

Modelización matemática en actividades estadísticas: Episodios clave para la generación de modelos

Mathematical Modeling in Statistical Activities: Key Episodes for Model Generation

Modelagem matemática em atividades estatísticas: Principais episódios para a geração de modelos

Àngels Aymerich ¹, Lluís Albarracín ¹

Received: Jun/7/2021 • Accepted: Set/1/2021 • Published: Jan/31/2022

Resumen

En nuestro trabajo estamos interesados en promover el aprendizaje de conocimientos estadísticos por parte de alumnos de educación secundaria a partir de la modelización matemática. En concreto, usamos problemas estadísticos abiertos donde se estudian fenómenos sociales reales a partir de grandes cantidades de datos que cumplen con los principios de diseño de las *model-eliciting activities* con alumnos del último curso de educación obligatoria en España (15 años). Las tareas se dirigen a promover el desarrollo del concepto de variabilidad y aplicarlo para entender la situación estudiada a partir de un modelo matemático. En este estudio nos centramos en identificar los episodios clave de la actividad en los que se avanza en la construcción de los modelos matemáticos y los elementos que los promueven. Para ello se desarrolla un análisis cualitativo a partir de las grabaciones del trabajo en grupo de los alumnos en el aula a partir del uso de diagramas de actividad modelizadora. 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 artículo tienen implicaciones para el diseño de tareas estadísticas, pero también para identificar el rol de la modelización matemática en el aprendizaje de conceptos estadísticos.

Palabras clave: modelización matemática; actividades promotoras de la modelización; diagramas de actividad modelizadora; educación secundaria; estadística; variabilidad

Abstract

This work is intended to assist 15-year-old secondary school students to learn about statistics using mathematical modeling. Specifically, open statistical problems are presented in which real social phenomena are studied by students using large data sets that comply with the design principles of *model-eliciting activities*. The tasks assigned are aimed at presenting the concept of variability and its application to understanding the situations studied using a mathematical model. The study focuses on identifying key episodes in the activities in which progress is made in the construction of mathematical models, and the elements that promote them. To do so, a qualitative analysis is carried out based on records of the students' group work in the classroom using Modeling Activity Diagrams. The results obtained show that decisions about the design of a problem, such as using large amounts of data, or the ambiguity of social concepts such as "fair taxation," are essential for promoting the development of mathematical models. The conclusions of this investigation have implications for the

design of statistical tasks, and also for identifying the role of mathematical modeling in the learning of statistical concepts.

Keywords: mathematical modeling; model-eliciting activities; Modeling Activity Diagrams; secondary education; statistics; variability

Resumo

Em nosso trabalho estamos interessados em promover a aprendizagem do conhecimento estatístico por estudantes do ensino médio a partir da modelagem matemática. Especificamente, utilizamos problemas estatísticos abertos onde fenômenos sociais reais são estudados a partir de grandes quantidades de dados que cumprem os princípios de desenho das *eliciting activities* com os alunos do último ano de educação obrigatória na Espanha (15 anos). As tarefas visam promover o desenvolvimento do conceito de variabilidade e aplicá-lo para compreender a situação estudada a partir de um modelo matemático. Neste estudo, focamos na identificação dos principais episódios da atividade em que se avança na construção de modelos matemáticos e de elementos que os promovem. Para isso, é desenvolvida uma análise qualitativa com base nas gravações do trabalho em grupo dos alunos em sala de aula a partir do uso de diagramas de atividade de modelagem. Os resultados obtidos mostram que algumas decisões de desenho do problema, como a grande quantidade de dados, ou a ambiguidade de conceitos sociais, como justiça e equidade em impostos, são essenciais para o desenvolvimento de modelos matemáticos. As conclusões do artigo têm implicações para o desenho de tarefas estatísticas, mas também para identificar o papel da modelagem matemática na aprendizagem de conceitos estatísticos.

Palavras-chave: modelagem matemática; atividades fomentadoras de modelagem; diagramas de atividade de modelagem; ensino médio; estatística; variabilidade

Introducción

Nuestra sociedad vive un auge del uso y necesidad de la Estadística. Los conocimientos y métodos estadísticos son parte de la capacidad de entender e interpretar informaciones estadísticas y actúan como una herramienta clave para formar futuros ciudadanos. Por ello, en tiempos recientes los currículos de varias partes del mundo (Burrill y Biehler, 2011) han incidido en potenciar los contenidos estadísticos, incluyendo recomendaciones específicas para su implementación y su trabajo en el aula (Gal, 2004).

En nuestro estudio tomamos el siguiente posicionamiento para enfrentarnos al reto de la enseñanza de la Estadística en las aulas de educación secundaria: promover que los alumnos construyan sus propios conceptos y métodos estadísticos para estudiar situaciones realistas 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 actividades en las que se proporcionan grandes cantidades de datos (Albarracín, Aymerich, y Gorgorió, 2017; Aymerich, Gorgorió, y Albarracín, 2017). Estos estudios ponen de manifiesto que es posible utilizar actividades promotoras de modelos para que los alumnos generen sus propios conceptos y métodos estadísticos para estudiar la variabilidad.

