Modelización matemática en actividades estadísticas: identificando elementos clave del diseño para promover la generación de modelos en el estudio de la variabilidad

Mathematical modelling in statistical activities: identifying key elements of design to promote model generation in the study of variability

Àngels Aymerich
Departament de Didàctica de la Matemàtica i les Ciències Experimentals
Universitat Autònoma de Barcelona
España
mariaangels.aymerich@e-campus.uab.cat

Lluís Albarracín
Departament de Didàctica de la Matemàtica i les Ciències Experimentals
Universitat Autònoma de Barcelona
España
lluis.albarracin@uab.cat

Resumen. En nuestro trabajo estamos interesados en promover el aprendizaje de conocimientos estadísticos a partir de la modelización matemática. Usamos problemas estadísticos donde se estudian fenómenos sociales reales a partir de grandes cantidades de datos y que cumplen con los principios de diseño de las *Modelling Eliciting Activities* con alumnos de Educación Secundaria (15 años). Las tareas se dirigen a promover el desarrollo del concepto de variabilidad y aplicarlo para entender la situación estudiada. En este estudio nos centramos en caracterizar los modelos matemáticos que desarrolla el alumnado en una actividad de modelización estadística intentando identificar por qué se integran elementos al modelo o se descartan. Para ello se desarrolla un análisis cualitativo a partir de las grabaciones del trabajo en grupo de los alumnos en el aula. Los resultados obtenidos muestran que algunas decisiones de diseño del problema, como la gran cantidad de datos, o la ambigüedad de conceptos sociales, como el de justicia y equidad en los impuestos, resultan esenciales para el desarrollo de los modelos matemáticos. Las conclusiones del estudio tienen implicaciones para el diseño de tareas estadísticas, pero también para identificar el rol de la modelización matemática en el aprendizaje de conceptos estadísticos.

Quadrante: Revista de Investigação em Educação Matemática
https://doi.org/10.48489/quadrante.23686
Received: February 2021 / Accepted: July 2021

Palabras-clave: modelización matemática; actividades promotoras de la modelización; educación secundaria; estadística; variabilidad.

Abstract. In our work we are interested in promoting the learning of statistical knowledge through mathematical modelling. We use statistical problems where real social phenomena are studied from large amounts of data and that comply with the design principles of the Modelling Eliciting Activities with secondary school students (15 years old). The tasks are aimed at promoting the development of the concept of variability and applying it to understand the situation studied using a mathematical model. In this study we focus on characterizing the mathematical models that students develop in a statistical modelling activity, trying to identify why elements are integrated into the model or discarded, as students conceptualize variability. For this purpose, a qualitative analysis is carried out on the basis of recordings of the students' group work in the classroom. The results obtained show that some problem design decisions, such as the large amount of data, or the ambiguity of social concepts, such as justice and equity in taxation, are essential for the development of mathematical models. The conclusions of the study have implications for the design of statistical tasks, but also for identifying the role of mathematical modelling in the learning of statistical concepts.
Keywords: mathematical modeling; modelling-eliciting activities; secondary education; statistics; variability.

Introducción

Nuestra sociedad vive un auge del uso y necesidad de la Estadística. Como ciudadanos somos tanto consumidores como generadores de datos. Las posibilidades que ofrece el análisis de datos han generado un aumento de las profesiones relacionadas con la estadística como ciencia y como profesión de futuro. Al mismo tiempo, los conocimientos y métodos estadísticos son esenciales para promover competencias que permitan entender e interpretar informaciones estadísticas y actúan como una herramienta clave para formar futuros ciudadanos. Por ello, en las últimas décadas los currículos educativos de enseñanza obligatorias a nivel internacional (Burrill & Biehler, 2011) han hecho hincapié en potenciar el trabajo en las aulas de los contenidos estadísticos, incluyendo recomendaciones específicas para su implementación (Gal, 2004).

En nuestro trabajo tomamos un posicionamiento específico para enfrentarnos al reto de la enseñanza de la Estadística en Secundaria: promover que los alumnos construyan por ellos mismos sus propios conceptos y métodos estadísticos para estudiar situaciones reales y complejas. En este artículo partimos de estudios previos en los que se corrobora que los alumnos de Educación Secundaria desarrollan sus propios modelos matemáticos para resolver tareas en las que se proporcionan grandes cantidades de datos (Albarracín et al., 2017; Aymerich et al., 2017). Estos estudios ponen de manifiesto que es posible utilizar actividades de modelización para que los alumnos generen sus propios conceptos y métodos estadísticos para estudiar la variabilidad.

Por ello, el propósito de nuestra investigación es identificar qué elementos ha de tener una tarea para que los alumnos de Educación Secundaria generen nociones propias sobre la variabilidad. Estas nociones serán las precursoras de la comprensión de los conceptos estadísticos formales, como la desviación, incluidos en el currículo. Ya que los estudiantes no sólo deben saber las fórmulas de su medida, sino también los procedimientos relacionados y su propósito dentro de un contexto (Makar & Confrey, 2005). Tomamos como referencia las Actividades Promotoras de Modelos (Lesh et al., 1997) entendiendo que será necesario dotar la actividad de elementos específicos en los enunciados, los materiales proporcionados y la gestión de aula por parte del profesor.

De esta forma, los objetivos de investigación de este estudio son los siguientes:

- Caracterizar los modelos matemáticos, según la forma de conceptualizar la variabilidad de los datos, que generan los alumnos de Educación Secundaria en una Actividad Promotora de Modelos orientada al estudio de una situación real en la que se proporciona una gran cantidad de datos.
- Identificar los elementos del diseño de la actividad que actúan como promotores del desarrollo del modelo matemático construido.

Modelización matemática

La modelización matemática como área de investigación en el campo de la educación matemática comenzó con el trabajo de Pollak (1969), quien introdujo la discusión sobre la relación entre las aplicaciones de las matemáticas y los procesos de enseñanza/aprendizaje. Posteriormente, el mismo Pollak (1979) presentó un primer marco teórico en el que interpretaba los procesos de modelización diferenciando entre el dominio matemático y el resto del mundo. Esta separación en dos dominios diferenciados nos lleva necesariamente al proceso de matematización de un fenómeno, pasando de la realidad al dominio matemático, y a la interpretación dentro del contexto real de los modelos generados en el dominio matemático, como forma de validación.

A partir de este posicionamiento inicial, la modelización matemática se estableció como un tema de investigación de interés en la educación matemática. Su finalidad principal es establecer actividades de aula que pongan de manifiesto la estrecha relación entre las matemáticas y el mundo que nos rodea (Blum, 2002). En las últimas décadas se han desarrollado una gran cantidad de investigaciones sobre la modelización matemática educativa y se han diversificado notablemente, tanto en lo que respecta a los objetivos de investigación como a los enfoques que toman los investigadores (Abassian et al., 2020; Kaiser & Sriraman, 2006). Recientemente, Blum (2015) reafirmó que la enseñanza de las aplicaciones y la modelización tiene una doble función. En primer lugar, el conocimiento de las matemáticas y sus aplicaciones es vital para el mundo real y su progreso, principalmente en lo que se refiere a resolver problemas reales y desarrollar proyectos complejos. Además, el mundo real y la forma en que se integra el conocimiento matemático es extremadamente importante como vehículo para dar sentido al aprendizaje de los conceptos matemáticos y, en general, de las matemáticas como disciplina.

