascarar el verdadero propósito del estudio (Sterkens

et al., 2020). Por eso los estudios de viñeta suelen utilizarse para el análisis de actitudes. No obstante, puede que los que respondan a las viñetas decidiendo contratar o promocionar a una mujer en lugar de un hombre o a una madre con hijos pequeños en lugar de alguien que no tiene hijos sean permeables a la corrección política, ya que, por mucho que el propósito del estudio no se explique y que no se advierta tan fácilmente como en una encuesta normal, no hay garantías de que el experto en recursos humanos no se percate de las diferencias clave entre las viñetas y que responda basándose en la deseabilidad social, en lugar de tal y como lo haría en una toma de decisiones real. Este mismo ejemplo puede extenderse a otros temas relacionados con el análisis de actitudes.

Kübler et al. (2018) aplicaron una solución a este posible sesgo en su estudio sobre la desigualdad de género a la hora de contratar a una persona. La solución consistía en que cada experto en recursos humanos evaluara un grupo de viñetas con candidatos del mismo género, en lugar de presentar candidatos de distinto género al mismo experto. En otra investigación muy imaginativa, Eifer y Petzold (2019) trataron de medir si el comportamiento real difiere de lo que se responde en las viñetas. Para ello, analizaron cómo se comportaron los conductores ante un coche que estaba delante de ellos parado en un semáforo en rojo y que seguía parado cuando el semáforo se ponía en verde. Después, aplicaron un estudio de viñetas a la misma población. El análisis comprobó que, en general, cuando se responde a las viñetas, la gente dice que se comportaría de una manera menos agresiva en comparación con cómo se comportaron en la situación real. En definitiva, todavía no existe evidencia empírica suficiente para dirimir los efectos de plantear situaciones hipotéticas y no reales.

4.2. El problema de la validez externa: ¿hasta qué punto pueden generalizarse los resultados de un estudio de viñeta?

Más allá de los problemas relacionados con la deseabilidad social y con el hecho de que las viñetas sean situaciones inventadas y no reales, existe otro problema que pasa muy desapercibido y que no se suele cuestionar en los estudios de viñetas aplicados a colectivos muy concretos, como son los responsables de la contratación y de la promoción de personal: la validez externa de estos experimentos derivada de las características del muestreo. ¿Hasta qué punto los descubrimientos de un estudio de viñetas sobre la contratación pueden generalizarse más allá de la muestra utilizada? ¿En qué medida las conclusiones pueden ser aplicadas a lo largo de diferentes poblaciones y situaciones?¹³

13. El metodólogo Jonh Gerring (2014: 106) advierte de que la distinción entre la validez interna y la externa es crucial en cualquier análisis, aunque no siempre es nítida: «Consideremos por ejemplo un estudio hipotético sobre un distrito escolar del estado de Nueva York que se basa en una muestra de estudiantes extraída en ese distrito y pretende identificar rasgos generales de todas las escuelas del Estado. Esto presenta tres niveles de validez potencial: uno, la muestra de estudiantes; dos, el distrito escolar; y tres, las escuelas del Estado (de varios distritos). La validez interna puede referirse a uno o a dos, mientras que la validez externa puede referirse a dos o a tres». Gerring advierte de que la pretensión de validez parece menos probable a medida que

Algunos investigadores han señalado que las características de los estudios de viñeta que pueden restringir su validez externa son la elección de los factores y de los niveles de las viñetas, su modo de presentación y lo que se pregunta (Karren y Barringer, 2002; Steiner et al., 2016). En realidad, se trata de grietas que pueden darse en cualquier estudio basado de encuestas. Es bien sabido que la forma en la que se pregunte, por lo que se pregunte, el orden en el que se pregunte y las categorías de las respuestas en las cuestiones cerradas pueden influir en lo que responda el entrevistado.

Las siguientes líneas remarcan la dificultad de los estudios de viñetas cuando se basan en encuestas *ad hoc* a colectivos muy específicos. Estas dificultades se concretan en recopilar datos mediante un procedimiento de muestreo igual de riguroso que en las encuestas habituales utilizadas en el mundo académico, lo que condiciona mucho, a nuestro entender, la validez externa de esta técnica en algunos análisis, como los centrados en la contratación o en la promoción de personal. Nuestro argumento se basa en las siguientes premisas, bien conocidas por las personas expertas en encuestas sociológicas y que siempre se deben considerar al analizar una encuesta. La primera indica que una encuesta no nos dice exactamente lo que piensa la población objeto de estudio, sino que nos lo dice *aproximadamente*, con un margen de error. Además, y esto es lo principal, nos lo dice *aproximadamente siempre y cuando* cumpla una condición imprescindible: que sea *representativa*. Para ello, si la población objeto de estudio está segmentada en grupos, la muestra debe tener el mismo porcentaje de cada grupo que nos podamos imaginar del que tiene esa población (Martínez Pastor, 2019). Aunque el puro azar, a través de un muestreo aleatorio simple, es lo que más garantiza que una muestra sea representativa, existen otros tipos de muestreos apropiados cuando no se dispone de un listado completo de la población que se quiere estudiar.¹⁴

Pese al avance de los estudios de viñetas, algunos inconvenientes condicionan el grado en el que podemos estar seguros de generalizar sus resultados. Estos inconvenientes no vienen dados por la técnica en sí —ya que lo que define a la viñeta, como se ha visto, es la forma en que se diseña la encuesta y no la manera de seleccionar la muestra—, sino por las limitaciones prácticas que las viñetas encuentran al ser aplicadas a muestras verdaderamente representativas cuando van dirigidas a colectivos muy concretos a los que es difícil localizar y convencer para que colaboren en la realización de la encuesta. Aunque esta técnica se aplica a muchos análisis dirigidos a la población general, no es menos cierto que también suele destinarse a colectivos muy reducidos como las personas encargadas de la contratación de personal, tema objeto de este artículo. En este caso, cabe señalar los siguientes aspectos en lo referente a la validez externa de este tipo de estudios de viñetas:

avanzan los círculos, pero ninguna es por completo plausible, y de que, desde esa perspectiva, no hay una posible demarcación entre la validez interna y la externa.