Por ello, el propósito de nuestra investigación es identificar actividades de aula en la que los alumnos de educación secundaria generan sus propios conceptos estadísticos para entender la variabilidad de datos como paso previo a la comprensión de los conceptos estadísticos incluidos en el currículo. Tomamos como referencia teórica las actividades promotoras de modelos (Lesh, 1997) entendiendo que será necesario introducir

características específicas en los enunciados, los materiales proporcionados y la gestión de aula por parte del profesor. De esta forma, el objetivo específico de este estudio es identificar aquellos episodios del proceso de modelización que permiten a los alumnos generar los elementos clave del modelo matemático que desarrollan para estudiar la variabilidad de un fenómeno real. Estos episodios clave son los que nos permiten concretar los aspectos del problema que realmente promueven la generación de modelos matemáticos por parte de los alumnos.

Marco teórico

En esta sección detallamos los distintos referentes que sustentan nuestra investigación. Entre ellos se encuentran el uso de investigaciones estadísticas para promover la alfabetización estadística entre los alumnos y la modelización matemática, de la que destacamos las actividades promotoras de modelos y el uso de los diagramas de actividad modelizadora como herramienta de análisis.

Investigaciones estadísticas y modelización matemática

La forma en 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 (Batenero, Estepa, Godino & Green, 1996; Ben-Zvi & Arcavi, 2001; Doerr & Tripp, 1999; Lehrer & Romberg, 1996; Lehrer & Schauble, 2000; Lesh, Amit & Schorr, 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 y 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. En los últimos años, los esfuerzos se han centrado en conjugarlos en un concepto denominado “sentido estadístico” (Batenero, Díaz, Contreras y Roa, 2013).

El concepto de alfabetización estadística ha sido desarrollado por diversos autores en las últimas décadas. Gal (2002) define la alfabetización estadística como la habilidad para interpretar y evaluar críticamente la información estadística con argumentos relacionados con datos o fenómenos estocásticos, que ocurren en diversos contextos; habilidad para discutir y comunicar las reacciones a esta información, así como entender el significado de la información y establecer opiniones sobre la implicación de estos resultados. Batenero, Díaz, Contreras y Roa (2013) consideran que un adulto está alfabetizado estadísticamente si tiene la capacidad de, por una parte, interpretar y evaluar críticamente información estadística y, por otra, discutir o comunicar su reacción a esta información, así como sus preocupaciones con respecto a la aceptabilidad de las conclusiones dadas.

Watson (2011) propone que al final de la educación obligatoria (16 años en España) la alfabetización estadística debe consistir en entender los problemas y terminologías básicas de la estadística, usarlos en el mundo real y cuestionar las conclusiones obtenidas. En este sentido, se ha insistido en que para promover la alfabetización estadística es necesario usar

datos reales o realistas, significativos para el alumnado y presentados en contextos motivadores, en lugar de usar datos simplistas, omitiendo deliberadamente todas las dificultades derivadas de los procesos de recolección o muestreo (Hahn, 2015; Muñiz-Rodríguez, Rodríguez-Muñiz y Alsina, 2020); aunque también se ha propuesto separar las actividades de investigación estadística en las que se recogen datos de aquellas en las que se procesan datos proporcionados a los alumnos (Ubilla, 2020).

El razonamiento estadístico se enfoca en el trabajo con datos y describe la habilidad de los estudiantes para establecer conexiones entre los conceptos estadísticos, crear representaciones mentales de los problemas estadísticos y explicar las relaciones entre los conceptos estadísticos (Garfield y Ben-Zvi, 2007; delMas 2002; 2004). Razonar estadísticamente requiere de usar coordinadamente más de un concepto estadístico y establecer conexiones entre ellos.

Finalmente, destacamos el concepto de pensamiento estadístico (Wild y Pfannkuchen, 1999) como una propuesta más ambiciosa desde el punto de vista del conocimiento y la competencia estadísticos. El pensamiento estadístico incluye el conocimiento de cómo y por qué utilizar un método en particular, una medida, un diseño o un modelo estadístico. Incluye entender en profundidad las teorías subyacentes a los procesos y métodos estadísticos, así como sus limitaciones y contrastes. El pensamiento estadístico también trata de cómo los modelos estadísticos se utilizan para simular fenómenos aleatorios, entender cómo los datos se producen, estimar probabilidades, reconocer cómo, cuándo y por qué existen y pueden ser utilizadas las herramientas de inferencia estadística, facilitar entender y utilizar el contexto de un problema o plan, evaluar las investigaciones y dar conclusiones. El pensamiento estadístico describe desde diferentes dimensiones el proceso y posicionamiento del estudiante en las investigaciones estadísticas y destaca la importancia del contexto (Wild y Pfannkuchen, 1999). Por este motivo, las investigaciones estadísticas conectan con los procesos de modelización matemática poniendo el foco de atención en la organización y representación de datos, la construcción de patrones y la búsqueda de relaciones (Lesh y Doerr, 2003), e involucran a los estudiantes en el razonamiento estadístico como la toma de decisiones, la inferencia y la predicción.

