



Universidad de  
**San Andrés**

UNIVERSIDAD DE SAN ANDRÉS

ESCUELA DE EDUCACIÓN

LICENCIATURA EN CIENCIAS DE LA EDUCACIÓN

Trabajo de graduación

**Encuestas de evaluación docente en educación superior:** un estudio de caso sobre las concepciones de estudiantes, docentes y directores de carrera.

Autor: María Eugenia Negri

Legajo: 27134

Mentor: Dra. María Mercedes Di Virgilio

Buenos Aires, diciembre de 2019

## AGRADECIMIENTOS

A mi familia, por haber apostado siempre por la educación.

A los *Edupeople*, por haber hecho del trayecto universitario algo inolvidable.

A Gon, por haberme apoyado en mis mil y una frustraciones los domingos por la tarde.

A mis amigos, por no haber reprochado mis (tantas) ausencias este último tiempo.

A la Universidad de San Andrés, por haberme dado la oportunidad de aprender de profesores tan inspiradores.

A mi mentora, Mercedes Di Virgilio, por haberme apoyado y alentado en los mil y un cambios de tema.

A todos aquellos que me prestaron un poquito de su tiempo (estudiantes, docentes y directores de San Andrés) e hicieron que esta investigación fuera posible.

A todos, infinitas gracias.



Universidad de  
**San Andrés**

## ÍNDICE

<b>ÍNDICE DE CUADROS</b> .....	<b>V</b>
<b>RESUMEN</b> .....	<b>6</b>
<b>INTRODUCCIÓN</b> .....	<b>7</b>
OBJETIVOS Y METODOLOGÍA.....	8
ORGANIZACIÓN DEL ESTUDIO .....	8
<b>CAPÍTULO 1</b> .....	<b>9</b>
1.1 LA UNIVERSIDAD EN LA SOCIEDAD DEL CONOCIMIENTO.....	9
1.2 LOS CUESTIONARIOS DE LA EVALUACIÓN DE LA DOCENCIA POR LOS ESTUDIANTES .....	11
1.3 APOYO Y OPOSICIÓN AL USO DE LOS CEDA.....	14
1.4 LA PREOCUPACIÓN POR ESTUDIAR LAS CONCEPCIONES .....	17
1.4.1 <i>Antecedentes</i> .....	18
1.4.1.1 Coordinadas temáticas.....	19
1.4.1.2 Coordinadas metodológicas.....	21
1.4.1.3 Coordinadas espacio-temporales .....	23
1.4.1.4 Hallazgos.....	23
Concepciones de los estudiantes.....	23
Concepciones de los docentes.....	26
Percepciones de los administradores .....	28
1.5 NUESTRO ESTUDIO .....	29
<b>CAPÍTULO 2</b> .....	<b>31</b>
2.1 LAS CONCEPCIONES .....	31
2.2 LA EVALUACIÓN DE LA DOCENCIA UNIVERSITARIA .....	32
2.2.1 <i>La dificultad de evaluar la enseñanza</i> .....	33
2.3 MODELOS DE EVALUACIÓN DOCENTE .....	34
2.3.1 <i>Finalidad de la evaluación</i> .....	35
2.3.1.1 Evaluación sumativa .....	36
2.3.1.2 Formativa .....	37
2.3.1.3 Evaluación híbrida .....	38
2.3.2 <i>Usos de los resultados: acciones High-Stakes y Low-Stakes</i> .....	38
2.3.3 <i>Momento de la evolución: durante o al final del proceso</i> .....	39
2.4 ESTRATEGIAS DE EVALUACIÓN .....	40
2.4.1 <i>Evaluación del desempeño docente a través de cuestionarios estudiantiles</i> .....	43
2.4.1.1 Cuestiones psicométricas .....	45
<b>CAPÍTULO 3</b> .....	<b>51</b>
3.1 ESTRATEGIA GENERAL.....	51
3.2 JUSTIFICACIÓN DEL DISEÑO Y ELECCIÓN DEL CASO.....	51
3.2.1 <i>Estudio de caso único: Universidad de San Andrés</i> .....	51
3.2.2 <i>Caracterización de los informantes</i> .....	52
3.3 ACERCA DEL DESARROLLO DE LA INVESTIGACIÓN.....	54
3.3.1 <i>Los preparativos</i> .....	54
3.3.2 <i>Recolección de datos</i> .....	54
3.3.3 <i>Registro y almacenamiento</i> .....	55
3.3.4 <i>Análisis y presentación de los datos</i> .....	56
3.4 NOTAS SOBRE EL TRABAJO DE CAMPO .....	59
<b>CAPÍTULO 4</b> .....	<b>61</b>
4.1 DEL PROPÓSITO .....	62
4.1.1 <i>... GENERAL Y LA INCLUSIÓN DE LOS ALUMNOS</i> .....	62
4.1.2 <i>... Y SU CONOCIMIENTO</i> .....	64
4.2 DE LOS USOS... ..	66
4.2.1 <i>...LA FORMACIÓN</i> .....	66
Utilidad formativa .....	66
Cambios.....	68
El caso de los comentarios .....	70
Limitaciones .....	72

4.2.2 ...Y EL CONTROL.....	73
Criterio relevante para decisiones high-stake .....	73
El caso de las eminencias .....	77
Consecuencias de utilizarla con fines sumativos .....	77
4.3 DE LA CONFIABILIDAD.....	80
4.3.1 ...Y LOS ESTUDIANTES.....	80
Seriedad y compromiso.....	80
Relación con la calificación esperada .....	83
Capacidad de evaluar la enseñanza .....	86
4.3.2 ...Y LOS PROFESORES.....	87
Compromiso con el aprendizaje de los alumnos.....	88
Estrategias de enseñanza/clase dinámica/contenido .....	89
Trato .....	90
4.3.4 ...LAS CARACTERÍSTICAS DEL CURSO .....	91
Tamaño de la clase .....	92
Área de interés primaria y electividad .....	93
4.3.5 ...Y LA ADMINISTRACIÓN DEL CUESTIONARIO .....	94
<b>CONCLUSIONES Y REFLEXIONES FINALES.....</b>	<b>97</b>
CONCLUSIONES .....	97
REFLEXIONES FINALES.....	103



Universidad de  
**San Andrés**

## ÍNDICE DE CUADROS

<b>CUADRO 1. QUIÉN EVALÚA LAS ACTIVIDADES DE DOCENCIA, INVESTIGACIÓN Y EXTENSIÓN .....</b>	<b>12</b>
<b>CUADRO 2. SISTEMATIZACIÓN DE ESTUDIOS SOBRE CONCEPCIONES SOBRE CEDA .....</b>	<b>19</b>
<b>CUADRO 3. TIPOLOGÍA DE EVALUACIÓN DOCENTE .....</b>	<b>35</b>
<b>CUADRO 4. MODELOS DE EVALUACIÓN DOCENTE .....</b>	<b>40</b>
<b>CUADRO 5. ESTRATEGIAS DE EVALUACIÓN DOCENTE SEGÚN PARTICIPANTES E INSTRUMENTO DE EVALUACIÓN .....</b>	<b>42</b>
<b>CUADRO 6. CARACTERIZACIÓN DE LOS INFORMANTES .....</b>	<b>53</b>
<b>CUADRO 7. DIMENSIONES Y SUB-DIMENSIONES DE ANÁLISIS .....</b>	<b>57</b>
<b>CUADRO 8. CANTIDAD DE FRAGMENTOS ANALIZADOS POR DIMENSIONES Y SUB-DIMENSIONES DE ANÁLISIS .....</b>	<b>61</b>



Universidad de  
**San Andrés**

## RESUMEN

La evaluación de la docencia es una de las cuestiones centrales en las universidades, en tanto la calidad del desempeño docente determina en gran medida la consecución de los fines de las casas de estudio. Una de las estrategias que se ha vuelto la más elegidas por las universidades para lograr este cometido ha sido la implementación de los cuestionarios de evaluación docente por parte de los alumnos (CEDA). La literatura que ha trabajado el tema, señala que no existen consensos acerca de su confiabilidad. Asimismo, su utilización ha despertado ciertas controversias en el ámbito de la educación superior. En este marco, este trabajo explora las concepciones que docentes, estudiantes y tomadores de decisiones de una universidad privada de Buenos Aires construyen en relación a este proceso. Desde un diseño de caso único y a través de entrevistas semi estructuradas se analizan sus percepciones a partir de tres dimensiones: su propósito, su uso y su confiabilidad. Entre los principales hallazgos se destaca que los tres grupos entrevistados no tienen claridad sobre el propósito y el uso de las encuestas en la Universidad. Todos ellos valoran que se incluya a los alumnos en esta instancia. Si bien los docentes señalan realizar cambios a partir de las encuestas cuando los comentarios de sus estudiantes son recurrentes, las modificaciones parecen ser imperceptibles para los estudiantes. En cuanto a su función sumativa, los docentes expresan reservas frente a su uso para orientar acciones *high-stake*. Los directores de carreras coinciden y argumentan que la encuesta constituye tan solo un indicador más a tener en cuenta en la evaluación de los y las docentes. En relación a la confiabilidad, se pone en dudas la idoneidad de los estudiantes para evaluar la enseñanza, especialmente por el impacto que puede tener la calificación recibida en el curso sobre su actitud a la hora de evaluar el docente. A pesar de ello, los estudiantes manifiestan adoptar una postura honesta y comprometida. Asimismo, los docentes encuentran que el tamaño del curso y su electividad resultan características a tenerse en cuenta a la hora de comparar los resultados. Se destaca, finalmente, la importancia de que las instituciones de educación superior elijan estrategias de evaluación docente acordes a los objetivos que orientan su quehacer. Asimismo, plantean la necesidad de que se desarrollen y lideren iniciativas para fomentar una cultura de la evaluación formativa orientada a la mejora.

**Palabras clave:** evaluación docente, encuestas, percepciones, creencias, CEDA.

## INTRODUCCIÓN

Al considerar los problemas actuales que atraviesa nuestro sistema educativo, vinculados con la búsqueda de la calidad, nos preguntamos cómo es que procesan los distintos actores del sistema de educación superior las evaluaciones de desempeño docente por parte de los estudiantes. Para ello, en este trabajo nos proponemos recoger y analizar las concepciones ó percepciones que tienen alumnos, docentes y directores de carreras de una universidad privada de la provincia de Buenos Aires sobre las encuestas como instrumento de evaluación del desempeño docente.

Las encuestas son, hoy en día, el instrumento más utilizado en el contexto de la educación superior para evaluar la *performance* docente. De hecho, existe una gran cantidad de literatura en relación a este tema, literalmente miles y miles de artículos según afirman Marsh y Dunkin (1992). A pesar de ello, no se ha llegado a un consenso acerca de la confiabilidad de este instrumento para la evaluación docente. Como resultado de lo anterior, muchos docentes han expresados sus reservas frente al uso de los cuestionarios, cuestionando principalmente su uso para fines sumativos. Es decir, para la toma de decisiones administrativas.

Bajo la idea de que las concepciones finalmente tienen un impacto en la práctica y, por tanto, influyen en la forma en que un sujeto atraviesa el proceso evaluativo, algunas investigaciones se han preocupado por indagar cuáles son las creencias y opiniones de los docentes acerca de las encuestas. Asimismo, entendiendo que en el proceso de evaluación docente intervienen otros actores, otros estudios se han interesado en estudiar las creencias de los estudiantes, en tanto evaluadores de los docentes, y administradores, en tanto responsables de la toma de decisiones.

A pesar de que el uso de las encuestas ha sido extendido a una amplia parte de las instituciones de educación superiores, todavía no se han realizado estudios que aborden esta temática en Argentina. Es así que, esta investigación se vio motivada por conocer qué es lo que piensan los docentes, estudiantes y tomadores de decisiones en el contexto local. Asimismo, es resultado de la preocupación que suscita entender el modo en el que los sujetos se vinculan con un instrumento de evaluación que ha sido impuesto en la región y en el país, a pesar del escaso desarrollo en materia de investigación. Finalmente, es consecuencia de nuestro interés personal de contribuir con el desarrollo de una evaluación docente más justa y provechosa para nuestra casa de estudios.

## Objetivos y metodología

Como objetivo general, entonces, nos proponemos explorar las concepciones de estudiantes, docentes y tomadores de decisiones de la Universidad de San Andrés sobre las encuestas, en particular sobre el papel de la encuesta en el proceso de evaluación docente. De este objetivo, se desprenden los siguientes objetivos específicos:

1. Conocer las concepciones que alumnos, docentes y directores de carrera de la Universidad de San Andrés tienen sobre el propósito, la utilización y la confiabilidad de las encuestas como instrumento de evaluación del desempeño docente.
2. Comprender los alcances y las limitaciones que los distintos actores atribuyen a la encuesta en el proceso de evaluación docente.

En cuanto al diseño metodológico, optamos por un estudio de caso único para llevar adelante la investigación. La institución seleccionada fue la Universidad de San Andrés y la recolección de datos se realizó a partir de un total de 25 entrevistas semi-estructuradas: 10 estudiantes, 10 docentes y 5 directores de carrera. En cuanto al análisis, estuvo guiado por la identificación de dimensiones y sub-dimensiones, asociadas a los objetivos de investigación, a las coordenadas teóricas y a los hallazgos de investigaciones anteriores.

### Organización del estudio

En primer lugar, luego de esta breve introducción, desarrollamos una descripción del problema de investigación escogido. Seguidamente, realizamos un recorrido por una sistematización de las investigaciones que son antecedentes de este estudio: presentamos los principales temas que abordaron, las metodologías que fueron utilizadas, y sus principales hallazgos. En el capítulo 2, desarrollamos las referencias teóricas que orientaron la investigación y el posterior análisis de la información recogida a través de las entrevistas. En el capítulo 3, exponemos y justificamos elección de la estrategia metodológica. Hacemos una descripción detallada del proceso de recolección de datos, y definimos las dimensiones y sub-dimensiones de análisis. En el capítulo 4 se despliega el análisis del material recolectado, en función de los objetivos específicos uno y dos. Para finalizar, exhibimos las conclusiones donde se condensan los principales hallazgos de la investigación, y nuestras reflexiones finales.

## CAPÍTULO 1

### PLANTEO DEL PROBLEMA

#### 1.1 La universidad en la sociedad del conocimiento

El tema de la calidad de la educación en general y de la educación superior en particular, ha comenzado a afirmarse en el escenario y en la agenda de la educación mundial desde hace varias décadas (Marsh, 1984; Seldin, 1993). Esto ha ocurrido en un tiempo donde se le ha dado una importancia considerable a la contribución económica de las universidades (Tomlinson, 2018). Ellas son las principales proveedoras de servicios y mano de obra altamente calificada destinadas a cubrir las demandas de un nuevo sistema económico caracterizado por tener el *conocimiento* como principal mercancía (Lyotard, 1993). En este contexto, la calidad de la educación superior, no solo determina cada vez más el potencial evolutivo de una sociedad, sino que, en términos económicos, afecta el crecimiento y la competitividad internacional (Alexander, 2000; Yorke y Vidovich, 2016).

Potenciado por la globalización, el advenimiento de la mercantilización del conocimiento ha evolucionado a través de los años, y con él, otras lógicas y visiones han ido emergiendo en el contexto global (Locke, 2015). Una de ellas ha sido la performatividad como discurso normativo, que ha impuesto la lógica de la eficiencia sobre los sistemas y que ha conferido “un valor prevalente a las actuaciones (desempeños) y los logros (resultados, rendimientos) de los sujetos (estudiantes, docentes u otros), así como de las instituciones [educativas] y sus proyectos” (Escudero Muñoz y Trillo Alonso, 2014: 84). Es así, que bajo esta ética “del gobierno de los números” (Ball, 2013), los estándares, las competencias y los indicadores han comenzado a constituir los resortes para determinar y requerir *performances*, así como para crear sistemas de rendición de cuentas que vigilen y sancionen sus realizaciones.

Uno de los efectos más visibles de ello ha sido la actual fiebre por tasar, indexar todo lo que vale: “las organizaciones y las personas son reducidas a datos, indicadores, a comparaciones y rankings” y el conocimiento vale en tanto externalizado y registrado, sometido al imperio de lo útil, práctico y eficiente (Ball, 2003). También eco de ello han sido las reformas educativas de las últimas décadas, que según Ball (2014), forman suerte de una “epidemia global” provocada por organismos internacionales como el Banco

Mundial, el FMI, la OCDE y la creciente presencia de grandes corporaciones transnacionales en la educación. Estas han venido a “racionalizar, disciplinar y someter los sistemas educativos a la horma mercantil” (Apple, 2002; Ball, 2013; Williamson, 2013; Ozga, 2014; en Escudero Muñoz y Trillo Alonso, 2014: 84).

Asociada a la lógica performativa y al discurso neoliberal, encontramos la noción de *new managerialism*, como forma de referirse a la adopción que han hecho las instituciones públicas y, entre ellas, las instituciones educativas de nivel superior; de las tecnologías, los valores, las prácticas y formas organizacionales comúnmente encontradas en el sector privado. En palabras de Trowler:

...bajo el nuevo pensamiento gerencialista aplicado a las universidades, la educación y el aprendizaje se mercantilizan, y el discurso de los clientes y consumidores, de gestión, de responsabilización, y mercado, se vuelve dominante. Las relaciones se vuelven poco confiables, por lo que siempre se percibe la necesidad de vigilar e intervenir de las agencias (2011: 20. Traducción propia).

Así, con la intención de adaptarse a este nuevo entorno caracterizado por un entramado de lógicas mercantilistas, muchas instituciones se han visto coaccionadas a definir nuevamente su misión y actividades, como proveedores de servicios que buscan atraer a nuevos clientes. De esta forma, el valor de la educación superior, inicialmente asociado a procesos de transformación y desarrollo personal e intelectual, hoy día se reduce a relaciones esencialmente transables entre individuos e institucional (Tomlinson, 2018).

En este nuevo mercado competitivo, la calidad se presenta como una meta principal de la educación superior: “los estudiantes que reciben una educación de alta calidad no solo obtendrán un retorno inmediato y futuro de su inversión, pero también adquirirán un nivel educativo que los preparara bien para la vida económica futura” (Tomlinson, 2018: 715. Traducción propia). Aquí la evaluación de la docencia juega un papel fundamental, en tanto pareciera ser que “un acuerdo generalizado en la universidad, es que los docentes tienen un papel determinante en la mejorar de la calidad de la enseñanza” (Luna Serrano y Torquemada, 2008: 2). Es así que, para garantizar un servicio de calidad a sus consumidores, las instituciones deben contar con sistemas aptos para monitorear y medir la *performance* docente (Douglas & Douglas, 2006).

El desenlace de este complejo cruce de intereses (Yorke & Vidovich, 2016) ha sido la implementación por parte de las IES, de programas de evaluación más rigurosos, aunque no necesariamente más sanos o acertados. La profesionalidad docente, una vez sometida a sospecha y denotada por el discurso neoliberal, ha sido reducida al

seguimiento de directrices externas y a evaluaciones de resultados cuantificables. En línea con esto último, los estándares, indicadores, sistemas de evaluación y de rendición de cuentas, pudiendo ser las evaluaciones docentes por parte de los estudiantes un ítem más de esta lista, han conquistado espacios de poder sin ser cuestionados (Moore & Kuol, 2007).

## 1.2 Los cuestionarios de la evaluación de la docencia por los estudiantes

La evaluación del profesorado universitario se ha basado históricamente en la performance en tres áreas: docencia, la investigación y el servicio (Centra, 1993). Aunque la precisión y la confiabilidad de su evaluación sea un reto común a las tres, particularmente la enseñanza es más difícil de evaluar que la investigación y el servicio. Tal es así que, con la intención de resolver la dificultad que supone evaluar una actividad tan compleja como la docencia (Rueda, Elizalde y Torquemada, 2003), se han desarrollado múltiples estrategias de evaluación, como la de pares, los portafolios docentes o las cartas de estudiantes, por poner algunos ejemplos. Sin embargo, a pesar de la existencia de numerosas herramientas de evaluación, en el ámbito de educación superior han primado las encuestas de evaluación docente por parte de los estudiantes. En otras palabras, las encuestas han sido, y son, el método más utilizado para evaluar la docencia (Stark & Freishtat, 2014: 1).

Seldin (1993) estima que por la década del 90' ya un 86% de las instituciones de educación superior estadounidenses hacían uso de los cuestionarios de evaluación de la docencia por los alumnos (CEDA)<sup>1</sup>. Estudios más recientes como los de Luna Serrano y Torquemada (2008) y García Garduño (2000), extienden los hallazgos de Seldin, y afirman que la evaluación docente con base en los puntajes de opinión de los estudiantes es la estrategia más utilizada no solo en Norteamérica, sino también en Europa y Asia. Una explicación posible, se corresponde con la producción de estudios en la década anterior, que habrían confirmado la validez transcultural de este instrumento, potenciando así su expansión más allá de fronteras geográficas y culturales.

En el contexto hispanohablante, el estudio de Rueda, Luna Serrano y Loredo Enríquez (2010) reveló que el cuestionario estudiantil es el instrumento que se emplea en la mayoría de las universidades mexicanas para evaluar el desempeño docente, alcanzando el 83.6% de las instituciones. En lo que respecta a la Argentina, “el uso de las

<sup>1</sup> La literatura anglosajona refiere a los CEDA como SET: Student Evaluation of Teaching. De ahora en más, nos referiremos a los cuestionarios de evaluación de la docencia por alumnos CEDA o encuestas, indistintamente.

encuestas, de resultados de aprendizaje (cantidad de graduados, por cohorte, calificaciones promedio, etc.), de resultados de investigaciones y la realizada por pares son las formas más frecuentes de evaluación de la docencia universitaria” (Fernández Lamarra y Coppola, 2008a: 99). De hecho, universidades como la Universidad de Buenos Aires, más específicamente el de la Facultad de Odontología, y la Universidad de Ciencias Sociales y Empresariales constituyen un ejemplo de lo mencionado. Ambas instituciones emplean encuestas que “son anónimas, confidenciales, no vinculantes y [que] aportan datos cuantitativos y cualitativos sobre aspectos del desempeño docente” (*ibíd.* :108). Es más, los casos anteriores ilustran una situación que pareciera ser recurrente en el contexto argentino pues, como podemos observar en el siguiente cuadro, una investigación que forma del proyecto internacional *The Changing Academic Profession*<sup>2</sup> ha relevado que los estudiantes son quienes están encargados de evaluar la docencia en más de la mitad de las instituciones de educación superior del país:

**Cuadro 1. Quién evalúa las actividades de docencia, investigación y extensión**

Docencia	Investigación	Extensión	
46.30	22.20	9.98	Los pares de su departamento o unidad académica
46.69	24.16	12.76	El director del departamento o unidad académica
13.14	21.81	6.35	Pares de otros departamentos o unidades académicas
25.79	21.60	17.16	Funcionarios de su institución
60.72	2.44	4.21	Los estudiantes
15.48	53.72	7.30	Evaluadores externos
42.16	24.91	14.15	Usted mismo (a través de una evaluación formal)
11.17	4.28	20.39	Por nadie dentro o fuera de la institución

Fuente: Fernández y Coppola (2008a)

Su popularidad se debe, quizás, al hecho de que “los promedios de las calificaciones de los estudiantes tienen un aire de objetividad por la simple virtud de ser numéricos”<sup>3</sup>; bien porque la comparación de los promedios es una tarea fácil, o tal vez porque la medición a través de este instrumento se vuelve una tarea sencilla y eficiente en términos de tiempo y esfuerzo (Stark y Freishtat, 2014: 1). Al mismo tiempo, el hecho de que cada estrategia

<sup>2</sup> The Changing Academic Profession Project (CAP), llevado adelante por una red de instituciones de 21 países: Estados Unidos, Alemania, Holanda, Francia, Reino Unido, Finlandia, Noruega, Portugal, Italia, República Checa, Japón, China-Hong Kong, India, Malasia, Rusia, Sudáfrica, Canadá, México, Brasil y Argentina. Al respecto, se pueden consultar las siguientes páginas: <http://www.capstudy.org>, <http://www.une.edu.au/chemp/projects/cap/index.php>, <http://www.usm.my/ipptn/fileup/CAP.pdf>, <http://en.rihe.hiroshima-u.ac.jp/hourouki.php>

<sup>3</sup> “Averages of SET ratings have an air of objectivity simply by virtue of being numerical” (Stark & Freishtat, 2014: 1)

de evaluación docente tenga sus propias limitaciones, “vuelve a las evaluaciones por parte de los estudiantes una de las herramientas más sobrevaloradas a la hora de evaluar la enseñanza” (Gravestock & Greenleaf, 2008, en Assasfeh, Al-Ebous, Khwaileh y Al-Zoubi, 2014). La apreciación de Haskell resulta oportuna, pues “quizás ningún otro método de evaluación se haya vuelto tan sacrosanto” (1997: 3).

Sin embargo, y a pesar de su inminente generalización, todavía existen interrogantes respecto al uso de los CEDA como la única fuente de evidencia sobre la enseñanza para mérito y promoción, así como también cuestionamientos en relación a la eficacia de las preguntas de evaluación y los distintos métodos de interpretación (Pounder, 2007, en Stark & Freishtat, 2014). A lo anterior se suman preguntas ligadas a cuestiones estadísticas, a los problemas suscitados a la hora de pensar cómo definir y medir la efectividad de la enseñanza, y a las distorsiones perniciosas que resultan del uso de los puntajes de los CEDA como *proxy* para la calidad de un profesor. En efecto, tal como sugieren Theall & Franklin (2001), “pocos asuntos de la educación superior son tan sensibles, divisorios y políticos como la evaluación del profesorado y, en particular, la calidad y el valor de la información proporcionada por los estudiantes en sus evaluaciones” (45).

Es así, que el debate interminable con respecto a las calificaciones de los estudiantes como medida de la enseñanza efectiva ha captado la atención de gran parte de la investigación en educación superior. Para dimensionar, Marsh y Dunkin (1992) señalan que existen literalmente, miles y miles de investigaciones. Feldman (1990) calcula que se han publicado alrededor de dos mil artículos y libros relacionados con su investigación, tan solo en el idioma inglés. Incluso, “probablemente existen más estudios en estas evaluaciones que de cualquier otro tipo de medio utilizado para evaluar al profesorado combinados” (Cashin, 1988 en Wachtel, 1998: 191). En definitiva, McKeachie no estuvo errado al reconocer que la evaluación de la docencia con base en los cuestionarios de opinión de los estudiantes “es reconocida como la estrategia de evaluación con mayor conocimiento conceptual y empírico acumulado en el contexto universitario” (1996, en Luna y Torquemada, 2008: 4)

Para ser más precisos, la investigación de las evaluaciones del profesorado por parte de los estudiantes y los factores que los afectan, se remontan a la década del 20, al trabajo pionero de Remmers (1928) y Brandenburg y Remmers (1927) en la Universidad de Purdue, donde desarrollaron *The Purdue Rating Scale for Instructors* y condujeron los primeros análisis psicométricos para demostrar lo que era, en ese entonces, un método

innovador de evaluación. Fue recién durante la década del 70' que sobrevino una nueva ola de investigaciones, que incluyó estudios que demostraban la validez y la utilidad de los CEDA y justificaban su uso para propósitos tanto formativos como sumativos. Llegada la década de los 80', se comienzan a ampliar y refinar los hallazgos de la investigación, incluyendo trabajos de meta-análisis que sintetizaron los resultados de muchos otros estudios (Centra, 1993).

Respecto a lo anterior, Marsh y Dunkin (1992) argumentan que a pesar de que muchas investigaciones metodológicamente defectuosas o polémicas hayan sido publicadas en la década del 70', a partir del 80', si bien el ritmo de publicación ha venido decreciendo, su calidad ha ido aumentando. Asimismo, Marsh (1984) sugiere que la validez de las evaluaciones fue lo suficientemente establecida y que el foco de la investigación ha virado en pos de estudiar las preocupaciones metodológicas y las características de fondo que pudieran atentar contra su validez. Por último, Algozzine *et al* (2004) proponen que durante la última década se ha expandido la investigación en las CEDA y se han incluidos temáticas como el uso de la tecnología y métodos alternativos, tales como la implementación de evaluaciones formativas, la alineación de las expectativas de estudiantes y alumnos, el uso de cartas de evaluación y el desarrollo de narrativas por parte del cuerpo de profesores.

Con todo, a pesar de que la investigación de los CEDA sea uno de los tópicos más antiguos y, muy probablemente, el más investigado en la educación superior; el estado de la investigación sobre los cuestionarios estudiantiles en Argentina es todavía incipiente (Fernández Lamarra y Coppola, 2008a).