14. Para un análisis de la importancia de la representatividad de las muestras, véase Martínez Pastor, 2019: capítulo 7.

1. Muchas de las muestras no son realmente aleatorias.
2. En el caso de aplicarse un muestreo aleatorio, no se tiene la seguridad de si el listado de donde se extrae la muestra es exhaustivo, es decir, si contiene a todo el colectivo que se quiere analizar.
3. Dada la especificidad de los colectivos que se suelen analizar, es muy difícil, casi imposible, estratificar la muestra o establecer cuotas en función de las características individuales de los entrevistados.
4. Problema del *n* pequeño. La mayoría de los estudios de viñetas dirigidos a colectivos muy específicos se basan en muy pocos participantes, en buena medida por la limitación de recursos.
5. La tasa de respuesta suele ser muy baja.

Para avalar estas ideas, nos serviremos de las investigaciones recopiladas por McDonald (2019), quien buscó a partir de bases de datos académicas —Google Scholar y JSTOR— estudios sociológicos sobre contrataciones laborales que hayan usado la técnica de las viñetas con muestras de empleadores o de profesionales de recursos humanos. La lista se compone de 16 estudios, a los que hemos añadido el de Fernández-Lozano et al. (2020). La siguiente tabla, elaborada por nosotros a partir de la lista recopilada por McDonald, recoge dichas investigaciones y expone las características de las muestras desglosadas en cinco apartados: a) a quién se dirige la muestra; b) el tipo de muestreo uti-

Tabla 3. Características muestrales de artículos con estudios de viñeta

	Muestra dirigida a	Tipo de muestreo	Tamaño de la muestra	Observaciones (viñetas totales evaluadas)	Tasa de respuesta (muestra finalmente usada)
Beek et al. (1997)	Mánager generales o personal con experiencia en selección de personal (salvo un 4 % sin experiencia en selección de personal). Holanda.	Muestra representativa de empresas holandesas a partir de un listado. Las empresas grandes están sobrerrepresentadas. No dicen si el muestreo fue aleatorio o no.	312	5.616	41 %. Se contactó con 758 y respondieron 475. Muestra finalmente usada: 312.
Wolf y Velden (2001)	Directores de departamentos de personal que contrataron a graduados de la Facultad de Ciencias Sociales de la Universidad de Utrecht.	Se deduce que es un muestreo aleatorio estratificado a partir de un listado de empresarios que tiene la universidad. La estratificación se hace por tamaño de la empresa, región y organizaciones con ánimo y sin ánimo de lucro. No queda claro cómo se ha creado el listado.	27	620	65 %
Biesma et al. (2007)	Empleadores de recién graduados de una organización de salud pública en Holanda.	Se deduce que es un muestreo aleatorio a partir de una lista de empleados potenciales de graduados en salud pública. No queda claro de dónde sacan el listado.	168	315	18 %. Se contactó por correo con 1.040 empleadores; respondieron 185 y finalmente usaron la información de 168.
Henkens et al. (2009)	2 muestras: 1) Los mánager de 4 grandes empresas multinacionales holandesas. 2) Estudiantes de escuelas de negocios.	Se seleccionaron —no se dice el procedimiento— a 26 mánager de 4 grandes empresas multinacionales holandesas y a 25 estudiantes de escuelas de negocios.	26 mánager y 25 estudiantes	611 y 610 viñetas en cada estudio	—

Tabla 3. Características muestrales de artículos con estudios de viñeta (continuación)

	Muestra dirigida a	Tipo de muestreo	Tamaño de la muestra	Observaciones (viñetas totales evaluadas)	Tasa de respuesta (muestra finalmente usada)
Karpinska et al. (2011)	2 muestras: 1) Mánager asistentes a un grupo de discusión sobre la vuelta al empleo de trabajadores mayores, y otros captados a través del método de bola de nieve. 2) Estudiantes de empresariales de la Universidad de Utrecht.	1) Muestra de mánager: bola de nieve a partir de los asistentes a un grupo de discusión. 2) Encuesta a estudiantes. No se especifica el tipo de muestreo.	20 mánager y 17 estudiantes	240 y 204 para cada estudio	–
Karpinska et al. (2013)	Supervisores incluidos en la <i>Longitudinal Internet Studies for the Social Sciences</i> de la Universidad de Tilburg.	Contactaron con los mánager de la encuesta mencionada.	238	1.190	No se sabe exactamente. Dicen que contactaron con 700 supervisores, que respondieron 515, pero que luego hicieron otro filtro para diferenciar a los mánager. En total participaron 238 mánager, pero no se sabe exactamente el dato sobre el que calcular la tasa de respuesta.
Di Stasio (2014a)	Reclutadores de las oficinas de recursos humanos de un sector de actividad concreto: código NACE=62 (<i>computer programming, consultancy, and related activities</i>). Italia.	Muestreo aleatorio. Se contactó con 148 empresas pertenecientes a ese sector registradas en la Cámara de Comercio.	59	1.062	40 %. Muestra una tabla de la representatividad de la muestra con respecto a la muestra diseñada y con respecto al sector completo en función del tamaño de la empresa y de la localización geográfica (norte, centro y sur).
Mulders (2014)	Los mánager en la <i>Longitudinal Internet Studies for the Social Sciences</i> de la Universidad de Tilburg.	Contactaron con los mánager de la encuesta mencionada.	443	2.215	No se sabe exactamente. Dicen que contactaron con 864 y que respondieron 698 (81 %), pero luego filtraron a los participantes en función de varias preguntas para asegurarse de que esos mánager habían estado involucrados en procesos de reclutamiento durante los últimos 10 años. Finalmente usaron datos de 443 individuos.
Humburg y Velden (2014)	Personas implicadas en procesos de reclutamiento en los últimos 5 años en alguno de siete sectores de actividad elegidos por los investigadores. Lo hacen para 9 países europeos.	Invitan a participar a todos los que cumplen las características anteriores a partir de un panel de consumidores registrados en una empresa internacional de investigación de mercados.	903	9.030	20 %, aunque no es una tasa de respuesta en el sentido estricto, puesto que, cuando se alcanzaban 100 respuestas en un país, dejaban de administrar la encuesta.
Di Stasio y Gërkhani (2015)	Profesionales en recursos humanos en el sector de ICT (información, comunicación y tecnología). Inglaterra.	Muestreo aleatorio a partir de empresas afiliadas a dos asociaciones de ese sector (una de pequeñas y medianas empresas, y otra que denominan de la «industria tecnológica»).	22	220	20 %. Contactaron con 172 empresas y participaron 34, pero a lo largo del artículo y en las tablas se señala que participaron en la encuesta 22. Ofrecen una tabla comparando el tamaño de las empresas con las que contactaron y el tamaño de las empresas que respondieron. Solo disponen de información para ofrecer esa comparación.
Karpinska et al. (2015)	Los mánager en la <i>Longitudinal Internet Studies for the Social Sciences</i> de la Universidad de Tilburg.	Contactaron con los mánager de la encuesta mencionada.	153	765	No se sabe exactamente. Dicen que contactaron con 700 supervisores y que respondieron 515 (74 %). Tras aplicar filtros para asegurarse de que eran mánager y no simplemente supervisores, se quedaron con 238, de los cuales participaron 153.