Actividades promotoras de la modelización

La modelización matemática es un proceso de resolución de problemas contextualizados en la que se elabora un modelo matemático para describir el fenómeno real estudiado. Lesh y Harel (2003) ofrecen una definición de los modelos matemáticos que generan los alumnos como sistemas conceptuales complejos 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.

En este trabajo consideramos una perspectiva teórica para interpretar las actividades de modelización matemática denominada Modelos y Modelización (M&M), donde 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 y English, 2003). Los problemas o proyectos 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: *model-eliciting activities*) y se centran tanto en el proceso de modelización como en 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, Lesh y English, 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. En consecuencia, los estudiantes tienden a pasar por secuencias de ciclos de interpretación-desarrollo en las que los datos, objetivos y procesos de solución relevantes se piensan de diferentes maneras. Una vez obtenido el modelo, los estudiantes deben determinar si este solo es aplicable a la situación particular que se presenta en el problema o si, por el contrario, el modelo puede compartirse, trasladarse, modificarse fácilmente y reutilizarse para otros tipos de problemas o situaciones.

Para diseñar este tipo de actividades nos basamos en los seis principios básicos de Lesh, Amit y Schorr (1997), que son los siguientes:

Principio de realidad: Responde a la pregunta de si el problema puede ocurrir en la vida real. Los estudiantes se verán motivados por experiencias y conocimientos personales. Los estudiantes deberán testar y confrontar sus ideas para encontrar la solución que desea el autor de la actividad.

Principio de 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 o predecir un sistema significativo.

Principio de autoevaluación: Debe haber 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.

Principio de 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.

Principio de generalización del modelo: El modelo que construyen debe poder ser aplicado o adaptado a un abanico mayor de situaciones.

Principio de 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 similares.

Diagramas de actividad modelizadora

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 & Leiss, 2006; Carreira, Amado & Lecoq, 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.

Ante la necesidad de describir los procesos de modelización que implementan los alumnos, Ärlebäck (2009) propone el uso de los diagramas de actividad modelizadora (DAM), basados en los protocolos de codificación de Schoenfeld (1985), como un gráfico bidimensional en el que se representa el tipo de actividad para cada intervalo de tiempo de una sesión de modelización. Originalmente, Schoenfeld (1985) considera actividades propias del proceso de resolución de un problema matemático, como leer el enunciado, explorar vías de resolución, implementar un plan de acción o verificar una respuesta. Modificando las actividades a considerar para ajustarlas al trabajo de modelización matemática, las actividades propuestas por Ärlebäck (2009) para caracterizar el proceso de modelización en un DAM son las siguientes: leer (R), desarrollar un modelo (M), estimar (E), calcular (C), validar (V) y redactar (W). La figura 1 muestra un DAM que pone de manifiesto la complejidad del proceso matemático desarrollado por los alumnos y que permite cualificar el trabajo de modelización de los alumnos como idiosincrático, en relación tanto a las características de los alumnos como protagonistas de la actividad matemática como por la propia naturaleza del problema que genera la actividad.

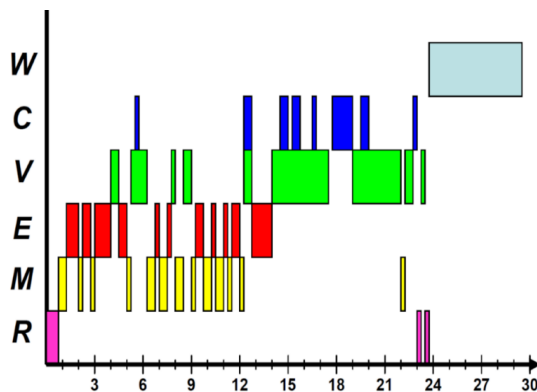


Figura 1. Diagrama de actividad modelizadora para la resolución del problema de la estimación de tiempo necesario para subir al Empire State Building a pie. Extraído de Ärlebäck (2009).

Los DAM se han adaptado a otras actividades de modelización introduciendo categorías de análisis y factores distintos de los originales (Albarracín, Ärlebäck, Civil y

Gorgorió, 2019), con lo que se abre la vía de adaptarlos específicamente a las características de cada proceso analizado, como planteamos en este estudio.

Diseño y justificación del problema utilizado

La actividad utilizada trata sobre la variabilidad del impuesto de circulación en España. El impuesto de circulación es municipal, con lo que depende de cada ayuntamiento y puede variar en función de las necesidades y características socioeconómicas del municipio. 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 su cantidad de población. Además del impuesto de vehículos de tracción mecánica (de turismo, tractores, camiones, remolques, ciclomotores, autobuses y motocicletas), proporciona: el impuesto sobre bienes 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).