### 1.3 Apoyo y oposición al uso de los CEDA

Después de casi nueve décadas de investigación acerca del uso de los cuestionarios de evaluación de la calidad de la docencia, autores como Wachtel han sugerido que “se puede afirmar con seguridad que la mayoría de los investigadores creen que las evaluaciones de los estudiantes son un medio válido, confiable y valioso para evaluar la enseñanza (por ejemplo, Centra, 1977, 1993; Cohen, 1981; Koon & Murray, 1995; Marsh, 1984; 1987; Marsh y Dunkin, 1992; McKeachie, 1990; Murray *et al*, 1990; Ramsden, 1991; Seldin, 1993)” (1998: 192). De hecho, referentes en el tema como Marsh (1987), han señalado que las evaluaciones de los estudiantes son el único indicador de la calidad de la enseñanza cuya validez ha sido, completa y rigurosamente, establecida

(Marsh, 1987). Sin embargo, y pese a lo expuesto anteriormente, las opiniones respecto a la utilidad de los CEDA han sido variadas e incluso contrapuestas, pues van desde confiables, válidas y útiles; a no confiables, inválidas e inútiles.

Como hemos anticipado, es posible señalar varios argumentos que apoyan el uso de las evaluaciones de la docencia por parte de los estudiantes ya sea por su utilidad, por su confiabilidad, o por su validez. Pongamos por caso los estudios de Cohen (1981) y Menges (1991) que han señalado la utilidad de este tipo de evaluación para la docencia, pues han encontrado que la retroalimentación provista por las encuestas puede ayudar a mejorar la enseñanza. Asimismo, estudios más recientes han revelado que la mayoría de los estudiantes proveen *feedback* significativo, ni abusivo ni ofensivo; ratificando así, la utilidad formativa propuesta por otros. Sin embargo, en este punto, otros han expresado ciertas preocupaciones, pues argumentan que, si bien están de acuerdo con la afirmación anterior, el efecto de la retroalimentación sobre la enseñanza no es automático, ni seguro, sin otros tipos de *feedback* (Seldin, 1989). Ya desde una perspectiva más institucional, investigadores como Aleamoni (1981) y McKeachie (1979), han encontrado que el uso de las CEDA aumenta la posibilidad de que la excelencia en la docencia sea reconocida y también recompensada.

Asimismo, otras investigaciones han demostrado que los CEDA se encuentran positivamente correlacionados con los niveles de aprendizaje y logro de los estudiantes; es decir, que los educandos califican mejor a aquellos profesores de quienes han aprendido más (Centra, 1977; Cohen, 1981; McKeachie, 1990). Además, autores como Centra (1979), Feldman (1989) y McKeachie (1979), atendiendo a cuestiones ligadas a la validez del instrumento, han encontrado que las calificaciones de los estudiantes están relacionadas positivamente con las calificaciones provistas por estudiantes egresados, poniendo en jaque el argumento que sostiene que, con el tiempo y, beneficiados por algunos años de sabiduría; los estudiantes cambiarían su postura respecto a los profesores (en especial de los exigentes). Por último, estudios como los de Feldman (1976) sostienen que, por lo general, los estudiantes y los profesores coinciden sobre cuáles son los componentes de una enseñanza de calidad y su relativa importancia.

No obstante, y a pesar de la extensa literatura que ha apoyado las evaluaciones de los estudiantes, muchos investigadores aún expresan sus reservas acerca de su uso (particularmente cuando se trata de decisiones de personal y tenencia); hasta se han opuesto a ellas directamente (Zoller, 1992). Así pues, hay abundante evidencia acerca de la hostilidad y el cinismo del profesorado cuando se trata de las CEDA (Franklin y Theall,

1989). A continuación, exponemos algunas de las principales preocupaciones y advertencias frente a su uso.

En primer lugar, Marlin (1987) sostiene que las evaluaciones por parte de los estudiantes pueden ser problemáticas, pues argumentan que esta idea de evaluación presupone un consenso que todavía no existe. Su reflexión, en este punto, resulta ilustrativa: “[¿]cómo podemos evaluar adecuadamente la efectividad de la enseñanza si ni siquiera podemos ponernos de acuerdo sobre lo que constituye una enseñanza efectiva[?]” (Marques *et al*, 1979). Además de no poder partir de una definición consensuada de lo que es una buena práctica pedagógica, hay quienes consideran que la enseñanza involucra muchas otras cuestiones, como el desarrollo de sistema de creencias, que no pueden ser fácilmente recogidas por estos instrumentos de evaluación (Ornstein, 1990).

Además de lo anterior, investigadores como Moore & Kuol (2005) han puesto en jaque la habilidad de los estudiantes de evaluar la *performance* de los docentes, principalmente discutiendo que el conocimiento que poseen de la enseñanza es limitado. Stark y Freishtat (2014) coinciden con lo anterior, pues afirman que, si bien “los estudiantes se encuentran en posición de observar algunos aspectos de la enseñanza como la claridad, el paso, la legibilidad, la audibilidad, y su propio entusiasmo (o aburrimiento) [...] no pueden evaluar la efectividad –sin importar sus intenciones”. De la misma forma, encontramos preocupaciones ligadas a la incapacidad de los estudiantes de realizar evaluaciones objetivas, alejadas de sus expectativas y sentimientos personales (Moore & Kuol, en Chan, Luk & Zeng, 2014).

Por otro lado, hay quienes han detectado algunas tensiones ligadas a la dimensión más práctica de la evaluación. Por ejemplo, Centra (1993), sugiere que el profesorado podría resentir la pérdida de tiempo de clase para la administración de la evaluación, además de que podría verse menos motivado para experimentar otras estrategias de enseñanza. Este punto, resalta la tensión que algunos argumentan existe entre la búsqueda por obtener mejores evaluaciones y la búsqueda por mejorar la enseñanza efectivamente. Pongamos por caso el estudio de Ryan (2015) que expone que el profesorado puede tender a reducir los estándares, la exigencia y la carga de trabajo como resultado de una evaluación obligatoria. Además, la existencia de la inminente posibilidad de experimentar una inflación de calificaciones, producida por la acción sistemática de aquellos profesores que, en el afán de obtener evaluaciones más favorables, tenderían a subir las calificaciones y a bajar la demanda del curso (Simpson & Siguaw en Chan *et al*, 2014:

277). Por último, autores como Franklin y Theall (1989) insisten sobre el poco conocimiento que tienen tanto profesores como tomadores de decisiones sobre la literatura existente en esta área, instalando la preocupación por las posibles distorsiones que podrían suscitarse a partir de una utilización inadecuada y poco informada de estos instrumentos.

Para finalizar, y si bien ahondaremos en esta cuestión más adelante, cabe destacar la existencia de investigaciones que sugieren que la validez de las evaluaciones podría verse amenazada por uno, o varios factores, que poco tienen que ver con el comportamiento del profesor o la calidad de su enseñanza, pero que impactan en la calificación del docente (Balam & Shannon, 2010). En palabras de Nasser & Fresko, “las evaluaciones de los estudiantes están a menudo influenciadas por factores que algunos investigadores consideran irrelevantes al proceso de enseñanza aprendizaje, y, consecuentemente, su validez puede ser fácilmente confusa o difícil de interpretar” (2002: 189).

#### 1.4 La preocupación por estudiar las concepciones

Como hemos visto hasta ahora, la investigación en torno a las CEDA ha sido un tópico recurrente en el contexto de la evaluación superior. A pesar de ello, muy pocas investigaciones se han avocado a estudiar las creencias, actitudes y opiniones de aquellos involucrados en el proceso de evaluación en sí (Marlin, 1987), a saber: los estudiantes, quienes están completando la evaluación; los profesores, sujetos de la evaluación, y aquellas personas que, a fines prácticos denominaremos administrativos, que utilizan las evaluaciones para diversos propósitos (Schmelkin-Pedhazur *et al*, en Nasser & Fresko, 2002). Ahora, ¿por qué resulta relevante estudiar sus concepciones?

Un repaso de la cuestión indica que los motivos por los cuales estudiar las concepciones de los actores respecto a los CEDA son varios. En primer lugar, algunos autores señalan que estudiar las percepciones que los estudiantes construyen del proceso de evaluación es importante, e incluso cuasi esencial, porque en última instancia son ellos quienes proveen el *feedback* acerca del proceso de enseñanza y aprendizaje (Abedin, Taib & Jamil, 2013; Smith and Carney en Asassfeh *et al*, 2014). Sojka, Gupta y Deeter-Schmelz (2002), coinciden señalando a la percepción como un componente crucial que ha sido descuidado, y añade que los estudiantes son quienes evalúan y quienes finalmente se benefician de las encuestas vía la mejora de la enseñanza. En este marco, es importante

tenerlos en consideración. En línea con lo anterior, las apreciaciones de Marlin (1987) en unos de los primeros trabajos dedicados a estudiar este objeto resultan fundantes, pues argumenta que hasta entonces este elemento del proceso de evaluación no había sido analizado en sí mismo. Por último, hay quienes señalan que los jóvenes han cambiado y no dudan en cuestionar las autoridades y, por tanto, “entender las percepciones de los estudiantes, exponer sus creencias y cimentar sobre su retroalimentación constructiva y sus sugerencias formativas no es más una opción” (Asassfeh *et al*, 2014: 121. Traducción propia). Este conocimiento esencial, sugieren los autores, podría servir, incluso, para facilitar nuevas vías de comunicación.

En segundo lugar, y como mencionamos anteriormente, los estudios no solo se han abocado a explorar las creencias de los estudiantes, sino también la de docentes e incluso administradores, en algunos casos. Los autores de dichas investigaciones, como la de Capa-Aydin, Cobanoglu & Gulmez-Dag (2017), señalan que la importancia de su estudio radica, por un lado, en la carencia “de estudios que hayan investigado las percepciones de los docentes de manera profunda, y dispuestos a recibir sugerencias acerca de la manera de mejorar las evaluaciones en un sentido amplio” (s/p). Por otro lado, sugieren que “la postura de los miembros del profesorado respecto a las evaluaciones de curso por parte de los estudiantes es crítica, pues ellos constituyen el grupo de interés primario afectado por los resultados de estas” (*idem*. Traducción propia). En el contexto asiático, también, se ha resaltado la importancia de contribuir a este cuerpo de conocimiento, pues aducen que conocer cómo los docentes interpretan y perciben la efectividad y el uso de los CEDA podría ser valioso para que los hacedores de políticas evalúen, construyan y refinen sus prácticas evaluativas actuales y otros mecanismos de evaluación de enseñanza y aprendizaje (Chan *et al*, 2014).

Para finalizar, cabe mencionar la existencia de estudios (Balam & Shannon, 2010; Campbell & Bozeman, 2007; Sojka *et al*, 2002; Abedin *et al*, 2013) que han encontrado contrastes significativos en la concepción de los distintos actores y que, por tanto, han puesto de manifiesto la importancia de estudiar las percepciones de distintos actores en paralelo, a fin de tener una mirada más acabada del asunto.

#### 1.4.1 Antecedentes

Habiendo dicho esto, presentamos a continuación una sistematización de los estudios que han abordado este tema, destacando las siguientes dimensiones: año, país,

los actores que se vieron involucrados, los aspectos que se analizan y la metodología utilizada. Este cuadro dará inicio a un breve recorrido por sus coordenadas temáticas, metodológicas y espacio-temporales:

**Cuadro 2. Sistematización de estudios sobre concepciones sobre CEDA**

Año	Autores	País	Actores	Temas	Metodología
1987	Marlin	Estados Unidos	Estudiantes Administradores	Propósito Uso	Encuesta
2002	Nasser & Fresko	Israel	Profesores	Confiabilidad Uso	Encuesta
2002	Sojka, Gupta & Deeter-Schemlz	Estados Unidos	Estudiantes Profesores	Confiabilidad Uso	Encuesta
2007	Campbell & Bozeman	Estados Unidos	Estudiantes Profesores Administradores	Propósito Uso	Encuesta Entrevista
2010	Balam & Shannon	Estados Unidos	Estudiantes Profesores	Confiabilidad	Encuesta
2011	Slocombe, Miller & Hite	Estados Unidos	Estudiantes	Uso Confiabilidad	Encuesta
2013	Abedin, Taib & Jamil	Malasia	Estudiantes Profesores	Propósito Uso	Encuesta
2014	Assasfeh, Al-Ebous, Jhwaileh & Al-Zoubi	Jordania	Estudiantes	Propósito Uso Confiabilidad	Encuesta
2014	Chan, Luk & Zeng	China	Profesores	Uso	Encuesta
2014	García Garduño y Medécigo Shej	México	Estudiantes	Propósito Confiabilidad	Encuesta de respuesta abierta
2017	Capa-Aydin, Cobanoglu & Gulmez-Dag	Turquía	Profesores	Confiabilidad Uso	Entrevista

Fuente: elaboración propia.

Conviene subrayar que el término coordenadas no es casual, pues lo elegimos con la intención de evidenciar que las investigaciones aquí retomadas orientan y posicionan nuestra investigación en el corpus de literatura existente.

#### 1.4.1.1 Coordenadas temáticas

Las preguntas que motivaron y se abordaron en distintos estudios fueron diversas. De hecho, Asassfeh *et al* (2014), sostiene que una revisión de la literatura existente indica que, generalmente, “se ha focalizado en la precisión de las percepciones de los estudiantes y la fuente de estas percepciones respecto a la efectividad de la enseñanza” (23). Sin embargo, como hemos anticipado, los temas que han emergido en la literatura han sido diversos, incluyendo cuestiones como: (a) los estándares adoptados en la evaluación, (b) las creencias acerca de los resultados de la evaluación, y (c) creencias acerca de los instrumentos metodológicos. A continuación, procedemos a detallar las principales preocupaciones que han abordado los estudios sobre percepciones, tomando como criterio de orden, los actores involucrados en los distintos estudios: estudiantes, profesores y administradores. Si bien las preocupaciones y preguntas derivadas de ellas se encuentran intrínsecamente relacionadas, creemos que este orden de exposición va a facilitar la lectura y comprensión de los argumentos.

Para comenzar, nos detendremos en las cuestiones que decidieron investigar los estudios que se abocaron a las percepciones de los estudiantes. Probablemente, una de las preguntas más frecuentes en estos estudios ha sido la cuestión por la importancia que los estudiantes le asignan a estas evaluaciones (Marlin, 1987; Asassfeh *et al*, 2014; García Garduño y Medécigo Shej, 2014; Abedin *et al*, 2013; Campbell & Bozeman, 2007). Más precisamente, se han preguntado cuáles son las percepciones que los estudiantes guardan en torno la utilización de los resultados, en dos sentidos: asociados al impacto de la enseñanza y a las consecuencias en la carrera docente. Por un lado, la cuestión del impacto de la enseñanza ha sido abordada en relación a las creencias que tienen los estudiantes respecto a si los docentes hacen cambios o no en base a las evaluaciones y el grado de esos cambios (Marlin, 1987; Asassfeh *et al*, 2014). Por el otro, en relación a los potenciales usos y las consecuencias ligadas a este proceso; un tema que ha sido estudiado muy frecuentemente es cómo creen los estudiantes que se utilizan efectivamente los resultados, es decir, cuáles son las decisiones que toma la institución respecto a la carrera del docente a partir de las encuestas (Slocombe *et al*, 2011; Campbell & Bozeman, 2007). Por último, y posiblemente la más abordada de todas, ha sido la cuestión de los estándares adoptados en la evaluación. En otras palabras, se han preguntado por la honestidad y la seriedad de los estudiantes y los factores que tienen más influencia o que pesan más a la hora de evaluar (Asassfeh *et al*, Slocombe *et al*, 2011). En relación a este último punto, se ha buscado contrastar lo encontrado con los vastos hallazgos de la literatura en relación a estas cuestiones, ligadas a dimensiones de validez y confiabilidad (García Garduño y

Medécigo Shej, 2014; Slocombe *et al*, 2011), como la relación con las características de los estudiantes y el curso, por poner un ejemplo (Balam & Shannon, 2010).

En segundo lugar, en el caso de los docentes, las dimensiones estudiadas han estado más relacionadas con las presuntas posturas -de apoyo u oposición, que hemos desarrollado anteriormente. Más precisamente, los estudios se han detenido en las actitudes que los docentes tienen en general respecto a la validez y la confiabilidad del instrumento (Nasser & Fresko, 2002), principalmente atendiendo a cuestiones como la capacidad de los estudiantes de evaluar justamente, y a los posibles sesgos supuestamente implicados en la evaluación (Capa-Aydin *et al*, 2017; Sojka *et al*, 2002). A su vez, han tenido en cuenta la cuestión de la utilidad de este instrumento para mejorar la enseñanza (Nasser & Fresko, 2002; Capa-Aydin *et al*, 2017), y cómo es que los docentes efectivamente responden a las encuestas (Sojka *et al*, 2002; Campbell & Bozeman, 2007).

Sin embargo, lo central ha sido revelar las preocupaciones de los docentes frente a la utilización de estos resultados con fines sumativos. Es así, que algunas preguntas se han dirigido hacia el uso de las encuestas como mecanismo primario y estandarizado para evaluar la docencia, y como principal medida informativa para tomar decisiones ligadas a la promoción y *tenure* (Chan *et al*, 2014; Nasser & Fresko, 2002; Abedin *et al*, 2013). Por último, se han presentado algunos interrogantes que atendían la potencial relación que podría haber entre las actitudes respecto a las encuestas y las características demográficas, como ser la experiencia y el género del profesor (Nasser & Fresko, 2002).

En síntesis, las cuestiones abordadas por la literatura han sido las siguientes: en primer lugar, la preocupación por las percepciones acerca de la importancia y el propósito general del instrumento de evaluación (propósito). En segundo lugar, las creencias acerca de la utilización de los resultados con fines sumativos o con fines formativos, es decir, para la mejora de la práctica docente (uso). Y, en tercer lugar, las cuestiones asociadas a los estándares que adoptan los estudiantes en la evaluación, y la preocupación por las cuestiones de validez y confiabilidad (confiabilidad).

#### *1.4.1.2 Coordinadas metodológicas*

Pareciera ser que, para estudiar este fenómeno, el paradigma que ha primado ha sido el cuantitativo. Es así que la mayoría de los estudios se valieron de cuestionarios con preguntas cerradas con escalas tipo Likert de cinco puntos, con escalas que incluyen el totalmente en desacuerdo en un extremo, y el totalmente de acuerdo, en el otro (Marlin,

1987; Nasser & Fresko, 2002; Sojka *et al*, 2002; Campbell & Bozeman, 2007; Balam & Shannon, 2010; Slocombe *et al*, 2011; Abedin *et al*, 2013; Asassfeh *et al*, 2014; Chan *et al* 2014). Estos cuestionarios fueron desde 16 (Abedin *et al*, 2013) a 71 ítems (Asassfeh *et al*, 2014). Solo algunos estudios combinaron cuestionarios de respuesta cerrada con algún tipo de respuesta abierta o espacio para comentarios (Marlin, 1987; Chan *et al*, 2014). Asimismo, hubo solo un caso, el del estudio de los mexicanos García Garduño y Medécigo Shej, donde se utilizó una encuesta, aunque con ítems de preguntas abiertas en su totalidad.

Como hemos anticipado, fueron pocos los casos que utilizaron metodologías mixtas. Uno de ellos fue el estudio de Campbell y Bozeman (2007), donde realizaron entrevistas grupales con los docentes post-encuesta y donde, además, condujeron entrevistas en profundidad para explorar las percepciones de los administradores/tomadores de decisiones. De la misma forma, Capa-Aydin *et al* (2017), condujeron entrevistas semi-estructuradas “para profundizar en las percepciones de los miembros de la facultad sobre la evaluación del curso estudiantil de una manera profunda” (s/p).

Por último, respecto a las dimensiones de la muestra, la mayoría de los estudios se llevaron a cabo en universidades de gran magnitud, por lo que las muestras fueron, en general, bastante amplias. En el caso de los estudios que exploraban las percepciones de los estudiantes, las muestras fueron como mínimo de 163 (García Garduño y Medécigo Shej, 2014) y como máximo de 968 estudiantes, en el caso de Balam & Shannon (2010). Por otra parte, en relación a los docentes, la selección más pequeña fue de 12 (Capa-Aydin *et al*, 2017) y la más extensa fue de 447 docentes, en el caso del estudio de Nasser & Fresko (2002). Cabe aclarar que, en el caso del estudio de Capa-Aydin *et al* (2017), la cantidad de informantes se encuentra asociada con el tipo de estudio conducido. En ese caso, no se buscó obtener una muestra estadísticamente significativa, porque se trataba de una investigación enteramente cualitativa, un estudio de caso único. Por consiguiente, no admitiría (o al menos sería por demás costoso) una muestra tan extensa como la de los estudios que utilizaron encuestas de respuesta cerrada, procesadas posteriormente por sistemas de análisis de datos como SPSS (Abedin *et al*, 2013).

#### 1.4.1.3 Coordinadas espacio-temporales

Por último, creemos pertinente hacer referencia a la temporalidad de los estudios. Si bien podemos ver que ya en 1987 Marlín comenzaba hacer de las percepciones objeto de estudio de este campo, recién para comienzos del siglo XXI se empieza a prestar más atención a las percepciones. Si observamos el **Cuadro 2**, podemos ver que el grueso de los estudios fue realizado en la última década.

De igual importancia es la dimensión geográfica, pues como podemos evidenciar en el **Cuadro 2**, una gran proporción de las investigaciones ligadas a las percepciones se han llevado a cabo en países orientales y en Estados Unidos. Es importante resaltar, en este punto, que han sido pocas las investigaciones desarrolladas en el contexto hispanohablante. Si bien hemos podido encontrar una fuerte línea de investigación en relación a la evaluación de la docencia universitaria en México, con autores como Rueda, Luna Serrano, Torquemada, por nombrar algunos. En relación a las concepciones, solo hemos dado con la investigación de García Garduño y Medécigo Shej (2014). En el contexto nacional no se han encontrado estudios relativos a esta temática puntual.

#### 1.4.1.4 Hallazgos

##### Concepciones de los estudiantes

Para comenzar, vale la pena detenerse en las creencias de los estudiantes en relación al propósito de las encuestas, es decir, respecto a la importancia que pueden atribuirle o no al proceso, su posición frente al mismo, entre otras cosas. Sobre los puntos anteriores, la mayoría de los estudios (Marlín, 1987; Abedin *et al*, 2013; Assasfeh *et al*, 2014) exponen que los estudiantes creen fuertemente que el proceso de evaluación del docente y del curso es importante y que ellos, no solo están calificados (Balam y Shannon, 2010), sino que es su responsabilidad completar evaluaciones formales de sus profesores (Campbell y Bozeman, 2007). Los hallazgos de Assasfeh *et al* (2014) resultan pertinentes para ilustrar lo anterior, pues los autores exponen que “las creencias más fuertes eran pertinentes para ver la evaluación como un derecho no solo para el estudiante sino también un deber que un estudiante debe cumplir” (129).

Con respecto a la utilidad formativa de las encuestas, a pesar de que, como observamos en el párrafo anterior, los estudiantes consideran que la evaluación docente es importante y que la encuesta es un medio efectivo para evaluar, los estudios han

demostrado que existe cierta incertidumbre respecto a su efecto en las prácticas pedagógicas. Por un lado, algunos estudios han revelado que la mayoría de los estudiantes sostienen que los profesores efectivamente realizan cambios a fin de mejorar su curso a partir de los comentarios (Marlin, 1987; Abedin *et al*, 2013). Sin embargo, otros han evidenciado que los alumnos consideran que las evaluaciones no contribuyen de manera significativa a la mejora de las prácticas docentes (Asassfeh *et al*, 2014), pues argumentan no haber percibido cambios en las clases basados en los comentarios (Sojka *et al*, 2002). Lo anterior, como sugieren Asassfeh *et al* (2014), podría motivar en los estudiantes la impresión de que las evaluaciones no son tomadas de manera seria y comprometida por parte de los docentes, convirtiéndose, de esta forma en una pérdida total de tiempo (Marlin, 1987).

En cuanto a la preocupación por los potenciales usos de estas encuestas, en general los estudios han demostrado que los estudiantes carecen de certezas respecto al posible impacto que pudieran tener sobre los docentes, así como también poseen poca noción sobre el rol que ocupan en los procesos de decisiones en general (Campbell & Bozeman, 2007; Sojka *et al*, 2002). De todas formas, los estudiantes han expresado que, en la mayoría de los casos, consideran que “es poco el efecto que un estudiante puede tener en la carrera de un profesor” (Marlin, 1987: 714; Sojka, *et al*, 2002;). Quizás, una razón de ello, sea la creencia de los estudiantes de que, pese a la existencia de una maquinaria para informar a los tomadores de decisiones, nadie les presta mucha atención a los resultados de la evaluación, o al menos no se evidencian cambios atribuibles a ella (Marlin, 1987). Una evidencia de ello pareciera ser que, según los estudiantes, las encuestas que resultan muy malas parecen no tener tanto impacto. En línea con esto, en la investigación de Assasfeh *et al* (2014), los estudiantes manifestaron que creen que la encuesta debe tener un rol activo en las decisiones que involucran la renovación de contratos o la promoción.

Por otra parte, como hemos anticipado, hay quienes, además de preocuparse por creencias que rondan los potenciales usos de las encuestas de evaluación docente, se preguntaron si los estudiantes se toman efectivamente las evaluaciones de manera seria y comprometida (Brandenburg *et al*, 1979; Marlin, 1987). Si bien siguiendo a Wachtel (1998), esta cuestión ha sido difícilmente abordada por la literatura existente, proponemos algunos estudios que así lo han investigado.

Marlin (1987) por ejemplo, encontró –a través de su encuesta– que más de la mitad de los estudiantes de su muestra declararon que se toman el tiempo suficiente para responder las encuestas y que intentan ser justos en sus calificaciones (710-714). Sin

embargo, el autor también advierte que los estudiantes tienden a ver las evaluaciones como “una ventana para desahogar el vapor estudiantil” (714). En otras palabras, se convierte en la oportunidad de los estudiantes para decir todo lo que piensan, realizando en algunos casos duras críticas y en otros comentarios complacientes, aunque siempre marcados por un cierto halo de impunidad. Los hallazgos en la mayoría de los estudios relevados parecen ser concluyentes, pues la mayoría de los sujetos parecen coincidir con que, no solo se sienten cualificados, sino que, también, se toman las evaluaciones con un alto grado de seriedad y compromiso (Sojka *et al*, 2002; Campbell & Bozeman, 2007; Abedin *et al*, 2013; Balam & Shannon, 2010). No obstante, conviene resaltar que, como muestra la investigación de Slocombe *et al* (2011), menos de la mitad de los estudiantes consideran que los otros estudiantes se las toman en serio.

Por otro lado, y dada la importancia que han adquirido las investigaciones que estudiaron las cualidades psicométricas de este instrumento de evaluación, hay quienes se han preguntado por la percepción de los estudiantes sobre aquellos factores que han sido catalogados como potenciales sesgos de la evaluación docente. Algunas de las cuestiones que han sido recurrentes en la literatura fueron la relación entre la evaluación del profesor con variables como: (i) la personalidad del docente, (ii) su género, (iii) la exigencia del curso y (iv) la expectativa de la calificación que tienen los estudiantes a la hora de evaluar. En relación a la primera variable, el estudio de Sojka *et al* (2002), sugirió que los estudiantes niegan calificar mejor a los profesores por ser ‘entretenidos’, aunque, por el contrario, otros estudios, arrojaron que la gran mayoría de los estudiantes están totalmente de acuerdo con que “un estilo de enseñanza emocionante” (Assasfeh *et al*, 2014), o que involucre el humor (Slocombe *et al*, 2011), se vuelven criterios a considerar. Por lo que refiere al género, los estudiantes no parecen estar de acuerdo con que los profesores hombres tienden a recibir mejores evaluaciones (Assasfeh *et al*, 2014; Slocombe *et al*, 2011)

Acerca de la tercera variable, algunos estudios han revelado que los estudiantes niegan dan peores calificaciones a aquellos cursos que implican más trabajo o aquellos profesores que dictan cursos más difíciles (Sojka, *et al*, 2002; Slocombe *et al*, 2011; Assasfeh *et al*, 2014). Por último, en lo que respecta a la hipótesis de la benevolencia, es decir, evaluar mejor a los docentes que dan calificaciones más altas; los hallazgos no han sido concluyentes. Por un lado, hay estudiantes que niegan calificar mejor a aquellos profesores que son benevolentes o asignan mejores calificaciones (Sojka, *et al*, 2002;

Slocombe *et al*, 2011; Assasfeh *et al*, 2014); mientras que otros admiten ser más positivos a la hora de evaluar si esperan tener una buena calificación (Abedin *et al*, 2013).