Tabla 3. Características muestrales de artículos con estudios de viñeta (*continuación*)

	Muestra dirigida a	Tipo de muestreo	Tamaño de la muestra	Observaciones (viñetas totales evaluadas)	Tasa de respuesta (muestra finalmente usada)
Damelang y Abraham (2016)	CEO o empresarios de la base de datos CORIS (<i>Cluster-Oriented Regional Information System</i>), Alemania.	Invitaron a todas las empresas de la base de datos mencionada.	187	997	11 %. Respondieron 187 de 1.474 empresas contactadas (12 %). Al final usaron los datos de 168 personas.
Di Stasio y Werfhorst (2016)	Reclutadores y profesionales de recursos humanos en empresas IT de Holanda e Inglaterra.	Inglaterra: Muestreo aleatorio a partir de empresas afiliadas a dos asociaciones de ese sector (una de pequeñas y medianas empresas, y otra que denominan de la «industria tecnológica») Holanda: Muestreo aleatorio. Se contactó con 148 empresas pertenecientes a ese sector registradas en la Cámara de Comercio.	38 en Holanda y 34 en Inglaterra	684 y 612 respectivamente	42 %. Contactaron con un total de 172 empresas. Respondieron 72. Dicen que hay un material suplementario en la revista para analizar la representatividad de la muestra, pero no hemos podido acceder a él con el enlace que señalan.
Liechti et al. (2017)	Los <i>mánager</i> de hoteles. Suiza.	Fueron invitados todos los miembros de la principal asociación empresarial de hoteles suizos.	Entre 243 y 232 dependiendo del análisis	Entre 967 y 931 en función del análisis	12 %. Invitaron a 1.982 <i>mánager</i> y participaron 238, aunque en función del análisis en las tablas se señala que participaron entre 243 y 232.
Oesch et al. (2017)	Los <i>mánager</i> de recursos humanos. Suiza.	Invitaron a 4.687 <i>mánager</i> de recursos humanos de la asociación más grande de gestión de recursos humanos de Suiza.	385	1.644	8 %. Invitaron a 4.687, respondieron 714 (15 %). Finalmente utilizaron la información de 385 reclutadores, tras filtrar por la falta de respuesta y por el hecho de que esos reclutadores estuvieran en activo.
Shi et al. (2018)	Reclutadores de vacantes de empleos reales anunciados en portales webs de la Suiza germano-parlante. Seleccionan ocupaciones relacionadas con mecánica, bancos y seguros, gastronomía, enfermería y con la tecnología de la información y la comunicación para incluir empleos de baja, media y alta cualificación.	Se deduce que invitaron a todos los reclutadores de las vacantes reales especificadas en la columna de la izquierda.	Entre 638 y 634 en función del análisis, aunque en una nota señalan que respondieron 580	Entre 6.338 y 6.298	27 %. Invitaron a 2.118 reclutadores y empezaron la encuesta 739; la acabaron 580.
Fernández-Lozano et al. (2020)	Supervisores de empresas de al menos 30 empleados en tres sectores en función del grado de feminización de la fuerza laboral: programación, consultoría y otras actividades relacionadas con la informática; comercio al por menor, y publicidad e investigación de mercado.	Muestreo aleatorio a partir de una base de datos de empresas adquirida a una empresa comercial.	71	426	14 %

Fuente: elaboración propia a partir del listado de publicaciones recopilado por McDonald (2019), añadiendo la investigación de Fernández-Lozano et al. (2020).

lizado; c) el tamaño de la muestra; d) las observaciones analizadas (o el número total de viñetas evaluadas por los participantes), y e) la tasa de respuesta.¹⁵ La elección de McDonald de centrarse en las investigaciones sobre contrataciones no es casual, puesto que, en los últimos años, los expertos en este tema han solido utilizar los estudios de viñeta.

Del análisis de las características de las muestras de esos 17 estudios, centrados en la contratación de personal, se pueden extraer varias conclusiones:

1. Holanda está sobrerrepresentada. Ocho de las investigaciones se han realizado con datos exclusivamente holandeses.
 2. A la hora de redactar los artículos, los autores se detienen mucho más en detallar el diseño de las viñetas, algo sin duda fundamental, que en explicar las características de la muestra. En algunos casos, el tipo de muestreo se tiene que deducir. En el caso de los muestreos aleatorios, no suele quedar claro qué tipo de listado utilizan para seleccionar a las empresas y, sobre todo, tampoco si son listados exhaustivos o no del colectivo estudiado.
 3. Diez investigaciones utilizaron listados de empresas o de profesionales pertenecientes a asociaciones. De ellas, se deduce que seis emplearon un muestreo aleatorio a partir de esos listados y que otras cuatro intentaron contactar con todas las empresas de los listados para que un *mánager* o profesional de los recursos humanos participara en el experimento.
 4. Otras tres investigaciones manejaron encuestas ya realizadas a la población general para contactar con todos los *mánager* o los profesionales de los recursos humanos que participaron en esas encuestas.
 5. El resto de las investigaciones se basaron en
 - a) Un muestro con el método de bola de nieve a partir de un grupo de discusión y otra muestra de estudiantes.
 - b) Un panel de consumidores.
 - c) Vacantes de empleos reales a partir de portales webs.
 - d) La información aportada por otra investigación no permite saber cómo se ha obtenido la muestra.
 6. La muestra más numerosa alcanza un tamaño de 903 individuos, mientras que la más pequeña solo llega a 22. La media del tamaño muestral de los 17 estudios es de 235 individuos.
 7. En algunos estudios no es posible saber exactamente cuál fue la tasa de respuesta. En los que está claro, las tasas oscilan entre el 8 % y el 65 %. La tasa de respuesta media (calculada sobre los casos finalmente utilizados en los análisis) es del 27 %.
15. McDonald (2019: apéndice 2) recoge en otra tabla el tamaño de la muestra. Las cifras con nuestra tabla no coinciden, puesto que nosotros hemos contabilizado los casos que realmente fueron incluidos en los análisis, y no los participantes que respondieron a la encuesta pero que luego no pasaron los filtros de los investigadores y no formaron parte de los análisis finales publicados.