También se ofrece a los alumnos una lectura en forma de artículo que analiza el hecho de que en poblaciones cercanas se pueden encontrar grandes diferencias en el valor de este impuesto. Se proveen ejemplos concretos de este hecho para ilustrar la situación. Con esta información, se pregunta a los alumnos qué municipios tienen un impuesto de circulación justo para sus habitantes. Dado el interés en recoger datos para este estudio, la tarea se divide en dos fases. 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 ciñéndose a las decisiones tomadas, generen resultados y redacten un informe para evaluar la situación y presentar propuestas en las que se valoren soluciones posibles para minimizar las desigualdades.

PROVINCIA	TARIFAS IMPUESTO DE CIRCULACIÓN 2018										
	TURISMOS (CVF)					OCIO	MOTOCICLETAS (C.C)				
	De menos de 8	De 8 a 11,99	De 12 a 15,99	Más 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	8,48	13,79	27,58	55,17	111,39
ALMERÍA	23,12	62,41	131,78	184,16	235,11	8,10	8,10	13,98	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
BADAJOS	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
CACERES	17,00	50,00	100,00	125,00	140,00	8,00	8,50	11,00	23,00	43,00	84,00
CADIZ	24,90	67,25	142,00	177,05	220,90	8,60	8,60	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	26,74	67,24	142,40	178,00	224,00	8,84	8,84	15,04	30,00	60,00	120,00

Figura 2. Tarifas del impuesto municipal de vehículos de tracción mecánica. Extraído de Automovilistas Europeos Asociados (2019).

La actividad sigue la propuesta de trabajos anteriores (Albarracín, Aymerich y Gorgorió, 2017), en la que la tarea cumple con las características que definen las MEA (Lesh, Amit y Schorr, 1997) al presentarse una situación real y contextualizada que debe ser modelada con el objetivo de presentar el modelo final como un producto útil en esta situación.

De esta forma, la actividad pretende apoyar la construcción de intuiciones correctas sobre conceptos estadísticos (Hernández-Solís, Batanero, Gea, y Álvarez-Arroyo, 2021). Al mismo tiempo, el trabajo propuesto a los alumnos exige un análisis de datos estadísticos en el que se genere un modelo matemático para evaluar la adecuación de los impuestos a la realidad, con lo que se trabaja en la dirección de que los alumnos desarrollen tanto su pensamiento estadístico al trabajar con datos generando conceptos propios (Garfield y Ben-Zvi, 2007; delMas 2002; 2004) como su razonamiento estadístico, al aprender qué conceptos son los adecuados en cada situación a estudiar (Wild y Pfannkuchen, 1999).

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 4º de ESO (15-16 años) que trabajan en grupos de 3 personas en sesiones de 60 minutos. El trabajo se desarrolla en diversas sesiones, variables en función del ritmo de cada grupo. En este artículo estudiamos la fase inicial en la que los alumnos deciden cómo dar respuesta al problema de forma grupal. 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, el documento en bruto y el redactado del informe final. Para el análisis se utilizan los datos de 3 de los grupos participantes, que denominamos grupo 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 de que permiten entender claramente las aportaciones de los alumnos durante toda la actividad sin generar posibles lagunas en el análisis.

Análisis

A partir de los datos recogidos, en este apartado mostramos el proceso de análisis de las grabaciones y la forma en la que se codifican las diferentes actividades de modelización de los DAM (Ärlebäck, 2009). En concreto, en esta sección nos interesa poner de manifiesto las decisiones tomadas por los investigadores al determinar la categoría (actividad) identificada en cada momento de la grabación, así como mostrar parte de los datos recogidos.

La primera decisión es la de elegir las actividades que conforman las categorías de análisis. Dado que la actividad utilizada se centra en el análisis estadístico de datos, tomamos las categorías del ciclo de investigación estadístico de Wild y Pfannkuchen (1999). De esta forma, las categorías de análisis son las siguientes: interpretar el enunciado, explorar los datos, planificar, implementar el análisis y generar conclusiones. Entendemos que la categoría de planificación de los procedimientos es aquella en la que los alumnos podrían desarrollar un modelo matemático al definir los procedimientos con los que se van a enfrentar al análisis de datos (Lesh y Harel, 2003). La Tabla 1 muestra la descripción de las categorías de análisis utilizadas.

Tabla 1
Categorías de análisis utilizadas.

Interpretar (I)	Leen, escuchan y ven. Adaptan el lenguaje a su vocabulario. Resumen. Comparan con otras situaciones similares. Conectan con ideas previas. Buscan información para entender mejor el enunciado.
Explorar datos (D)	Exploración visual en la que se fijan en lo que la experiencia o sus prejuicios les dicen. Cálculos exploratorios con datos.
Planificar (P)	Trazan un plan de acción que puede incluir en mayor o menor medida la estadística. Modificaciones del plan de acción.
Implementar el análisis (A)	Cálculos, organización de los datos, generación de gráficos, toda actividad que produzca resultados.
Generar conclusiones (C)	A partir de la observación e interpretación de los datos y los resultados del análisis, generan conclusiones y las redactan.