En nuestro trabajo consideramos que una tarea es una actividad de modelización cuando los alumnos generan o utilizan modelos matemáticos para describir o analizar fenómenos reales. Adoptamos la definición de modelo matemático propuesta por Lesh y Harel (2003), que definen los modelos matemáticos como sistemas conceptuales que describen otros sistemas. Estos autores entienden que los modelos están formados por: (a) un conjunto de conceptos para describir o explicar los objetos matemáticos relevantes del fenómeno estudiado, así como las relaciones, acciones, patrones y regularidades que se atribuyen a la situación de resolución de problemas; y (b) los procedimientos que los acompañan para generar construcciones útiles, manipulaciones o predicciones para el logro de objetivos claramente reconocidos. Este sistema que conforma el modelo puede ser expresado de formas múltiples, ya sea a partir de relaciones algebraicas, símbolos escritos, gráficos o esquemas, lenguaje natural o metáforas basadas en la experiencia.

La modelización matemática es un proceso de resolución de un problema real en el que intervienen conceptos, métodos y resultados matemáticos (Blum & Niss, 1991). Para ello, los objetos, los datos y las relaciones que se dan en realidad se trasladan al mundo de las matemáticas, obteniendo así un modelo matemático. A continuación, se aplican métodos matemáticos a este modelo para llegar a una solución matemática, que debe ser interpretada y validada en el mundo real en el que se enmarca el problema, dando lugar a una solución real. La forma en la que los estudiantes elaboran modelos matemáticos en procesos de resolución de problemas para describir fenómenos o realizar predicciones es objeto de discusión y existen diferentes posiciones teóricas al respecto (Borromeo Ferri, 2006). En general, se acepta en la literatura que los procesos de modelización tienen una naturaleza cíclica (Blum & Leiβ, 2006; Carreira et al., 2011; Doerr & English, 2003; Galbraith & Stillman 2006; Greefrath, 2011; Kaiser & Stender, 2013). Aunque existen diversas formas de concretar la estructura de un proceso de modelización, se entiende que se puede dividir en diferentes fases que los alumnos recorren. El ciclo de modelización, como propuesta para caracterizar la actividad de modelización matemática, ha sido utilizado tanto en estudios que inciden en el comportamiento de los alumnos en actividades de modelización (Ramírez-Montes et al., 2021) como en el diseño de actividades (Hernández-Sabaté et al., 2020). Al mismo tiempo, dada la complejidad del proceso de modelización, diversos autores (Aymerich & Albarracín, 2016; Blum & Borromeo Ferri, 2009) coinciden en la no linealidad de los ciclos de modelización, especialmente en estudiantes noveles, con lo que se constata que es una actividad compleja y cognitivamente demandante para los estudiantes.

Análisis de datos y modelización matemática

La forma en la que los estudiantes se enfrentan a situaciones reales en las que es necesaria la modelización de datos ha sido un foco de atención en la investigación educativa desde finales del siglo XX. Diversos estudios constatan que estas actividades proporcionan un contexto distintivo para observar el trabajo estadístico de los estudiantes en situaciones abiertas (Batanero et al., 1996; Doerr & Tripp, 1999; Lehrer & Schauble, 2000; Lesh et al., 1997). Ante la necesidad de hacer estadísticamente competentes a los estudiantes, los investigadores han desarrollado diversos conceptos que tratan de definir las necesidades estadísticas de los ciudadanos (Crites & Laurent, 2015). De esta forma, aparecen en la literatura términos que provienen de marcos teóricos distintos y que no son excluyentes, sino que se complementan (Ubilla, 2019), como son la alfabetización estadística, el razonamiento estadístico, y el pensamiento estadístico.

El tipo de actividades que utilizamos en este estudio conectan con los conocimientos estadísticos. La Estadística como ciencia por sí misma tiene unas características que la diferencian de las matemáticas, pero en las investigaciones estadísticas resalta la importancia del contexto y de la variabilidad. Cobb y Moore (1997) explican el valor del contexto como dominio en el que interpretar sus resultados. Así un modelo estadístico se define por su propósito. Para Dvir y Ben-Zvi (2018) un modelo matemático aplicable a situaciones estadísticas debe incluir la variabilidad y el uso de consideraciones probabilísticas. Cuando nos referimos a estudiantes de Secundaria, de modeladores noveles, hemos de entender que, como no han adquirido todavía las destrezas necesarias para cuantificar la probabilidad, el estudio de los modelos creados y de los ciclos de modelización es complejo. Por eso en nuestro trabajo pretendemos identificar el modelo matemático que generan, ya que nuestro alumnado no ha trabajado el desarrollo del pensamiento estadístico ni los conceptos necesarios para cuantificar la incertidumbre en un estudio estadístico.

Actividades promotoras de la modelización

En este trabajo consideramos una perspectiva teórica para interpretar las actividades de modelización matemática denominada Modelos y Modelización (M&M). Se asume que el proceso de modelización incluye múltiples ciclos de interpretaciones, descripciones, conjeturas, explicaciones y justificaciones que se redefinen y reconstruyen iterativamente a través de la interacción entre los estudiantes (Doerr & English, 2003). Los problemas o contextualizados en el mundo real y con características y demandas que los convierten en actividades de modelización se denominan Actividades Promotoras de la Modelización

(MEAs, de sus iniciales en inglés, *Modelling-Eliciting Activities*) y se centran tanto el proceso de modelización como el producto final (modelo matemático).

Lesh (1997) caracteriza las MEAs como aquellos problemas basados en la vida real que, cuando son resueltos por los estudiantes, ellos construyen un modelo que representa un sistema del mundo real, presentado en el problema, y que describe, explica o predice el comportamiento del sistema. Así pues, el enfoque del problema debe permitir a los estudiantes establecer criterios apropiados que les ayuden a decidir qué solución es la adecuada entre un conjunto de alternativas diferentes. Por lo general, al proponer una MEA, se pide a los estudiantes que trabajen en pequeños grupos (Zawojewski et al., 2003) y se enfrenten a una situación problemática significativa y relevante para ellos, y para la cual necesitan crear, ampliar y mejorar sus propias construcciones matemáticas. Para diseñar este tipo de actividades nos basamos en los seis principios básicos que definen una MEA, propuestos por Lesh, Amit y Schorr (1997; Tabla 1).