De los puntos anteriores, destacan los siguientes problemas. En los muestreos aleatorios no se suele saber si el listado de las empresas con las que se contacta es exhaustivo o no. Lo mismo sucede con las asociaciones a las que los *mánager* o los profesionales de los recursos humanos están afiliados. Si las listas de las que parten las muestras no son exhaustivas, cabe preguntarse si existe algún tipo de sesgo en las características de las empresas o de los reclutadores de personal que forman parte de dichas listas. Si ese sesgo afecta a las preferencias de contratación, el análisis estará necesariamente sesgado.

A eso se añade la imposible tarea de estratificar la muestra en función de las características de los *mánager*, de los profesionales de recursos humanos o de los encargados de seleccionar o contratar al personal. Esto es así puesto que, a diferencia de poblaciones generales, cuyas características básicas pueden saberse a partir de los censos o de los padrones de población, no suelen existir registros exhaustivos con las características de todos los individuos pertenecientes a una ocupación determinada.¹⁶ Algunos de los estudios analizados estratifican la muestra en función del tamaño de la empresa, de la región o de si las empresas tienen o no ánimo de lucro, pero, si a la hora de contratar a alguien entran en juego los prejuicios, es verosímil que estos no solo tengan que ver con las características de la empresa, sino también con las características de los que seleccionan al personal. Al no saber realmente las características básicas de ese grupo de profesionales, no se puede estratificar por esos rasgos individuales y, de hecho, no se hace.

Otra cuestión que limita la validez externa de los estudios de viñeta en el campo de la contratación del personal es el tamaño de la muestra, que suele ser muy limitado. No es extraño, puesto que los colectivos a los que se dirigen estos estudios son muy concretos y especiales, como los *mánager*, por lo que no suele ser fácil contactar con ellos y, si se consigue, tampoco es sencillo que acepten participar en este tipo de experimentos. Como se ha señalado, el tamaño medio de la muestra de los 17 estudios analizados es de 235 individuos. Una vez recopilada la información, los investigadores analizan los datos mediante un análisis multinivel y a veces mediante otras técnicas de regresión, con casos suficientes para llevarlos a cabo e incluso para hacer interacciones entre variables. Es necesario remarcar, como ya se indicó, que dada la estructura multinivel de la base de datos resultante, y a diferencia de la mayoría de los análisis basados en encuestas, los casos en esta técnica de viñetas no son los individuos que evalúan las viñetas, sino el número de viñetas evaluadas. De esta manera, los casos para analizar aumentan notablemente, lo que permite llevar a cabo los análisis con una cierta robustez estadística. En los 17 estudios citados, el número medio de observaciones se eleva hasta las 1.938, pero desde el punto de vista del muestreo lo relevante no es el número de viñetas evaluadas, sino la representatividad de los individuos que las han evaluado. En este sentido,

16. En algunas ocupaciones muy numerosas, podría hacerse una aproximación de las características de los ocupados a partir de encuestas con muestras muy grandes, como la EPA en España.

7 de los 17 análisis aquí estudiados no llegan a los 100 participantes, y otros 3 superan los 100 pero no llegan a los 200. Aunque el colectivo analizado sea muy concreto, se trata de números muy alejados de las encuestas que suele utilizar la sociología académica para llegar a conclusiones fiables.¹⁷

Un último problema con el muestreo en los estudios de viñeta aplicados a colectivos muy concretos es la baja tasa de respuesta. Los académicos solemos utilizar encuestas con muestreos probabilísticos, generalmente realizadas por instituciones públicas debido a su elevado coste y al acceso que tienen estas instituciones a los listados de la población sobre los que diseñar la muestra. Una vez diseñada, es muy importante lograr una tasa de respuesta alta, ya que de lo contrario puede darse un sesgo de autoselección de los entrevistados. En España es obligatorio responder a todas las encuestas incluidas en el Plan Estadístico Nacional, que son todas las del INE y, para la sociología, algunas del CIS. La importancia de lograr altas tasas de respuesta es tal que las instituciones que realizan estas encuestas tratan de contactar varias veces con los individuos seleccionados antes de darlos por perdidos. En ocasiones, se emplean sustitutos con perfiles sociodemográficos idénticos a los de los individuos con los que finalmente no se puede contactar tras varios intentos.

Para hacernos una idea, la tasa de respuesta en España de una encuesta con un muestreo probabilístico sin sustitución, como es *La encuesta social europea*, ha sido de casi el 70 % desde la tercera hasta la octava ola. En las dos primeras, se situó entre el 50 % y el 60 %; en la octava, la tasa de respuesta media para los 23 países que participaron fue del 55 % (CIS, 2019). Como se ha mencionado en las características de las muestras, de los 17 estudios de viñeta analizados, la tasa de respuesta en los muestreos aleatorios es mucho menor y, tal y como se aprecia en la tabla 3, las muestras con tasas de respuestas altas tienen un número de casos muy pequeño. En efecto, las encuestas aleatorias que lograron las tasas más altas, por encima del 40 %, entrevistaron a pocos individuos; la que más, a 72.