Nota: Fuente propia de la investigación.

Para analizar las discusiones de los alumnos, las grabaciones se transcriben anotando cada una de las intervenciones y su autor. Cada una de las tres grabaciones analizadas pasa a ser un listado de intervenciones ordenadas en el tiempo. Para dar significado a las intervenciones, se agrupan en paquetes que tratan sobre el mismo objeto de discusión, dando lugar a episodios como intervalos de tiempo en los que se discute una misma idea o proceso. A partir de esta estructuración en episodios se procede a su categorización, identificando la idea central de la discusión en cada uno de los casos.

A continuación, mostramos un de estos episodios y su análisis. Entre las marcas de tiempo 16:12 y 16:51 el grupo G1 se replantea la forma en la que deben analizar los datos y tienen la siguiente conversación:

- I: Entonces ¿ahora hay que mirar todos?
- S: Sí, y comparar.
- I: Vale, eh.
- S: O sea, yo diría que sí.
- I: No, yo creo que todos... No hay que mirarlos todos.
- N: No hay que mirarlos todos, lo que tenemos que hacer es mirar algunos. Los que queramos.
- I: Cogemos los [impuestos de] turismos de yo qué sé.
- S: Y cogemos un poco del norte, del sur y un poco de Extremadura.

En el episodio anterior los alumnos han descubierto la dificultad de trabajar con todos los datos proporcionados, y en este episodio se replantean usar una parte de ellos, a partir de

elegir una muestra. Las últimas intervenciones evidencian sus primeras ideas sobre cómo deben elegir los valores de la muestra, hecho que cambia el planteamiento previo que tenían y les permite elegir una nueva vía de trabajo. Por ello hemos categorizado este episodio como un proceso de planificación (P). Una vez caracterizados todos los episodios, se genera el DAM usando TaskTimeTracker (Pla-Castells y García-Fernández, 2020): un software diseñado específicamente para representar DAM de forma ágil.

Resultados: DAM de cada grupo

A continuación, mostramos como resultados del análisis cualitativo desarrollado los DAM obtenidos a partir de las grabaciones de los tres grupos de alumnos.

Primer grupo (G1)

El DAM elaborado para caracterizar la actuación del grupo G1 es el que se muestra en la figura 3.

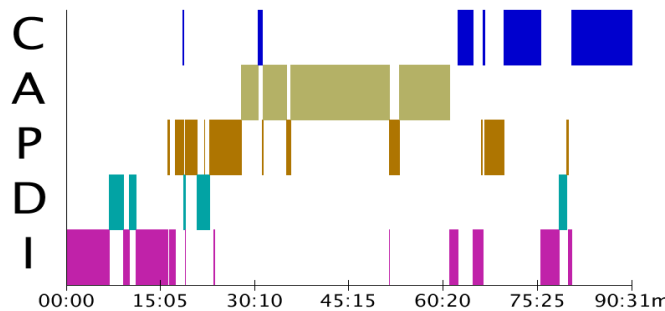


Figura 3. DAM para el grupo G1. Fuente propia de la investigación.

Los estudiantes del grupo G1 empiezan leyendo e interpretando el enunciado, hasta el minuto 6:58 de la grabación. No solo leen el enunciado literalmente, sino que también discuten sobre los significados de cada elemento. Este episodio se ve interrumpido por la primera fase de exploración de datos en la que aparece una idea para entender mejor el enunciado (minuto 9:06), y vuelven a él logrando entender una parte que no les había quedado clara. La exploración inicial de los datos les permite entender el enunciado y, por eso, van volviendo a él, hasta que en el minuto 15:07 deciden trabajar con todos los datos (P); pero momentos después, 15:14 minutos, se dan cuenta de que les es imposible (con los recursos de que disponen) y deciden escoger una pequeña muestra al azar para trabajar a partir de pequeñas simulaciones. 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. Vuelven a discutir sobre lo que pide el enunciado, redefiniendo sus objetivos. Esto provoca que redefinan la muestra con la que van a trabajar y decidan elegir una que pueda ser representativa. Deciden que en la muestra no escogerán provincias al azar, sino que se centrarán en el impuesto de un tipo

concreto de coche. Los alumnos no definen de ninguna forma qué entienden por “muestra representativa”, pero pasan a escoger provincias de diferente población y repartidas por el territorio, así como los impuestos de los vehículos más comunes.