Tabla 1. Descripción de los principios que definen una MEA (Lesh et al., 1997)

Principio	Descripción
Realidad	Responde a la pregunta si el problema corresponde a una situación que puede ocurrir en la vida real. Los estudiantes se verán motivados por experiencias y conocimientos personales.
Construcción de un modelo	La tarea debe crear la necesidad de construir un modelo, modificarlo, extenderlo o refinarlo. La tarea debe involucrar construir, explicar, manipular, predecir un sistema significativo.
Autoevaluación	Deben presentarse unos criterios claros para la evaluación de las diversas soluciones alternativas. Los estudiantes deben poder juzgar por sí mismos si las soluciones dadas son bastante buenas y los resultados son los que necesitan.
Documentación del modelo	Los estudiantes deben hacer explícito lo que piensan sobre la situación, y el tipo de sistemas (objetos matemáticos, relaciones, operaciones, patrones, regularidades) en los que han pensado.
Generalización del modelo	El modelo que construyen debe poder ser aplicado o adaptado a un abanico mayor de situaciones.
Prototipo simple	La tarea pone al estudiante en una situación tan sencilla como sea posible para crear la necesidad de un modelo significativo, pero también provee al estudiante de un prototipo útil (o metáfora) para interpretar situaciones estructuralmente similares.

Metodología

En esta sección se describen los diferentes elementos que caracterizan esta investigación. Se incluye la descripción y justificación del problema utilizado, las características de los estudiantes participantes, los datos recogidos y el análisis desarrollado.

Diseño y justificación de la actividad

La actividad utilizada trata sobre la variabilidad del Impuesto de Circulación en España y tiene como objetivo que los estudiantes comprendan la necesidad de medir la variabilidad en una gran cantidad de datos cuantitativos. El Impuesto de Circulación depende de cada Ayuntamiento (como se ilustra en el recorte representado en la Figura 1), así que puede variar en función de las necesidades y características socioeconómicas de cada municipio. La tarea empieza con la lectura en forma de un artículo real publicado en un periódico que analiza el hecho de que en poblaciones cercanas se pueden encontrar grandes diferencias en el valor de este impuesto. En el artículo se dan ejemplos concretos de este hecho para ilustrar la situación. Con esta información se plantea a los alumnos la necesidad de decidir qué provincias son más justas respecto a los impuestos mencionados. Por otro lado, se les pide un informe que evalúe la situación, justificando y dando propuestas para decidir si se debe de continuar del mismo modo o el Estado tiene que intervenir.

TARIFAS IMPUESTO DE CIRCULACIÓN 2018

PROVINCIA	TURISMOS (CVF)					CICLO	MOTOCICLETAS (C.C)				
	De menos de 8 De menos de 8	De 8 a 12 De 8 a 11,99	De 12 a 16 De 12 a 15,99	Más de 16 De 16 a 19,99	Más de 20		Hasta 125	125 a 250	250 a 500	500 a 1.000	Más de 1.000
A CORUÑA	18,95	60,80	128,35	179,20	224,00	6,65	6,65	13,50	27,00	60,60	121,15
ALBACETE	23,98	64,75	136,69	173,84	217,28	7,87	7,87	13,47	26,97	53,92	107,83
ALICANTE	22,28	60,47	128,37	162,32	204,75	7,43	8,49	13,79	27,58	55,17	111,39
ALMERÍA	23,12	62,41	131,78	164,16	205,11	8,10	8,10	13,88	27,74	55,46	110,88
ÁVILA	20,69	57,93	124,45	163,09	200,48	8,08	8,13	14,53	29,54	60,58	119,34
BADAJOZ	19,85	53,61	113,16	140,96	176,18	6,95	6,95	11,91	23,83	47,65	95,29
BARCELONA	23,47	64,06	136,69	172,05	217,28	8,39	8,39	14,38	28,78	58,16	117,53
BILBAO	25,08	67,92	(*)	268,32	343,20	10,08	10,08	17,88	37,08	74,88	155,04
BURGOS	23,19	62,62	132,20	164,66	205,80	8,13	8,13	13,91	27,84	55,66	111,32
CÁCERES	17,00	50,00	100,00	125,00	140,00	8,00	8,50	11,00	23,00	43,00	84,00
CÁDIZ	24,90	67,25	142,00	177,05	220,90	8,80	8,80	14,85	29,90	59,85	119,60
CASTELLÓN	24,99	67,48	142,44	177,43	221,76	8,75	8,75	14,99	30,00	59,97	119,95
CEUTA	12,60	34,10	71,95	89,60	112,00	4,40	4,40	7,55	15,15	30,30	60,60
CIUDAD REAL	25,24	68,16	143,88	179,22	224,00	8,84	8,84	15,14	30,30	60,58	121,16
CÓRDOBA	24,22	65,40	143,88	179,22	224,00	8,57	8,57	14,69	29,39	58,76	117,53
CUENCA	22,09	59,64	125,89	179,22	224,00	7,74	7,74	13,25	26,51	53,01	106,01
GIRONA	25,24	67,24	142,40	178,00	224,00	8,84	8,84	15,04	30,20	60,24	121,16

Figura 1. Tarifas del Impuesto Municipal de Vehículos de Tracción Mecánica. Fuente: Automovilistas Europeos Asociados (AEA)

En la primera fase se plantea el problema y se pide a los alumnos que desarrollen el modelo que van a usar para dar respuesta a las preguntas formuladas. En la segunda fase se pide que trabajen con los datos proporcionados, generen resultados, redacten un informe para evaluar la situación y presentar propuestas en las que se valoren soluciones posibles para minimizar las desigualdades. A continuación, mostramos un extracto del enunciado proporcionado a los alumnos (Figura 2).

El artículo deja claro que cada Ayuntamiento decide cuál es el valor del Impuesto y que el Estado sólo dicta qué automóviles están exentos.

- En el artículo se explica que, dentro de la misma comunidad, hasta, en la misma provincia, las diferencias son evidentes. ¿Qué provincias son más justas, y, por tanto, el hecho de matricular en una población u otra no supone un agravio comparativo para los propietarios de los vehículos?
- Hacienda os pide un informe para evaluar esta situación y decidir si el estado tiene que intervenir. Tiene que incluir un recopilatorio de propuestas que valoren diferentes soluciones.

Figura 2. Enunciado de la MEA propuesta

A los alumnos se les proporciona una hoja de cálculo con los valores de todos los impuestos según el tipo de vehículos de cada uno de los municipios españoles publicados por Hacienda. Este documento contiene información sobre los impuestos a cargo del municipio, provincia a la que pertenece y la población que tiene. Además del Impuesto de Vehículos de Tracción Mecánica (turismos, tractores, camiones, remolques, ciclomotores, autobuses y motocicletas), nos proporciona: el Impuesto sobre Bienes e Inmuebles (IBI), el Impuesto sobre Actividades Económicas (IAE), el Impuesto sobre el Incremento del Valor de los Terrenos de Naturaleza Urbana (IIVTNU), y el Impuesto sobre Construcciones, Instalaciones y Obras (ICIO).