Es sabido que la tasa de respuesta en los muestreos no probabilísticos, como el de cuotas, suele ser aún más baja. Este asunto preocupa a las empresas que recogen datos. El Pew Research Center (PRC), una de las instituciones estadounidenses más prestigiosas que realizan encuestas, analizó con detenimiento este asunto, al constatar un drástico descenso en la tasa de respuesta de las encuestas telefónicas del 36 % al 6 % entre 1997 y 2018 (Kennedy y Hartig, 2019). El PRC concluyó que los sesgos que produce la falta de respuesta pueden minimizarse a través de cuotas y de ponderaciones convenientes. De hecho, suelen ponderar sus encuestas por bastantes variables, como el género,

17. Otra cuestión es el abuso que se produce en las encuestas cuando se manejan submuestras de determinados colectivos y que contienen errores muestrales mayores que los incluidos en las fichas técnicas, ya que esas fichas se basan en el tamaño muestral total y no en los subconjuntos de la muestra. Es cierto que, en otros campos de investigación relacionados con la medicina, la psicología o la neurociencia, los experimentos suelen hacerse con un grupo muy reducido de individuos y se dan por válidos. No sucede lo mismo con la sociología convencional que estudia poblaciones generales.

la edad, la región, el tamaño del municipio, el nivel de estudios, el país de origen y la etnia.

En las encuestas llevadas a cabo a la población general, es posible ponderar los casos en función de estas variables si se saben las características de la población, pero, como se ha advertido, no puede hacerse en los estudios de viñetas aplicados a colectivos muy concretos, puesto que no se suelen conocer sus características. Incluso ponderando por todas esas variables, se sabe que, al menos en Estados Unidos, los que responden a encuestas son los que más participan en la vida pública de su comunidad, lo que puede acarrear sesgos con respecto al objeto de estudio.

En definitiva, los estudios de viñeta llevados a cabo para colectivos específicos plantean el problema de la posibilidad de extrapolar los resultados, incluso entre los miembros de dicho colectivo. Se trata de un problema de cualquier investigación basada en encuesta aplicada a un colectivo muy particular, por lo que esta crítica a la validez externa de los estudios de viñeta no es aplicable si esta técnica va destinada a poblaciones generales con un procedimiento de muestreo habitual.

4.3. Posibles soluciones o resignaciones

¿Tienen estos problemas, que pueden afectar a la representatividad de la muestra, alguna solución? Una de ellas, ya mencionada, sería estratificar las muestras o incluso ponderarlas a partir de las características del grupo que participa en el experimento. Esta solución requiere saber dichas características, lo cual no es fácil.¹⁸ Además, si se dispusiera de tales características, se podrían cotejar con la muestra finalmente obtenida en un ejercicio de transparencia. Algo parecido a dicha comparación hacen Di Stasio (2014a) y Di Stasio y Gërkhani (2015), pero no comparando las características de los participantes en las viñetas, sino las características de las empresas donde trabajan.

Otra solución ideal sería incrustar las viñetas en una encuesta normal representativa del colectivo que se quiere analizar. Esto tampoco suele ser fácil en colectivos muy específicos, como el de los *mánager* o los profesionales de recursos humanos, por la ausencia de encuestas representativas dirigidas a ellos, aunque como veremos hay excepciones. A falta de dichas encuestas específicas, lo que suelen hacer los investigadores, tal y como se ha indicado en la tabla 3, es aprovechar la muestra de una encuesta tradicional para contactar con el grupo que se quiere investigar (Karpinska et al., 2013; Mulders et al.,

18. Por ejemplo, en las investigaciones sobre el comportamiento de los profesionales de recursos humanos, una opción sería acudir a la EPA. No obstante, hasta ahora el INE ofrece la EPA con las ocupaciones detalladas a 3 dígitos, no a 4. Esto supone un inconveniente, puesto que los profesionales de los recursos humanos tienen el código 1212 en la Clasificación Nacional de Ocupaciones de 2011, la vigente en la fecha de publicación de este artículo. Así, en los microdatos disponibles para los investigadores, estos profesionales quedan subsumidos en la categoría 121 «Directores de departamentos administrativos». En la EPA del cuarto trimestre de 2020, había 319 casos en dicho código.

2014; Karpinska et al., 2015). El problema de utilizar un subconjunto de una muestra es el limitado número de casos de una ocupación determinada que hay en la muestra original, obstáculo que se agudiza porque el experimento no está incrustado en la propia encuesta, sino que los sujetos de esa submuestra deben responder a la invitación y querer participar posteriormente en el estudio de viñetas. Así pues, la muestra final suele ser todavía más reducida, de modo que se complica el problema del n pequeño y del posible sesgo de selección.

El análisis de Steiner et al. (2016) constituye un ejemplo de un estudio de viñetas que sortea el problema del n pequeño, al diseñar un muestreo idéntico a los más habituales en sociología (un muestreo por cuotas de sexo y edad) y al aplicar las viñetas a todos los entrevistados, que fueron un total de 980. El estudio trata sobre las actitudes hacia la brecha salarial de género. Los investigadores pudieron utilizar esas cuotas con garantías porque el universo de su estudio no era un colectivo muy específico, sino la población vienesa de entre 18 y 65 años, cuya estructura por sexo y edad puede conocerse con facilidad.

La muestra utilizada por Kübler et al. (2018) representa una excepción entre los estudios de viñetas dirigidos a las personas encargadas de contratar personal. Incrustaron las viñetas en una ola de una encuesta anual de panel realizada por el Instituto Federal de Formación Profesional (BIBB) de Alemania. Esa encuesta la responden anualmente los propietarios de las empresas o los directivos de recursos humanos. Tal y como detallan en su investigación, la muestra se obtiene a partir de una base de datos de la Agencia Federal de Empleo, y contiene todas las empresas con al menos un empleado que cotiza a la seguridad social. Los investigadores no pasaron las viñetas a todas las empresas de la muestra (3.450), sino que seleccionaron aleatoriamente a 680, con una tasa de respuesta del 98 %, que fueron las que finalmente participaron en el experimento. De este modo, existe una mayor seguridad de que sus hallazgos son válidos para el conjunto de las empresas alemanas, dado el diseño de la muestra original y la alta tasa de respuesta sobre la muestra diseñada aleatoriamente a partir de la original. La ventaja de esa investigación consiste en incrustar el estudio de viñetas en una encuesta realizada a una muestra representativa de todas las empresas alemanas, que además es respondida por los propietarios o los encargados de la contratación, colectivos muy específicos e ideales para los propósitos del estudio.