Este grupo es muy coherente y constante en su trabajo: cuando deciden una forma de trabajar y eligen un concepto, lo mantienen firmemente en su trabajo. Su primera propuesta de muestra se detalla posteriormente para determinar una submuestra, siguiendo unos criterios (de aleatoriedad intuitiva) similares. Estos criterios son muy discutidos por el grupo a lo largo del proceso. El análisis de los datos les permite alcanzar los primeros resultados, pero en el minuto 19 del audio observan que necesitan un análisis más detallado, e introducen un análisis de la variabilidad a partir de estudiar diferencias entre los valores. Discuten esta nueva vía hasta el minuto 27 que empiezan a llevarla a cabo y se replantean la elección inicial de la muestra con la que trabajar. En este proceso, notamos que al introducir una mejora en su método vuelven a iniciar todo el trabajo, estableciendo diversos ciclos en los que seleccionan una muestra, observan las diferencias, obtienen resultados y conclusiones, y planifican las siguientes muestras. Es un grupo disciplinado y que evalúa los resultados para saber si tienen suficiente coherencia y, por ello, periódicamente revisan el enunciado. Aunque observamos que, a partir del minuto 72, las conclusiones se asientan y son redactadas en su plenitud ya con pocas interrupciones.

Segundo grupo (G2)

El DAM presentado en la figura 4 muestra la categorización del trabajo del grupo G2.

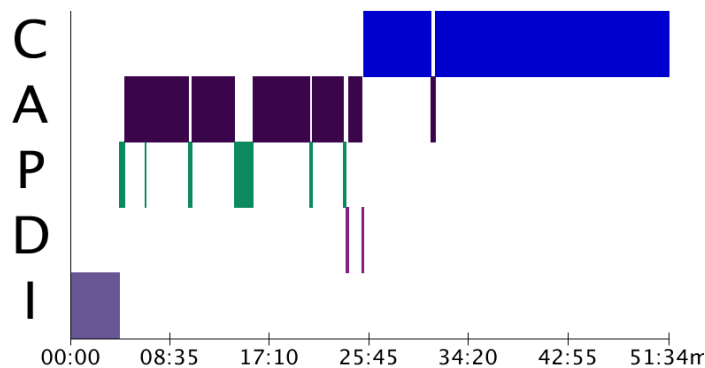


Figura 4. DAM para el grupo G2. Fuente propia de la investigación.

El grupo G2 establece un primer modelo para valorar la adecuación de los impuestos en la primera lectura del enunciado. En concreto, se plantean un modelo que utiliza como concepto la media de los valores máximos y mínimos de los impuestos de cada provincia (minuto 4:30). Los impuestos por encima de este valor serán calificados como injustos, y los que estén por debajo como justos.

Este grupo utiliza todos los datos proporcionados, que trataron usando la hoja de cálculo. Mientras trabajan con la hoja de cálculo, observan que obtener un valor concreto 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 un intervalo centrado en la media calculada previamente. 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. Entonces metódicamente clasifican los datos según se hallen dentro del intervalo o fuera del intervalo, con lo que se la calificación de la adecuación de los impuestos pasa a tener tres valores (por debajo, dentro del intervalo-justo, por encima). Durante este proceso los alumnos alternan entre la planificación del trabajo y el análisis de datos hasta aproximadamente el minuto 25 de la grabación. Cuando empiezan a sacar conclusiones para algunos resultados, se dan cuenta de la necesidad de establecer una muestra sobre la que trabajar, ya que generar un informe completo les parece inviable. Después de que todos los miembros del grupo hicieran propuestas para elegir un método de determinación de la muestra, decidieron. Todos los miembros del grupo aportaron ideas para construir el modelo, pero se veía un vínculo entre hacia dónde iba la conversación y la herramienta elegida.

Tercer grupo (G3)

El DAM que caracteriza la actuación del grupo G3 se muestra en la figura 5.

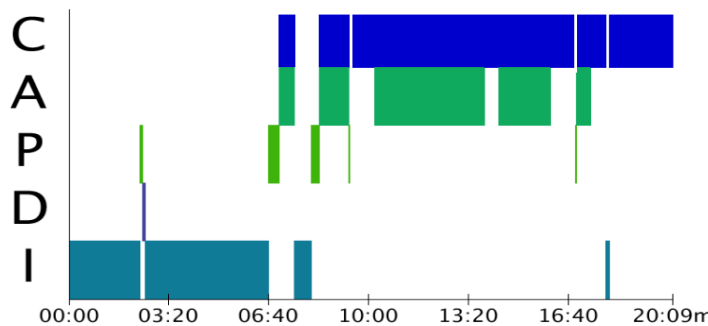


Figura 5. DAM para el grupo G3. Fuente propia de la investigación.

El grupo 3 determina un esquema base para su modelo en la primera etapa del trabajo (los primeros 6 minutos) después de una discusión en la que contemplan diversas opciones y las discuten. A partir de ese momento, lo revisan añadiendo elementos que van necesitando cuando se enfrentan a los datos concretos. Esta forma de actuar se explica por la determinación que aporta una de las integrantes del grupo que tiene muy clara la forma de trabajar y marca con ímpetu las actividades a seguir.