### Concepciones de los docentes

Como hemos adelantado, las opiniones que ha expresado el profesorado frente a este tema, van desde el apoyo y la aceptación, hasta la oposición y el rechazo completos (por ejemplo: Braskamp y Ory, 1994; Centra, 1993; Seldin, 1993). Sin embargo y, a pesar de lo anterior, a menudo existe una impresión general de que los profesores tienden a objetar el uso de este instrumento de evaluación y las calificaciones provistas por los estudiantes (Franklin & Theall, en Wachtel, 1998). Una de las principales causas de lo anterior, ha sido la desconfianza que los profesores han expresado respecto a calidad y legitimidad de esta medida (Sojka *et al*, 2002; Abedin *et al*, 2013; Chan *et al*, 2014; Capa-Aydin *et al*, 2017). En este punto, uno de los argumentos más recurrentes ha sido la sospecha frente a las actitudes o los estándares que los estudiantes pudieran adoptar a la hora de completar las evaluaciones.

La cuestión de la seriedad con la que los profesores perciben que los estudiantes se toman las encuestas constituye un ejemplo claro de lo anterior. En este punto, si bien estudios como los de Abedin *et al* (2013) han detectado que casi un 70% de los docentes están de acuerdo o muy de acuerdo con que los estudiantes los evalúan seriamente, estudios como los de Sojka *et al* (2002) y Nasser & Fresko (2002), han encontrado que los profesores están bastante en desacuerdo con esta afirmación. Aquellos que expresan desacuerdo, hacen referencia a la existencia de otros factores que podrían atentar contra la objetividad de la evaluación, pues la mayoría coinciden en que los estudiantes evalúan a los profesores basándose en su personalidad o entusiasmo, “ven la evaluación por parte de los estudiantes como un indicador de cuán simpático [o entretenido] es un profesor” (*ibídem*: 192). Asimismo, han expresado cierta inquietud frente a los sesgos que, afirman, intervienen en el proceso de evaluación; como ser la dificultad del curso, el interés previo por la materia, la cantidad de trabajo requerido, y la insatisfacción con la calificación obtenida, entre otros (Capa-Aydin *et al*, 2017; Chan *et al*, 2014). En síntesis, los estudios sugieren que los profesores tienden a desacreditar las encuestas, incluso más que los estudiantes, bajo el argumento de que no son confiables ni tampoco válidos (Balam & Shannon, 2010).

A pesar de lo anterior, los estudios han encontrado que, en general, los docentes tienen una actitud relativamente positiva frente a la utilidad de estos instrumentos (Nasser & Fresko, 2002; Abedin *et al*, 2013), pues argumentan que los insumos provistos por la encuesta les resultan provechosos para mejorar su enseñanza (Balam & Shannon, 2010). Particularmente, se han mostrado especialmente positivos frente a la utilidad de los comentarios escritos o las preguntas abiertas (Nasser & Fresko, 2002). Pese a ello, aparentemente pocos docentes han hecho uso de la retroalimentación para realizar cambios, o al menos cambios que considerasen significativos (Nasser & Fresko, 2002; Sojka *et al*, 2002), pues el único estudio donde los docentes han reportado hacer cambios para ajustar sus cursos ha sido el de Chan *et al* (2014). Una de las dificultades por las cuales los docentes podrían verse confundidos a la hora de decidir si hacer cambios o reorganizar sus cursos, es la extrema variación que existe en las percepciones de los estudiantes: “un estudiante que alaba completamente la enseñanza frente a otro que la critica severamente” (Capa-Aydin *et al*, 2017: 2). En general, la decisión de hacer cambios o mejorar como respuesta al *feedback* de los estudiantes es compleja y lejos de depender de un factor, depende de varios factores; por ejemplo, como la cantidad de estudiantes que reportaron la misma experiencia, la importancia relativa que un profesor le otorga a cierto aspecto del curso.

Si bien hasta ahora hemos observado actitudes variadas hacia la efectividad de los CEDA como una medida de la enseñanza, la preocupación frente su uso para la toma de decisiones *high-stake* parece ser unánime. Los estudios han revelado que los docentes expresan ciertas reservas frente al uso dominante y crítico de estos puntajes para decisiones de promoción y *tenure*. Esto es, manifiestan cierta preocupación frente a la idea de que este sea el único mecanismo de evaluación del desempeño docente (Chan *et al*, 2014). Uno de los argumentos centrales que subyace esta creencia es la idea de que la introducción de evaluaciones obligatorias podría potencialmente motivar a los profesores a reducir los estándares, la exigencia y a ser a ser más benevolentes con las calificaciones para obtener mejores evaluaciones (*ídem*). A lo anterior, se le suman los hallazgos de Nasser & Fresko (2002), que acaban por complejizar este asunto aún más, sugiriendo que la oposición a las encuestas no es solo una cuestión ideológica sino una suerte de mecanismo de racionalización de interés propio. En otras palabras, “los docentes apoyan las evaluaciones cuando les resultan aduladoras, y las oponen cuando son criticados” (*op. cit.*: 192. Traducción propia).

En definitiva, los profesores manifiestan actitudes positivas y negativas frente al uso de los CEDA. Mientras que algunos se encuentran muy a favor de las evaluaciones y abogan por la necesidad de ser evaluados por sus estudiantes para mejorar la calidad de su práctica pedagógica (Nasser & Fresko, 2002; Abedin *et al*, 2013; Capa-Aydin *et al*, 2017). Otros expresan reticencias frente al uso de este instrumento de forma crítica para la toma de decisiones. En resumidas cuentas, las palabras de Nasser & Fresko ilustran el estado de la cuestión, pues sugieren que “[l]a variabilidad de la respuesta de los profesores es consistente con la controversia que rodea la validez y el mérito de la calificación de estudiantes que se encuentra en la literatura académica” (2002: 192. Traducción propia).

### Percepciones de los administradores

Son pocas las investigaciones que se han avocado a explorar las percepciones de los tomadores de decisiones (Campbell y Bozeman, 2007; Marlin, 1987). Incluso, las que lo han hecho, lo han realizado de manera secundaria, como un complemento más a la hora de entender mejor las percepciones de otros actores, sean profesores o estudiantes.

Lo que los trabajos han demostrado, hasta el momento, es que los tomadores de decisiones valoran la práctica de las encuestas y consideran que “tener un proceso que le da voz a los estudiantes es vital para la integridad institucional” (Campbell y Bozeman, 2007: 21. Traducción propia). Sin embargo, pese a que muchos de ellos indican que miran y les prestan atención a los resultados, creen que el impacto de las encuestas en la enseñanza depende “de la capacidad del instructor de ser conscientes sobre los comentarios que reciben y de efectivamente tener ganas de tomar esos comentarios para mejorar tu performance” (Marlin, 1987: 22. Traducción propia)

Por otro lado, en relación a las potenciales consecuencias derivadas de la utilización de este instrumento para la toma de decisiones, los administradores consideran que los resultados de las encuestas no impactan directamente sobre el salario de los docentes, aunque sí indirectamente, de tenerse en cuenta para decisiones de promoción o *tenure* (Marlin, 1987; Campbell & Bozeman, 2007). Pese a lo anterior, los entrevistados en el estudio de Campbell y Bozeman, afirman que las consecuencias dependen del tipo de contrato del profesor en cuestión, sugiriendo que “la información provista tiene un efecto mínimo en los profesores *full time*” (2007: 23. Traducción propia). Finalmente, cabe destacar que la administración que participó en el estudio de estos autores, afirmó no confiar exclusivamente en las encuestas a la hora de tomar decisiones en torno al

desempeño docente, “sino que los *ratings* eran tomados en cuenta como una fuente de información más” (*idem*. Traducción propia).

### 1.5 Nuestro estudio

Como se ha venido sugiriendo, todavía no existen consensos acerca de la utilización de los CEDA para la evaluación de la docencia en el contexto de la educación superior. De hecho, los estudios sobre las percepciones respecto a su propósito, su uso y su confiabilidad, han demostrado que las opiniones son variadas e incluso contrapuestas. Sin embargo, y a pesar de ello, su utilización se ha extendido mundialmente. Argentina, no ha escapado a ello.

Hacer un estudio de esta naturaleza en el contexto nacional resulta menester pues no se han encontrado antecedentes que aborden esta temática en el país. El estudio de caso en cuestión, constituye tan solo un ejemplo de la utilización generalizada de los cuestionarios de evaluación docente, cuya lógica ha permeado en las instituciones –a pesar del desconocimiento de la literatura o del poco desarrollo en materia de investigación de la temática que existe en el contexto nacional. Asimismo, como afirman Fernández Lamarra y Coppola:

en términos comparados con el resto de la región (...) en Argentina, el tema de la evaluación de la docencia universitaria constituye un área de vacancia que, a la vez, limita la interpretación de los estudios que toman como objeto los cambios recientes en nuestro sistema de educación superior (2008a: 97).

En otras palabras, la importancia de este trabajo de investigación radica en la necesidad de resolver algunos interrogantes que surgen a partir de la utilización de la encuesta, como instrumento de evaluación que se ha puesto en práctica en la región sin poder ser cuestionado. A nuestro juicio, el desconocimiento acerca de las creencias, suposiciones y actitudes que los grupos de interés construyen alrededor de su uso atenta, o al menos limita, la legitimidad y por tanto el potencial de los CEDA.

Para finalizar, sostenemos que si se tiene la intención de hacer de este instrumento de evaluación una herramienta útil para la mejora de la calidad de la enseñanza en educación superior y evitar, por el contrario, que se produzcan deformaciones del valor de la experiencia universitaria –cargadas de significaciones vacías y sensaciones de descontento, conocer los puntos de vistas de los grupos de interés involucrados en los procesos de evaluación se vuelve un paso fundamental. Las creencias y concepciones de los actores serán los cimientos sobre los cuáles se deberán construir las bases de una

nueva cultura de la evaluación. Explorar las concepciones implícitas de los sujetos nos ayuda a determinar y conocer las prácticas de evaluación y así trazar líneas para su mejora (Hidalgo y Murillo, 2017).



Universidad de  
**San Andrés**

## CAPÍTULO 2

### ENCUADRE TEÓRICO

La pregunta sobre el papel de la encuesta en la evaluación docente es la pregunta por el por qué y para qué de esta instancia de evaluación. En otras palabras, es la indagación de aquello qué se entiende por evaluación y las decisiones que se derivan de ellas. En este sentido, a fin de conceptualizar lo encontrado en el campo, resulta útil recorrer los distintos tipos y modelos de evaluación posible. Además, en tanto nuestra investigación se enfoca en la utilización de un instrumento de evaluación docente concreto, hacemos una aproximación a las estrategias de evaluación docente y luego precisamos el caso de los cuestionarios estudiantiles. Sin embargo, primero que todo, nos detenemos a definir que entendemos por concepción.

#### 2.1 Las concepciones

Entendemos a las concepciones como “un sistema organizado de creencias originadas en las experiencias del individuo y desarrolladas a través de las distintas interacciones que tiene la persona con su entorno (Coll y Remesal, 2009, en Hidalgo y Murillo, 2017). De esta manera, dado que se construyen y fundamentan en espacios compartidos con otros, las concepciones tienen un carácter marcadamente social. Asimismo, las concepciones que cada individuo construye, inciden en la forma en que se comporta e interacciona a lo largo de su vida. Por lo tanto, aquellas concepciones que tenga cada persona, determinan en gran medida su forma de actuar y la intencionalidad de dichos actos (*ídem*).

Aplicadas al ámbito educativo, las concepciones pueden definirse como la estructura de creencias y vivencias que tienen los distintos actores educativos y que, construidas en un contexto social, determinan en cierta medida la práctica educativa. Precisamente, la importancia de las concepciones radica en este último punto, pues en tanto las concepciones determinan las prácticas, conocerlas, nos permitirá establecer vinculaciones con las actuaciones y las actitudes adoptadas en los procesos de evaluación.

Por último, cabe aclarar que, en esta investigación, se entienden a las concepciones de manera amplia, como “una estructura mental general que abarca creencias, significados, conceptos, posiciones, reglas, imágenes mentales, preferencias, etc. que configuran la forma en que las personas nos enfrentamos a distintos fenómenos”

(Thompson, 1992, en Hidalgo y Murillo, 2017: 109). De esta forma, a pesar de que muchas investigaciones han diferenciado todos los términos comprendidos en la definición anterior, aquí se considerarán como sinónimos, y se utilizarán de manera indistinta.

## 2.2 La evaluación de la docencia universitaria

En las últimas décadas, la evaluación de la docencia ha ocupado un lugar importante en los debates educativos, principalmente en los que tienen como contexto la educación superior. Si entendemos a la evaluación como algo más que una tarea técnica, “como un proceso público-político dotado de repercusiones sociales (Dias Sobrinho y Balzan, 1995 en Calderón *et al*, 2018: 267)”, podemos afirmar que necesariamente asumirá diversas connotaciones. Así es que, a pesar de que existe un consenso acerca de que la evaluación de la docencia será esencial para lograr una mejora en la calidad de los docentes (Rueda, 2008), aún está abierto el debate acerca de su definición, los componentes de sus procesos, y las estructuras metodológicas que la pudieran orientar (Fernández y Coppola, 2008b: 97). Eco de lo anterior, resultan las divergencias que exponen las distintas tendencias en materia de evaluación de la docencia universitaria, que acorde a Fernández Lamarra y Coppola (2008b), han sido tres:

1. Orientada a resultados, siendo la evaluación de la docencia parte de la evaluación institucional, relacionada con el control o rendición de cuentas. Este tipo de evaluación es efectuada, generalmente, por agencias o por pares externos a las instituciones, tomando de base las autoevaluaciones institucionales.
2. La evaluación orientada a la formación, donde la de la docencia aparece como un proceso formativo, sistemático y permanente; integrada a la actividad cotidiana de la enseñanza, operando al interior de las cátedras y contribuyendo a mejorarlas.
3. La evaluación de la docencia relacionada con políticas institucionales, para el desarrollo los recursos humanos. Vincula con la carrera docente y su relación con el ingreso, formación, permanencia, promoción, incremento salarial, incentivos a la investigación, etc. (98)

Como podemos observar, el autor propone tres orientaciones de la evaluación, con fines claramente distintos. La primera, hace referencia a una evaluación docente pensada para el ámbito central de la universidad, en el marco de un sistema de regulación o de aseguración de calidad externo, y posiblemente asociada a una instancia de acreditación. La segunda, a una evaluación pensada en el marco de los sistemas de regulación al interior de las universidades, orientada directamente a los profesores como principales usuarios y beneficiarios de esta instancia. Por último, una tercera, inscrita en medio de marcos inter

e intrainstitucionales, y profundamente ligada a procesos de desarrollo de la carrera docente.

### 2.2.1. La dificultad de evaluar la enseñanza

Ahora bien, sea cual fuere el uso o la perspectiva adoptada por una institución, existen algunos problemas que resultan comunes a la evaluación del desempeño docente, que se derivan, principalmente, de la complejidad que entraña la práctica pedagógica (Murphy, MacLaren y Flynn en Surgenour 2013). El primero de ellos, nos advierten Rueda *et al* (2010), se deriva de la dificultad de definir las cualidades o características de un “buen profesor”. Asociado a esto, la dificultad que supone la naturaleza subjetiva de lo que puede considerarse una buena enseñanza (Cohen en Surgenour, 2013), pues no existe suficiente evidencia para establecer cuáles son las dimensiones de la docencia efectiva y cómo éstas se interrelacionan (Abrami, 1989).

En relación a lo anterior, y, en segundo lugar, el problema que constituye la falta de una definición explícita de las funciones o expectativas asignadas por la institución a la actividad docente. Teniendo en cuenta que normalmente la evaluación del profesorado se ha basado en la performance en tres áreas básicas –la docencia, la investigación y el servicio (Marlin, 1987)–, no parece quedar del todo claro cuáles son las responsabilidades de un docente en cuanto a la práctica pedagógica. Asimismo, la vaguedad de las definiciones de lo que implica ser un ‘buen profesor’ o brindar una ‘enseñanza de calidad’, no parecen echar luz a la cuestión.

Al mismo tiempo, un problema creciente es la selección de indicadores que sean precisos, claros y confiables. En este punto, Tejedor Tejedor y García-Valcárcel sostienen que “[l]a mayoría de los indicadores que se utilizan para evaluar al profesorado están relacionados con la productividad investigadora y muy pocos están referidos a la calidad de la docencia” (2003: 162). Es probable que esa sea la consecuencia lógica de la mayor dificultad que resulta evaluar el proceso de enseñanza, frente a otras cuestiones como la investigación, donde posiblemente baste con ver la cantidad de artículos publicados. Asimismo, y en estrecha relación con los indicadores, esta también el desafío que reviste la selección de una estrategia de evaluación que permita captar la complejidad del acto educativo. Para Marsh (1987), desafortunadamente, la mayoría de los instrumentos dedicados a medir el desempeño docente no han podido reconocer la naturaleza multidimensional de la docencia.

Como hemos visto hasta ahora, varios obstáculos se presentan a la hora de evaluar la docencia. Aún así, no se debe dejar de reconocer que, como sugiere Fernández Lamarra (2010), la evaluación es:

una oportunidad para las instituciones y para los sistemas educativos para promover la mejora en las instituciones, para que cuenten con información y mantengan un estrecho seguimiento en cuanto al logro de los fines y objetivos propuestos, así como el grado de pertinencia y trascendencia de estos para garantizar que efectivamente responden a lo que ofrecen y a lo que la sociedad espera de ellas. (35)

En otras palabras, desde nuestra perspectiva, la evaluación de la docencia resulta una instancia clave para generar información que pueda ser utilizada posteriormente para el crecimiento de los docentes, pero también de las universidades.

### 2.3 Modelos de evaluación docente

Hablar de evaluación docente, como se ha venido sugiriendo, es esencialmente una tarea compleja. Tal como proponen Calderón, Barrera, Marshal y Wandercil (2018), parte de su complejidad se debe esencialmente a las potenciales connotaciones que pudiera adoptar, sea “tanto para justificar políticas de acreditación, evaluación a gran escala y ranqueamiento, como para propiciar espacios formativos y procesuales para la mejoría del desempeño docente en su dimensión académica pedagógica” (Calderón *et al*, 2018: 250). De igual modo, no se puede desconocer la diversidad de coordenadas que se deben tomar en cuenta a la hora de tomar decisiones sobre qué evaluar, cómo, y cuándo; pues éstas se corresponderán a cada universidades, carrera o programa, configurándose así diferentes formas de evaluar la labor docente (Rueda, 2008). Como resultado, siendo los procesos e instrumentos de evaluación docente adoptados por las instituciones de lo más diversos, también lo han sido sus categorizaciones.

Algunos proponen clasificar la evaluación con criterios temporales, acorde al momento en el que se realiza la evaluación; otros, lo hacen acorde a su función. También se puede hacer una tipología según “los componentes que conforman el proceso de evaluación” (Elola *et al*, 2010: 62) o según las decisiones de los docentes sobre el proceso de evaluación. Es decir, en la literatura sobre evaluación hay diversas categorizaciones que, además, pueden tener el mismo nombre y presentar características diferentes o tener nombres distintos y presentar características similares.

Con todo, y con la intención de esclarecer nuestro análisis, elaboramos un cuadro de tipologías propias. Para dicha categorización tomamos, por un lado, la categorización que hacen Calderón *et al* (2018), porque es una tipología de referencia en la literatura

contemporánea sobre evaluación docente. Por el otro, las categorías de estrategias evaluativas de Amarante y Daura (2012), porque permiten un abordaje mucho más amplio que las propuestas por Calderón *et al* (2018). El resultado, se puede evidenciar en el cuadro que presentamos a continuación:

**Cuadro 3. Tipología de evaluación docente**

<b>Tipologías de evaluación docente</b>	<b>Finalidad de la Evaluación</b>	Evaluaciones Formativas
		Evaluaciones Sumativas
		Evaluaciones Híbridas
	<b>Usos de los resultados</b>	Acciones <i>Low-Stakes</i>
		Acciones <i>High-Stakes</i>
	<b>Momento de la evaluación</b>	Evaluaciones durante el proceso
		Evaluaciones al final del proceso
		Autoevaluación
	<b>Estrategias evaluativas</b>	Heteroevaluación
		Evaluación Mixta

Fuente: elaboración propia a partir de Calderón et.al. (2018) y Amarante y Daura (2012)

Conviene subrayar que la tipología no implica determinismo ni agota las posibilidades de las prácticas educativas. Los límites entre los tipos de evaluación suelen ser difusos y las instituciones de educación superior tienen la posibilidad de evaluar sus docentes utilizando los distintos tipos de evaluación y, más aún, utilizando una combinación entre ellos. Con esta salvedad, este cuadro constituye una aproximación a los distintos tipos de evaluación que serán caracterizados a continuación. Es decir, el **Cuadro 3**, servirá de eje organizador del capítulo.

### 2.3.1 Finalidad de la evaluación

La evaluación educativa puede definirse como un juicio sobre el valor educativo de un sistema, una institución o un programa, o de cualquier realidad educativa (Camillioni, s/f: 8-9). En este sentido, la práctica evaluativa resulta indisociable de un propósito o una finalidad, así como también de un cierto posicionamiento (Vain, 1998). Dada la infinidad de posibilidades posibles, partiremos de la propuesta de Scriven (1967) que ya, hace casi cinco décadas, distinguía según la finalidad de la evaluación, dos grupos: de control/sumativa o de mejora/formativo.

Antes de detenernos en la caracterización de estas dos grandes matrices referenciales, es conveniente mencionar que, por un lado:

independientemente del tipo de evaluación docente adoptada, no existe a priori una estrategia mejor o menos eficaz. Existe la necesidad de adecuarlas a los objetivos institucionales y a la realidad institucional, y hacer uso adecuado de los resultados siguiendo los procesos decisorios internos de cada IES y los objetivos institucionales (Calderón, *et al*, 2018: 265).

Por el otro que, es necesario advertir los peligros que podría suscitar el hecho de encontrarse con matrices didácticamente divididas, pues “en determinados contextos acaban generando pasiones que acaban por dicotomizarlas de forma maniqueísta” (*ibidem*: 264). Con todo, a pesar de que hayamos tomado la decisión de tratarlos separadamente, lo hemos hecho con intenciones prácticas, lejos de intentar alentar este tipo de comportamiento reduccionista y simplista del que nos advierten Calderón y otros (2018).

#### 2.3.1.1 Evaluación sumativa

Históricamente, se toma a la evaluación sumativa para caracterizar a las prácticas que se encuentran asociadas a la certificación o acreditación. Según Fernández Lamarra y Aiello (2012), “[l]a evaluación sumativa puede definirse como una “práctica administrativa” que tiene por objeto el control de las acciones educativas desarrolladas” (170). De esta forma, este tipo de evaluación está dirigida a individuos y productos objetivando medir el desempeño del sujeto y la identificación de sus resultados de trabajo. De manera similar, Dias Sobrinho (2003, en Calderón *et al*, 2018) advierte que la evaluación sumativa:

es una evaluación burocrática, de prestación de cuentas y responsabilización. Sigue la lógica del control y objetiva recoger información que sirvan de soporte para la toma de decisiones relacionadas a la gestión eficientista, especialmente en lo que se refiere a las cuestiones contractuales del profesor, a las promociones y bonificaciones por mérito y productividad (168)

Podemos vislumbrar en este tipo de evaluación la presencia de un espíritu clasificatorio, signado por la intención de establecer, a veces de manera dicotómica, el éxito o el fracaso de un docente con respecto a un comportamiento o un desempeño esperado. Bajo esta premisa, se asegura que determinado individuo ha cumplido -o no- los propósitos acordados, o los criterios a alcanzar; tomando como resultado, para un caso u otro, decisiones con “consecuencias directas [sobre el individuo], propias del proceso administrativo que representa” (Fernández Lamarra y Aiello, 2012: 170). En

consecuencia, autores como Perrenoud (2010), han adoptado una postura crítica, y han tildado este tipo de evaluaciones como ‘perjudiciales’. El autor, se ha demostrado en contra de la generación de jerarquías de excelencia, propias de un tipo de evaluación que “[construye] un ranking de los sujetos en función de algún patrón o norma” (Elola *et al*, 2010: 77).

### 2.3.1.2 *Formativa*

A diferencia de la evaluación sumativa, la evaluación *formativa*:

se presenta como una relación colaborativa, crítica y constructiva, debiendo orientarse a contribuir a la mejora de las prácticas educativas realizadas a través de juicios de valor que toman la forma de recomendaciones o sugerencias de mejora” (Fernández Lamarra y Aiello, 2012, :172)

Así es que, subsumida bajo el compromiso de la mejora continua, este tipo de evaluación tiene por objeto “asegurar la progresión de cada individuo (...) para aportar mejoras o correctivos apropiados” (Scallon, 1988:155).

En otras palabras, en este tipo de evaluación, lejos de orientarse a la consecución de una medida de éxito o fracaso, se privilegia el valor del proceso. Como resultado, la evaluación formativa no puede circunscribirse a un momento único, pues de acuerdo con Freitas, Conta y Miranda (2014, en Calderón *et al*, 2018), “es realizada a lo largo del proceso, es continua, y proporciona parámetros al profesor para verificar si los objetivos fueron alcanzados, pudiendo interferir en lo que puede estar comprometiendo el aprendizaje” (87). De esta forma, se privilegia el intercambio entre los actores y “sus juicios de valor y sus recomendaciones pretenden convencer sobre la conveniencia estratégica de ser aplicadas, más que la aplicación por su peso normativo” (Fernández Lamarra y Aiello, 2012: 172).

En síntesis, la evaluación docente entendida desde una dimensión formativa implica “reafirmar el valor del proceso y de las relaciones pedagógicas, [...] privilegiar las redes intersubjetivas que dan consistencia al esfuerzo colectivo, aunque no homogéneo de construcción de una universidad social y cualitativamente relevante” (Dias Sobrinho y Balzan, en Calderón *et al*, 2018: 267). En otras palabras, implica entender la evaluación como un proceso formativo sistemático y permanente, que no puede ser sino por la colaboración democrática de los distintos actores involucrados.

### 2.2.1.3 Evaluación híbrida

Como se ha venido sugiriendo, el uso efectivo de la evaluación dependerá del grado de adecuación que tengan con los objetivos y la realidad institucional. Esta visión puede ser englobada en un movimiento teórico que aboga por abordajes más eclécticos en el ámbito de la evaluación, como es el de las evaluaciones híbridas. En esta línea, Worthen, Sanders y Fitzpatrick proponen que, lejos de preocuparse por identificar cual abordaje es el mejor, “[e]scogiendo y combinando conceptos de abordajes evaluativos para adoptarlos a cada situación, usando partes de los diversos abordajes cuando parecen apropiados” (2004: 260), todos los abordajes son útiles.

Desde la perspectiva anterior, se sostendrá aquí que las concepciones teóricas no deberían circunscribirse a la devoción de una modalidad específica de evaluación. Calderón y Borges (2013, en Calderón *et.al.* 2018), de manera prudente, nos alertan sobre los peligros existentes de ser ‘discípulos irracionales’ de un determinado modelo de evaluación, convencidos de que:

solamente [el abordaje híbrido] sirve para toda y cualquier situación; invitándonos a escoger y combinar diversos conceptos, modelos, métodos y técnicas para adaptarlos a la situación que va a ser evaluada, usándolas siempre que parezcan más apropiadas para alcanzar los objetivos de la evaluación de forma justa y coherente (265)

### 2.3.2 Usos de los resultados: acciones *High-Stakes* y *Low-Stakes*

Como hemos adelantado, las estrategias de evaluación adoptadas por las universidades permiten identificar la coexistencia de concepciones de evaluación que se fundamentan en diferentes matrices teóricas y desencadenan usos específicos de los resultados.

Por un lado, tenemos las evaluaciones que desencadenan acciones *high-stakes*, que usualmente se encuentran relacionadas con un proceso de responsabilización, es decir, de acciones dirigidas a una clasificación y valorización individual. Siguiendo a Allen, “las evaluaciones *high-stakes* se encuentran dominadas por las demandas de la rendición de cuentas, la acreditación y el rigor científico” (2012: 641. Traducción propia) y, en este punto, es que puede verse asociada con el paradigma sumativo. Por contraste, las evaluaciones *low-stake*, son evaluaciones cuya repercusión o uso se ve orientado por acciones dirigidas a una evaluación formativa y reflexiva sobre el accionar docente, y cuyas consecuencias “son más simbólicas” (Brook, 2013, en Calderón *et al.*, 2018). Asimismo, tienen la reputación de ser más gentiles, y menos críticas (Allen, 2012). En

otras palabras, una vez que los resultados no implican acciones de responsabilización directas de los profesores, o sea, que de la evaluación o por los resultados insatisfactorios que los profesores pudieran tener no resultan acciones de punición y tampoco, a la inversa, fuertes acciones de estímulo para la mejoría del desempeño por medio de beneficios para los docentes, estaríamos ante la presencia de evaluaciones *low-stake*.