La actividad está diseñada para cumplir con los principios de diseño de las MEAs (Lesh et al., 1997). A continuación, justificamos estos aspectos de diseño:

- *Principio de realidad:* El contexto de la actividad responde a un debate que se daba en la sociedad en aquel momento. La edad del alumnado y la materia en que se hacía la actividad nos ofrecía ese interés por el tema. Los impuestos y su marco legal son necesarios para un ciudadano adulto y ellos ya estaban dando sus primeros pasos en esa responsabilidad.
- *Principio de construcción de un modelo*: Existen diversos modelos útiles para resolver esta tarea, se trata de cuantificar la variabilidad de los datos. Se puede tratar de trabajar utilizando todos los datos disponibles o bien escoger una muestra. Una vez decidido, se pueden utilizar representaciones gráficas, parámetros de centralidad y dispersión.
- *Principio de autoevaluación*: Cuando se les pide el informe es para que reflexionen sobre los resultados obtenidos y puedan evaluarlos, de hecho, se espera que sean momentos de reflexión que comporten un cambio en el modelo si fuese necesario.
- *Principio de documentación del modelo*: Las dos tareas implican generar un informe en el que se describa el tratamiento de datos y el modelo desarrollado.

- *Principio de generalización del modelo*: El modelo de un estudiante experimentado pasa por el análisis de la dispersión de los datos, que responde a un gran número de situaciones reales.
- *Principio de prototipo simple*: El análisis de diferencias entre los datos, sujeto al contexto propuesto le sirve al estudiante como situación concreta que puede abrir la puerta al análisis de otras situaciones en las que se dan diferencias entre localidades, provincias, estados u otros.

Participantes y datos recogidos

La recogida de datos se desarrolla en un centro concertado de una ciudad del área metropolitana de Barcelona. Participan 30 estudiantes de 4.º de ESO (15-16 años) que trabajan en grupos de tres personas en sesiones de 60 minutos. Estos alumnos cursan la materia optativa de economía, el perfil profesional que domina en sus estudios posteriores o en un bachillerato social o bien en un ciclo formativo. Así mismo en el centro educativo en el ámbito matemático se trabaja la resolución de problemas, pero no se hace explícito qué es la modelización matemática ni su ciclo. Respecto a los contenidos estadísticos, tanto en 1º de ESO como en 3.º de ESO se trabajan los conceptos estadísticos que marca el currículum español. Se empieza con estadística unidimensional: organización en tablas, representación y cálculo de parámetros de centralidad. Se realiza algún estudio estadístico contextualizado, normalmente relacionado con conceptos de población, muestra, recogida de datos y el planteamiento de preguntas interesantes sobre la población a estudiar. Este grupo en particular no pudo realizar el proyecto por falta de tiempo. En tercero de ESO se amplía a los parámetros de dispersión, siempre trabajando en estadística unidimensional. Si bien en otras materias se pueden tratar contenidos estadísticos, el tratamiento se limita al uso de fórmulas y a lectura de gráficos estadísticos.

La tarea se desarrolla en diversas sesiones, variable en función del ritmo de cada grupo. La intervención del profesor es mínima puesto que el objetivo es identificar los elementos del diseño de la tarea que promueven la necesidad de cuantificar la variabilidad. Todas las discusiones de los alumnos se graban en audio y se recogen todos los materiales elaborados durante la actividad. Estos materiales son los ficheros de Excel en los que manipulan los datos, los borradores y documentos en bruto y el redactado del informe final. Para el análisis se utilizan los datos de tres de los grupos participantes, que denominamos grupos G1, G2 y G3. Se eligen estos grupos por dos motivos: por el interés del trabajo de modelización desarrollado y la calidad de los audios recogidos, en el sentido que permiten entender claramente las aportaciones de los alumnos durante toda la actividad sin generar posibles lagunas en el análisis.

Análisis

Durante el proceso de modelización los alumnos construyen versiones parciales (por incompletas o no definitivas) del modelo que desarrollan. Para cada una de estas versiones parciales del modelo construimos una representación de los elementos que lo componen y que están considerando los alumnos. Este tipo de representación se basa en unos diagramas que nos permiten mostrar los conceptos y procedimientos que constituyen el modelo matemático en un momento del proceso de modelización, y nos permiten analizar su evolución. Fueron utilizados en Albarracín, Aymerich y Gorgorió (2017) y han sido reconstruidos para simplificar su uso. Concretamente, representamos los conceptos y procedimientos que utilizan los alumnos con circunferencias con la codificación del elemento en su interior. Si los utilizan conjuntamente los agrupamos y si los utilizan interrelacionados los unimos con flechas.

Para ejemplificar el método de análisis mostramos a continuación la elaboración de uno de estos diagramas. Corresponde a un grupo que descartamos ya que las grabaciones no tenían la calidad suficiente. El grupo 4, en el minuto 4:57 de la resolución, después de leer el enunciado, observar los datos y discutir entre ellos qué pide el problema, cuando se encuentran en la fase de interpretación, tienen el siguiente diálogo:

A:	Vale chicas, hay que mirar cuál es. Mirad, para cada comunidad autónoma, si los precios están igualados.
C:	Lo apuntamos en papel.
B:	Y ¿por qué no hacemos Catalunya y ya está? Que son 4 provincias….
A:	¡Nooooooooo! Lo hacemos por provincias.

Han decidido estudiar las diferencias de los impuestos por provincias, aunque todavía no saben cómo. Cuando hablan de mirar si están igualados, en el diálogo, el concepto que representa son las diferencias en el valor de los impuestos entre las diferentes provincias (df). Para este grupo, si son suficientemente diferentes, los impuestos son injustos, si, por el contrario, están más igualados, son más justos. Deciden, también, utilizar todos los datos disponibles, lo que sería estudiar a toda la población (P). Identificamos así el primer modelo parcial (Figura 3), calcular las diferencias en los impuestos de todas las provincias.

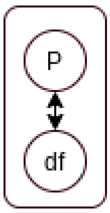

Figura 3. Primer modelo del Grupo 4

Posteriormente, a partir del minuto 06:54, estudian la hoja de cálculo completa y se encuentran con dificultades para implementar el modelo pactado anteriormente. La dificultad principal se encuentra en el hecho que deben determinar respecto a qué provincia concreta hacen sus cálculos de diferencias, porque descartan calcular las diferencias de todas las provincias respecto a todas ellas. A partir de este momento (9:17) cambian el modelo a implementar centrándose en identificar la provincia con un valor menor de los impuestos (min) como criterio. Esta provincia será la que utilicen como referencia para determinar de forma intuitiva la distancia de una provincia a esa provincia de referencia.