Pueden citarse otros remedios más clásicos para los problemas de la validez externa de los estudios de viñeta, soluciones que no son exclusivas para esta técnica. Una de ellas consiste en replicar los estudios con otras muestras, en el mismo o en otros países y, si tomamos como ejemplo las contrataciones de personal, eligiendo otros sectores de actividad. En la medida en que exista un conjunto de investigaciones que obtengan resultados parecidos, más generalizables pueden llegar a ser las conclusiones.

Steiner et al. (2016: 76) van un paso más allá y señalan que los que no se fién de la validez externa de cualquier investigación, no solo de los estudios de

viñetas, sino también de las muestras no probabilísticas, pueden desdeñar las inferencias estadísticas y leer los resultados simplemente como las estimaciones de los parámetros descriptivos de la muestra. Llevado a un escepticismo extremo, puede concluirse que cada hallazgo es externamente válido, pero solo bajo un conjunto determinado de condiciones (Keuschigg y Wolbring, 2016), dicho de otro modo: los resultados son válidos para los sujetos analizados en las condiciones en que han sido analizados.

Pese a las dudas sobre hasta qué punto pueden generalizarse los resultados de los estudios de viñetas o de cualquier experimento no aplicado a una muestra representativa, puede argumentarse, por último, que este tipo de diseños contribuyen a aumentar el conocimiento sociológico, puesto que, desde la perspectiva de algunos autores como Falk y Heckman (2009), la tarea fundamental de las ciencias sociales es saber el efecto causal entre fenómenos sociales, esto es, que lo que se descubra tenga una fuerte validez interna o que lo que se diga sea cierto. En este sentido, si hablamos de certidumbre sobre las relaciones causales, tendremos que convenir que los diseños experimentales, o cuasiexperimentales, ganan la partida.

5. Conclusiones

Las encuestas factoriales (*factorial surveys*), a menudo conocidas como estudios de viñeta (*vignette studies*), son un tipo de estudios cuasiexperimentales basados en encuesta que presentan dos ventajas principales. La primera es que, mediante un abordaje indirecto de la pregunta (a través de la presentación de una situación ficticia), son más adecuadas para evaluar prejuicios, estereotipos u otras actitudes sujetas a problemas de deseabilidad social en comparación con las encuestas tradicionales. La segunda se refiere a que, mediante la manipulación de los escenarios ficticios presentados, los encuestados evalúan combinaciones de variables que raramente se dan en la realidad. Por ello, uno de los usos habituales de esta técnica ha sido la evaluación de preferencias o discriminación en el mercado laboral. Aunque esta técnica se aplica a poblaciones generales, en no pocas ocasiones se analizan colectivos muy específicos, como los encargados de la contratación de personal, lo que condiciona la validez externa de los hallazgos, ya que su participación en los estudios de viñeta no suele basarse en un muestreo representativo. Esta limitación, conviene subrayarlo, la padece cualquier encuesta dirigida a una población muy específica con problemas de muestreo. El estudio de viñeta es, en todo caso, una técnica creativa y poco costosa que presenta *a priori* una gran versatilidad para acceder a opiniones, actitudes e incluso a potenciales comportamientos discriminatorios.

Agradecimientos

Agradecemos a los dos evaluadores anónimos sus valiosos comentarios, gracias a los cuales ha mejorado la última versión del artículo.

Financiación

Este artículo forma parte de los proyectos de investigación del Plan Nacional I+D+I financiados por el Ministerio de Economía y Competitividad con referencias PID2020-114702RA-I00 y PDI2020-119339GB-C21; y del proyecto «Men in Care: Workplace Support for Caring Masculinities» (VS/2018/0417), financiado por la Comisión Europea.

Referencias bibliográficas

- AUSPURG, Katrin y HINZ, Thomas (2015). *Factorial Survey Experiments*. Los Ángeles: Sage Series Quantitative Applications in the Social Sciences, n.º. 175.
<<https://doi.org/10.4135/9781483398075>>
- AUSPURG, Katrin y JÄCKLE, Annette (2015). «First Equals Most Important? Order Effects in Vignette-Based Measurement». *Sociological Methods & Research*.
<<https://doi.org/10.1177/0049124115591016>>
- BEEK, Krijn van; KOOPMANS, Carl y PRAAG, Bernard van (1997). «Shopping at the labour market: A real tale of fiction». *European Economic Review*, 41, 295-317.
<[https://doi.org/10.1016/S0014-2921\(96\)00037-2](https://doi.org/10.1016/S0014-2921(96)00037-2)>
- BENARD, Stephen y CORRELL, Shelley J. (2010). «Normative Discrimination and the Motherhood Penalty». *Gender & Society*, 24, 616-46.
<<https://doi.org/10.1177/0891243210383142>>
- BIESMA, Regien; PAVLOVA, Milena; MERODE, Frits van y GROOT, Wim (2007). «Using Conjoint Analysis to Estimate Employers Preferences for Key Competencies of Master Level Dutch Graduates Entering the Public Health Field». *Economics of Education Review*, 26 (3), 375-386.
<<https://doi.org/10.1016/j.econedurev.2006.01.004>>
- BOOTH, Alison y LEIGH, Andrew (2010). «Do employers discriminate by gender? A field experiment in female-dominated occupations». *Economics Letters*, 107 (2), 236-238.
<<https://doi.org/10.1016/j.econlet.2010.01.034>>
- CASTILLA, Emilio J. y BERNARD, Stephen (2010). «The paradox of Meritocracy in Organizations». *Administrative Science Quarterly*, 55 (6), 543-576.
<<https://doi.org/10.3905/JOI.2010.19.1.032>>
- CIS (2019). *La encuesta social europea 9, sesión de formación*. Centro de Investigaciones Sociológicas, octubre de 2019, presentación disponible en <https://www.cis.es/cis/export/sites/default/-Archivos/Formacion/BRIEFING_ESE9.pdf>.
- CORRELL, Shelley J.; BENARD, Stephen y PAIK, In (2007). «Getting a Job: Is There a Motherhood Penalty?». *American Journal of Sociology*, 112 (5), 1.297-1.339.
<<https://doi.org/10.1086/511799>>
- CORTINA, Clara; RODRIGUEZ, Jorge y GONZÁLEZ, M. J. (2021). «Mind the Job: The Role of Occupational Characteristics in Explaining Gender Discrimination». *Social Indicators Research*, 156, 91-110.
<<https://doi.org/10.1007/s11205-021-02646-2>>
- CUDDY, Amy J. C.; FISKE, Susan T. y GLICK, Peter (2004). «When Professionals Become Mothers, Warmth Doesn't Cut the Ice». *Journal of Social Issues*, 60 (4), 701-718.
<<https://doi.org/10.1111/j.0022-4537.2004.00381.x>>