En un estudio realizado por Rueda (2008), en el contexto mexicano, se indica que solo excepcionalmente los resultados de la evaluación se emplean para generar acciones de formación y capacitación para los docentes. Por el contrario, se indica que los resultados de evaluación se emplean regularmente para dar información personalizada a los docentes e informar a las instancias administrativas. Si bien una generalización sería apresurada, autores como Rueda (2006, 2008) sostienen que, en el discurso del contexto de educación superior, prevalece el uso de la evaluación con la intención de controlar las acciones de los profesores o bien para recompensarlos a través de programas existentes. Aunque a simple vista parecieran distinguirse, la compensación y el control suponen caras de una misma moneda, cuya lógica está lejos de ser la de una evaluación formativa preocupada auténticamente por la mejora de la actividad.

Sin lugar a dudas, una actitud crítica o de permanente revisión de los procesos de evaluación se vuelve condición indispensable si es que se quiere evaluar al docente con el fin de formarlo en su accionar pedagógico. Para nuestra mala fortuna, y tal como señalan Calderón *et al*, aún nos queda un largo camino de estudio y profundización de las evaluaciones *high-stakes* en pos de “verificar hasta qué punto realmente pueden ser inductores de la mejoría del desempeño docente” (2018: 267).

### 2.3.3 Momento de la evolución: durante o al final del proceso

En cuánto a los momentos de evaluación, podemos identificar dos bien claros, la evaluación durante o al final del proceso. En cuanto a la primera, es una evaluación que se realiza durante la concreción de la propuesta de enseñanza-aprendizaje (De Ketele y Roegiers, 1995), es decir, que regula la acción o el estado del proceso al momento en el que se lleva a cabo para corregir los procedimientos, a fin de mejorarlos. Respecto a la evaluación al final del proceso, es una práctica que permite hacer un balance final. Corresponde más bien a la fase de síntesis y su objetivo es verificar y valorar lo realizado (Carlino, 2003). Por último, al realizarse al concluir una etapa o un proceso, no admite la posibilidad de cambio o réplica, sino hasta el inicio del siguiente ciclo.

A pesar de que, como hemos evidenciado, existen múltiples tipos de evaluación y combinaciones posibles, sostenemos que es posible advertir la existencia de dos grandes correlatos. Para su definición y evitando caer en las categorizaciones ya propuestas, que darían cuenta de una sola dimensión de la evaluación (finalidad, concepción, uso o momento) dimos uso a la nominación propuesta por Ricco (2001). El resultado fue la agrupación de las distintas categorizaciones en dos grandes modelos de evaluación (Ver Cuadro 4):

**Cuadro 4. Modelos de evaluación docente**

Dimensiones/Modelo	Tecnológico y productivista	Participativo
Finalidad	Control	Formación
Concepción	Sumativa	Formativa
Uso de los resultados	<i>High-Stake</i>	<i>Low-Stake</i>
Momento	Final del proceso	Durante el proceso

Fuente: elaboración propia

Por un lado, el *tecnológico y productivista*, marcado por la eficiencia y la productividad, teniendo modelo de referencia la empresa. Desde este modelo, la evaluación es concebida como instrumento de control y de regulación. Asimismo, se sobrevalora la medida estadística analizándose los productos o resultados e ignorándose los procesos, las especificidades y las complejidades institucionales. Este tipo de evaluación habilita decisiones de alto impacto y de responsabilizarían individual, dónde la mejora pedagógica no es una preocupación fundamental. Por el otro, el modelo de *evaluación participativa*, que supone una reflexión colectiva de sus miembros sobre la propia institución universitaria. Más que poner énfasis en los resultados, esta evaluación trata de contextualizar la acción de referencia en el medio socioeconómico, realzar la identidad del proyecto educativo y centrarse sobre la mejora de la oferta educativa. Generalmente, la evaluación participativa conlleva decisiones más bien simbólicas, destinadas a la consecución de una mejora durante el proceso mismo.

## 2.4 Estrategias de evaluación

Así como existen distintas concepciones y potenciales usos de los resultados de la evaluación docente que pueden ser considerados, la pregunta de quién y cómo evaluar puede suscitar la existencia de una extensa variedad de técnicas o instrumentos de evaluación. En efecto, así como se pueden adoptar distintas concepciones y formas de

puesta en marcha de los procesos de evaluación, también se definen diferentes roles para los distintos actores de la comunidad educativa: docentes, alumnos, directivos, evaluadores y expertos, agencias y para la sociedad. En este punto, cabe resaltar la complejidad que introduce a la evaluación, la selección y la ubicación de un actor predominante en el acto de evaluar, pues:

para cada actor la evaluación se referencia a un objeto distinto: para los académicos se refiere a los saberes; para los empleadores a competencias; para los estudiantes a la empleabilidad; para la sociedad a ciudadanos respetables y competentes; para el Estado, según la concepción que asuma, puede variar de aspectos vinculados con el desarrollo social y humano a la eficiencia, a los costos y a los requerimientos de capital humano (Fernández Lamarra, en Fernández Lamarra y Copolla, 2009: 144).

De esta forma, a la hora de seleccionar qué evaluar, se debe tener en cuenta la potencial amenaza que supone a la práctica de evaluación docente su indefinición, pues nunca resulta del todo claro cuáles son los componentes de la praxis educativa que deben de incluirse o tenerse en cuenta en dicho proceso. Al mismo tiempo, la selección de un tipo de actor determinado supone asumir, según Vain (1998), ciertos aportes y limitaciones:

- a. **Docentes:** como gestores principales de la enseñanza, poseen una mirada particular acerca de las condiciones de trabajo, sobre los alumnos, la universidad, etc. A su vez, sus ideas y creencias sobre la enseñanza y sobre la docencia como su principal función pueden contribuir a erradicar el carácter anónimo o aparentemente desvinculado de su situación concreta que parece primar en las evaluaciones. El autor nos advierte sobre una técnica en particular, la autoevaluación, porque considera que, en el contexto de la carrera docente, habida cuenta de las posibles consecuencias que ésta podría tener en la estabilidad laboral, podría “adquirir cierto grado de autocomplacencia” (89).
- b. **Estudiantes:** Aunque más adelante ahondaremos en esta cuestión, diremos por ahora que las posibilidades se vinculan, por un lado, a su situación de no pertenencia a la corporación docente, lo que los convierte en potenciales críticos de esta actividad. Por el otro, a su carácter de consumidores, pues son los principales implicados por las consecuencias de la enseñanza y quienes establecen una relación cotidiana con el profesor. Sin embargo, puede mencionarse como limitaciones la ausencia de una cultura de la autoevaluación en el propio alumno, la disparidad de niveles de inserción institucional entre

los ingresantes y los alumnos de los últimos años o el desconocimiento - en la mayoría de los casos - sobre las condiciones de trabajo de cada docente.

- c. **Pares:** el principal obstáculo es el escaso desarrollo de la cultura evaluativa, al cual se le suma el aporte de la proximidad con quienes serán evaluados, que puede tener implicancias en términos afectivos. No obstante, este mismo factor resulta en posibilidad, en tanto ambos (evaluador y evaluado), comparten un mismo contexto, y por tanto el conocimiento sobre esa realidad es más profundo y puede dar cuenta de una dimensión más procesual, así como de la historia.
- d. **Autoridades institucionales o gestores académicos:** su posición les permite visibilizar un marco de evaluación más amplio, que posibilita situar a la acción docente en relación con los planes estratégicos de la institución. Las limitaciones pueden darse, de situarse en contextos institucionales caracterizados por estructuras organizativas muy verticalizadas, y relacionados con estilos de gestión más autocráticos y de escasa participación

Por otra parte, si atendemos a los instrumentos y a los participantes de la evaluación docente, entonces daremos, según Amarante y Daura (2012) con tres grandes tipos de evaluación: la *Autoevaluación de la cátedra*, la *Heteroevaluación* y la *Evaluación Mixta*.

**Cuadro 5. Estrategias de evaluación docente según participantes e instrumento de evaluación**

Tipo de evaluación	Instrumento de evaluación	Participantes
<b>Autoevaluación</b>	Frases incompletas	Equipo docente
	Registros textuales	
	Problemas focalizados	
	Opiniones breves	
<b>Heteroevaluación</b>	Trabajos grupales	Grupo de alumnos
	Cuestionarios abiertos	
	Escalas de valoración	
<b>Evaluación Mixta</b>	Se aplican unos y otros instrumentos	Docentes y estudiantes

Fuente: Amarante y Daura (2012)

A partir de la propuesta de los autores y, atendiendo especialmente a los participantes, podemos encontrar tres estrategias de evaluación de evaluación docente. En primer lugar, la autoevaluación, que refiere a la evaluación llevada a cabo por los

mismos docentes. En segundo lugar, la heteroevaluación, que es llevada a cabo por los estudiantes. En tercer y último lugar, la evaluación mixta, que es aquella que conjuga los instrumentos de evaluación de los anteriores.

#### 2.4. 1 Evaluación del desempeño docente a través de cuestionarios estudiantiles

Los cuestionarios de evaluación de la docente por los alumnos (CEDA) son el método más tradicional y utilizado para que los estudiantes evalúen a los docentes. En ausencia de otros mecanismos estandarizados, los CEDA constituyen el instrumento de heteroevaluación más utilizado en el mundo (Seldin, 1993; Surgenour, 2013). Estos cuestionarios son utilizados para que los estudiantes evalúen el desempeño docente, utilizando escalas de valoración “tipo cafetería” (Algozzine *et al*, 2004: 135). Los sistemas de evaluación a través de cuestionarios estudiantiles presentan las siguientes características:

- a. Se desarrolla un instrumento que comprende una serie de preguntas abiertas y cerradas con escala tipo Likert, acerca del curso y del desempeño docente,
- b. al menos un ítem se refiere al desempeño general,
- c. se asume y se asegura el anonimato de las respuestas,
- d. las respuestas se obtienen al final del semestre y en la ausencia del docente
- e. los datos obtenidos se sintetizan en relación a docentes, departamentos y facultades como una evidencia del desempeño docente, habilitando ciertas decisiones de personal,
- f. y las diferencias entre las características de los estudiantes, cursos y docentes son ignoradas en el análisis

Asimismo, una cuestión importante en relación a los cuestionarios estudiantiles es la forma en la que son realmente utilizados. Siguiendo a Marsh (1989), este instrumento evaluación puede ser utilizado con cuatro propósitos: (a) para proveer *feedback* formativo a los docentes acerca de su desempeño con la intención de mejora, (a) como una medida sumativa del desempeño docente a utilizar en decisiones, (c) como una fuente de información para que los estudiantes elijan sus profesores y cursos, y (d) como una fuente de información para la investigación en enseñanza (Marsh y Dunkin, 1997)

De acuerdo con Surgenour (2013), de utilizarse con intenciones formativas, los CEDA se vuelven un instrumento invaluable para el desarrollo académico y profesional

de los docentes. De utilizarse de forma sumativa, resultan en un mecanismo rápido y de bajo costo, que ofrecen una fotografía con información cuantitativa de un curso o un docente en un momento determinado (Gray & Berman, 2003, en Surgenour, 2013), que habilita la toma de decisiones de personal fundadas. Al mismo tiempo, ha provisto a las instituciones de educación superior de un método para responder a los mandatos de rendición de cuentas y aseguramiento de la calidad (Campbell & Bozeman, 2007).

Sin embargo, como hemos anticipado en el Capítulo 1, la utilidad y la utilización de los CEDA, fuera para uno u otro propósito, ha sido puesta en cuestionamiento. Los argumentos en contra del uso de los CEDA, a menudo plantean cuestiones relacionadas con: (a) la dificultad que plantea encontrar un consenso acerca de lo que constituye una enseñanza efectiva, (b) los problemas que se suscitan a partir de la utilización de estos instrumentos con fines sumativos, y (c) la potencial amenaza que podrían causar aquellos factores que, “aún pudiendo no estar relacionados con la habilidad de enseñar”, impactan en las evaluaciones (Chan, *et al* 2014: 276). Es así que, a pesar de que muchos han desafiado este argumento sosteniendo que el efecto de estas variables es legítimo y no implicaría entonces resultados sesgados (por ejemplo Marsh, 1989), se ha desarrollado “una enorme cantidad de literatura sobre los CEDA que se enfoca principalmente en explorar su calidad psicométrica (por ejemplo, confiabilidad y validez), su relación con otras variables y posibles factores de sesgo (por ejemplo, Cashin, 1995; Feldman, 1993; Wachtel, 1998)” (Nasser & Fresko, 2002: 187).

Con todo, dada la importancia que se le ha dado a estos factores en la literatura, y visto que uno de nuestros objetivos específicos es explorar las concepciones que alumnos, docentes y directores de carrera tienen en relación a la confiabilidad del instrumento, relevamos a continuación los principales argumentos que han puesto en duda la confiabilidad de los CEDA. Para facilitar su lectura, expondremos los razonamientos conforme a la categorización propuesta por Wachtel (1998), a saber: (a) Características asociadas a la administración de las encuestas, (b) Características del curso, (c) Características del profesor, y (d) Características de los estudiantes. Asimismo, dedicaremos al final un pequeño apartado con el fin de atender ciertas cuestiones estadísticas que podrían influenciar “o hacer tambalear [la confiabilidad de] los resultados” (Stark & Freishtat, 2014:3. Traducción propia).

### 2.4.1.1 Cuestiones psicométricas

#### Características asociadas a la administración de las encuestas

**Tiempo/momento:** Feldman (1979) encontró que el tiempo en que las encuestas son administradas, sea en el medio de la cursada, antes, durante o después del examen final; no tiene efecto alguno en los resultados. Sin embargo, contrario a lo propuesto por Feldman, Aleamoni (1981) sugiere que los resultados podrían verse distorsionados si la evaluación es administrada antes o después del examen final, bajo la hipótesis de que la calificación recibida podría impactar en la manera en que los estudiantes perciban el curso o el docente.

**Anonimato:** Feldman (1979) sugiere que existen ciertas circunstancias que podrían interactuar con el anonimato. A modo de ejemplo, propone si los estudiantes completan la encuesta antes o después de conocer sus calificaciones y en tal caso si los estudiantes creen que existe una posibilidad de confrontación con el profesor. Las preocupaciones de otros estudios pasan por sí los estudiantes sienten o no que pueden confiar en las promesas de anonimidad y confidencialidad y, en todos los casos, recomiendan que así se mantenga (Braskamp *et al*, 1984; Centra, 1993).

**Declaración del propósito:** Algunos estudios han concluido que las evaluaciones podrían ser mejores si el propósito que se declara está asociado con cuestiones de promoción y *tenure* (Braskamp *et al*, 1984; Feldman, 1979). Sin embargo, otros han encontrado que este factor no tiene una influencia significativa. En todos los casos, Braskamp *et al*, (1984) sugiere que se comunique a los estudiantes si las evaluaciones van a utilizarse con fines sumativos.

#### Características del curso

**Electividad:** el grado de 'electividad' de una clase ha sido definido en las investigaciones como el porcentaje de estudiantes de una clase que la están tomando como una electiva u optativa (Feldman, 1978). En relación a este factor, los investigadores han encontrado desde pequeñas a significativas relaciones positivas entre la electividad de un curso y los ratings (Brandenburg *et al*, 1997; Feldman, 1978, McKeachie, 1979). En otras palabras, sostienen que los profesores de cursos electivos u no obligatorios tienen a recibir mejores calificaciones que los que dictan cursos no

obligatorios. Esto último podría deberse a un grado de interés inicial menor en los casos de los cursos obligatorios versus los no obligatorios.

**Nivel del curso:** la mayoría de los estudios han concluido que los cursos de niveles más altos, es decir, aquellos dictados ya avanzada la carrera, tienden a recibir calificaciones más altas (Feldman, 1978; Marsh, 1987). Sin embargo, Wachtel (1998) pone en cuestión esta asociación al sugerir que la relación positiva entre el nivel del curso y la calificación recibida podría verse confundida por variables como el tamaño de la clase, la electividad, entre otras.

**Tamaño del curso:** se le ha prestado una atención considerable a la relación entre la cantidad de alumnos en un curso y las calificaciones. La mayoría de los autores han encontrado que las clases más pequeñas tienden a recibir mejores calificaciones (Feldman, 1978; McKeachie, 1990). Marsh (1987) sugiere que el efecto de esta variable se debe principalmente a ciertas dimensiones de la enseñanza efectiva, en particular, a la interacción con el grupo y el *rapport*. Sin embargo, Abrami (1989) en su revisión de Marsh, sostiene que este argumento no puede ser utilizado para sustentar la validez de las encuestas. A pesar de ello, sostiene que deberían ser contempladas en caso de que las encuestas fueran utilizadas con fines sumativos.

**Área/disciplina:** las investigaciones han encontrado que el factor disciplina tiene un efecto en las evaluaciones. Es más, los trabajos han concluido que particularmente las evaluaciones de cursos de matemática y ciencias han obtenido sistemáticamente las calificaciones más bajas (Cashin, 1990; Feldman, 1978). Por consiguiente y, en tanto las diferencias entre disciplinas sean lo suficientemente grandes, hay quienes sugieren que no se deberían hacer comparaciones de las evaluaciones interdisciplinarios o interdepartamentales (Centra, 1993).

**Cantidad de trabajo:** En el estudio hecho por Greenwald y Gilmore (1997) se concluyó que las evaluaciones docentes más altas estaban correlacionadas con cargas de trabajo más bajas (en Sojka *et al*, 2002). Si esta hipótesis fuera cierta, sería de esperar que, como sugiere Ryan (2015), la introducción de estas evaluaciones de manera obligatoria conducirá al profesorado a reducir significativamente la carga del trabajo en clase. Ampliaremos esta cuestión más adelante.

#### Características del profesor

**Personalidad y el "Efecto del Dr. Fox":** la inquietud por estudiar la relación entre la personalidad de un profesor y la evaluación recibida puede atribuirse, en parte, al estudio producido por *Naftulin et al* en 1973. En el estudio original, un actor profesional, conocido como “Dr.Fox”, dio una clase altamente expresiva y entusiasta aunque carente de contenido y como resultado obtuvo altas calificaciones. Esto supuso demostrar que un profesor altamente carismático y entusiasta puede seducir a la audiencia a otorgar calificaciones altas, aunque poco meritorias. A pesar de que este estudio propuso un fuerte argumento para desestimar las encuestas como una medida válida del desempeño docente, la crítica de algunos estudios posteriores –como el de Marsh (1987)– pusieron fin al fuerte apoyo que tenía en la literatura. Con todo, estudios como los de Abrami (1982) encontraron que las evaluaciones de los estudiantes eran mucho más sensibles a la expresividad y al carisma del profesor que al contenido de la clase y es por eso que sería prudente caracterizar a la personalidad como un potencial sesgo.

**Género:** La preocupación frente al hecho de sí el género constituye un factor de sesgo para las evaluaciones docentes ha recibido bastante atención en la literatura. Sin embargo, los resultados no han sido concluyentes (*Sojka et al*, 2002). Por un lado, hay varios estudios, como los de Basow (1994), Martin (1984) y Rutland (1990), que sugieren que las evaluaciones se encuentran sesgadas contra las mujeres (en Wachtel, 1998). A pesar de ello, en el meta-análisis producido por Feldman (1992, 1993), se encontró que la correlación positiva género-evaluación no era significativa. Concretamente, en una revisión de los estudios realizados bajo condiciones experimentales (Feldman, 1992), el autor no encontró diferencias entre las evaluaciones globales de profesores hombres y mujeres. Al mismo tiempo, su repaso de los estudios en situaciones normales<sup>4</sup> reportó nuevamente que la mayoría de los estudios no parecen haber encontrado diferencias significativas entre los géneros (Feldman, 1993).

#### Características de los estudiantes

**Área de interés primaria:** la evidencia sugiere que aquellos estudiantes con un mayor grado de interés en el área del curso tienden a dar evaluaciones más favorables (Feldman, 1977). A pesar de esto, Marsh y Dunkin (1992) sugieren que el interés previo no constituye un sesgo, aunque admiten que, de usarse con fines sumativos, esta

<sup>4</sup> El autor se refiere a los *classroom studies*, como estudios conducidos en situaciones de clase real; en contraposición con las clases dictadas bajo condiciones experimentales.

influencia podría ser una fuente de injusticia, en el sentido de que es una cuestión del curso y no del profesor. Siguiendo lo anterior, sería importante controlar por esta variable en los casos donde la mayoría de los alumnos se encuentran cursando de manera obligatoria.

**Calificación esperada y la hipótesis de la benevolencia:** El efecto de la calificación esperada sobre la evaluación del profesorado ha sido uno de los temas más controversiales en la literatura sobre los CEDA. Actualmente, existe un consenso de que existe una correlación positiva, aunque moderada, entre la calificación esperada y las evaluaciones de los estudiantes. En otras palabras, aquellos estudiantes que esperan calificaciones más altas evaluarán más favorablemente a sus docentes (Braskamp & Ory, 1994; Centra, 1979; Marsh, 1987; entre otros). La controversia, sin embargo, por la interpretación de esta asociación persiste. Marsh (1987; Marsh & Dunkin, 1992) sugiere tres interpretaciones plausibles:

- a. La hipótesis de la benevolencia: los profesores que son más condescendientes a la hora de calificar reciben mejores evaluaciones, i.e. un profesor puede "comprar" mejores evaluaciones poniendo mejores calificaciones. Esto sugiere que la relación calificación-evaluación constituye un sesgo y una amenaza para la validez de los CEDA.
- b. La hipótesis de la validez: los mejores profesores hacen que los estudiantes trabajen más y que, por lo tanto, aprendan más y obtengan mejores calificaciones. En este caso, la relación apoyaría la validez de las encuestas.
- c. La hipótesis de las características del estudiante: características previas del estudiante (como el interés por la disciplina) afectan ambas variables.

En síntesis, “a pesar de que se haya encontrado una correlación entre las calificaciones y las evaluaciones, es imposible determinar si mejores evaluaciones reflejan son producto de más aprendizaje, de características preexistentes de los estudiantes o de un proceso de calificación benevolente (Marsh & Roche 1997, 2000)” (en Sojka *et al*, 2002: 44. Traducción propia)

**Expectativas:** autores como McKeachie (1979) sugieren que el factor más fundamental afectando las evaluaciones docentes son las expectativas de los estudiantes. En decir, aquellos estudiantes que esperan que el profesor va a ser bueno (o malo), usualmente lo corroboran.

Como hemos visto hasta ahora, numerosos estudios se han concentrado en investigar la relación que guardan distintos factores con las evaluaciones docentes, a fin

de demostrar su validez o, por el contrario, de atentar contra ella. Sin embargo, son pocos los casos donde se ha arribado a un consenso o a resultados consistentes. Es así que, la reflexión propuesta por Wachtel (1998), uno de los meta-análisis más referidos en la literatura, resulta pertinente:

Cabe señalar aquí que la mera existencia de una correlación entre una variable de fondo y los puntajes de calificación no constituye necesariamente un sesgo o una amenaza para la validez de las calificaciones de los estudiantes. Por ejemplo, si la calificación esperada del estudiante en el curso se asocia con la calificación que el estudiante le da al instructor, no necesariamente sigue eso y el instructor puede obtener una calificación más alta simplemente dando calificaciones más altas. Las explicaciones alternativas incluyen la posibilidad de que una enseñanza más efectiva inspire a los estudiantes a trabajar más duro y obtener mejores calificaciones. Es necesaria una investigación cuidadosa para determinar la naturaleza de la relación entre la variable de fondo y las calificaciones de los estudiantes (194. Traducción propia)<sup>5</sup>

Como propone el autor, todavía queda un largo camino para determinar si las direcciones y las magnitudes de las relaciones propuestas son lo suficientemente significativas como para determinar la invalidez o descartar la confiabilidad de este tipo de instrumento de evaluación.

### Cuestiones estadísticas

Como habíamos adelantado, además de los posibles sesgos que hemos mencionado, Stark & Freishtat (2014) proponen ciertos aspectos, aunque ahora ligados a cuestiones estadísticas, que podrían influenciar los resultados.

En primer lugar, los autores se preguntan ‘quién responde’, haciendo referencia a la muestra. Especialmente hacen énfasis en la representatividad estadística de la muestra, poniendo bajo consideración que aquello que usualmente motiva a los usuarios a responder es el enojo o la disconformidad. De esta manera, de no ser obligatoria la evaluación de la docencia para todos los estudiantes, los resultados de la encuesta se verían sesgados por un alto nivel de o descontento o gratitud, general. Asimismo, expresan ciertas preocupaciones frente a escenarios de baja tasa de respuesta, sugiriendo que en las clases chicas las evaluaciones van a ser más extremas, esto es, o muy buenas o muy malas.

<sup>5</sup> It must be noted here that the mere existence of a correlation between a background variable and rating scores does not necessarily constitute a bias or a threat to the validity of student ratings. For example, if the student's expected grade in the course is found to be associated with the rating which the student gives to the instructor, it does not necessarily follow that an instructor can obtain higher ratings merely by giving higher grades. Alternative explanations include the possibility that more effective teaching will inspire students to work harder and earn better grades. Careful research is necessary in order to determine the nature of the relationship between the background variable and student ratings (Wachtel, 1998: 194)

En segundo lugar, problematizan la cuestión de los promedios. Desde su perspectiva, en los procesos de revisión de personal la comparación de los docentes o de los departamentos que utilizan los promedios carecen de sentido, pues argumentan que quienes toman las decisiones:

presumen que la diferencia entre 3 y 4 significa lo mismo que la diferencia entre 6 y 7. Presumen que la diferencia entre 3 y 4 significa lo mismo para diferentes estudiantes. Presumen que 5 significa lo mismo para diferentes estudiantes y para estudiantes en diferentes cursos. Presumen que un 3 "equilibra" un 7 para hacer dos 5. Para las evaluaciones de enseñanza, no hay ninguna razón por la cual cualquiera de esas cosas sea cierta (Stark & Freishtat, 2014. Traducción propia)

En otras palabras, los autores expresan su descontento frente a la utilización de promedios para las calificaciones generales de los docentes. Para contrarrestar los problemas que evidencian, sugieren, por un lado, convertir las escalas ordinales en escalas categóricas con descripciones que eviten mejor la ambigüedad. Por el otro, plantean que resultaría mucho más informativo comparar distribuciones y no promedios. De esta forma, un excelente y un malo no significarían a fin de cuentas un regular...no se perdería información.

Por último, bajo el título *Comparando lo incomparable*, los autores hacen referencia a lo problemático que podría resultar comparar cursos de naturaleza tan distinta. Pues, si tenemos en cuenta las características que mencionamos en el apartado anterior, sería injusto comparar cursos de distintos tamaños, para distintas audiencias, con distintos intereses. Aún así, resaltan que la comparación justa todavía es una deuda pendiente de la literatura.

El conjunto de argumentos que hemos expuesto constituye una base de las principales preocupaciones que han caracterizado las posturas que desestiman el uso de las CEDA y, por tanto, se vuelven necesarios como herramientas teóricas para echar luz a las concepciones que los actores involucrados en esta investigación construyen acerca de la confiabilidad del instrumento.

## CAPÍTULO 3

### ENCUADRE METODOLÓGICO

#### 3.1 Estrategia general

Para el estudio de la problemática en cuestión, nos posicionamos desde el paradigma cualitativo en tanto nos proponemos “describir e interpretar experiencias subjetivas” (Sautú, 2005: 18) de estudiantes, docentes y directores de carrera<sup>6</sup> de una universidad privada de la Provincia de Buenos Aires, respecto al propósito, al uso y la confiabilidad de las encuestas como instrumento de evaluación del desempeño docente.

A fin de ser consistentes con el paradigma elegido, se realizaron 25 entrevistas semi-estructuradas entre los meses de octubre y noviembre de 2019. Asimismo, se optó por elaborar una matriz de análisis con dimensiones propias, que permitiera conjugar lo hallado en el campo, con el estado de la cuestión y el marco teórico propuesto. De esta forma, la presente investigación pretende dar continuidad a las investigaciones ya propuestas, al ampliar su universo de estudio a la percepción de estudiantes, docentes y directores de una universidad argentina.

#### 3.2 Justificación del diseño y elección del caso

##### 3.2.1 Estudio de caso único: Universidad de San Andrés

Como dijimos anteriormente, esta investigación es de carácter exploratorio, pues supone construir un antecedente acerca de las creencias y actitudes de los distintos actores involucrados en el proceso de evaluación en el contexto nacional. Desde este punto de partida, la decisión de optar por un estudio de caso único pareció apropiada, pues en palabras de Chetty (1996), un estudio de caso es una metodología rigurosa que “[p]ermite explorar en forma más profunda y obtener un conocimiento más amplio sobre cada fenómeno, lo cual permite la aparición de nuevas señales sobre los temas que emergen” (en Martínez Carazo, 2006: 175). De la misma forma, además de dar lugar a cuestiones

<sup>6</sup> Es preciso aclarar que en el estado de la cuestión hemos hecho referencia a las percepciones de los administradores. Esta ha sido la traducción del término *administrator* o *chairperson*, que se utiliza en la literatura anglosajona para referirse a aquellas personas con autoridad para tomar decisiones *high-stake* sobre los docentes. En el caso en cuestión, esta figura estuvo representada por los directores de carrera, en tanto ocupan un cargo con autoridad sobre los docentes, con potestad de decisión a partir de los resultados de las encuestas.

que pudieran emerger en el desarrollo mismo de la investigación, un diseño de estas características resulta idóneo en tanto el objeto de nuestra investigación se corresponde más al mundo de lo intangible, lo tácito.