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 lo 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 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. En una situación similar, en el

minuto 9:05 deciden que Ceuta es una localidad justa porque es una ciudad autónoma fuera de la península ibérica, con las dificultades económicas que eso supone. Conforme van estudiando casos concretos, van observando la complejidad de la situación y añaden nuevas variables. Esta forma de trabajar, que no incluye cálculos específicos, permite que avancen rápidamente en la resolución del problema, hecho que refuerzan al dividir el trabajo y que uno de los integrantes del grupo se dedique en exclusiva a redactar el informe mientras están trabajando sobre los datos.

Resultados: Episodios clave en la modelización

En esta sección identificamos los resultados obtenidos del análisis basado en el DAM de la actividad de los alumnos. En primer lugar, corroboramos que la actividad propuesta promueve realmente la modelización matemática (Lesh, 1997). 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 y Harel, 2003) pero sin utilizar procedimientos que involucren cálculos específicos, con lo que trabajan a partir de un modelo informal.

En segundo lugar, la actividad de los alumnos se focaliza mayoritariamente en medir las diferencias entre los valores de los impuestos proporcionados, con lo que comprobamos que responde a la necesidad de promover el pensamiento estadístico (Garfield y Ben-Zvi, 2007; delMas 2002; 2004) y el razonamiento estadístico (Wild y Pfannkuchen, 1999). De esta forma, observamos que la resolución del problema conjuga actividades de modelización matemática y de pensamiento y razonamiento estadísticos. Interpretamos que este hecho se debe realmente a que estos procesos de actividad comparten una base común, y que se diferencian mayormente en aspectos que se han esquivado con el diseño del problema presentado.

El análisis mostrado en los DAM de los tres grupos nos permite identificar aquellos episodios en los que los alumnos han tomado decisiones que les han permitido desarrollar con más precisión el modelo matemático que construían. En concreto, los elementos que provocan que los alumnos incidan en el modelo matemático que van a usar para justificar sus respuestas son los siguientes:

Gran volumen de datos: El alumnado se plantea formas efectivas o viables de analizar una cantidad de datos de gran volumen como la que se les ha proporcionado. Los alumnos pueden ser conscientes de esta dificultad desde que leen el enunciado, o puede ser que se den cuenta de ello mientras exploran los datos o incluso en una fase posterior, mientras los están analizando. 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 substituyen las dificultades de recolectar datos por las de tratar y manipular datos (Hahn, 2015; Muñoz-Rodríguez, Rodríguez-Muñoz y Alsina, 2020, Ubilla, 2020).

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 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.

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 con una situación compleja en la que observan que toda reducción de la complejidad cambia el problema, por lo que deben tratar de trabajar sin reducirla. Este hecho se muestra en las dificultades que presentan los alumnos al tratar de determinar muestras de los datos para obtener resultados, 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.

Discusión y conclusiones

En este estudio hemos analizado el trabajo de alumnos de 4º curso de educación secundaria obligatoria en España al enfrentarse a una tarea en la que deben tratar con una gran cantidad de datos. Para ello hemos usado una adaptación específica para las actividades con contenidos estadísticos de los DAM (Ärlebäck, 2009) como herramienta de análisis. Los resultados nos muestran que la actividad promueve efectivamente los dos elementos para los que ha sido diseñada: la modelización matemática como actividad principal y el tratamiento de datos estadísticos.

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 innovando en la forma de combinarlos para conseguir sus objetivos. También observamos que los alumnos identifican elementos matemáticos clave para dar forma a su modelo, aunque no conozcan un procedimiento específico para especificarlo formalmente. De esta forma los alumnos generan modelos matemáticos propios para resolver una situación que requiere de conocimientos matemáticos específicos (Aymerich, Gorgorió, y Albarracín, 2017). 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 compleja que claramente es multifactorial y que provoca que no sea sencillo elegir muestras representativas para trabajar.

Dada la naturaleza de la actividad desarrollada para este estudio, se decidió que el profesor diera un soporte básico, tratando de no influir en la toma de decisiones de los alumnos. Los resultados muestran que uno de los grupos desarrolló un modelo matemático sin procedimientos de cálculo específicos. De esta forma se pone de manifiesto que la actividad, en la forma en la que se presentó a los alumnos, no los fuerza a tomar un posicionamiento analítico. La forma de conseguir que los alumnos trabajen en la actividad a partir de la modelización matemática se puede conseguir vía la gestión de aula del profesor. Una opción es desarrollar previamente alguna actividad de menor impacto en el tiempo disponible para sentar las bases del trabajo. Otra posibilidad es monitorizar en el aula el tipo

de actividades que desarrollan los alumnos y sugerir cambios para adecuarlos a los objetivos de aprendizaje propuestos.

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 no solo habrán aprendido a modelizar grandes cantidades de datos contextualizados y desarrollar su razonamiento estadístico general (Wild & Pfannkuchen, 1999); 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.