Por tanto, están descartando mirar las diferencias (df) a cambio de utilizar el mínimo (min) de los datos dados. Así el grupo continúa el proceso de modelización incorporando o eliminando elementos, fruto de revisiones y relecturas del enunciado, hasta conseguir un modelo definitivo. De esta forma, este grupo cambia el modelo propuesto y pasa a seguir el modelo descrito en la Figura 4.

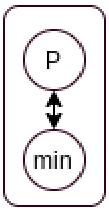

Figura 4. Segundo modelo del Grupo 4

Resultados: modelos desarrollados por cada grupo

A continuación, mostramos la caracterización en forma de representaciones de la evolución de los modelos generados por cada uno de los tres grupos.

Primer grupo (G1)

Los estudiantes del grupo G1 empiezan leyendo e interpretando el enunciado, hasta el minuto 6:58 de la grabación, no sólo leen el enunciado literalmente, sino que también discuten sobre el artículo y lo que pide la tarea. Este episodio se ve interrumpido por la primera fase de exploración de datos para entender mejor el enunciado, al observar los datos entienden que el impuesto cambia según la población (minuto 9:06). La exploración de los datos les permite entender el enunciado, hasta que en el minuto 15:07 toman su primera decisión: trabajar con todos los datos, por tanto, deciden estudiar toda la población, aunque todavía no saben cómo. Lo representamos en la Figura 5.

Figura 5. Primera versión del modelo - Grupo 1

Pero momentos después, 15:14 minutos, se dan cuenta de que les es imposible (con los recursos de que disponen) y deciden estudiar los caballos fiscales más frecuentes. Así su criterio de representatividad sería el tipo de vehículo que utiliza la mayoría de las personas estudiadas, decidiendo que éste sería el turismo de más de 20 caballos fiscales. Una integrante del grupo también propone escoger municipios repartidos por toda España, pero se centran en la elección de los caballos fiscales ignorando la propuesta. Así que se elimina el concepto de Población (P) del modelo que construyen e incorporan el uso de una muestra (m_1). La necesidad de tomar una muestra aparece cuando toman conciencia de la gran cantidad de datos que presenta el problema y las dificultades que anticipan en el uso de herramientas digitales para tratarlas. No son conscientes de lo que supone estadísticamente que las conclusiones extraídas con la muestra luego sean extrapoladas a la población. Podemos identificar que están simplificando (estructurando) la tarea. La Figura 6 muestra el estado parcial del modelo construido al considerar trabajar usando una muestra de los datos.

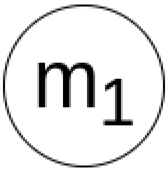

Figura 6. Segunda versión del modelo - Grupo 2

Vuelven a discutir sobre lo que pide el enunciado redefiniendo sus objetivos. Descubren que en la hoja de cálculo están los municipios clasificados por provincias y es imposible estudiar todos los municipios de todas las provincias. Esto provoca que redefinan la muestra escogida con la que van a trabajar y decidan elegir una submuestra que pueda ser representativa (siempre de manera intuitiva). Los alumnos no definen de ninguna forma qué entienden por una muestra representativa, pero pasan a escoger provincias repartidas por el territorio, y de ellas escogen tres municipios. Por lo tanto, de la primera muestra escogen una submuestra. El siguiente modelo (Figura 7) queda coordinado de este modo, donde m_1 es la categoría de caballos fiscales predominante y de esta categoría escogen unos municipios determinados.

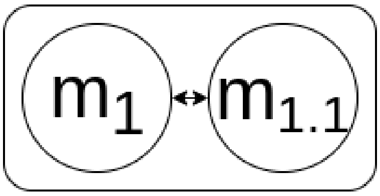

Figura 7. Tercera versión del modelo - Grupo 1

En el minuto 16:12, mientras seguían planificando, decidiendo cómo escogerían la muestra, surge la primera idea de qué harán con su muestra: comparar. Pero en el minuto

19:08 del audio observan que necesitan un análisis más detallado e introducen un análisis de la variabilidad a partir de estudiar diferencias entre los valores, aunque no concretan cómo lo harán. Están matematizando la tarea sin concretar el método exacto. Discuten esta nueva vía hasta el minuto 27:14, momento en el que empiezan a ejecutar su plan escogiendo los municipios que utilizarán para el análisis. Todavía sin concretar cómo calcularán esta diferencia. Así que el modelo final sería el presentado en la Figura 8.

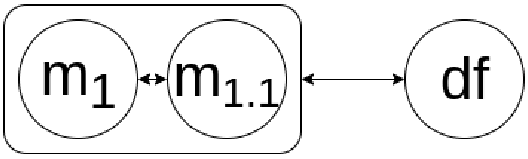

Figura 8. Versión final del modelo - Grupo 1

La submuestra $m_{1.1}$ no está fijada inicialmente, simplemente deciden sobre la marcha los municipios que les parecen más representativos según sus criterios. Buscan municipios que sean diferentes, aunque a veces son diferentes debido a algún prejuicio inicial sobre ese municipio en concreto.

Hasta que no han de redactar conclusiones, momento que han de validar su modelo, no definen qué es para ellos ver que sus datos son diferentes, en el minuto 78:37 de trabajo, concretan diferencia como si la muestra recogida (los impuestos en tres municipios de algunas provincias y de la capital de provincia) se parecen entre ellos. Así el modelo queda completado introduciendo el cálculo de las diferencias en las muestras escogidas, aunque el cálculo no sea explícito.

Segundo grupo (G2)

El grupo G2 establece un primer modelo para valorar la adecuación de los impuestos ya en la primera lectura del enunciado, en la fase de interpretación de la tarea (Figura 9). En concreto, plantean un modelo que utiliza como concepto la media, M, de los valores máximos y mínimos de los impuestos de cada provincia (minuto 4:30).

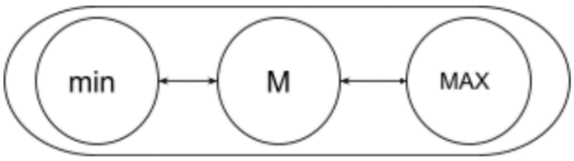

Figura 9. Primera versión del modelo - Grupo 2

Cuando se ponen a trabajar con los datos, el grupo G2 utiliza todos los datos proporcionados, sobre el que trabajan usando la hoja de cálculo (Figura 10). En ningún momento discuten sobre la posibilidad de reducir el conjunto de datos a utilizar, como si no fuera una opción. De esta forma, el estudio de toda la población es una elección implícita y sucede mientras están llevando a cabo los cálculos, utilizando las matemáticas.