En relación al caso abordado, se optó por la Universidad de San Andrés, una universidad privada, ubicada en el partido de Victoria, Buenos Aires. Parte de la decisión estuvo fundada por la accesibilidad y el conocimiento previo del campo. Más precisamente, elegimos este caso por la garantía que supone ser parte de esta casa de estudios, en cuanto al interés y predisposición por parte de los informantes. Por el otro, por la ventaja que brinda la familiaridad con el contexto, en cuanto al contacto con los sujetos a entrevistar y la viabilidad en la coordinación de una amplia cantidad de entrevistas en un tiempo relativamente acotado. Por último, pero no menos importante, se seleccionó la Universidad de San Andrés por el firme conocimiento de la existencia y la utilización de un sistema de encuestas estandarizado para la evaluación del desempeño docente. De hecho, en el informe realizado por CONEAU (2019) de la Universidad de San Andrés, se afirma que “un insumo importante para las evaluaciones de desempeño [docente] esta constituido por las encuestas que se realizan a los estudiantes al finalizar el cursado de cada asignatura” (54).

### 3.2.2 Caracterización de los informantes

A la hora de seleccionar los sujetos a entrevistar, se procuró en todo momento contactar a aquellos sujetos que permitieran conformar un espectro con la mayor diversidad posible. En primer lugar, para el caso de los estudiantes se tuvieron en cuenta los potenciales sesgos mencionados en la literatura, como ser género, año de cursada (como *proxy* de edad) y carrera (como *proxy* de disciplina). En segundo lugar, para el caso de los docentes, se tuvieron en cuenta atributos como el género, tipo de dedicación en la universidad, departamento (como *proxy* de disciplina), el grado académico alcanzado y la tendencia en las encuestas. Para los directores de departamento se tuvieron en cuenta las mismas características que los docentes, con la diferencia de que no se pregunto por la dedicación (sería redundante dadas las incumbencias del puesto) y la tendencia en las encuestas estuvo referida a la media de departamento. En este caso, la variedad en la selección estuvo limitada por el número total de directores existentes. A continuación, presentamos una caracterización de la totalidad de los entrevistados:

**Cuadro 6. Caracterización de los informantes**

<b>Características</b>	<b>Estudiantes (n=10)</b>	<b>Docentes (n=10)</b>	<b>Directores (n=5)</b>
<b>Género</b>			
<i>Masculino</i>	4	7	1
<i>Femenino</i>	6	3	4
<b>Disciplina</b>			
<i>Negocios</i>	2	2	1
<i>Sociales</i>	2	1	2
<i>Educación</i>	2	1	0
<i>Economía</i>	1	1	1
<i>Humanidades</i>	2	3	1
<i>Derecho</i>	1	0	0
<i>Matemática</i>		2	
<b>Año</b>			
<i>Primero</i>		0	
<i>Segundo</i>		3	
<i>Tercero</i>		4	
<i>Cuarto</i>		3	
<b>Grado académico alcanzado</b>			
<i>Doctor</i>		8	4
<i>Magister</i>		2	1
<i>Otro</i>		0	0
<b>Dedicación</b>			
<i>Exclusiva</i>		7	
<i>Parcial</i>		3	
<i>Simple</i>		0	
<b>Tendencia en las encuestas respecto a la media de la universidad</b>			
<i>Por debajo de la media</i>		2	1
<i>Cercana a la media</i>		4	1
<i>Por encima de la media</i>		3	3

Fuente: Elaboración propia.

Resulta preciso hacer algunas aclaraciones en relación a la categorización utilizada para la caracterización de los informantes. En primer lugar, cabe señalar que para delimitar las disciplinas utilizamos la propuesta departamental de la universidad. De esta forma, la intención fue encontrar una categoría que pudiera englobar las

especificidades de las materias dictadas por los profesores y a las carreras dirigidas en la universidad bajo un criterio común: la pertenencia a un departamento (o disciplina en este caso). En segundo lugar, en cuanto al tipo de contrato, procuramos englobar todas las particularidades que existen dentro de la Universidad dentro de una categoría que hiciera alusión a su dedicación. Por último, la tendencia en las encuestas fue informada por los actores en esos términos, por lo que somos incapaces de diferenciar cuán por encima de la media o por debajo de la media efectivamente se encontraba cada uno.

### 3.3 Acerca del desarrollo de la investigación

#### 3.3.1 Los preparativos

Una vez decidido que la investigación se desarrollaría en la Universidad de San Andrés y recibido el consentimiento por las autoridades de la institución, procedimos a contactar a los directores de carreras y docentes a través de correo electrónico. El contacto con los directores se estableció primeramente a través de sus secretarías, pero en el caso de los docentes se los contactó directamente a su mail personal. En el caso de los estudiantes, se los contactó a través del Centro de Estudiantes de la Universidad. En los casos donde la respuesta fue favorable, se pautó una fecha para la entrevista. Todas las entrevistas tuvieron lugar en la universidad durante los meses de octubre y noviembre del 2019 y en todas las instancias se comunicó el objetivo de la investigación con antelación.

#### 3.3.2 Recolección de datos

En concordancia con el paradigma en el cual nos inscribimos, se seleccionó la entrevista semiestructurada como técnica de recolección de datos. En palabras de Corbetta, el objetivo último de este instrumento es “conocer la perspectiva del sujeto estudiado, comprender sus categorías mentales, sus interpretaciones, percepciones y sentimientos” (2003: 344). Siguiendo la propuesta del metodólogo, la entrevista en el paradigma cualitativo se define como una conversación, guiada por el investigador, que se realiza con sujetos seleccionados de acuerdo a un plan de investigación a partir de una serie de preguntas flexibles y no estandarizadas.

De este modo, por lo que se refiere al diseño de la guía de entrevista, se contempló, en primer lugar, una presentación del estudio y la solicitud de algunos datos de presentación del/la entrevistado/a. Para cada grupo particular, los datos de presentación

fueron ajustados según las características mencionadas anteriormente. Por otra parte, los distintos ítems de la guía se orientaron a cubrir las 3 grandes temáticas detectadas en las investigaciones relevadas en los antecedentes y las conceptualizaciones teóricas propuesta en el Capítulo 2. Asimismo, dada las características de las entrevistas semi estructuradas y el carácter exploratorio de la investigación, se procuró hacer un cuestionario breve con preguntas amplias, que dieran lugar a cuestiones emergentes sin que la duración del encuentro se volviera engorrosa. El cuestionario finalmente constó de 16 preguntas para los alumnos, 16 para los docentes y 16 los para directores de carrera (Ver Anexo 1).

### 3.3.3 Registro y almacenamiento

Con el propósito de recuperar las palabras e ideas de los informantes con el mayor grado de precisión posible, la totalidad de las entrevistas fueron registradas de manera digital. Además, se llevó un registro de las reflexiones suscitadas en el encuentro en un cuaderno de campo, también digital. Todos los registros de audio se almacenaron con un código alfanumérico referido al tipo de actor y al orden cronológico en el que se desarrollaron las entrevistas.

Las transcripciones se hicieron a medida que pasaban las entrevistas, a fin de tener presentes las sensaciones que suscitaban los diferentes encuentros. De esta forma, se volvió más sencilla la interpretación de algunas entonaciones o características del lenguaje oral que, aunque difícilmente trasladadas al texto escrito, fueron tenidas en cuenta al momento de analizar la información. Asimismo, todas las entrevistas fueron transcritas en orden cronológico a un archivo único que no sufrió ningún tipo alteración posterior. Paralelamente, se conservaron todos los registros de audio digital como fuente de consulta constante.

Además de contar con un archivo único textual, se produjeron tres archivos de cualidades más analíticas. En primer lugar, se separaron las entrevistas siguiendo los códigos alfanuméricos, es decir, conforme al actor en cuestión y al orden cronológico, lo que dio como resultado tres archivos diferenciados: entrevistas de estudiantes, de docentes y de directores. Una vez hecho esto, se realizó una primera caracterización de los fragmentos según las dimensiones analíticas a la que harían referencia. Este último paso se hizo con el objetivo de agilizar el análisis más profundo que se haría posteriormente.

Con todo, cabe resaltar que este primer proceso de estructuración y organización de los datos según las dimensiones analíticas propuestas inicialmente nos ayudó, tal como señalan Strauss y Corbin (1990, en Schettini y Cortazzo, 2015), por un lado, a comprender mejor el problema de investigación y, por el otro, a comprender la importancia de dar lugar a nuevas dimensiones o sub-dimensiones analíticas que inicialmente no habíamos tenido en cuenta.

### 3.3.4 Análisis y presentación de los datos

Como hemos anticipado, se optó por elaborar dimensiones de análisis propias, que nos permitieran conjugar lo hallado en el campo, con las teorías centrales y los hallazgos de la literatura de la manera más articulada y pertinente posible. Es así, que la selección de las dimensiones que organizaron nuestro análisis partió de las principales temáticas que emergieron de la sistematización de las investigaciones sobre el tema (Ver capítulo 2) y de las preguntas que dieron origen a nuestros objetivos de investigación. Entender cómo los actores perciben la utilización de un instrumento de evaluación, es preguntarse cuáles son sus creencias, actitudes y opiniones acerca de su propósito, acerca de sus potenciales usos, y acerca de la fiabilidad de la medición. En cuanto a las sub-dimensiones, se correspondieron, por un lado, con las precisiones teóricas enunciadas y, por el otro, con las cuestiones que, aunque inicialmente no fueron contempladas, se mostraron recurrentes en el trabajo de campo. De esta forma, las dimensiones y sub-dimensiones presentadas a continuación, son resultado de la investigación en sí misma y no fueron definitivas sino hasta concluido la totalidad de un proceso de análisis.

**Cuadro 7. Dimensiones y sub-dimensiones de análisis**

<b>Dimensiones</b>	<b>Sub-dimensiones</b>
Del propósito...	<i>...general y la inclusión de los alumnos</i> <i>...y su conocimiento</i>
De los usos...	<i>...y la formación</i> <i>...y el control</i>
De la confiabilidad...	<i>...y los estudiantes</i> <i>...y los profesores</i> <i>...y los factores externos</i> <i>...y la administración de la encuesta</i>

Fuente: elaboración propia.

Conviene subrayar que su carácter definitorio hace eco de la necesidad de dar con ciertos límites para la conclusión de esta investigación, pero de ninguna manera agota las dimensiones de análisis imaginables y posibles.

### **Del propósito...**

La pregunta que da origen a esta dimensión de análisis es el por qué de la utilización de la encuesta como mecanismo de evaluación docente. Se encuentra subdividida en dos:

*...General y la inclusión de los alumnos*

Esta sub-dimensión recoge las concepciones sobre el propósito general de la encuesta, así como también la importancia que le atribuyen los distintos actores a esta instancia. Asimismo, corresponde a cuestiones asociadas a la inclusión de los alumnos como agentes evaluadores de la docencia.

*...su conocimiento*

En esta sub-dimensión si indaga hasta qué punto los actores involucrados conocen la finalidad de esta instancia de evaluación desde un punto de vista institucional. Al

mismo tiempo, se profundiza sobre existencia de las instancias de comunicación del propósito.

### **De los usos...**

Luego de preguntarnos el por qué, ahora nos preguntamos el para qué de las encuestas de evaluación docente. Así, esta dimensión agrupa las cuestiones ligadas a la finalidad de la evaluación y a las potenciales acciones derivadas de ella.

#### *...y la formación*

Esta sub-dimensión agrupa las concepciones ligadas a la utilidad de la encuesta con fines formativos, a la realización de cambios (o no) a partir de la información recogida, y a los criterios para estos cambios. Finalmente, hace alusión a las limitaciones de este instrumento cuando se propone esta finalidad.

#### *...y el control*

Esta sub-dimensión agrupa las concepciones ligadas a la utilidad de la encuesta con fines sumativos. Recoge las concepciones respecto a la relevancia de las calificaciones de las encuestas como criterio para la toma de decisiones *high-stake*. Al mismo tiempo, recoge las preocupaciones relacionadas a la utilización de la encuesta con estos fines.

### **De la confiabilidad...**

Por último, cabe preguntarnos sobre las concepciones que guardan los actores en relación a la confiabilidad de este instrumento de evaluación. Es así que en esta dimensión se abordan todas las cuestiones ligadas a los factores que atentan contra la confiabilidad de los resultados producidos por las encuestas.

#### *...y los estudiantes*

Refiere principalmente a las actitudes que los estudiantes podrían tener a la hora de completar la encuesta. Principalmente se retoman cuestiones ligadas a la seriedad con la que se completa la encuesta, y a la relación que podría tener o no la calificación obtenida en el curso con la evaluación docente.

#### *...y los profesores*

Se retoman todas las concepciones de los distintos actores acerca de cuáles son los atributos que valoran los alumnos de un docente.

#### *...y los factores externos*

Refiere a los factores que son externos al docente y podrían impactar en los resultados de las encuestas que buscan medir su desempeño.

*...y la administración del cuestionario*

Por último, esta sub-dimensión agrupa las cuestiones ligadas a la administración del cuestionario, como ser el tiempo, la obligatoriedad, etc.; y que podrían ir en detrimento de la obtención de datos confiables.

Para el proceso de análisis propiamente dicho, se construyó una matriz de análisis organizada en función de: las dimensiones y sub-dimensiones pensamos inicialmente y nuestros tres grupos de interés. A su vez, se dio lugar a una columna destinada específicamente a un código que hace referencia a la entrevista y al fragmento particular en dicha entrevista. Es así que cada fragmento fue codificado y ubicado dentro de la dimensión y sub-dimensión a la que hiciera referencia. Para resumir, se trabajó con un total de 193 fragmentos, que dieron lugar a la base de datos que guió nuestro análisis.

El trabajo analítico culminó con una integración final de lo revelado en el total de la investigación, mediante un análisis profundo, con el propósito de interpretar las relaciones encontradas entre las dimensiones establecidas con base en el desarrollo teórico y los datos obtenidos, con el fin último de comprender –al menos parcialmente, el fenómeno estudiado. Dicho análisis, estuvo a su vez guiado por una expresa intención comparativa derivada de nuestros objetivos de investigación, que se vio facilitada por la elección de un cuadro de doble entrada para la organización y el análisis de los datos.

### 3.4 Notas sobre el trabajo de campo

A los efectos de observar la confiabilidad de análisis y las conclusiones a las que arribemos, consideramos menester hacer algunas aclaraciones acerca de las limitaciones que se nos presentaron a la hora de desarrollar el trabajo de campo.

La primera limitación se corresponde con la selección de los informantes, que más que una selección, fue una disposición. En otras palabras, no es que pudimos seleccionar del todo a nuestros informantes, sino que nos adecuamos a la disposición de aquellos que finamente quisieron ser entrevistados. De esta forma, creemos que un potencial riesgo que corremos es el sesgo que podría introducir la voluntariedad: los participantes que se voluntarearon, seguramente hayan tenido alguna inquietud particular por este tema o una postura bien formada. Creemos que este problema se vuelve principalmente una limitación en el caso de los profesores, pues en el caso de los directores este sesgo puede

haber sido salvado por la representatividad de los informantes, pues se entrevistaron a más de la mitad de los directores de carrera. En el caso de los estudiantes podría ser una fuerte limitación, en tanto no pudimos ser capaces de recolectar las percepciones de alumnos que no tienen interés alguno en la temática, pero aún así forman parte de la gran masa crítica que completa las encuestas.

Además, otra cuestión a tener en cuenta es la distinción que existe entre docentes y directores de departamento. El problema podría radicar en que muchos directores de departamento también ejercen la docencia, por lo que sus respuestas podrían estar referidas a sus vivencias y creencias como docentes o como directores, indistintamente.

Otro punto a considerar ha sido la falta de relevamiento y contextualización del instrumento desde el discurso institucional. En decir, no nos detuvimos a caracterizar el instrumento y su utilización desde una perspectiva institucional, con todo lo que ello conlleva: normativa, perspectiva de las autoridades, entre otras. Si bien esto no forma parte del objetivo de investigación, creemos que podría servir para contextualizar los hallazgos del trabajo, y otorgar al lector un panorama más completo de la cuestión. Con todo, algunas inferencias pueden hacerse a partir de los testimonios recuperados.

Por último, cabe destacar las dificultades que trae aparejada la transcripción de las entrevistas, en tanto se vuelve difícil captar las marcas propias de la oralidad, que podrían cambiar totalmente el sentido de una oración. En este sentido, corremos el riesgo de que el lector malinterprete el texto, como consecuencia de una transcripción que no ha logrado captar del todo las pausas y las entonaciones a través de los recursos gráficos, propios del lenguaje escrito.

## CAPÍTULO 4

### ANÁLISIS

En este capítulo presentamos el análisis, realizado a partir de la realización de 25 entrevistas, con la intención de dar cuenta de las percepciones que construyen alumnos, docentes y directores de carrera de la Universidad de San Andrés, alrededor de la encuesta como instrumento de evaluación docente.

Tal como lo indican nuestros objetivos específicos, si bien nuestro propósito es indagar en las concepciones de cada grupo en particular; también es nuestro objetivo hacer un ejercicio comparativo. La organización del análisis responde entonces, por un lado, a las categorías propuestas en el capítulo cuatro y, por el otro, a nuestra intención de encontrar puntos de encuentro y desencuentro entre las concepciones de los grupos involucrados. Es por ello que presentamos el análisis por temática y no por actor.

De forma preliminar, desplegamos a continuación, un cuadro de frecuencia con la cantidad de fragmentos analizados por dimensión y sub-dimensión:

**Cuadro 8. Cantidad de fragmentos analizados por dimensiones y sub-dimensiones de análisis**

Dimensiones	Sub-dimensiones	Estudiantes	Docentes	Directores
<b>Del propósito...</b>	<i>...general y la inclusión de los alumnos</i>	4	8	4
	<i>...y su conocimiento</i>	5	6	3
<b>De los usos...</b>	<i>...y la formación</i>	9	23	5
	<i>...y el control</i>	5	12	6
	<i>...y los estudiantes</i>	18	20	9
<b>De la confiabilidad...</b>	<i>...y los profesores</i>	7	10	5
	<i>...y las características del curso</i>	5	14	6
	<i>...y la administración de la encuesta</i>	5	2	2

Fuente: elaboración propia

Cabe mencionar, que la disparidad entre las frecuencias responde a la centralidad que le dio cada uno de los actores, a alguna sub-dimensión en particular. Asimismo, la selección de fragmentos responde a dos criterios: la riqueza y la representatividad del fragmento. En el Anexo I incluimos la selección de la totalidad de los fragmentos tenidos en cuenta para la elaboración de este análisis.

#### 4.1 Del propósito...

##### 4.1.1 ...general y la inclusión de los alumnos

El primer tema que emerge en nuestro análisis refiere a las concepciones que rodean el propósito de las encuestas y, particularmente, a la participación que se les otorga a los alumnos en este proceso. En otras palabras, nos preguntamos sobre la creencia general alrededor del para qué de las encuestas y el por qué de la selección de los estudiantes como principales agentes evaluadores.

En relación a los directores de carrera, los fragmentos que presentamos a continuación sirven para dar cuenta de la heterogeneidad de creencias respecto al propósito del sistema de encuestas en la Universidad. Mientras que para algunos resulta muy claro, para otros no lo es tanto:

“Para mí hay un propósito muy claro que es la evaluación de los cursos y de los docentes. Me imagino que los chicos también lo toman así. Evaluación ex post de lo que ha sido la experiencia, la performance del profesor, de la materia; creo que es un elemento bastante bueno” (Director 3, fragmento 2)

“Muchas veces está poco claro [el propósito], por eso resulta un instrumento tan ambiguo y cuestionado muchas veces. Me parece que es poco claro. Me parece que parte del propósito es saber cuál es la opinión de los estudiantes y detectar problemas que puede haber...lo que pasa es que una cosa es detectar problemas y otra cosa es evaluar a otra persona” (Director 1, fragmento 1)

En el discurso de los directores podemos detectar como se le atribuye a la encuesta dos propósitos marcadamente disímiles. Por un lado, tenemos la noción de la encuesta como un instrumento de evaluación de los docentes y de los cursos, marcado por una concepción más bien sumativa de la evaluación, que busca medir desempeños. Por el otro, si bien se hace referencia a alguna dimensión del control en tanto sirve para “detectar problemas”; la finalidad pareciera estar orientada a recoger información útil para llevar a cabo correcciones a partir de ello.

En cuanto a las creencias de los docentes, también se observan visiones de las encuestas orientadas más bien por una concepción formativa y otras orientadas por una

concepción sumativa o de control. Por ejemplo, para un docente, “la pertinencia de este tipo de evaluaciones es si logran mejorar las prácticas educativas o generar más aprendizajes” (Docente 6, fragmento 4), mientras que para otro “es un insumo importante, sobre todo para detectar una cuestión relativamente más bien gruesa” (Docente 2, fragmento 3).

En este punto, si bien entendemos que la caracterización podría ser algo forzosa, nuestra intención es demostrar que distintas finalidades de evaluación se encuentran en el caso de los CEDA. Tal como expresa el Director 1, observamos que el propósito de la encuesta resulta ambiguo para la mayoría de los actores, pues esta marcado por esta doble finalidad: la evaluación para el control y la evaluación para la formación. En relación la superposición antedicha, la reflexión de un docente resulta atinada: “El riesgo de esto es que se burocraticen y ahí su sentido es muy limitado. Pero si se logra focalizar el instrumento, bueno ‘buscamos que tengas mayores aprendizajes’.” (Docente 6, fragmento 4). En efecto, las encuestas tienen un potencial formativo en tanto su foco este puesto en la mejora de los aprendizajes. Sin embargo, esta potencialidad puede verse eclipsada por su inminente burocratización. Sobre ella un docente señala: “Pero como mínimo, a mí me gustaría erradicar la frase: como tuvo malas encuestas es mal docente...que acá la dicen con una liviandad.” (Docente 5, fragmento 4). Esta preocupación, nos remite a las reflexiones de Calderón *et al* (2018), quienes se pregunta por la capacidad de las evaluaciones sumativas de inducir mejorías en el desempeño.

Por lo que refiere a la heteroevaluación, es decir, a que los alumnos sean los encargados de evaluar el desempeño docente, identificamos que los distintos grupos por lo general mantienen posturas positivas, aunque algunos con matices más críticas que otros. Por un lado, están quienes creen que la consideración de los estudiantes para evaluar la docencia es esencial:

“Me parece clave que los estudiantes tengan voz en la evaluación de las prácticas de enseñanza. Soy firme defensor, creo que la voz de los estudiantes es súper aguda, es clave, me parece que es la voz más importante en cuanto la retroalimentación se que se pueda esperar.” (Docente 6, fragmento 1)

“Y me parece bien. Me parece que la relación de la docencia es una relación que se construye y que el input de los alumnos es importante.” (Director 2, fragmento 2)

Por otro lado, hay quienes opinan que, si bien la devolución de los estudiantes es importante, tienen una postura más crítica respecto a su interpretación como algo fundamental. Esto es, si bien les parece bien, importante y útil, la voz de los estudiantes es algo más a tener en cuenta. Asimismo, quienes son más críticos resaltan que los CEDA

son instrumentos que no resultan del todo precisos y que deben de utilizarse con precaución:

“Hay muchas maneras de evaluar estas cosas. Pero me parece que tener una devolución de los estudiantes de lo que les pareció en general la materia, me parece bien. Que es otro dato más. Estoy de acuerdo con ellas.” (Docente 10, fragmento 13)

“Que se involucre a los estudiantes en el proceso de evaluación docente me parece bien. Esa palabra dice que es un factor a tomar en cuenta en la evaluación de desempeño docente y estoy de acuerdo con que un factor es la opinión de los estudiantes” (Docente 5, fragmento 1)

“Sí son útiles. Creo que si hay que preguntarles a los alumnos. Pero también creo que son una cosa imperfecta que captura de manera imperfecta lo que el docente hace. Después está la discusión de cómo preguntas y qué preguntas. Pero para el instrumento en general, estoy de acuerdo en que estén, y de que se usen, pero que se usen con criterio”. (Docente 7, fragmento 6)

Por último, están quienes relacionan los CEDA con una disposición universal que, extendida a otros ámbitos además del educativo, se caracteriza por ser una tendencia donde “se involucra a los stakeholders...que tiene en cuenta la opinión de todas las partes interesadas” y que busca asegurarse que tengan una voz activa en todas las partes del proceso (Director 1, fragmento 4).

Por su parte, la visión de los estudiantes ha sido la más homogénea y categórica. La totalidad de los entrevistados considera que es una instancia fundamental y que su visión es elemental para entender como es el desempeño de un docente, en tanto son ellos los que “reciben el servicio” (Estudiante 6, fragmento 1) y son “los alumnos los protagonistas del proceso de aprendizaje (Estudiante 10, fragmento 1):

“La evaluación docente por parte de los alumnos para mi lo es todo, realmente lo es todo. [...] y es la opinión que más vale, sin lugar a dudas, así que el propósito de las encuestas, es decir preguntarles a los alumnos sobre los profesores es absolutamente esencial” (Estudiante 5, fragmento 1)

“Pienso que es bueno que se empodere a los alumnos como agentes evaluadores, es valioso que se invite a que ‘alcen su voz’, a que calibren y den su opinión. De hecho, me parece un espacio indispensable que merece ser conservado.” (Estudiante 7, fragmento 2)

#### 4.1.2 ...y su conocimiento

Para nuestra sorpresa, cuando se les preguntó a los docentes si tenían conocimiento del uso que la institución hacía de la encuesta, la gran mayoría de los docentes manifestaron no saber:

“No, no sé. Se que las ve, y que por ahí ...si sistemáticamente hay un desfase con respecto a la media hará preguntas...Pero no, no tengo idea como se evalúa, ni quién las lee, o que se hace, no tengo ni idea” (Docente 1, fragmento 12)

“No. No sé...no tengo la menor idea si otros las ve, me imagino que los directores de departamento las ven. Yo mucho drama no me hago...Hace 10 años que estoy en la Universidad, entonces debe ser eso. Pero no sé si toman decisiones en función de eso...” (Docente 3, fragmento 11)

“No, yo se que la Universidad cada tanto hace un relevamiento. No sé que miradas usan o toman; pero se que las van mirando...después dentro de cada departamento yo se que cada director de carrera las ve, también para ir pensando cómo está la carrera y que ajustes hacer. Yo sé que eso se hace” (Docente 4, fragmento 7)

“En mi caso, jamás me senté con el director de la carrera, el director del departamento, el director del ciclo de fundamento. Jamás en 20 años me senté a hablar con nadie de las encuestas. Puede ser que es porque la cosa funciona más o menos bien. Si conozco casos de docentes a los que no les fue bien, que estuvieron muy por debajo del promedio y que tuvieron conversaciones con el director del departamento” (Docente 2, fragmento 10)

Como podemos ver, en muchos casos el desconocimiento del uso de las encuestas se ve asociado a la ausencia de situaciones críticas o alarmantes, es decir, a que su desempeño no ha sido lo suficientemente malo como para que la Universidad haya tomado medidas al respecto. En los pocos casos donde los docentes manifiestan tener alguna noción de su uso, generalmente se hacen referencias a cuestiones anecdóticas.

Más sorprendente aún fue constatar que muchos directores de carrera tampoco conocen el manejo institucional de las encuestas:

“Bueno, esa parte a mi no me toca [...] pero sí, obviamente se ven las encuestas, eso se que se hace. Es parte de los elementos que se utilizan en la evaluación” (Director 5, fragmento 8)

“Yo tampoco sé muy bien. Se porque soy directora que impacta en el premio a la docencia...cuánto impacta en otros ámbitos o lugares de decisión no lo sé. Para que se usan concretamente no lo sé.” (Director 4, fragmento 15)

“Como yo no tomo decisiones de promoción, entonces no se si usan para eso” (Director 2, fragmento 12)

En este caso, la causa del desconocimiento fue atribuida al hecho de que no son partícipes de instancias de toma de decisión *high-stake*.