- DAMELANG, Andreas y ABRAHAM, Martin (2016). «You Can Take Some of It with You! A Vignette Study on the Acceptance of Foreign Vocational Certificates and Ethnic Inequality in the German Labour Market». *Zeitschrift Für Soziologie*, 45 (2), 91-106.
<<https://doi.org/10.1515/zfsoz-2015-1005>>
- DI STASIO, Valentina (2014a). «Education as a Signal of Trainability: Results from a Vignette Study with Italian Employers». *European Sociological Review*, 30 (4), 796-809.
<<https://doi.org/10.1093/esr/jcu074>>
- (2014b). *Why education matters to employers: a vignette study in Italy, England and the Netherlands*. Amsterdam: University of Amsterdam.
- DI STASIO, Valentina y GËRKHANI, Klarita (2015). «Employers' Social Contacts and Their Hiring Behaviour in a Factorial Survey». *Social Science Research*, 51, 93-107.
<<https://doi.org/10.1016/j.ssresearch.2014.12.015>>
- DI STASIO, Valentina y WERFHORST, Herman G. van de (2016). «Why Does Education Matter to Employers in Different Institutional Contexts? A Vignette Study in England and the Netherlands». *Social Forces*, 95 (1), 77-106.
<<https://doi.org/10.1093/sf/sow027>>
- DÜLMER, Hermann (2007). «Experimental Plans in Factorial Surveys. Random or Quota Design?». *Sociological Methods & Research*, 35 (3), 1-28.
<<https://doi.org/10.1177/0049124106292367>>
- EIFLER, Stefanie y PETZOLD, Knut (2019). «Validity Aspects of Vignette Experiments: Expected "What-If" Differences Between Reports of Behavioral Intentions and Actual Behavior». En: LAVRAKAS, Paul J.; TRAUOGOTT, Michael W.; KENNEDY, Courtney; HOLBROOK, Allyson L.; LEEUW, Edith D. de y WEST, Brady T. (eds.). *Experimental Methods in Survey Research: Techniques that Combine Random Sampling with Random Assignment*. John Wiley & Sons, Inc.
<<https://doi.org/10.1002/9781119083771.ch20>>
- ESCOT, Lorenzo; BELOPE-NGUEMA, Sabina; FERNÁNDEZ-CORNEJO, José Andrés; POZO-GARCÍA, Eva del; CASTELLANOS-SERRANO, Cristina y CRUZ-CALDERÓN, Selene Fabiola (2021). «Can the legal framework for prostitution influence the acceptability of Buying sex?». *Journal of Experimental Criminology*.
<<https://doi.org/10.1007/s11292-021-09465-y>>
- FALK, Armin y HECKMAN, James J. (2009). «Lab Experiments are a Major Source of Knowledge in the Social Sciences». *Science*, 326, 535-538.
<<https://doi.org/10.1126/science.1168244>>
- FERNÁNDEZ-LOZANO, Irina; GONZÁLEZ, M. José; JURADO-GUERRERO, Teresa y MARTÍNEZ-PASTOR, Juan-Ignacio (2020). «The Hidden Cost of Flexibility: A Factorial Survey Experiment on Job Promotion». *European Sociological Review*, 36 (2), 265-283.
<<https://doi.org/10.1093/esr/jcz059>>
- GADDIS, Michael S. (2015). «Discrimination in the credential society: An audit study of race and college selectivity in the labor market». *Social Forces*, 93 (4), 1.451-1.459.
<<https://doi.org/10.1093/sf/sou111>>
- GERRING, John (2014). *Metodología de las ciencias sociales*. Madrid: Alianza.
- GNIZA, Jan; MARTIN, Abraham; LEHMAN, Isabella y LOREK, Kerstin (2022). «Expecting equality or efficiency? A factorial survey on the justive perception of regional redistributive policies». *Regional Studies*.
<<https://doi.org/10.1080/00343404.2021.2018129>>