Agradecimientos

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, España) y 2017 SGR 497 (AGAUR, Generalitat de Catalunya).

Conflicto de intereses

Los autores declaran no tener algún conflicto de interés.

Declaración de la contribución de los autores

Todos los autores afirmamos que se leyó y aprobó la versión final de este artículo.

El porcentaje total de contribución para la conceptualización, preparación y corrección de este artículo fue el siguiente: À.A. 50% y L.A. 50 %.

Declaración de disponibilidad de los datos

Los datos que respaldan los resultados de este estudio serán puestos a disposición por el autor correspondiente [À.A.], previa solicitud razonable.

Referencias

- Albarracín, L., Arleback, J., Civil, E. & Gorgorió, N. (2019). Extending Modelling Activity Diagrams as a tool to characterise mathematical modelling processes. *The Mathematics Enthusiast*, 16(1), 211-230.
- 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.
- Ärleback, J. (2009). On the use of realistic Fermi problems for introducing mathematical modelling in school. *The Mathematics Enthusiast*, 6(3), 331-364.

- Aymerich, À., Gorgorió, N. & Albarracín, L. (2017). Modelling with statistical data: Characterisation of student models. In *Mathematical Modelling and Applications* (pp. 37-47). Springer.
- Batanero, C., Díaz, C., Contreras, J.M. & Roa, R. (2013). El sentido estadístico y su desarrollo. *Números*, 83, 7-18.
- 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.
- Ben-Zvi, D. & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representation. In J. Garfield, D. Ben-Zvi, & C. Reading (Eds.), *Background Readings of the Second International Research Forum on Statistical Reasoning, Thinking, and Literacy* (pp. 73-110). Centre for Cognition Research in Learning and Teaching, University of New England.
- Blum, W., & Leiß, 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.
- Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *ZDM*, 38(2), 86-95.
- 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.
- 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.
- Crites, T., & Laurent, R. T. (2015). *Putting essential understanding of statistics into practice, Grades 9-12*. National Council of Teachers of Mathematics.
- delMas, R. C. (2002). Statistical literacy, reasoning, and thinking: A commentary. *Journal of Statistics Education*, 10(3). doi.org/10.1080/10691898.2002.11910674
- delMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79-95). Kluwer.
- 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.
- Doerr, H. M. & Tripp, J. S. (1999). Understanding how students develop mathematical models. *Mathematical Thinking and Learning*, 1, 231-254.
- Hahn, C. (2015). La recherche internationale en éducation statistique: état des lieux et questions vives. *Statistique et Enseignement*, 6(2), 25-39.
- Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review*, 70(1), 1-25.

- Gal, I. (2004). Statistical Literacy: Meanings, Components, Responsibilities. En D. Ben-Zvi, & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 47-78). Kluwer Academic Publishers.
- Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modeling process. *ZDM*, 38(2), 143-162.
- Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372-396.
- Greefrath, G. (2011). Using Technologies: New Possibilities of Teaching and Learning Modeling – Overview. In G. Kaiser, W. Blum, R. Borromeo Ferri, R. & G. Stillman (Eds.), *Trends in teaching and Learning of Mathematical Modeling* (pp. 301-304). Springer.
- Hernández-Solís, L., Batanero, C., Gea, M., & Álvarez-Arroyo, R. (2021). Comparación de probabilidades en urnas: Un estudio con estudiantes de educación primaria. *Uniciencia*, 35(2), 1-18. <https://doi.org/10.15359/ru.35-2.9>
- Lehrer, R. & Romberg, T. (1996). Exploring children's data modeling. *Cognition and Instruction*, 14, 69-108.
- Lehrer, R. & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade children's classification models. *Mathematical Thinking and Learning*, 2, 51-74.
- Lesh, R. (1997). Matematización: La necesidad "real" de la fluidez en las representaciones. *Enseñanza de las Ciencias*, 15(3), 377-391.
- Lesh, R., & Harel, G. (2003). Problem solving, modeling, and local conceptual development. *Mathematical Thinking and Learning*, 5(2-3), 157-189.
- Lesh, R. E., & Doerr, H. M. (2003). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Lawrence Erlbaum Associates Publishers.
- 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.
- 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.
- Kaiser, G. & Stender, P. (2013). Complex modeling problem in cooperative learning environments self-directed. 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. 277-294). Springer.
- Pla-Castells, M. & García-Fernández, I. (2020). TaskTimeTracker: A tool for temporal analysis of the problem solving process. *Investigación en Entornos Tecnológicos en Educación Matemática*, 1, 9-14.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.

- 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.
- Watson, J. M. (2011). Foundations for improving statistical literacy. *Statistical Journal of the IAOS*, 27(3-4), 197-204. <https://doi.org/10.3233/SJI-2011-0728>
- Wild, C. J., & Pfannkuchen, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-248.
- 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*. Lawrence Erlbaum Associates, Inc.



Uniciencia is protected by [Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/)