En cuanto a los estudiantes, en base al análisis realizado se evidencia que, en su mayoría, desconocen el uso efectivo de las encuestas:

“No sé. Creo que uno, se usan para detectar si hay algún problema puntual con un profesor. Si en el trato con el alumno, claramente todos ponen 1, entonces es porque claramente el profesor no está haciendo algo bien...eso del lado institucional. [...] Y después, para el profesor, si quiere variar sus clases, o ve que tiene un 60% desaprobados en el final quizás puede en las encuestas ver lo que pasó” (Estudiante 4, fragmento 9)

“No, no tengo ni idea. Lo único que sé, que me han dicho los profesores o cualquier persona que hablo de eso, nomas que es después un balance del curso con los profesores. Pero más que eso nada” (Estudiante, 3, fragmento 9)

“Tengo la sensación de que, si te fue muy mal, recién ahí un comité evaluador va a ver los comentarios y si también son muy malos” (Estudiante 6, fragmento 8)

“Yo tengo entendido que es bastante, subjetivo...hasta donde yo tenia entendido sube hasta los directores de departamento que si no me equivoco son los que contrataron a los directores...” (Estudiante 1, fragmento 5)

El carácter hipotético de sus afirmaciones, combinado con la multiplicidad de orientaciones distintas que tomaron las respuestas, dan cuenta de la ausencia de una instancia de información o declaración del propósito de esta instancia por parte de la institución. De hecho, en base a las entrevistas, pudimos identificar que varios docentes se encuentran preocupados ante esta situación: “Yo no sé cuánto valora el estudiante la encuesta. Tampoco sé cuánto percibe el estudiante la importancia que se le dan a las encuestas” (Docente 2, fragmento 17). Asimismo, un docente advierte una suerte de falta de cultura de la evaluación y hace referencia a la falta de contextualización o información que se le provee a los alumnos sobre esta instancia:

“A veces creo que los estudiantes que recién empiezan son estudiantes con una mentalidad muy argentina que quizás no entienden bien la importancia de la encuesta, como te hacía la comparación con el alumno norteamericano que está acostumbrado y lo usa de una manera más responsable. Entonces en algunos casos percibo que el estudiante a veces no se lo toma demasiado en serio; que lo toma como una formalidad que tiene que completar y chau. A veces puede obrar por venganza a tal docente, entonces me parece que también debería haber una concientización a los estudiantes respecto a que es una encuesta, que funciones cumple, como debe hacerse” (Docente 4, fragmento 12)

Ciertamente, este desconocimiento del propósito de la encuesta, queda evidenciado en el testimonio de un estudiante: “creo que nunca nos hicieron como una buena introducción” (Estudiante 6, fragmento 10). Esta referencia a la falta de contextualización o información de los alumnos sobre la finalidad del cuestionario, podría ser un indicio de la falta de canales de comunicación efectivos o sobre la falta de intención de la institución de garantizar que los actores involucrados sepan de que va esta instancia.

## 4.2 De los usos...

### 4.2.1 ...la formación

#### Utilidad formativa

A partir de las entrevistas realizadas, podemos decir que muchos docentes valoran la utilidad formativa de las encuestas. En otras palabras, además de que le atribuyen importancia a la instancia, creen que es un instrumento de gran utilidad, en tanto les permite “ver que estamos haciendo bien o que estamos haciendo mal...que podemos cambiar, que podemos corregir, que podemos agregar” (Docente 2, fragmento 7). En línea con lo antedicho, otros docentes plantean:

“Pero muchas buenas ideas vienen de ahí. Siempre las he leído con muchísima atención. Me parecen muy importantes y muy útiles.” (Docente 8, fragmento 8)

“Pero si me parecen súper útiles. Me parece bueno tenerlo, porque uno saca cosas interesantes de lo que a los estudiantes les gusta, que es lo que percibieron más logrado, que es lo que habría que mejorar.” (Docente 10, fragmento 8)

“A mi me sirvieron y pude introducir cambios a partir de las encuestas.” (Docente 2, fragmento 9)

“Entonces mi idea general es que puede ser algo útil para cada docente, para ver qué respuesta hay del otro lado respecto a lo que está haciendo y en todo caso después el docente calibrará si tiene que hacer ajustes o no... si darle crédito a lo que se dice o como tomarlo. Pero como principio general me parece que son útiles.” (Docente 4, fragmento 2)

Lo anterior parece coincidir con una de las finalidades de los CEDA, que proponen autores como Marsh (1989), la de ser un instrumento de evaluación que provee retroalimentación formativa a los docentes, con la intención de mejora. Además, se puede evidenciar un cierto compromiso por parte de los docentes con la mejora de sus cursos a través de la información provista por estos instrumentos, lo que evidencia una concepción de evaluación formativa, en tanto se utilizan para aportar las mejoras o los correctivos pertinentes (Scallon, 1988). Lo anterior coincide con los hallazgos de los estudios de Nasser & Fresko (2002), y Balam & Shannon (2010), donde los docentes no solo mantienen una actitud positiva, sino que argumentan que resultan provechosos para mejorar su enseñanza.

Por otro lado, un director añade que la utilidad formativa —es decir, la oportunidad que ofrece la encuesta como instancia de evaluación concentrada en la mejora del desempeño y no en los resultados— pasa, en parte, por la representatividad que logra. En otras palabras, porque son todos los alumnos los que están opinando:

“Uno ve un feedback general mientras esta dando clase, pero eso no necesariamente es representativo [...] entonces está bueno tener las encuestas para calibrar un poco como vieron eso” (Director 2, fragmento 3)

Sin embargo, a pesar de que a algunos les parezca una propuesta potente y que promueva a la reflexión formativa para los profesores (Estudiante 9, fragmento 4), hay quienes expresan sus reservas acerca de la cuestión:

“Por último, temo que las encuestas no son leídas crítica y constructivamente por los propios docentes. Intuyo que, a esa altura del curso, son pocos los docentes que se detienen a reflexionar en qué aspectos creen que pueden crecer en pos de plantear la próxima materia a dictar.” (Estudiante 7, fragmento 6)

El fragmento anterior muestra como existe cierta desconfianza por parte de los estudiantes acerca de la utilidad formativa, en tanto creen que son pocos los docentes que toman el tiempo para ver de que forma mejorar. La reflexión propuesta por Assasfeh *et al* (2014) resulta pertinente en este punto, pues los autores sugieren que de no percibir los alumnos cambios en las clases basados en los comentarios, podría motivar ellos la

impresión de que las evaluaciones no son tomadas de manera seria y comprometida por parte de los docentes. Asimismo, sostenemos que esto podría estar relacionado con la naturaleza propia de las concepciones, que son fruto de las interacciones del sujeto con su entorno (Hidalgo y Murillo, 2017). En tanto la única evidencia de la utilidad formativa de las encuestas sean los cambios concretos que un docente pueda producir y mientras que los alumnos no sean capaces de percibir o vivencias esos cambios, creemos será natural que los alumnos expresen cierta desconfianza.

### Cambios

Como hemos anticipado, una cuestión que es fundamental a la hora de pensar la evaluación desde una concepción formativa, es la idea del cambio. En otras palabras, para que la evaluación sea formativa debe existir la posibilidad de cambiar el rumbo a partir de la devolución obtenida. A pesar de lo que expresaba el estudiante en el párrafo anterior, varios docentes han manifestado la oportunidad de realizar cambios a partir de la devolución de sus alumnos:

“Sí, muchísimos. (...) Siempre estamos haciendo cambios a la materia. Hemos, por ejemplo, invertido los contenidos que se daban en clases magistrales y en tutoriales. Y la última modificación, que ahora quiero ver las encuestas para ver cómo dan, porque a veces decían algunos que no veían el vínculo entre magistrales y tutoriales, y eso que es algo que nos matamos pensando. (...). Así que sí, me han servido, varias cosas...por ejemplo esto de normativa pasarlo al campus. Así que bueno eso...introducimos el examen parcial.” (Docente 10, fragmento 9)

“Sí, yo tengo un estilo de docente bastante abierto en cuanto al planteo del trayecto que van a hacer los chicos, y bueno en el feedback de las evaluaciones de la maestría me decían que se sentían desorientados frente a esta falta de claridad. Entonces a partir de eso empecé a ser más claro en cuanto a los productos esperados, o mis expectativas, y no fui definiendo durante el curso que íbamos a hacer al final del curso. Entonces a partir de eso me di cuenta que tenía que brindarles más certidumbres. No en cuanto a los procesos cognitivos, sino en cuanto a los productos esperados.” (Docente 6, fragmento 8)

“Si, por ahí algunos temas empezar a dedicarle más tiempo y a otros temas menos tiempo...fui como reestructurando la clase y...ese es quizás el cambio más importante que hice [...] sistema de evaluación no...porque tampoco tuve un comentario. Pero si en eso de ‘este tema era fácil...este tema era difícil’ ahí sí fui cambiando cuánto tiempo dedicarle a qué cosa” (Docente 1, fragmento 9)

“Si, miro la parte donde los alumnos se toman un tiempito mas de aclarar alguna cosa, tanto los que le parece bien, como los que le parece mal y corregible, y ahí entonces me fijo si quizás tal libro debería reemplazarlo por otro, o incluir ciertas cosas o no. Me permite ir pensando ajustes...nunca de manera radical, pero sí pequeñas cosas para ver” (Docente 4, 3)

A partir de los fragmentos presentados, podemos evidenciar que los cambios introducidos por los docentes son de los más diversos. Van desde la dinámica de la clase, a la cantidad de tiempo dedicado para cada tema, hasta el tipo de bibliografía utilizada. En este punto,

un director percibe que los cambios pasan más por la estrategia, “cómo abor das al alumno, qué metodologías estas usando, como buscas herramientas para que los temas se comprendan mejor, como los ayudas a que hagan un mejor proceso de aprendizaje” (Director 5, fragmento 7), que por cosas más centrales como la estrategia de evaluación.

Resulta interesante aquí, volver sobre las percepciones de los alumnos. Una cuestión relevante en relación a lo cambios, es que ellos no perciben que todos los docentes les presten atención a las encuestas y, por consiguiente, lo evidencien realizando cambios puntuales, de hecho:

“Yo creo que si le prestan atención no es lineal, hay docentes que si le prestan y hay docentes que no. En promedio, quizás si le prestan atención, no lo sé. *Pero creo que las cuestiones estructurales que son criticables de un docente no suelen ser cambiadas*, para mi, a partir de una encuesta.” (Estudiante 10, fragmento 4)

“Ponele que el 75% de los docentes le prestan atención, bueno igualmente *son muy pocos los que cambian las cosas más relevantes*, más importantes. De hecho, a muchos docentes con el correr de las cámaras se les critican las mismas cosas y dudo que sea la encuesta por si sola, sea lo que haga que el docente impulse un cambio por sí mismo.” (Estudiante 9, fragmento 5)

La percepción de los alumnos parece adecuarse a lo expresado por el director o mismo por los propios docentes. En su experiencia, no han podido evidenciar cambios “estructurales”, de hecho, son muy pocos los que cambian las cosas “importantes”. Esto parece concordar con los hallazgos de estudios anteriores, que han revelado que los alumnos han percibido que las encuestas no contribuyen de manera significativa a las prácticas docentes (Marlin, 1986; Abedin, *et al*, 2013; Assasfeh *et al*, 2014). Una última cuestión a señalar, es que algunos suponen que la cuestión de darle un uso formativo a las encuestas “depende un poco del profesor y la capacidad de hacer autocrítica” (Estudiante 3, fragmento 6), lo que nos da a entender que la mejora pareciera ser algo más coyuntural, que una consecuencia directa de esta instancia de evaluación.

Probablemente, una cuestión que explicaría lo anterior –esto es, la falta de cambios que sean realmente estructurales o centrales– sean los criterios que los docentes explicitan como válidos para introducir variaciones:

“Generalmente en las encuestas...yo por lo menos no es que le doy bolilla a una encuesta, sino voy viendo como fue el historial de las encuestas y en función de eso voy tratando de cambiarlo, adaptarlo a lo que se pueda hacer.” (Docente 1, fragmento 7)

“Si cuando vemos, 2,3,4, comentarios de que un texto es un plumazo, lo quitamos; o nos gusto muchísimo mas clase 'deberíamos tener más tiempo este tema” (Docente 3, fragmento 9)

“Ahí tienes tres cosas que tomé, porque se repiten. Cuando uno ve que se repiten, ahí es como que empieza a darle valor. Sino es una idea que alguien tira, que la puedo tomar o no. Pero cuando uno ve que hay varias cosas que se van repitiendo, yo por lo menos prendo la luz amarilla y digo, bueno acá estaba pasando algo.” (Docente 2, fragmento 8)

A partir de los fragmentos presentados, podemos decir que uno de los criterios que aparecen más fuertes para introducir cambios es el tema de la recurrencia, es decir, que lo digan “no en una opinión, no en dos, pero en varias” (Docente 7, fragmento 5). Esto podría verse asociado a la confusión que produce, en muchos casos, la extrema variación que existe entre las evaluaciones de los estudiantes. Ante la polarización de estos comentarios y la existencia de posturas totalmente contrapuestas en relación a algo, la recurrencia parecería ser la puerta de escape. De hecho, uno de los alumnos expresa que la recurrencia ayuda a evitar opiniones o cuestiones personales:

“Yo siento que esas encuestas si son tenidas en cuenta, o eso quiero creer, y que los profesores las leen, también quiero creer. Entonces también creo que les sirve no solo a [...] sino a cada profesor para ver... bueno, si es la opinión de uno será una cuestión personal; pero si ya son varias personas que opinan lo mismo sobre esta cuestión negativa a cambiar, nada. Siento que por ahí lo ayuda” (Estudiante 2, fragmento 2)

Además, otra cuestión que podemos evidenciar, es que los docentes no se encuentran dispuestos a ceder y cambiar algo solo por el hecho de que muchos alumnos lo hayan expresado. En muchos casos, lo que prima es una suerte de convicción de qué eso es realmente relevante para el aprendizaje del alumno:

“Para mi es importante. No me parece...un tema que me aparece siempre en las encuestas es mucho trabajo. Lo lamento. [...] Digo...lo tomo, pero también es cierto que es el costo que yo considero que se paga por el aprendizaje” (Docente 2, fragmento 4)

“Pero si a veces hago algunos cambios a partir de las encuestas. No soy de los que dicen tengo que cambiar todo lo que los estudiantes objetan, porque creo que el docente debe darse cuenta de que los estudiantes opinan desde cierto lugar que...de no conocimiento de lo que es dar una materia, o de lo que el docente le se propone en genera al dar tal cosa o no dar otra; entonces sabiendo eso, *trato de percibir que es lo que es un buen señalamiento*, que me podría servir para mejorar la materia y hacerlo” (Docente 4, fragmento 4)

En general, como han concluido estudios anteriores y como pudimos evidenciar con lo anterior, la decisión de hacer cambios como respuesta al feedback de los estudiantes es compleja. Lejos de responder a un factor, depende de varios, siendo la recurrencia y el sentido común o las convicciones del profesor las que priman en la mayoría de los casos. Con todo, nos preguntamos en este punto qué sucede cuando los factores anteriores se encuentran. Una hipótesis, en base al último fragmento (Docente 4, fragmento 4) es que finalmente termina prevaleciendo la opinión del profesor.

#### El caso de los comentarios

Una cuestión que ha sido recurrente a la hora de fundamentar la utilidad formativa de las encuestas, ha sido el caso de los comentarios. De la misma forma que en el estudio conducido por Nasser & Fresko (2002), los docentes se han mostrado fundamentalmente

positivos frente a la utilidad de los comentarios escritos, facilitados por las preguntas abiertas:

“Yo creo que las dos son importantes. Y en general veo como cierta alineación entre los datos cuantitativos y los comentarios. Me parece que son dos registros de un mismo fenómeno y es raro que pase alguna cosa muy disonante entre una cosa y otra. Me parece que los dos...lo que es el registro cualitativo para este tipo de instrumentos me parece más rico que el registro número, que te aporta una foto. En cambio, el otro, los alumnos en general te especifican que cosas les gustaron o que metodología son que cosas no repetirían. Entonces en ese sentido, a nivel pedagógico es más rico el cualitativo” (Docente 6, fragmento 7)

“Si, porque los números...ósea, generalmente siempre va siendo los mismo; y hay un par de números que es como un comentario. Pero sobre todos los comentarios es cuando leo más. De vuelta, cuando te preguntan las cosas si re programaron clase, eso yo lo sé, no le doy tanta bolilla” (Docente 1, fragmento 8)

“En los comentarios hay pistas. Los comentarios algunos son muy buenos (...) hay comentarios que aprecian. Hay de todo y son muchos sí, son varios” (Docente 10, fragmento 10)

“A mi me resulta siempre lo más interesante de la encuesta y lo más útil, son los comentarios y no las evaluaciones numéricas. Para mi son bastante poco explicativas o me parece que *tienen demasiadas posibilidades de sesgo por un montón de cosas*, la evaluación numérica a diferencia de la cualitativa” (Docente 5, fragmento 6)

Como hemos adelantado, típicamente las encuestas de evaluación docente se componen por escalas de valoración tipo *Likert*. De hecho, la encuesta utilizada en la institución en cuestión presenta una sección de este estilo que provee a los docentes datos numéricos, pero también una sección de pregunta abierta con lugares para comentarios. En todos los casos, los docentes han manifestado la utilidad particular de este tipo de información cualitativa, frente a la provista por los números. Una de las cuestiones centrales que recalca el último fragmento (Docente 5, fragmento 6) es la ventaja de la especificidad de los comentarios, y a la inversa, la desventaja de los sesgos especialmente en las escalas de valoración.

En línea con lo que expresan los docentes, un director añade “me consta a mi personalmente y a muchos colegas les ha ayudado mucho la cuestión de las sugerencias” (Director 3, fragmento 8). Asimismo, otro agrega la importancia de esta instancia se debe “[m]ucho menos por los números que arrojan, por cuanto lo que dicen los estudiantes abajo, los comentarios (...) Porque en esa opinión, hay una palabra no hay un número” (Director 4, fragmento 1). En concordancia, un alumno cree que los docentes “deben mirar muy en general y a algunos números no le deben dar tanta bolilla”, pero que por el contrario cree “que los comentarios [...] sí son tenidos en cuenta” (Estudiante 6, fragmento 5). Otro estudiante reflexiona: “Al decirlo con tus palabras, es algo que escribió una persona y no tuvo que encasillar en un numero; y te dice particularmente quizás un aspecto que no está tenido en cuenta en esta valoración” (Estudiante 2,

fragmento 4). Esto último, comienza a delinear las limitaciones de este instrumento, en las que nos detendremos a continuación.

### Limitaciones

A pesar de que a lo largo de las entrevistas pudimos detectar cierto consenso acerca de la utilidad formativa de este instrumento de evaluación, también pudimos detectar ciertos aspectos que podrían atentar contra ella. Como todo instrumento, la encuesta utilizada en la Universidad de San Andrés, presenta ciertas limitaciones en relación a su dimensión formativa. En otras palabras, hay características propias del instrumento y su administración que limitan las oportunidades de los docentes en relación a la mejoría de su desempeño.

Una de las más mencionada, esta relacionada con el momento de la evaluación. Como hemos expuesto en el capítulo 2, una evaluación llevada a cabo al final del proceso, tiene como objetivo valorar lo realizado (Carlino, 2003) y, al mismo tiempo, no admite la posibilidad de cambio o réplica:

“El problema que tiene eso, es que vos recibís la opinión de los que ya cursaron, que no sabes si va a similar a la opinión de los que siguen, pero bueno...maso menos permite a ajustarlo” (Docente 3, fragmento 3)

“Yo les pido que llenen las encuestas y les digo "ustedes ya no van a ver cambios con respecto a estas encuestas, pero háganlo por los estudiantes que vienen después” (Director 2, fragmento 5)

“Aunque sí creo que no debería ser la encuesta solamente al final de la cursada. Porque si hay errores muy *heavys*, como por ejemplo que te están dictando una materia que ya sabes que no tuvo nada que ver con el parcial, la [...] que este a último momento es que es contraproducente, porque no se puede hacer nada, no se puede hacer un cambio o lo va a ver solamente la siguiente camada” (Estudiante 6, fragmento 4).

Los fragmentos exponen una suerte de tensión entre la dimensión formativa de la encuesta y su momento de aplicación. Lo lógico, si atendemos a los modelos de evaluación, sería que, de tener una finalidad formativa, la evaluación se realizase durante la propuesta, así los docentes tendrían la oportunidad de corregir los procedimientos y mejorarlos. Esta tensión, acaba redundando en una situación injusta para los docentes, cuya voluntad de realizar cambios se ve constreñida por el momento de evaluación:

“No puedo esperar al final del semestre para que me digan tu clase es muy aburrida...me hubieras dicho antes, así me dabas tiempo para recuperarme” (Docente 2, fragmento 6)

Por otro lado, también se han destacado cuestiones ligadas al contenido de la encuesta propiamente dicho, es decir, a las preguntas concretas que se encuentran en el

cuestionario. Algunos docentes y estudiantes señalan que las preguntas de la encuesta no permiten recoger información que sea del todo útil:

“En segunda instancia, cada vez que realizo una encuesta me sorprende de que las preguntas/consignas no cambien desde que entré a la Universidad. Pienso que sería bueno que sean revisadas y reestructuradas. A veces, me da la sensación de que no logran obtener una objetiva respuesta de lo que significó o comprendió el curso” (Estudiante 7, fragmento 4).

“A veces suena que es muy tabulada. Aunque entiendo que para el procesamiento debe ser así muy tabulada. Lo ideal sería que diera más oportunidad de expresarse. Y a veces es muy...no tiene en cuenta la diferencia entre las materias de los alumnos y de los alumnos de las materias” (Docente 5, fragmento 4).

“A ver, no es la información que yo preguntaría. Me parece que es una formación más general o hablan más del curso. En cierto sentido está bueno, porque me da como la visión más general del proceso. Para el día a día no me sirve, porque para cuando me lo dan ya termino el curso. Entonces, en algún sentido tiene como ese límite ese tipo de instrumento” (Docente 6, fragmento 6).

Aquí, se exponen nuevamente tensiones propias de una evaluación que pretende ser formativa, pero que presenta rasgos propios del modelo tecnológico y productivista. Rasgos como la estandarización de las preguntas y la tabulación resultan fundamentales en las evaluaciones cuya finalidad es la de controlar y comparar, pero resultan contraproducentes para evaluaciones cuya finalidad es retroalimentación para la mejora continua.

#### 4.2.2 ...y el control

Como se ha evidenciado en el sub-apartado 5.1, el propósito de la encuesta parece ser doble. Por un lado, resulta un instrumento y una instancia con fines formativos, en tanto recupera el feedback de los alumnos en relación a las prácticas de los docentes. Por el otro, en cambio, adquiere una función más de control, en tanto provee información a los directores de carrera para tomar decisiones *high-stake*. Es decir, relacionadas a instancias más bien administrativas. Respecto a esto último, un repaso por la literatura indica que la preocupación de los docentes en relación a este tipo de uso es unánime.

#### Criterio relevante para decisiones *high-stake*

Como hemos dicho, los cuestionamientos al uso de la encuesta, usualmente vienen asociados a la utilización de sus resultados para orientar a acciones ligadas a el salario, la promoción, la permanencia y el reconocimiento de los docentes. Ahora bien, ¿qué sucede en el caso de la Universidad de San Andrés?

En primera instancia, pareciera ser que, entre los docentes de este estudio, no hay un conocimiento claro de la utilización de las encuestas con fines sumativos en la universidad. Este desconocimiento se puede inferir a través de su lenguaje, pues no parecen tener una opinión muy fundada respecto al tema. Quizás, lo anterior resulte del escaso contacto que han tenido con situaciones que involucraran a las encuestas y que, por tanto, hayan repercutido en su carrera docente:

“La verdad que no conozco ningún caso puntual. En mi caso, a la inversa, no veo que haya redundado en ningún plus salarial, ni en ningún ofrecimiento...Por la inversa me di cuenta de que no funciona tan así...por lo menos en el caso positivo” (Docente 4, fragmento 8)

“No. Mira, hubo una sola vez que un director me dijo que una asistente que yo tenía, me dijo que las clases eran flojas y que bueno eso le preocupaba y que era una pena porque de alguna manera eso bajaba el nivel del departamento. Y a mi me molesto eso. Me molesto porque era una asistente que había dado dos clases. y yo le dije, uno se forja dando clases. Ella está empezando y a mi no me afecta en nada que el promedio baje un poco por esto. (...) Entonces ahí dije, esto de las encuestas. Como que me parecía que él estaba midiendo mucho que pasaba en un departamento y que pasaba en otro y bueno, la evaluación era un poquito mezquina. Salvo que haya un error o haya alguien que no cumple...y después fluyó. Nunca nadie más me dijo nada” (Docente 9, fragmento 2)

En el primer fragmento se evidencia que, en su experiencia como docente de la Universidad, las encuestas no se han tomado para disponer acciones *high-stake*. El segundo fragmento, ya sí nos ofrece información más valorativa, asociada a una creencia o una concepción que parece ser negativa. A partir del relato de su experiencia, el docente expresa cierto descontento frente a la actitud del director de tomar una decisión tan determinante a partir de la encuesta, a partir de un número. En relación a esto último, en cuanto al lugar que deben o deberían ocupar las encuestas en este tipo de decisiones, las opiniones son diversas, aunque la mayoría parecen converger en que las encuestas son *un* factor, “[n]o son ni el factor más importante ni el factor definitivo” (Director 2, fragmento 11). Al mismo tiempo, algunos docentes opinan que la encuesta puede ser un criterio, pero que se debe hacer un uso razonable, con todos los matices que ello pudiera tener:

“Sí, como un indicador para ese tipo de decisiones desde ya. No el único. No puede ser el único, pero sí es un indicador muy útil sobre todo si se diseña como corresponde y se toman los recaudos mínimos para evitar la demagogia y las cosas que no pueden ser” (Docente 8, fragmento 12)

“Y sí, me parece que sí...si hay otro mejor, mejor. Pero si es el instrumento que se tiene, está bien. Siempre que se haga una lectura razonable de esos datos. No vas a estar diciendo ‘venias siempre en 9 y caíste a 7, bajaste el desempeño y te vamos a cambiar de materia’, eso es una tontera. Pero bueno, no tengo la menor idea de cómo se toman las decisiones y cuáles son esos criterios” (Docente 3, fragmento 12)

Sin embargo, a pesar de desconocer si efectivamente las encuestas son utilizadas con esta finalidad, algunos pocos docentes se muestran totalmente reacios a que lo sean:

“Para mí no. Porque de vuelta, es la visión del alumno. Y primero la visión del alumno está sesgada por si aprueba o no...o, a veces, al alumno le falta ver cuál es el objetivo final del docente...Por ejemplo, en mi caso de [...], uno puede explicar como hacer un ejercicio un millón de veces [...] pero por ahí hay otra parte que no se muestra, que es la parte más conceptual y es más difícil...y eso no se evalúa, y el alumno no lo sabe. Cuando uno da clase, especialmente de [...], tiene que tomar ciertas decisiones y eso el alumno no lo sabe. Entonces al no saber eso, su va a estar juzgada por eso; que le tienen que enseñar una cosa, cuando en verdad le están enseñando otra” (Docente 1, fragmento 13)

En este caso, el docente señala que las encuestas se vuelve un criterio poco válido en tanto se corresponden con la visión de un alumno, que puede estar muy sesgada y que puede desconocer las intenciones de fondo o el quid de la cuestión de la materia. Esta postura, aunque poco frecuente en el presente estudio, es la actitud que han adoptado los docentes en general, según la literatura (Nasser & Fresko, 2002; Chan *et al*, 2014). Fundamentalmente, las razones para desestimar el uso de las encuestas vienen asociadas, como podemos evidenciar en el fragmento, a cuestiones que rondan la confiabilidad y que analizaremos con mayor profundidad en el apartado que corresponde a la dimensión tres.

En cuanto los directores de carrera, ciertamente quienes son partícipes de la toma de decisiones, las posiciones también han sido diversas, aunque han estado siempre marcadas por un halo de reflexión:

“Me parece que, si alguien se opone, que explique por qué. Me parece que es un elemento más como para analizar cómo están yendo los cursos, los profesores. Es un indicador más. No es que con solo la base de eso uno puede obtener un diagnóstico completo. Pero ciertamente, un poco como decimos los economistas, es ver lo que opina la demanda. Es como que te diga tener una encuesta de satisfacción de los consumidores, o que opinan los consumidores es siempre muy valioso” (Director 3, fragmento 4)

En este caso, como expresa el Director 3, se puede evidenciar la reivindicación del valor de las encuestas, en tanto son los estudiantes los consumidores últimos del servicio provisto por la universidad. Esto se puede asociar con lo que propone Vain (1993) y las potencialidades que tiene elegir a los estudiantes como actores principales de la evaluación docente. Para el autor, al ser los estudiantes los principales implicados por las consecuencias de la enseñanza, los vuelve ciertamente calificados para convertirse en críticos de esta actividad. Sin reconocer lo anterior, otros autores advierten sobre las cargas de concebir a los estudiantes como consumidores y los peligros del advenimiento de la lógica de mercado sobre los sistemas educativos (Alexander, 2000; Ball, 2013).