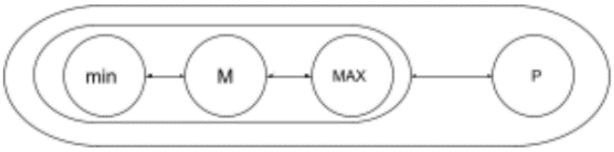

Figura 10. Segunda versión del modelo - Grupo 2

Mientras trabajan con la hoja de cálculo, observan que obtener un valor concreto que separe las dos opciones que se plantean para decidir si un impuesto es justo no les permite representar adecuadamente la variabilidad de los datos dentro de cada provincia. Por ello, deciden modificar el modelo utilizado (minuto 14:35) e incorporan al modelo un intervalo, I, centrado en la media calculada previamente (Figura 11). La elección de la anchura de este intervalo es subjetiva, no depende de un cálculo específico y responde a la observación de las semejanzas o diferencia entre los datos.

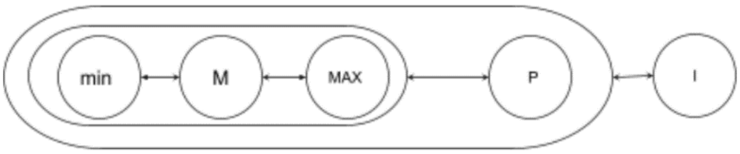

Figura 11. Versión final del modelo - Grupo 2

A partir de este momento, los integrantes del grupo G2 clasifican metódicamente los datos según pertenezcan o no al intervalo. Así aplican las matemáticas utilizando una hoja de cálculo hasta aproximadamente el minuto 25. Al cabo de cinco minutos en que el grupo está escribiendo el informe se dan cuenta que no lo han hecho de todas las columnas y deciden dejar ahí los cálculos y, con lo que tienen, generar conclusiones. En el informe escrito revelan cómo han utilizado este margen para poder decidir que la situación no es regular en toda España. Los alumnos proponen diferentes soluciones, pero en una de ellas se intuye cómo ese margen lo entienden como una necesidad, no todos los impuestos deben ser iguales si la distribución de riqueza en el territorio no lo es.

Tercer grupo (G3)

En el grupo G3 durante los seis primeros minutos leen el artículo, miran algunos datos y las preguntas de la tarea. Para determinar si las tarifas son justas o no deciden ir mirando cada municipio que aparece en la tabla del artículo. El primer elemento añadido al modelo es el estudio de todos los datos de la población (P). Aunque en este momento del trabajo el grupo G3 solo está considerando los datos disponibles en el artículo, donde aparece una tabla con las ciudades capital de provincia. No tienen en cuenta la hoja de cálculo aportada en la actividad, por tanto, para ellos todos los datos son los datos ofrecidos en el artículo. El primer criterio para decidir si una provincia es justa es la renta per cápita (RC). El siguiente episodio es una muestra de su discusión:

D:	No… Alicante, tampoco… Ávila, menos. Badajoz, sí.
V:	Vale, apunta Badajoz, vale. Porque el impuesto de esta ciudad es… ¿qué ponemos? ¿Es justa? ¿Por qué?
D:	Porque pertenece a una comunidad autónoma... con la renta per cápita (RC) más baja.

De esta forma acuerdan que a menor renta per cápita, más justa es la provincia en consideración, pero sin explicitar una vía para decidir el valor del impuesto necesario para que deba considerarse justa. En este punto del análisis de datos, los alumnos trabajan con una idea que comparten pero que no han explicitado ni para la que desarrollan un método de cálculo que permita una concreción: que debe existir una relación entre la renta per cápita en una provincia y los impuestos que se aplican a sus ciudadanos (Figura 12).

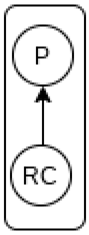

Figura 12. Primera versión del modelo - Grupo 3

En la discusión de diversos elementos relevantes del contexto estudiado van observando la complejidad de la situación y consideran ampliar el conjunto de criterios a considerar. El modelo se amplía añadiendo criterios geográficos (Geo), después de observar que los impuestos en las zonas insulares les parecen más justos, así como la situación de las ciudades autónomas. Consideran relevante también la existencia de ciudades en las que determinados tipos de vehículos están exentos del pago de impuestos (Ex) y también son tomadas como más justas. El número de habitantes (Hb) es el último criterio socioeconómico utilizado en el modelo.

Este grupo desarrolla un modelo basado en establecer relaciones entre el valor del impuesto y diversas variables socioeconómicas de cada provincia, pero esquiva el análisis matemático de los datos sin utilizar procedimientos de cálculo de ningún tipo. En su lugar, utilizan unas asignaciones intuitivas de los que supone un impuesto con un valor alto o bajo a partir de comparar los datos proporcionados en las tablas. Por ejemplo, en el minuto 7:30 deciden que Badajoz es una provincia justa respecto a su impuesto de circulación porque pertenece a una comunidad autónoma con una de las rentas per cápita más bajas del estado español y el valor del impuesto se ajusta a esa valoración (df). En una situación similar, en el minuto 9:05 deciden que Ceuta es una ciudad justa porque es una ciudad autónoma fuera de la península Ibérica, con las dificultades económicas que eso supone (df). De esta forma, el modelo con el que trabajan se muestra en la Figura 13.

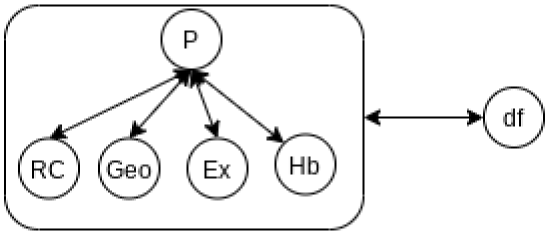

Figura 13. – Versión final del modelo - Grupo 3

Conforme van estudiando casos concretos van observando la complejidad de la situación y añaden nuevas variables, que no utilizan de manera relacionada, es decir seleccionan una opción por uno u otro índice. Esta forma de trabajar, que no incluye cálculos específicos permite que avancen rápidamente en la resolución del problema. Hecho que se refuerza al dividir el trabajo: dos estudiantes miran los datos y dictan al tercero qué han decidido. Así que las fases del ciclo quedan solapadas una vez han entendido la tarea, interpretando los datos bajo su criterio y validando la conclusión entre ellos.

Resultados: elementos clave de la actividad para la generación de modelos

Observamos que los grupos G1 y G2 desarrollan modelos matemáticos usando procedimientos específicos para dar respuesta a la problemática planteada. Por su parte, el grupo G3 desarrolla los conceptos que darían forma a un modelo matemático (Lesh & Harel, 2003) pero sin utilizar procedimientos que involucren cálculos específicos, con lo que trabajan a partir de un modelo informal. En este sentido, los grupos G1 y G3 identifican un conjunto de provincias con impuestos más bajos en relación con otras características relevantes (población o renta per cápita). En cambio, el grupo G2 no proporciona un listado concreto, se centra en generar un criterio complejo por el que poder determinar si los impuestos en una provincia pueden considerarse justos.