- GUTFLEISCH, Tamara; ROBIN, Samuel y SACCHI, Stefan (2021). «The application of factorial surveys to study recruiters' hiring intentions: comparing designs based on hypothetical and real vacancies». *Quality & Quantity*, 55, 775-804.
<<https://doi.org/10.1007/s11135-020-01012-7>>
- HAINMUELLER, Jens; HANGARTNER, Dominik y YAMAMOTO, Teppei (2015). «Validating vignette and conjoint survey experiments against real-world behavior». *PNAS*, 112 (8), 2.395-2.400.
<<https://doi.org/10.1073/pnas.1416587112>>
- HENKENS, Kène; SOLINGE, Hanna van y COZIJNSEN, Rabina (2009). «Let Go or Retain? A Comparative Study of the Attitudes of Business Students and Managers about the Retirement of Older Workers». *Journal of Applied Social Psychology*, 39 (7), 1.562-1.588.
<<https://doi.org/10.1111/j.1559-1816.2009.00495.x>>
- HUMBURG, Martin y VELDEN, Rolf van der (2014). *Skills and the Graduate Recruitment Process: Evidence From Two Discrete Choice Experiments*. Research Centre for Education and the Labour Market.
<<https://doi.org/10.1016/j.econedurev.2015.07.001>>
- JACOBS, Jerry A. y GERSON, Kathleen (2016). «Unpacking americans' views of the employment of mothers and fathers using national vignette survey data: SWS Presidential Address». *Gender & Society*, 30 (3), 413-441.
<<https://doi.org/10.1177/0891243215597445>>
- KARPINSKA, Kasia; HENKENS, Kène y SCHIPPERS, Joop (2011). «The Recruitment of Early Retirees: A Vignette Study of the Factors That Affect Managers' Decisions». *Ageing and Society*, 31 (4), 570-589.
<<https://doi.org/10.1017/S0144686X10001078>>
- (2013). «Retention of Older Workers: Impact of Managers' Age Norms and Stereotypes». *European Sociological Review*, 29 (6), 1.323-1.335.
<<https://doi.org/10.1093/esr/jct017>>
- KARPINSKA, Kasia; HENKENS, Kène; SCHIPPERS, Joop y WANG, Mo (2015). «Training Opportunities for Older Workers in the Netherlands: A Vignette Study». *Research in Social Stratification and Mobility*, 41, 103-112.
<<https://doi.org/10.1016/j.rssm.2015.03.002>>
- KARREN, Ronald J. y BARRINGER, Melissa Woodbard (2022). «A Review and Analysis of the Policy-Capturing Methodology in Organizational Research: Guidelines for Research and Practice». *Organizational Research Methods*, 5 (4), 337-361.
<<https://doi.org/10.1177/109442802237115>>
- KENNEDY, Courtney y HARTIG, Hannah (2019). «Response rates in telephone surveys have resumed their decline». *Pew Research Center*.
<<https://pewrsr.ch/2XqsgTT>>
- KEUSCHNIGG, Marc y WOLBRING, Tobias (2016). «The Use of Field Experiments to Study Mechanisms of Discrimination». *Analyse & Kritik. Zeitung für linke Debatte und Praxis*, 38 (1), 179-202.
<<https://doi.org/10.1515/auk-2016-0109>>
- KING, Gary; KEOHANE, Robert O. y VERBA, Sidney (2005). *El diseño de la investigación social. La inferencia científica en los estudios cualitativos*. Madrid: Alianza.
- KOOTSTRA, Anouk (2016). «Deserving and Undeserving Welfare Claimants in Britain and the Netherlands: Examining the Role of Ethnicity and Migration Status Using a Vignette Experiment». *European Sociological Review*, 32 (3), 325-338.
<<https://doi.org/10.1093/esr/jcw010>>

- KÜBLER, Dorothea; SCHMID, Julia y STÜBER, Robert (2018). «Gender discrimination in hiring across occupations: a nationally-representative vignette study». *Labour Economics*, 55, 215-229.
<<https://doi.org/10.1016/j.labeco.2018.10.002>>
- LAENEN, Tigs; HOOTEGEM, Arno van y ROSSETTI, Federica (2022). «The multidimensionality of public support for basic income: a vignette experiment in Belgium». *Journal of European Public Policy*, online.
<<https://doi.org/10.1080/13501763.2022.2055112>>
- LIECHTI, Fabienne; FOSSATI, Flavia; BONOLI, Giuliano y AUER, Daniel (2017). «The Signalling Value of Labour Market Programmes». *European Sociological Review*, 33 (2), 257-274.
<<https://doi.org/10.1093/esr/jcw061>>
- MARTÍNEZ PASTOR, Juan Ignacio (2019). *Los datos sin tapujos. Cómo interpretar y difundir las estadísticas sociales*. Madrid: Los Libros de la Catarata.
- MCDONALD, Patrick (2019). «How Factorial Survey Analysis Improves Our Understanding of Employer Preferences». *Swiss Journal of Sociology*, 45 (2), 237-260.
<<https://doi.org/10.2478/sjs-2019-0011>>
- MULDERS, Jaap; DALEN, Hendrik van; HENKENS, Kène y SCHIPPERS, Joop (2014). «How Likely Are Employers to Rehire Older Workers after Mandatory Retirement? A Vignette Study Among Managers». *De Economist*, 162, 415-431.
<<https://doi.org/10.1007/s10645-014-9234-8>>
- OESCH, Daniel; LIPPS, Oliver y MCDONALD, Patrick (2017). «The Wage Penalty for Motherhood: Evidence on Discrimination From Panel Data and a Survey Experiment for Switzerland». *Demographic Research*, 37, 1.793-1.824.
<<https://doi.org/10.4054/DemRes.2017.37.56>>
- RUDMAN, Laurie A. y MESCHER, Kris (2013). «Penalizing men who request a family leave: Is flexibility stigma a femininity stigma?». *Journal of Social Issues*, 69 (2), 322-340.
<<https://doi.org/10.1111/josi.12017>>
- SHI, Lulu P.; IMDORF, Christian; SAMUEL, Robin y SACCHI, Stefan (2018). «How Unemployment Scarring Affects Skilled Young Workers: Evidence From a Factorial Survey of Swiss Recruiters». *Journal for Labour Market Research*, 52 (7), 1-15.
<<https://doi.org/10.1186/s12651-018-0239-7>>
- STEINER, Peter, M.; ATZMÜLLER, Christiane y SU, Dan (2016). «Designing Valid and Reliable Vignette Experiments for Survey Research: A Case Study on the Fair Gender Income Gap». *Journal of Methods and Measurement in the Social Sciences*, 7 (2), 52-94.
<<https://doi.org/10.2458/jmm.v7i2.20321>>
- STERKENS, Philippe; BAERT, Stijn; ROOMAN, Claudia y DEROUS, Eva (2020). «As If it Weren't Hard Enough Already: Breaking Down Hiring Discrimination Following Burnout». *IZA Institute of Labor Economics, Discussion Paper Series No. 13514*.
<<https://doi.org/10.2139/ssrn.3660260>>
- WALLANDER, Lisa (2009). «25 years of factorial surveys in sociology: A review». *Social Science Research*, 38 (3), 505-520.
<<https://doi.org/10.1016/j.ssresearch.2009.03.004>>
- WOLF, Inge de y VELDEN, Rolf van der (2001). «Selection Processes for Three Types of Academic Jobs: An Experiment Among Dutch Employers of Social Sciences Graduates». *European Sociological Review*, 17 (3), 317-330.
<<https://doi.org/10.1093/esr/17.3.317>>