Por otra parte, hay testimonios que evidencia que el valor de las encuestas se corresponde con el hecho de considerar a las encuestas como único elemento o como un instrumento de muchos para evaluar la docencia. En este sentido, un director se pregunta:

“Es un instrumento ¿es un instrumento de muchos, o es el único instrumento para tomar decisiones respecto a las carreras de los docentes?, y ahí cambia.” (Director 1, fragmento 19)

En esa misma línea, el Director (4) recalca la inexistencia de un criterio mejor o alternativo y reflexiona acerca de las limitaciones que supone la primacía de este instrumento, y de los números en particular:

“Si yo tuviese otro instrumento diferente para comparar, te diría bueno, me parece que este es mejor que este. Pero no tengo otro, siempre han sido los números. Es evidentemente el número, pesa a la hora de las decisiones políticas, cualquiera fuera. Por ejemplo, pesa a la hora de elegir el premio a la docencia...es el número ¿Qué otra forma hay? Me encantaría que hubiera otra. Porque muchas veces las diferencias entre el profesor 1 y el profesor 2 son dos centésimos ¿Y qué hacemos con el de 9,51 que también es extraordinario? ¿Todos los años va a perder con 9,50? ¿qué otra forma de evaluar hay? entonces que el número pese en las decisiones políticas de la universidad y... no hay otro. Pero yo pondría en duda su completa eficacia. Muéstrénnos otro instrumento posible de evaluación y vemos” (Director 4, fragmento 13)

Es relevante además resaltar que se hace referencia esta vez a la utilización de las encuestas para orientar acciones *high-stake*, como es el Premio a la Docencia. Este reconocimiento, justamente se vale de los números, aunque el Director (1) expresa sus preocupaciones frente a las limitaciones de dicho criterio. En última instancia, se evidencia cierto sentimiento de desconsuelo de cara a la ausencia de una instancia alternativa para evaluar el desempeño docente.

Por último, respecto a las creencias de los estudiantes, pareciera ser que están en línea con la pluralidad de posturas en los otros dos grupos. Por un lado, hay estudiantes que han demostrado tener posturas radicalmente a favor de la utilización de encuestas con fines sumativos: “Es más, te diría que deberían ser cien por ciento tenidas en cuenta para tanto echar, o para charlar con profesores o demás” (Estudiante 5, fragmento 10). Sin embargo, por otro lado, hay estudiantes con posturas algo más conciliadoras, que entienden que la encuesta es relevante aun cuando debe estar acompañada de una lectura correcta de sus datos y de la instauración de una cultura de la evaluación más reflexiva:

“No me parece que la encuesta tenga que tener tanto poder por sí misma, como para echar un docente. Hay que ver que hay en el trasfondo. Pero por eso me parece importante hacer una correcta interpretación de las respuestas de los alumnos...me parece fundamental hacer preguntas correctas, charlar con el profesor. Todo ese proceso me parece fundamental” (Estudiante 10, fragmento 9)

“Yo creo que sí, que se podría tomar de ese modo, pero que la universidad debería enfocarse en crear una cultura en la que el alumno sepa que tiene ese poder y que sepa usarlo bien. Creo que, si haces eso ahora, con cómo la gente se toma las encuestas, podría llegar a ser un desastre.” (Estudiante 4, fragmento 10)

### El caso de las eminencias

Sumado a lo anterior, vale la pena detenerse en los dichos de algunos estudiantes que evidencian como el criterio de utilizar la encuesta para orientar acciones *high-stake* varía según quién sea el profesor involucrado. Para estos estudiantes, el impacto de la encuesta no solo depende de la gravedad de lo que debe, sino de la figura o la autoridad del profesor. En otras palabras, existen casos de profesores reconocidos, eminencias que, a pesar de haber recibido fuertes llamados de atención en las encuestas, han permanecido en el cargo sin ningún tipo de responsabilización por lo sucedido:

“Yo creo algunos profesores que no...ósea yo por ejemplo curse economía con [...]. La verdad que [...] me pareció la anti pedagogía y me pareció que es algo que viene hace años, me dijeron que no cambia. Y no me imagino que las personas le pongan excelente en las encuestas, y la verdad que me parece que no le importa...no recibe ningún tipo de reprimenda.” (Estudiante 3, fragmento 5)

“Creo que depende mucho el caso, de la gravedad del asunto. No debería depender del profesor, pero muchas veces depende de quién es el profesor. Si es una eminencia se le permite más cosas. Y supongo también depende mucho de cuales sean las perspectivas del director o la directora de la carrera. Si le recortan el salario, si toman cartas en el asunto...Pero yo creo que hay casos en los que sí, toman cartas en el asunto, y casos en los que no. Pero conozco muchos rumores de profesores que aparentemente se vieron afectado por las encuestas. Y también conozco casos de ‘impunidad’. Por lo cual...no sé, hay de todo” (Estudiante 10, fragmento 7)

### Consecuencias de utilizarla con fines sumativos

Una de las cuestiones que ha sido abordada en la literatura son los posibles desenlaces que podría tener utilizar las encuestas con fines sumativos, esto es, como una instancia decisiva y de responsabilización individual, focalizada solamente en los resultados obtenidos. Aunque como hemos dicho anteriormente, la encuesta no parece utilizarse con esta finalidad en la universidad en cuestión, algunos docentes señalaron cuales podrían ser las consecuencias. Por un lado, están quienes consideran que, de ocupar la encuesta un lugar central en las decisiones de la universidad, traería aparejada la valoración de la enseñanza:

“Yo creo que, si se valorizan las encuestas, parecería que sería razonable que los docentes le den más importancia al *teaching* o a la enseñanza, que eso me parece un *outcome* positivo. Si eso sucede, estaría cómodo con eso. Sería difícil, si la encuesta está bien hecha, que el docente trabaje para la encuesta y los chicos no aprendan más” (Docente 6, fragmento 13)

Esta postura nos remite a los hallazgos de algunas investigaciones (Aleamoni, 1981; McKeachie, 1979) que han encontrado que el uso de las CEDA incrementa las chances de que la excelencia en la docencia sea reconocida a nivel institucional y, por tanto, también recompensada. Sin embargo, la gran mayoría de los docentes coinciden con que la valoración de las encuestas podría conducir a los docentes a cambiar su proceder, a

en el afán de obtener evaluaciones más favorables: “Entonces para mí hay una tensión con la encuesta. Si uno pone la encuesta por delante de todo, termina trabajando para las encuestas, y muchas veces eso no es lo mejor” (Docente 5, fragmento 20). En línea con lo anterior, otro docente plantea:

“Y si vos le pones demasiado peso a eso, tiendes a distorsionar el comportamiento del docente, que empieza a enfocarse a sacarse buenas notas en esa medida imperfecta y parcial, es desmedro de otras cosas que también son importantes o que te gustaría que el docente hiciera el frente del curso. Ese sería mi visión sobre las encuestas [...] Quiero decir, imagínate que hay decisiones realmente importantes detrás de esto, bueno el incentivo a ser más contemplativo con la nota como una forma de que después te evalúen mejor está ahí, y uno no puede desconocerlo” (Docente 7, fragmento 2)

En este fragmento, el docente comienza a dar pistas sobre el significado que se le atribuye a comportarse o trabajar “para las encuestas”, como el ser más contemplativo con la nota. Para otros docentes, trabajar para las encuestas significa bajar los estándares de evaluación, bajar la carga de lecturas y de trabajo, ser más entretenidos:

“Puede ser que los docentes sean medio... "alumnistas", viste que hagan "alumnismo" académico. Que sea pura atención de la demanda de los estudiantes, entonces bajar los estándares de evaluación, bajar la lectura...en principio uno podría decir que ese es el efecto que se genera” (Docente 3, fragmento 13)

“En parte depende de para que se usen las encuestas después. En un punto diría, si yo se que mi trabajo depende de la encuesta...me hago un favor a mí misma. Hago las clases fáciles, nuestro películas, canto, bailo...” (Director 1, fragmento 27)

Lo dicho por los entrevistados nos lleva a pensar que, debajo de estas creencias, subyacen ciertas hipótesis, como la de la benevolencia (Braskamp y Ory, 1994; Centra, 1997; Marsh, 1987), que indica que los alumnos ponen mejores calificaciones en las encuestas, a quienes esperan los van a calificar mejor, por ejemplo. Si bien indagaremos sobre estos puntos más adelante, es preciso resaltar que estas actitudes son propias de los docentes que tienen algunas creencias sobre el potencial comportamiento de los alumnos, y que esperan que los califiquen mejor si ellos adoptan actuaciones como las mencionadas.

Volviendo sobre lo anterior, al efecto de centralidad de las encuestas en la toma de decisiones sobre el comportamiento docente, algunos expresaron que puede reducirse a una cuestión de incentivos:

“A ver, es control de gestión. Si vos pones donde pones el incentivo, ahí va a ir la persona definitivamente. ¿Todo el mundo actúa así? No, somos complejos, tenemos distintas formas. Pero si vos decís, por esto me miden, y por todas estas cosas que hago, nadie las ve, y no hay un premio; estas definitivamente conduciendo a que se haga algo por lo que te miden. Vos tenés distintas formas de lograr que las cosas pasen. Entonces si tu objetivo organizacional es encuestas 10, y lo único que miras y premias, entonces, y comunicas que eso vale, entonces la gente va a exigirse sólo para la encuesta 10 y van a ver de qué manera conseguir un 10; y si no te sale por un camino...por el camino que salga. Entonces empezás a ser más condescendiente con el alumno” (Director 5, fragmento 12)

“Para mi, mirar eso, sin mirar, sin correlacionarlo con las calificaciones de los alumnos me parece en general que es muy inapropiado. No inapropiado, muy inapropiado. Primero que todo, por los incentivos hacia los profesores. El incentivo es hacia la demagogia, a poner alta calificaciones y recibir altas calificaciones. Entonces los evalúan mejor...esa es una de las cosas más terribles que hay.” (Docente 8, fragmento 1)

Podemos vislumbrar como los argumentos de los docentes y directores encuentran un correlato en la naturaleza de la evaluación sumativa, para el control. Si la intención de la organización es establecer el éxito u fracaso de un docente respecto a un desempeño esperado (Fernández Lamarra y Aiello, 2012), a través de la encuesta, entonces resultará ineludible que el individuo que esta siendo evaluado adecue su comportamiento para la consecución de una evaluación satisfactoria. Tal como sugieren Chan y otros (2014) en su investigación, bajo estas creencias subyace la idea de que la introducción de evaluaciones de estas características podría, potencialmente, motivar a los profesores a reducir la exigencia, los estándares, con la intención de obtener mejores evaluaciones.

Ahora bien, lo expuesto nos lleva a preguntarnos cuál es el comportamiento actual, “real”, de los docentes de la Universidad de San Andrés. En este punto, algunos docentes afirman no conocer ningún caso en que eso haya acontecido, de hecho, dicen que: “eso podría ser una hipótesis, pero no veo evidencia empírica de que eso haya sucedido” (Docente 3, fragmento 13). De hecho, un docente atestigua que:

“Siempre lo que trato de hacer es no caer en la demagogia de, tengo que satisfacer a los estudiantes sí o sí, y ser el más querido. Que es un peligro además en el que el sistema lo lleva a uno, y ese es el lado oscuro de las encuestas porque...sí las encuestas sirven para determinar si un docente continúa o no al frente de un trabajo, ahí hay un peligro. Porque si el docente eso lo percibe, va a correr tras de esta demagogia que no es bueno.” (Docente 4, fragmento 6)

Sin embargo, hay quienes sospechan que muchos profesores si se comportan “para las encuestas”:

“Yo tengo la sospecha de que muchos profesores lo hacen, sí. Que, en algún punto, no sé, te puedo decir que quizás es entendible. Me parece que quizás es distinto el profesor *full-time* que tiene un compromiso distinto con la universidad. Pero el profesor *part-time*, que te pagan por venir y dar clase, hasta que punto tengo que ir en contra de todo para que encima después en la encuesta me critiquen porque me pusieron tal nota. Entonces sabes que...vamos a simplificarnos la vida. Y yo creo que muchas veces termina teniendo consecuencias que no están buenas para el aprendizaje, para las materias siguientes...así que yo creo que sí, eso puede tener impacto, y puede haber docentes que decidan hacer determinadas concesiones a los estudiantes para tener mejor las encuestas.” (Docente 5, fragmento 22)

La reflexión de este docente resulta importante para ilustrar todo lo anterior, pues es de esperar, “es entendible” que los docentes se comporten de cierta manera si los incentivos son los de calificar bien en las encuestas y más aún si esto impacta de manera categórica en su carrera. Tal como afirma Vain (1998), toda evaluación es indisociable de un propósito y de un cierto posicionamiento. En ese sentido, resulta importante reflexionar

sobre los potenciales incentivos que podría motivar la utilización de la encuesta con ciertas intenciones.

### 4.3 De la confiabilidad...

Como hemos mencionado en diversas ocasiones, los argumentos principales que esgrimen quienes se oponen a las encuestas o quienes desconfían de ellas se desprenden de cuestiones relacionadas a su confiabilidad. En otras palabras, las creencias versan sobre las limitaciones que presenta este instrumento a la hora de dar una medida válida del desempeño docente. Ya en fragmentos anteriores se han podido evidenciar estas creencias, como la idea de que los estudiantes califican mejor a las clases más fáciles o a los profesores más entretenidos o, la noción de que los alumnos utilizan las encuestas para desquitarse por malas calificaciones, o que no se encuentran calificados para evaluar la enseñanza, entre muchas otras. Sin embargo, dada la importancia que le ha dado la literatura a estos factores que atentan contra la confiabilidad de las encuestas y el lugar que ocupan en la construcción de ciertas concepciones (y por tanto de ciertos comportamientos) decidimos parte del trabajo a su profundización.

#### 4.3.1 ...y los estudiantes

Para comenzar, una cuestión que vuelve problemática o al menos poco confiable a las CEDA es la actitud con las que los estudiantes completan las encuestas. Aunque bien la heteroevaluación presenta algunos beneficios, también es posible encontrar algunas preocupaciones ligadas a la incapacidad de los estudiantes de realizar evaluaciones objetivas, serias y alejadas de sus preferencias, expectativas y emociones particulares.

“A veces hay críticas injustas también porque uno sabe que las responden alumnos que no vienen mucho a clase, o que están enojados por la nota que se sacaron, o que no trabajaron, o que les molesto...también hay cosas que no son reales” (Docente 5, fragmento 15)

#### Seriedad y compromiso

La primera cuestión indagada en relación a las actitudes de los estudiantes refiere a las creencias en relación a la seriedad con la que asumen la evaluación. Frente a este interrogante, las posturas de docentes y directores han sido de lo más diversas y van desde los que no tienen una opinión muy formada a los que presentan posturas muy determinantes.

En primer lugar, están quienes consideran que las respuestas de los chicos evidencian un alto nivel de honestidad y dedicación:

“En general vos ves las respuestas de los chicos y son realmente genuinas, bastante sinceras.” (Director 3, fragmento 1)

“Yo creo que se lo toman en serio, sí; que por lo menos la retroalimentación que recibí yo siempre fueron todas súper razonables. No vi nada que me llamó la atención [...] En mi experiencia no creo que tengan incentivos para decir una cosa que no es cierta.” (Docente 6, fragmento 16)

En segundo lugar, y con concepciones totalmente contrarias al grupo anterior, se encuentran los docentes y directores que desconfían de la actitud de los estudiantes a la hora de completar las encuestas, pues según su concepción los estudiantes son deshonestos, mienten, o ponen cualquier cosa:

“Se les pregunta a los chicos si [los profesores] están disponibles...puede ser que sí, puede ser que no. Los chicos igual en general ponen cualquier cosa” (Director 1, fragmento 21)

“A mi me ha pasado que generalmente, en todos los años que llevo acá, que hay estudiantes que mienten. En esos casos donde se pregunta si faltó, hay estudiantes que dicen que sí faltó; y yo lo sé, yo no falte nunca. Entonces bueno, ¿porque contestan eso? Se han sacado mala nota en el primer parcial...no lo sé. La materia les habrá resultado horrible, porque son contadores y no lograron engancharse. Pero yo también sé que a mi me ha ido muy muy bien...entonces me pregunto, ¿por qué aparecen expresiones negativas? ¿por qué aparecen además diciendo cosas que no son verdad? Me parece que las encuestas habilitan este tipo de comentarios.” (Director 4, fragmento 16)

“No, yo creo que no se las toman en serio, que hay cosas de falta de honestidad. No por falta de honestidad por malos, sino porque bueno...no les gusta la materia entonces no son honestos con lo que se les pregunta. Si puedo matarte te mato, para que no esté más esta materia, esa es una. O porque fuiste muy exigente. Entonces en lugar de valorar la exigencia, porque uno puede ser muy exigente y a la vez ser un gran docente...entonces dicen 'no porque fue muy exigente' y entonces te tiran a matar” (Docente 10, fragmento 16)

“A veces hay críticas injustas también porque uno sabe que las responden alumnos que no vienen mucho a clase, o que están enojados por la nota que se sacaron, o que no trabajaron, o que les molesto...también hay cosas que no son reales” (Docente 5, fragmento 15)

Este grupo concibe a los estudiantes como un grupo deshonesto frente a la evaluación, que sabe perfectamente como ejercer su poder a través de este tipo de instrumentos para lograr cometidos que los favorezcan.

Por último, en una posición más intermedia, están los directores y docentes que no pueden afirmar con certeza si los estudiantes se toman las encuestas con un alto nivel de compromiso o con ninguno. Estos comprenden al grueso de los entrevistados y pueden afirmar que sus actitudes, en general, son variadas:

“Supongo que hay varianza, que algunos se la toman muy en serio, y otros no tanto; y que maso menos en el promedio se la toman maso menos en serio.” (Director 2, fragmento 18)

“Hay de todo, pero sí, hay muchos que sí, ósea. En todo tenés de todo. El universo es grande y hay algunos que sé que ni siquiera la responden y pasa. Pero después tenés otro donde se lo toman a pecho, hacen comentarios que son súper valiosos. Yo creo que en general el alumno las responde a conciencia.” (Director 5, fragmento 14)

“Una intuición que tengo es que los alumnos no contestan cualquier cosa, no contestan el valor del medio. Eso sería 0 seriedad. [...] Pero eso no significa que piensen cada respuesta. Pensá como contestamos encuestas habitualmente que se yo, y ellos tienen que llenar 4-5 por semestre...con las mismas preguntas, entonces...” (Director 4, fragmento 19)

“Eso no te podría decir. [...] Supongo que habrá un poco y un poco, pero no me puedo dar cuenta. Que son ‘fiacas’ para hacerlo sí, pero eso no quiere decir que se lo tomen en chiste” (Docente 9, fragmento 12)

Así es que, para la mayoría de los docentes, los estudiantes adoptan una postura intermedia: cierta seriedad y cierto compromiso. La situación con los comentarios parece sustentar esta postura, pues hay quienes los consideran como un *proxy* del interés de los alumnos por la evaluación de la docencia. Aducen que: “en este momento la encuesta para muchos es como un trámite. Completo, completo maso menos y sigo. Porque están poniendo menos comentarios” (Docente 5, fragmento 23). Asimismo, la reflexión del Director 4 ilustra esta idea:

“De 80, ¿cuántos chicos ponen comentarios? 15 como mucho. Entonces ¿qué hacen los otros 65? Supongo que hacen ‘tiki tiki tiki’, ponen la media. ¿Cómo contesto yo las encuestas de AySA? Maso menos puedo proyectar eso. No creo que los chicos pongan mucha dedicación” (fragmento 20)

Lo que perciben los docentes y directores como “seriedad y compromiso promedio” no se aleja mucho de los testimonios de los estudiantes. De hecho, es posible constatar que lo percibido por este último grupo es lo que tiene un correlato más acertado con la realidad:

“Intenté que la encuesta sea un instrumento confiable, pero creo que podría ser un poco más. La parte de comentarios muy pocas veces fui muy detallista. O porque me daba fiaca o porque no tenía una opinión muy formada. Pero de cada 5 encuestas que llené, la mitad las hice muy detalladas y el resto las hice muy por arriba. Que no me enorgullece y que siento que el docente por más que esta buena que las vean, preferiría que se fíen mas de otras” (Estudiante 6, fragmento 13)

Como podemos ver, en general este nivel de compromiso promedio se encuentra asociado a la “fiaca” que acarrea llenar los cuestionarios. Es decir, a lo engorroso que resulta tomarse el tiempo y tener las ganas para completar las encuestas. Ahora bien, ¿cuál es el criterio de los estudiantes para detenerse a llenar una encuesta con dedicación o hacerla por arriba para sacársela de encima?

Un hallazgo interesante es la relación que existe entre compromiso con el que los estudiantes asumen la evaluación docente y el carácter extremista de la experiencia, en sentido tanto positivo como negativo. En otras palabras, los estudiantes se toman el tiempo de completar la encuesta con mayor rigurosidad cuando los profesores han tenido un desempeño sobresaliente o pésimo:

“A ver...con algunos profesores que me interesa destacar cosas sí, pero hay algunos otros que no, como que pasó más rápido (...) A veces como que no te quedas pensando...entre un 4, 5 o un 3..y en donde hay partes de qué aspectos destacas; difiere mucho de algún profesor en los que realmente quiero destacar algo con otro que no...la extensión de eso que quiero destacar, sea positivo o negativo” (Estudiante 2, fragmento 7 y 8)

“Yo también trato de hacerlo bastante extendido cuando puedo. Creo que igual, cuanto más mediano sea el profesor [ni muy bueno ni muy malo], si es demasiado malo lo voy a tener que escribir y decir porque es malo; y si es muy bueno trato que escribir que es bueno y porque es bueno. Pero cuando no son ni buenos ni malos, suele ser cuando menos escribo, y los puntajes son un poco arbitrarios siempre. Con el dedo elijo un número sin mucho fundamento...pero bueno, las encuestas son así.” (Estudiante 1, fragmento 9)

“Creo que me lo tomo 3 [en una escala del 1 al 5], ósea si tomamos la escala de la encuesta si. Pasa que al menos que sean realmente materias muy puntuales que me parece que hay que cambiar tal cosa...después si mas o menos me gustó le pongo a todo 5 y ni lo pienso. Entonces, y si no me gusto a todo 1...ósea soy re extremista digamos, no es que realmente lo pienso.” (Estudiante 4, fragmento 12)

Los docentes y directores parecen advertir esta relación, pues afirman que quién completa la encuesta usualmente es el que está enojado (Director 1, fragmento 9):

“Yo creo que algunos se lo toman en serio pero otros no. Pero, es más, los que más se lo toman en serio son los extremos, los que quieren criticar mucho y los que quieren decir que fue bueno. Eso pasa.” (Docente 8, fragmento 21)

“Que en general los comentan son los que les gusto mucho y lo quieren poner, o están muy enojados y lo quieren manifestar. Ósea que los comentarios vienen más bien de las puntas.” (Docente 5, fragmento 23)

“Es como el que va a votar. ¿Quién va a votar? va el que tiene más preocupación por el resultado, lo cual es otra razón para tener cuidado con las encuestas. Porque típicamente lo que vas a ver son los extremos, para un lado y para otro; y quizás los otros que pusieron todo en el medio, para no pensar muy bien en discriminar, ni comparar con otro docente.” (Docente 7, fragmento 5)

A pesar de que en esta Universidad la encuesta sea obligatoria, la analogía propuesta por el Docente 7 resulta de utilidad para argumentar que, en tanto la radicalidad de la experiencia se encuentre íntimamente asociada con la seriedad con la que los estudiantes se toman las encuestas, constituye un sesgo importante que debe tenerse bajo consideración.

### Relación con la calificación esperada

Otra cuestión que ha recibido mucha atención en la literatura es el tipo de relación que existe entre la calificación que reciben o esperan recibir los estudiantes y su actitud para con las encuestas. En nuestras entrevistas, pudimos detectar que muchos docentes creen en la “hipótesis de la benevolencia”, pues: “si son justos o no...esto también

depende de cómo les fue en la materia” (Docente 1, fragmento 17). Esto es, tanto docentes y directores sostienen, en su mayoría, que aquellos estudiantes que esperan recibir calificaciones más altas evaluarían más favorablemente a sus docentes (Braskamp & Ory, 1994; Centra, 1979; Marsh, 1987; entre otros):

“Es cierto que ya saben como les fue en el parcial, con nota y todo. Tienen una apreciación de cómo les fue en el final, y pueden pronosticar con cierto grado de precisión que se van a sacar en la materia. Y también es cierto que hay formas de ver tu nota antes de la encuesta. Y lo que la evidencia ha demostrado, es que eso pasa. Los chicos reaccionan, lo que ponen no es independiente de la calificación que se sacan.” (Docente 7, fragmento 1)

“Si a muchos les fue mal, las encuestas van a ir más bajo. Y muchos sobre todo en los comentarios por ahí van a ser un poco más agresivos. Pero eso no quita que quizás a muchos les fue mal pero tienen razón. Tampoco uno lo tiene que tomar personal” (Docente 1, fragmento 18)

“En general adoptan una actitud vengativa cuando les va mal en los exámenes. Pero eso es un poco natural en las personas, eso es de esperar.” (Docente 8, fragmento 16)

“Y cuando el alumno al responder, si tiene esto de que me fue muy mal y estoy con mucha bronca, responde con más agresividad. O me fue bien, estoy contenta, si está todo fantástico, esta todo bien. No por eso no se la tomen en serio; pero creo que siempre hay algo fuera de cómo se dió la materia o no, que a ellos les influye en contestar una cosa o la otra” (Docente 1, fragmento 19)

“Yo creo que puede haber un momento de enojo post examen que te puede impactar. Creo que hay algo en ese momento...Yo estudié un montón y rendí un montón, y he visto situaciones de enojo. Ósea si vos respondes las encuestas post examen, y el examen no cumplió tus expectativas, creo que impacta de manera negativa. Y generalmente el examen no cubre tus expectativas por distintos motivos” (Director 5, fragmento 18)

De esta forma, la calificación y el examen final se encuentran relacionados con la encuesta en tanto despiertan sentimientos de enojo o cierta actitud de responsabilización en los estudiantes, de culpar o premiar al docente por su propio desempeño como alumnos: “en las encuestas muchas veces pasa que el que aprobó, va a estar más de acuerdo de cómo fue la materia y el que desaprobó más en desacuerdo” (Docente 1, fragmento 6). Son pocos los docentes que creen en hipótesis alternativas como la de la validez, que esgrime que los buenos profesores hacen que los estudiantes aprendan más y por tanto tengan mejores calificaciones: “Pero ojo, una cosa quizás está relacionada con la otra [...] quizás a muchos les fue bien porque tuvo un muy buen cuatrimestre dando clases, entonces está todo bien. [...] como que una cosa va llevando a la otra” (Docente 1, fragmento 18). O que creen, como el Docente 9, que el desempeño del estudiante y la calificación en el curso no impactan en la evaluación del profesor: “No necesariamente, no se ve tan reflejado. Pero he tenido gente que ha recursado en los últimos años (...) Pero eso no me parece que se haya reflejado tanto en las encuestas, por lo menos que yo sepa” (fragmento 11).