En segundo lugar, la actividad de los alumnos se focaliza mayoritariamente en medir las diferencias entre los valores de los impuestos proporcionados. Por lo tanto, genera la necesidad de cuantificar la variabilidad de los datos. De esta forma, observamos que la resolución de esta actividad sigue las fases del ciclo de modelización y genera la necesidad de estudiar nociones estadísticas clave para la comprensión de conceptos fundamentales en la estadística, en este caso, la variabilidad.

El análisis del proceso, juntamente con los modelos creados de los tres grupos, nos permite identificar aquellos elementos de la actividad, propios de su diseño, que han obligado a los estudiantes a tomar decisiones que les han permitido desarrollar con más precisión el modelo matemático que construían. En concreto, los elementos del diseño de la actividad que provocan que los alumnos inciden en el modelo matemático que van a usar para justificar sus respuestas son los siguientes:

- *Gran volumen de datos proporcionados*: Los grupos de trabajo se plantean formas efectivas o viables de analizar una cantidad de datos de gran volumen como la que se les ha proporcionado. Se puede observar en el análisis del trabajo de los dos primeros grupos. El grupo G3 no se percató de la existencia de los datos proporcionados en la hoja de cálculo y, por tanto, la cantidad de datos a tratar era mucho menor. Los alumnos pueden ser conscientes de esta dificultad desde que leen el enunciado, procurando entender la tarea asignada, o puede ser que se den cuenta de ello mientras exploran los datos o incluso en una fase posterior. En todos estos casos la gran cantidad de datos afecta al trabajo de los alumnos que, al no conocer métodos específicos para enfrentarse a esta situación, buscan estrategias para esquivar el problema o replantearse el método usado. De esta forma, se sustituyen las dificultades de recolectar datos por las de tratar y manipular datos (Hahn, 2015; Muñiz-Rodríguez et al, 2020; Ubilla, 2020) al introducir una característica distinta, la dificultad de gestionar una gran cantidad de datos.
- *Indefinición del concepto de justicia*: El enunciado de la actividad incluye la expresión "impuestos justos", pero en este contexto el concepto de justicia no está claramente definido, con lo que los obliga a los alumnos a generar su propia concepción de lo que se entiende por "impuestos justos". Los alumnos discuten también esta concepción durante la actividad, tanto cuando interpretan el enunciado como en momentos de validación de la tarea realizada. Y a raíz de las discusiones se replantean la estrategia a seguir. La indefinición del concepto es la que promueve que los alumnos exploren diferentes vías para responder a la demanda de la actividad, pero, al mismo tiempo, les conduce hacia el concepto de variabilidad de datos. De esta forma, se presenta implícitamente la necesidad de estudiar la variabilidad de los datos proporcionados sin la necesidad de proporcionar una herramienta específica para ello, como podría ser el uso de la desviación típica. En este sentido la actividad cumple con el objetivo de promover la creación de modelos matemáticos que puedan servir como base para interpretar los conceptos habituales del estudio de la variabilidad y podrán usarse como referencia, a modo de andamiaje.
- *Complejidad de la realidad estudiada*: Los datos proporcionados se presentan acompañados de un contexto complejo en el que no se proporciona información específica sobre los factores que influyen en la variabilidad de los datos. Esto supone enfrentar a los alumnos una situación compleja en la que observan que toda reducción de la complejidad cambia el problema, con lo que deben tratar de trabajar sin reducirla. La actividad propuesta se presenta de forma similar a la que tendría que estudiar un profesional de la Estadística y son los propios alumnos los que deben tomar las decisiones para simplificarla. Este hecho se evidencia en las dificultades

que presentan los alumnos al tratar de determinar muestras de los datos para obtener resultados en las fases de estructurar o matematizar la tarea, ya que consideran diversas variables socioeconómicas y observan que deben considerar un gran número de ellas para asegurar que la muestra puede representar adecuadamente a cada conjunto de datos estudiados.

Conclusiones

En este estudio hemos analizado el trabajo de alumnos de 4.º curso de Educación Secundaria Obligatoria en España al enfrentarse a una tarea en la que deben tratar con una gran cantidad de datos y cuantificar la variabilidad de estos datos. Para ello hemos usado un análisis que nos permite estudiar la evolución de los modelos matemáticos generados para identificar las características de la actividad de modelización que promueven la generación de modelos.

Los resultados del estudio muestran que los alumnos generan sus propios modelos matemáticos para resolver la situación planteada empleando conceptos matemáticos que conocen de su experiencia anterior. Pero lo hacen innovando en la forma de combinarlos para conseguir sus objetivos. Es el caso de los parámetros de centralidad y la generación de intervalos. También observamos que los alumnos identifican elementos matemáticos clave para dar forma a su modelo, aunque no conozcan un procedimiento específico para concretarlo formalmente. De esta forma, los alumnos generan modelos matemáticos propios para resolver una situación que requiere conocimientos matemáticos específicos. Los elementos clave que provocan esta forma de proceder de los alumnos tienen diferente naturaleza. Por una parte, se encuentran los elementos propios del diseño de la actividad, como proporcionar una gran cantidad de datos o referirse a un concepto ambiguo como es el de una asignación justa de impuestos, que promueve las discusiones entre los alumnos. Por otro lado, identificamos un elemento clave al elegir el contexto en el que basar el trabajo proponiendo una temática real y compleja que claramente es multifactorial y que provoca que no sea sencillo elegir muestras representativas para trabajar.

Las investigaciones futuras deberían ayudarnos a decidir si estos alumnos han empezado a desarrollar conocimientos propios sobre el papel de las herramientas estadísticas en el estudio de datos. La hipótesis que guía nuestro trabajo es que los alumnos se habrán iniciado con éxito en la modelización de grandes cantidades de datos contextualizados. Al haber generado sus propios modelos han identificado aspectos que definen el concepto de variabilidad de datos para concretarlo posteriormente de forma significativa con las herramientas curriculares propias de la estadística y conectar su experiencia en esta actividad con esos conocimientos.

Agradecimentos

Lluís Albarracín es profesor Serra Húnter en la Universitat Autònoma de Barcelona y agradece el soporte de los proyectos EDU2017-82427-R (Ministerio de Economía, Industria y Competitividad, Spain) y 2017 SGR 497(AGAUR, Generalitat de Catalunya).

Referencias

Abassian, A., Safi, F., Bush, S., & Bostic, J. (2020). Five different perspectives on mathematical modeling in mathematics education. *Investigations in Mathematics Learning, 12*(1), 53-65. https://doi.org/10.1080/19477503.2019.1595360

Albarracín, L., Aymerich, À., & Gorgorió, N. (2017). An open task to promote students to create statistical concepts through modelling. *Teaching Statistics, 39*(3), 100-105. https://doi.org/10.1111/test.12136

Aymerich, À., & Albarracín, L. (2016). Complejidad en el proceso de modelización de una tarea estadística. *Modelling in Science Education and Learning, 9*(1), 5-24. https://doi.org/10.4995/msel.2016.4121

Aymerich, À., Gorgorió, N., & Albarracín, L. (2017). Modelling with statistical data: Characterisation of student models. In G. Stillman, W. Blum, & G. Kaiser (Eds.), *Mathematical modelling and applications* (pp. 37-47). Springer. https://doi.org/10.1007/978-3-319-62968-1_3

Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27*, 151–169. https://doi.org/10.5951/jresematheduc.27.2.0151

Blum, W. (2002). ICMI Study 14: Applications and modelling in mathematics education–Discussion document. *Educational Studies in Mathematics, 51*(1), 149-171. https://doi.org/10.1023/A:1022435827400

Blum, W. (2015). Quality teaching of mathematical modelling: What do we know, what can we do? In S. J. Cho, (Ed.), *Proceedings of the 12th International Congress on Mathematical Education* (pp. 73-96). Springer. https://doi.org/10.1007/978-3-319-12688-3_9

Blum, W., & Borromeo Ferri, R. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application, 1*(1), 45-58.

Blum, W., & Leiβ, D. (2006). How do students and teachers deal with modeling problems? In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical Modeling (ICTMA12): Education, Engineering and Economics* (pp. 222–231). Horwood Publishing. https://doi.org/10.1533/9780857099419.5.221

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects – State, trends and issues in mathematics instruction. *Educational Studies in Mathematics, 22*(1), 37-68. https://doi.org/10.1007/BF00302716

Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *ZDM–The International Journal on Mathematics Education, 38*(2), 86-95. https://doi.org/10.1007/BF02655883

Burrill, G., & Biehler R. (2011), Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics, Challenges for teaching and teacher education* (pp. 57-69). Springer. https://doi.org/10.1007/978-94-007-1131-0_10

Carreira, S., Amado, N., & Lecoq, F. (2011). Mathematical modeling of daily life in adult education: Focusing on the notion of knowledge. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modeling* (pp. 199-210). Springer.

Cobb, G., & Moore, D. (1997). Mathematics, statistics and teaching. *The American Mathematical Monthly, 104*(9), 801–823. https://doi.org/10.2307/2975286

Crites, T., & Laurent, R. T. (2015). *Putting essential understanding of statistics into practice, Grades 9-12*. National Council of Teachers of Mathematics.

Doerr, H. M., & English, L. D. (2003). A modeling perspective on students' mathematical reasoning about data. *Journal for Research in Mathematics Education, 34*(2), 110-136. https://doi.org/10.2307/30034902

Doerr, H. M., & Tripp, J. S. (1999). Understanding how students develop mathematical models. *Mathematical Thinking and Learning, 1*, 231–254. https://doi.org/10.1207/s15327833mtl0103_3

Dvir, M., & Ben-Zvi, D. (2018). The role of model comparison in young learners' reasoning with statistical models and modeling. *ZDM Mathematics Education, 50*(7), 1183-1196. https://doi.org/10.1007/s11858-018-0987-4

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Springer. https://doi.org/10.1007/1-4020-2278-6_3

Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modeling process. *ZDM–The International Journal on Mathematics Education, 38*(2), 143-162. https://doi.org/10.1007/BF02655886

Greefrath, G. (2011). Using technologies: New possibilities of teaching and learning modeling – Overview. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modeling* (pp. 301–304). Springer. https://doi.org/10.1007/978-94-007-0910-2_30

Hahn, C. (2015). La recherche internationale en éducation statistique: État des lieux et questions vives. *Statistique et Enseignement, 6*(2), 25-39.

Hernández-Sabaté, A., Albarracín, L., & Sánchez, F. J. (2020). Graph-based problem explorer: A software tool to support algorithm design learning while solving the salesperson problem. *Mathematics, 8*(9), 1595. https://doi.org/10.3390/math8091595

Kaiser, G., & Sriraman, B. (2006). A global survey of international perspectives on modelling in mathematics education. *ZDM–The International Journal on Mathematics Education, 38*(3), 302-310. https://doi.org/10.1007/BF02652813

Kaiser, G., & Stender, P. (2013). Complex modelling problems in co-operative, self-directed learning environments. In G. A. Stillman, G. Kaiser, W. Blum, & J. Brown (Eds.), *Teaching mathematical modeling: Connecting to research and practice. International perspectives on the teaching and learning of mathematical modeling* (pp. 277-293). Springer. https://doi.org/10.1007/978-94-007-6540-5_23

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning, 2*, 51–74. https://doi.org/10.1207/S15327833MTL0202_3

Lesh, R. (1997). Matematización: La necesidad "real" de la fluidez en las representaciones. *Enseñanza de las Ciencias, 15*(3), 377-391.

Lesh, R., Amit, M., & Schorr, R. Y. (1997). Using "real-life" problems to prompt students to construct statistical models for statistical reasoning. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 65–84). IOS Press.

Lesh, R., & Harel, G. (2003). Problem solving, modeling, and local conceptual development. *Mathematical Thinking and Learning, 5*(2-3), 157-189. https://doi.org/10.1080/10986065.2003.9679998

Makar, K., & Confrey, J. (2005). Variation talk: Articulating meaning in statistics. *Statistics Education Research Journal, 4*(1), 27-54.

Muñiz-Rodríguez, L., Rodríguez-Muñiz, L. J., & Alsina, Á. (2020). Deficits in the statistical and probabilistic literacy of citizens: Effects in a world in crisis. *Mathematics, 8*(11), 1872. https://doi.org/10.3390/math8111872

Pollak, H. O. (1969). How can we teach applications of mathematics? *Educational Studies in Mathematics, 2*, 393-404. https://doi.org/10.1007/BF00303471

Pollak, H. O. (1979). The interaction between mathematics and other school subjects. In H. G. Steiner & B. Christiansen (Eds.), *New trends in mathematics teaching IV* (pp. 232-248). UNESCO.

Ramírez-Montes, G., Carreira, S., & Henriques, A. (2021). Mathematical modelling routes supported by technology in the learning of linear algebra: A study with Costa Rican undergraduate students. *Quadrante, 30*(1), 219-241. https://doi.org/10.48489/quadrante.23721

Ubilla, F. (2019). Componentes del sentido estadístico identificados en un ciclo de investigación estadística desarrollado por futuras maestras de primaria. In J. M. Marbán, M. Arce, A. Maroto, J. M. Muñoz-Escolano, & Á. Alsina (Eds.), *Investigación en Educación Matemática XXIII* (pp. 583-592). SEIEM.

Ubilla, F. (2020). ¿Qué rol juegan los datos en el ciclo de investigación estadística? *UNO, 91*, 63-68.

Zawojewski, J., Lesh, R., & English, L. D. (2003). A models and modeling perspective on the role of small group learning activities. In R. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics teaching, learning, and problem solving* (pp. 337-358). Lawrence Erlbaum Associates.