Las afirmaciones de los estudiantes acompañan de cierto modo las impresiones de los docentes y directores. Lo llamativo es que las actitudes vengativas o la poca seriedad aparecen cuando los entrevistados se refieren al comportamiento de otros estudiantes:

“Es que sí, uno en general, y lo hablaba el otro día con mis compañeros, y decían este profesor es buenísimo; y yo les decía, es buenísimo porque te fue bien, el día que te vaya mal va a ser malísimo. Y dicen, bueno yo los mido según como me va, es obvio. Y te lo decían así, me entendés...yo no lo pienso así, pero hay gente que lo piensa así.” (Estudiante 8, fragmento 7)

“Generalmente se lo toman como un trámite. Creo que lo veo como una forma de descarga. Se enojan en el parcial porque fue muy difícil entonces en la encuesta le voy a poner todo mal. Entonces no sé hasta qué punto se puede confiar en las encuestas, cuando muchas veces expresan más emociones que racionalmente.” (Estudiante 9, fragmento 14)

Esto coincide con los hallazgos del estudio de Slocombe *et al* (2011), donde aproximadamente la mitad de los estudiantes consideran que el resto se las toma en serio. Por el contrario, cuando se refieren a si mismos, las actitudes suelen ser diversas. Por un lado, están quienes niegan relacionar la calificación esperada con la forma de evaluar a los docentes:

“Creo que no. Quizás pongo muy buena nota, aunque no me haya ido bien. Lo que pasa es que generalmente las que me gustan muchísimo me suele ir muy bien, le pongo todo 5 porque creo que la materia fue perfecta. Pero después hay materias que me cuestan y no es que digo 'ay les voy a poner todo 2', porque la realidad es que es mi problema.” (Estudiante 4, fragmento 15)

“Y no creo nada en eso de si ‘me reprueba, cambia...’. Yo no creo mucho en eso. Hay que mirar a grandes rasgos la clase entera.” (Estudiante 5, fragmento 5)

“En mi caso, la única forma que tenga que ver la calificación con mi evaluación es si hay algo en particular de la evaluación que merece ser destacado en la encuesta. Por ejemplo, si considero que fueron muy injustos conmigo, lo voy a reflejar en la encuesta, lo voy a decir. No sé donde, calculo que en los comentarios.” (Estudiante 10, fragmento 13)

Por el otro, se encuentran quienes reconocen la existencia de esta correlación, aunque afirman que en su caso intentan tratar de separar la calificación de la evaluación:

“Si, ósea siento que existe [correlación entre nota y encuesta]. Siento que quizás, muchas personas, y me incluyo, tratamos de separarlo. Pero a veces es muy difícil...quizás uno no hace autocrítica” (Estudiante 2, fragmento 10)

“A ver, para ser sumamente claro, en mi caso, intento que una disconformidad no tiña todas las dimensiones que pueden evaluarse de un docente. Si la calificación estuvo bien puesta, a mi criterio, no hay ninguna forma de que la encuesta hable sobre la calificación” (Estudiante 8, fragmento 15)

“Yo creo que la correlación existe. Y supongo que hasta uno que trata de ser objetivo...si te fue mal, empezás a buscarle más defectos, por esto, el sesgo este...de autovalidación: si te va mal, tratas de encontrar algo que excuse que te haya ido mal, y si te fue bien tratas de que sea mérito tuyo. Acá es re evidente eso, y eso creo que tiene mucho peso” (Estudiante 1, fragmento 8)

En estos casos es evidente la tensión que existe entre la calificación y la intención de evaluar objetivamente al docente. Los estudiantes, a pesar de tener en claro sus intenciones de evaluar justamente al profesor, señalan la dificultad que entraña diferenciar

sus sentimientos –muchas veces motivados por la calificación, del desempeño del docente:

“Necesito 5 minutos de meditación antes de llenar la encuesta, y pensarlo bien. Pero, si a mi no me sirve de nada desquitarme con el profesor por algo que ni siquiera es culpa suya” (Estudiante 3, fragmento 14)

“Yo creo que intento ser objetiva, pero sí confieso que me ha pasado de llenar los formularios y tener un examen muy malo, y decir que fiaca que ya llene la encuesta, porque pondría esto, pondría mucho enojo.” (Estudiante 6, fragmento 12)

### Capacidad de evaluar la enseñanza

Siguiendo con la idea anterior, se puede notar que los alumnos expresan cierta voluntad por lograr ser objetivos y honestos en las encuestas. En efecto, la noción de objetividad se encuentra atravesada por la capacidad de discernir entre el desempeño del profesor propiamente dicho y el interés propio por la materia, el divertimento o calificación obtenida:

“Yo intento ser lo más objetiva posible, y como no considerar si me gusto la clase o no, ósea enteramente por diversión o por si me gusta la materia; sino más bien por rasgos del profesor...ósea como da la clase, si pudo llevar bien la clase, también la buena predisposición con los estudiantes. Me parece que las preguntas están bien, y lo único que intento es ser objetiva.” (Estudiante 3, fragmento 10)

“Ahora, es importante esto...ponerlo en el lugar donde va. Yo me quejaría de la nota donde me piden comentarios de más. Pero atendería a cada dimensión que me pregunta la encuesta. No voy a poner que el tipo llega tarde o que tiene mal manejo del conocimiento si no es verdad, solo porque considero que mi examen merecía otra nota o la consigna no fue clara.” (Estudiante 10, fragmento 14)

Asimismo, algunos señalan las dificultades que atraviesan la evaluación de la docencia. En primer lugar, hacen referencia al desconocimiento de la materia, de lo qué es efectivamente ser un buen docente. Al mismo tiempo, y como consecuencia de lo anterior, no se creen del todo idóneos para evaluar ciertas cualidades podrían ser fundamentales en a la docencia:

“Y también cómo establecer parámetros normales, porque capaz que para mi que un docente sea excelente no es lo mismo que para otra persona. Mismo que seas de la carrera de educación influye, no es lo mismo que seas contador” (Estudiante 6, fragmento 7)

“Yo también creo que claramente es importante, pero por ahí, la complejidad entra en que claramente nosotros no somos los más calificados para dar la opinión fundamental...para definirlo. Creo que tiene que ver con que somos los clientes que consumimos su servicio. Hay un tipo de bien que estudiamos en economía que se dice que el cliente no sabe la calidad del bien aún cuando lo está consumiendo...y esto me parece un ejemplo claro de eso. En el que nosotros somos clientes que lo estamos consumiendo al servicio que nos da la educación, pero que tal vez tenemos una percepción sesgada y no es lo que realmente definiría la capacidad de un profesor.” (Estudiante 1, fragmento 1)

Esto nos remite a los escollos que supone la evaluación de la docencia en general. Es decir, si bien cabe reconocer que los estudiantes están limitados en cierto sentido (por ejemplo, para evaluar la profundidad del contenido dictado), la subjetividad de lo que implica ser un excelente docente resulta un problema ineludible, incluso si son otros los evaluadores o si es otro el instrumento utilizado (Cohen en Surgenour, 2013).

Por su parte, de la misma manera que el Estudiante 1, los docentes objetan la capacidad de los alumnos, al menos, de manera parcial, de evaluar su desempeño. Desde su perspectiva su carácter de aprendices les impide darse cuenta si ciertas cosas son apropiadas o no, si los van a ayudar a aprender o no:

“Y bueno, a veces también, ellos...es como la opinión pública, porque ellos no saben si dominas o no dominas un tema, sino si parece que dominas o no dominas un tema.” (Docente 3, fragmento 19)

“El alumno tiene una opinión y es importante saberla, pero el alumno no sabe para donde va el curso, cual es el grado de profundidad del curso, o porque hay algunos temas que se ven o que no se ven...eso el alumno no lo sabe, y por eso ahí puede haber cortocircuitos.” (Docente 1, fragmento 14)

“Me parece que hay una parte que los alumnos pueden ver, pero hay otra parte que los alumnos no pueden ver y que para mí es muy importante...que tiene que ver con que el programa elegido [...] con qué profundidad se dan las cosas...Me parece que cuando uno está aprendiendo algo, todavía es muy difícil entender como se evalué eso que uno esta aprendiendo. Entonces me parece que hay una parte que el alumno todavía no puede ver, porque todavía no lo sabe, porque esta aprendiéndolo: dónde está el punto clave, y qué es lo que uno necesita y cómo se llega a eso. Entonces como toda esa parte todavía no la puede ver, no tiene manera de decir y de darse cuenta si el profesor dio todo lo que necesitaba...” (Docente 5, fragmento 7)

#### 4.3.2 ...y los profesores

“¿Buen docente es el que explica bien? ¿o el que entretiene? ¿o el que...? Cada alumno puede tener expectativas distintas también.” (Director 1, fragmento 22)

“Debe haber una conjunción de factores que deben ir desde la afinidad personal, hasta la capacidad personal, los conocimientos, las dinámicas que genera, los aprendizajes que se llevan los chicos, la experiencia que están viviendo ellos en su vida. Imagino que todos esos factores influyen la decisión que toman sobre la calificación.” (Docente 6, fragmento 15)

Los fragmentos anteriores introducen una cuestión que resulta problemática a la hora de evaluar la docencia que es: *¿qué evaluamos a la hora de evaluar la docencia?* Como hemos mencionado repetidas veces, resulta difícil definir las características de un “buen profesor” (Rueda *et al*, 2010). Así es que existen tantos criterios como sujetos evaluadores posibles, lo que trae como resultado una evaluación un tanto sesgada. A pesar de que la encuesta presenta ciertos ítems que orientan y suponen una evaluación alrededor de ciertos criterios determinados, nos interesa saber cuáles son los factores que los estudiantes tienen en cuenta a la hora de evaluar un docente. Asimismo, nos interesa

indagar acerca de las concepciones que tienen los docentes y directores sobre lo que valoran los estudiantes, pues tenemos la intención de compararlas con la de los estudiantes y evaluar si se encuentran alineadas.

### Compromiso con el aprendizaje de los alumnos

A partir de las entrevistas realizadas, se encontró que las actitudes o las características que valoran los estudiantes de un profesor son diversas. Sin embargo, hay una cuestión que ha sido recurrente y que parece englobar muchas: los alumnos califican mejor a los docentes que manifiestan un expreso *compromiso con el aprendizaje de los alumnos*:

“Creo que hay que diferenciar el talento del esfuerzo acá también. Y que uno a veces no es tan consciente, pero yo trato de siempre pensar 'el profesor puede ser una luz, puede saber explicarte'; pero después el trabajo más allá de dar la clase puede ser vago: me parecía admirable pero a la materia le faltaba ir más allá...y eso yo lo pienso como cuánto se esfuerza el profesor para que te sirva la materia: el orden...darte un programa, hacer guías...que cada cosa esté bien explicada y armada” (Estudiante 1, fragmento 11)

“Yo creo que es por la manera de explicar, en cuanto a los métodos que utiliza para explicar. Hay gente como que aprovecha esa posición y te explican de una manera que...y hay profesores que se matan explicándote de una manera más sencilla y pedagógica. La manera de explicar me parece determinante” (Estudiante 2, fragmento 12)

Un grupo de docentes y directores, al preguntarle por los factores que creen que los alumnos tienen en cuenta hicieron referencia a este compromiso con el aprendizaje que, aunque puede ser una categoría amplia, tiene un componente pasional asociado al esfuerzo y la voluntad del docente por querer que sus alumnos aprendan:

“Yo creo que lo primero que califican es la capacidad y la voluntad del docente de querer realmente que el alumno aprenda. Uno puede saber mucho, y no importarle demasiado si el alumno aprende, entonces es más lo que tiene que hacer el alumno por sí mismo” (Docente 9, fragmento 7)

“Yo creo que tiene que ver con la claridad de la exposición, si preparas o no preparas las clases, si dominas el tema, el trato en el aula y fuera del aula” (Docente 3, fragmento 16)

“Por un lado me parece que el estudiante valora mucho el haber hecho un buen proceso de aprendizaje, y eso incluye desde: el material, todos los medios que pone el profesor, y cómo eso se ve reflejado en el examen, todas las instancias de interacción que sean coherentes. El alumno que arranca sabiendo nada y que dice 'uy tengo todo esto', eso se valora” (Director 5, Fragmento 16)

“En principio tienen que responder a los ítems que la encuesta les pide. Y dentro de cada ítem a los estudiantes les importa como es el docente en la clase, qué tipo de actitud tiene frente a la materia que da, si maneja fluidamente, si tiene conocimiento, si le pone entusiasmo a lo que da. Si está atento a lo que los estudiantes les demandan, si los oyen...eso me parece que son cosas que les importan. Que no viene simplemente a cumplir el horario, sino que aporta algo más...” (Docente 4, fragmento 13)

“Mucho menos por lo que dicen los números, pero por lo que dicen los comentarios abajo, la pasión con que ellos perciben las. El compromiso con la materia, se percibe eso...el esfuerzo la dedicación, la pasión.” (Director 4, fragmento 18)

### Estrategias de enseñanza/clase dinámica/contenido

Como hemos adelantado, relacionado con la voluntad por hacer que la clase comprenda, los estudiantes valoran el saber explicar y las estrategias que utiliza el docente para que el contenido se vuelva interesante o ameno:

“[Si es bueno, o malo, gracioso o aburrido, para mí va en segundo plano (...) Si se entiende lo que explica...va más allá” (Estudiante 3, fragmento 13)

“Después otra cosa que me fijo es si tiene manejo del contenido, del conocimiento. y Si encuentra formas de que trascienda de una explicación de manual” (Estudiante 10, fragmento 12)

“Creo que lo que yo considero es, uno, si el contenido se me hizo interesante...más allá de si me gustara o no, si el docente lo lograba hacer interesante. No sé, me acuerdo que en [...] no me resultaba apasionante, pero al ver que el docente lo daba de una manera interactiva y abierta a que el alumno hable y eso, se me hacía ameno y de hecho lo puntúe bien, a pesar de que la evaluación fue bastante catastrófica. Pero otras materias que a mí me resultaban interesantes, como [...], pero que el pibe las dió súper mal, o muy estructurada, con “cero voz” del alumno. Me fijo esas cosas...si el contenido era interesante, la forma de darlo, o la intención de ver el pensamiento del alumno, o si la evaluación tenía sentido en relación a la cursada y lo que el docente intento promovernos a lo largo de la cursada” (Estudiante 6, fragmento 11)

Algunos docentes y directores, también creen que los estudiantes valoran el dinamismo de la clase, incluso casi al mismo nivel que los contenidos: “se valora tanto la dinámica como los contenidos. Creo que el texto puede ser muy aburrido, pero si la clase es dinámica, la clase maso menos fluye. Tienen un peso relativamente similar.” (Docente 2, fragmento 15).

Sin embargo, una preocupación que parece estar latente en un extenso grupo de profesores y directores es que los estudiantes valoran la personalidad por sobre los contenidos o el aprendizaje:

“Uno ve cómo trabajan otros docentes. Y por eso te digo, a veces hay docentes que son súper responsables, que trabajan un montón, que son organizados para armar el cronograma, que tienen criterio a la hora de pensar los exámenes y de corregir, y que sus clases quizás no son súper entretenidas...y tienen encuestas no tan buenas. Y ves otros docentes que es un despiste, pero tiene clases divertidas y las encuestas son mejores. (...) en ese punto me parece que la encuesta genera distorsión, porque para mí influye mucho en la encuesta, lo divertido lo entretenido del docente y cómo genera el ambiente en la clase y el nivel de exigencia...que si es más bajo a los alumnos le va a parecer mejor en general” (Docente 5, fragmento 27)

“No lo sé, está toda esta onda de que todo debería ser lúdico y demás. Pero todo no se puede hacer lúdico...en algún momento hay que sentarse. Y yo no sé, la verdad es que no lo sé. Tiendo a pensar que chicos de esa edad y demás, si cae alguien que es un buen *entertainer*, lo van a calificar bien, aún cuando 3 años después se den cuenta que no aprendieron nada. Mientras que viene otro, los sienta a hacer ejercicios durante 1.40 dos veces por semana lo van a detestar, y capaz después 5 años después lo aprecian, lo que fue. Quizás estoy equivocado yo, y la gran mayoría de los chicos son recontra maduros y pueden separar lo que es el entretenimiento de lo que es el verdadero contenido y pueden hacer una evaluación muy fina. Pero como te decía, si hay cosas que recién

las podés ver en perspectiva es difícil de hacerlo, entonces el tipo que es *entertainer* aún cuando quizás no tenga el contenido le termina yendo mejor, pero eso no lo sé” (Docente 7, fragmento 3)

Lo expuesto en el fragmento nos remite al estudio del “Dr. Fox”, que supuso demostrar que un profesor altamente carismático y entusiasta puede seducir a la audiencia a otorgar calificaciones altas, aunque poco meritorias. Aunque bien otros estudios como los de Marsh (1978) supieron dar fin a este argumento para desterrar las encuestas como una medida válida del desempeño docente, podemos ver que la idea de que lo que interesante, dinámico y entretenido es indisociable de lo banal está fuertemente instalada.

Ahora, si bien resulta difícil definir que se entiende por personalidad y que resulta natural que se la asocie a la figura del *entertainer*, no logramos detectar, en ningún caso, que los estudiantes hicieran referencia al carisma y a la actitud recreativa del docente. De hecho, sostenemos que la noción de personalidad que orienta a los estudiantes entrevistados en la evaluación docente está mucho más cerca de parecerse al involucramiento argumentado por el Docente 8, que a la noción propuesta en los estudios del “Dr. Fox”:

“Yo creo que valoran mucho, fundamentalmente, y algunos conscientemente y otros inconscientemente, la personalidad del profesor en cuanto a su involucramiento en que ellos aprendan. Si se fijan en lo divertido o no divertido, no lo puedo sacar de lado totalmente. El problema es enseñar cuando no es divertido y al que no le divierte. (...) Si ellos perciben un interés en el profesor por enseñarles.” (Docente 8, fragmento 14)

Por último, fueron muy pocos los casos donde los directores manifestaron creer de manera más determinante que los estudiantes califican a un profesor pura y exclusivamente en base a su personalidad:

“Bueno, la evaluación docente, en general, tiene que ver con el concepto. Y en el concepto entran un montón de variables subjetivas que... si es simpático digamos” (Director 1, fragmento 20)

“Para mí algo muy importante, muy importante, es si les caes bien. Mi sensación con las encuestas, y eso no tiene tanto que ver con si sos simpático [...], con llegar horario, estar a tiempo... cosas que tienen que ver con construir la relación. Pero si uno le cae bien a los estudiantes, sea por los motivos que sea [...] casi que te perdonan cualquier cosa [...] si no te llevas bien con los estudiantes por los motivos que sea, te lo hacen notar” (Director 2, fragmento 17)

En otras palabras, estos fragmentos evidencian como algunos directores creen que los alumnos son mucho más sensibles al carisma de un profesor, que a las propuestas realmente formativas. En este marco, sería razonable que los actores con concepciones similares respecto a la actitud y al juicio de los estudiantes, consideraran que las encuestas no son confiables por encontrarse altamente sesgadas por la personalidad del profesor (Abrami, 1982).

Trato

En un segundo plano aparecieron cuestiones ligadas al trato de los docentes. En este caso, parece haber concordancia, pues tanto los alumnos como los docentes lo identifican como un factor importante a considerar:

“Me parece que el trato con los alumnos afectivo. Si hay maltratos o no, me parece algo importante. Después algo muy relevante es la concepción pedagógica que tiene, a los que él no está pudiendo llegar. Qué piensa de los que les va mal...cuál es su trato...si busca encontrar una forma para que lo entiendan, eso me parece importante.” (Estudiante 10, fragmento 10)

“Yo no creo que sea algo de segundo plano...la violencia del profesor me parece súper relevante, y me parece que lo teneés que poder penalizar más allá de que sea brillante” (Estudiante 1, Fragmento 12)

“En las tutoriales si me enseñaron, se sentaron conmigo...y ese trato más afín los alumnos lo valoran bastante y lo tienen en cuenta” (Docente 1, fragmento 16)

"Bueno a ver...evalúan siempre si hay un trato cordial o no, eso lo tienen muy en cuenta. Opinan si saben, o si aprendieron o no, y si lo que ellos aprendieron les va a servir o no. Eso lo consideran bastante" (Docente 1, fragmento 15)

#### 4.3.4 ...las características del curso

Otra cuestión que se ha indagado ha sido las características que se encuentran asociadas a la naturaleza de los cursos y como eso impacta en la evaluación de la docencia. Entendiendo que la encuesta es estandarizada y que en muchos casos se utiliza con fines comparativos, muchos estudios han sugerido la disconformidad de los docentes de utilizar los mismos criterios de evaluación para cursos de características muy distintas, produciéndose así ciertas desventajas.

Al explorar este tema dimos con algunas variables que los docentes y directores consideran que, de no ser contempladas, impactan injustamente en la evaluación del desempeño de un docente:

“Si sí, no es lo mismo un curso muy temeroso muy atemorizado, que un curso mas chico o un curso más avanzado.” (Director 3, fragmento 16)

“El tamaño del curso y la obligatoriedad para mi son factores bastante decisivos. El ciclo de fundamentos a los de contabilidad les debe parecer horroroso. Las materias de arte literatura evaluada por chicos de contabilidad...Pero la obligatoriedad de estas materias eventualmente es un factor que disminuye la correcta evaluación, y que sean muchos estudiantes tampoco es fácil” (Director 4, fragmento 21)

“Matemática para un tipo que está haciendo humanidades le va a resultar árida, le va a ser cuesta arriba. Si logró aprender algo ese chico, y si pudo poner algo positivo vale 3 veces más que si es un estudiante de humanidades lo que diga de filosofía. Eso tiene que estar contemplado. Hay como un ranking de arranque en la cantidad y en la disciplina” (Docente 9, fragmento 14)

“Ósea yo digo que no puede ser lo mismo la encuesta de un curso de muchos estudiantes. Seguro hay que tomar en cuenta si el curso es del ciclo de fundamentos, si es específico de tu carrera...porque un curso específico de tu carrera, de último año, con 10 estudiantes, la encuesta no va a decir mucho. Si la encuesta ahí es mala, seguro que dice mucho del docente, ahora...si vos agarras un curso de primer año, con distintas carreras que todavía no entienden porque están, y

que les resulta difícil, y bueno me parece que se tiene que mirar de otra manera la encuesta” (Docente 5, fragmento 28)

Como podemos observar, las características que aparecen más recurrentemente son el tamaño del curso, es decir, si un curso tiene muchos o pocos estudiantes y el área o disciplina de interés primario asociado a la obligatoriedad del curso.

### Tamaño de la clase

En relación al tamaño de la clase, los docentes creen que quienes dictan en clases muy numerosas se encuentran en desventaja frente a los que dictan cursos con pocos estudiantes, porque no es lo mismo el tipo de atención o la posibilidad de generar *rapport* que se habilita en clases pequeñas que en clases grandes:

“Definitivamente. Yo doy [...], un curso que lo toman 10/11 y doy [...] que es un curso de 100/90/80, y me es imposible pasar el nueve y pico; la dispersión es muy grande y siempre hay alguien que se queda atrás o heridos en una materia tan grande” (Docente 3, fragmento 21)

“Muy importante la cantidad de alumnos que pueda llegar a tener un docente. El tema hasta 30 un profe puede atenderlos...en cuanto salimos de eso, va bajando la calidad necesariamente. Es otra atención.” (Docente 8, fragmento 24)

“Porque la capacidad de atención es menor, entonces el número no ayuda. Y si finalmente ese esfuerzo que tiene que hacer el profesor tiene que ser el doble, y a veces lo logra y a veces no. Pero hasta qué punto pesó todo lo que está alrededor...Ese es un problema muy importante a la hora de comparar cursos. Sabes que, un curso de doce les hago bailar la chacarera. Pero tengo 70, y tuve 100, y con una materia poco afín.” (Docente 9, fragmento 13)

“En las tutoriales cuando son menos alumnos, como el docente está más cerca y tiene un trato más personal...entonces las encuestas les favorecen porque tienen mayor llegada. En cambio, en una clase magistral de 60 personas, al no tener tanta llegada, entonces no tenes tanta afinidad con el docente de la magistral, entonces por ahí te juzgan de otra forma [...] los alumnos generalmente valoran mucho la llegada, el trato.” (Docente 1, fragmento 20)

“Si. Clases más grandes son más difíciles de enseñar generalmente, y eso hace que la calidad de la docencia sea un poco menor, porque le prestas menos atención al estudiante. [...] entonces eso puede afectar la calidad de la enseñanza.” (Director 2, fragmento 20)

“Si, si, por supuesto que no es lo mismo un curso que tiene 50 alumnos que tiene 15. [...]” (Director 3, fragmento 3)

Asimismo, vemos como los docentes y directores expresan cierta preocupación por la disminución de la calidad de la docencia en aquellos cursos con muchos alumnos, íntimamente asociada a la atención que un profesor le puede dirigir a algunos alumnos, pero que se limita cuando exceden los 30. En relación a esto, Marsh (1987) ha sugerido que el impacto de esta variable se debe a algunas ciertas dimensiones de la enseñanza efectiva, como la interacción y el vínculo con los alumnos. La observación de Abrami (1987) parece sensata en este punto pues ha expresado que, si bien no puede ser un

argumento para invalidar las encuestas, sería prudente que se tuvieran este tipo de variables en cuenta a la hora de orientar acciones *high-stake*.

### Área de interés primaria y electividad

Otra inquietud que fue repetida en diversidad ocasiones fue la comparación entre cursos electivos y no electivos, pues según los docentes “no es lo mismo estudiar lo que te dicen que tenés que estudiar que lo que te gusta y lo que elegís vos” (Docente 8, fragmento 20). De hecho, muchos de ellos aluden a la diferencia de desempeñarse en cursos donde todos, o la gran mayoría de los alumnos, estudian carreras afines a los temas que se dictan o donde, por el contrario, casi todos los estudiantes se encuentran allí por obligación. A modo de ejemplo: “No es lo mismo enseñar una materia para gente que quiere estudiar Ciencias Sociales, que para gente que quiere hacer economía, una materia como matemática” (Docente 1, fragmento 21). Otros, de forma similar, expresan:

“A veces no tiene en cuenta la diferencia entre las materias y los alumnos de las materias. No es lo mismo en mi caso, una materia que doy para el ciclo de fundamentos, para una mayoritaria proporción de alumnos de negocios digitales y de administración, que para alumnos de humanidades. No es lo mismo lo que va a pasar en una materia a la que a mí se me impone que da clases magistrales virtuales, que a una que no tengo.” (Docente 4, fragmento 10)

“En los casos de grado, bueno, la materia no me interesa entonces veo todo negro. Entonces puede tener un sesgo no deseado en grado, o al menos en los primeros años” (Docente 10, fragmento 6)

Al mismo tiempo, la experiencia de un docente parece confirmar las creencias expresadas con antelación:

“Nosotros notamos algo muy interesante. Antes era obligatoria para todos, y ahora es optativa para las que lo quieran hacer. Y nos mejoró muchísimo la evaluación. Entonces yendo a tu pregunta, me parece que si hay un factor que está influyendo en la evaluación.” (Docente 3, 22)

Si bien las creencias de los docentes encuentran su correlato en algunas investigaciones que han encontrado relaciones positivas entre las encuestas y la obligatoriedad del curso (Feldman, 1978), también ha habido estudios que han concluido que no son significativas (McKeachie, 1979). Las creencias de muchos otros se identifican más bien con estos otros estudios, pues han argumentado que no es la obligatoriedad o la no obligatoriedad lo que se pone en juego, sino las expectativas de los estudiantes:

“Es verdad que las materias que son obligatorias generalmente reciben puntajes más bajos, pero es verdad que, si el alumno elige hacer una materia, también tienen más expectativas ... ‘yo elegí esta materia y no resultó como quería’, es distinto a ‘bueno la tengo que hacer y la tengo que hacer’ (Director 2, fragmento 21)

“No necesariamente, porque puede ser que yo quiero hacer esta materia porque me gusta y de repente me doy cuenta que no me gustaba, y es una frustración. Entonces, este profesor me sacó

el interés por esta materia que a mí me gustaba y es un desastre. O quizás es al revés. Uy esta materia que pensé que era un embole, uh ahora entiendo porque iba esto [...] no creo que la obligatoriedad o la no obligatoriedad influya tanto...puede influir, pero no por mejor o peor...puede influir, pero de otra forma” (Docente 1, fragmento 22)

En otras palabras, para estos docentes y directores la electividad se pone en juego, pero a través de un mecanismo distinto con resultados contrapuestos. En este caso, las materias electivas corren el riesgo de no cumplir con las expectativas iniciales de quienes se inscribieron, produciéndose así frustración e implicando desventaja frente a los cursos obligatorios, donde las expectativas son nulas o no son tales.

#### 4.3.5 ...y la administración del cuestionario

Las características asociadas a la administración del cuestionario fueron un tópico recurrente para explicar las actitudes adoptadas a la hora de completar la encuesta. Una de las principales cuestiones que se puso de manifiesto en las entrevistas es la preocupación por el momento en que las encuestas son administradas y el mecanismo de doble ciego<sup>7</sup> que se utiliza como incentivo para que, por un lado, los estudiantes completen las encuestas y, por el otro, para que los docentes carguen las notas. De acuerdo con lo expuesto por los estudiantes, este mecanismo podría atentar contra la responsabilidad para con las encuestas, pues la ansiedad por ver las calificaciones obtenidas podría superar la intención de tomarse esta instancia con el compromiso que merece:

“Siento que puede haber un montón de personas que hacen esto, con el incentivo de lo haces porque no te dejas ver la nota, entonces lo haces ‘ta ta ta’ quiero ver mi nota. Entonces, muchas veces la quiere ver y siento que es una instancia importante, y eso puede jugar muy en contra o muy a favor y no define lo bueno o lo mala que esta una materia” (Estudiante 2, fragmento 11)

“También hay que tener en cuenta, y lo admito, que muchas veces te ponen esto de hacer la encuesta porque sino no te dejamos ver las notas. Y se de muchos casos en donde se hacen volando, donde no se toma el tiempo necesario para hacerlo, para poder ver la nota. Porque es una condición para hacerlo, y no creo que se tome el tiempo.” (Estudiante 9, fragmento 6)

“No me parece que puedas confiar en el 60, 70% de la facultad y en su criterio para las encuestas, no. Más que nada porque estás en medio de finales, y querés ver las notas.” (Estudiante 4, fragmento 11)

“En este sentido, hipotetizo que las tres consignas a escribir no deben ser respondidas con la dedicación, el compromiso que, quizá, sería bueno que tengan. En muchos de los casos, estamos ansiosos por ver la nota y los completamos al ‘tun, tun’.” (Estudiante 7, fragmento 5)

<sup>7</sup> El acceso a las calificaciones finales se utiliza como incentivo para que los estudiantes completen las encuestas, es decir, solo pueden acceder a las notas cuando han completado todas las encuestas. De la misma forma, los docentes solo pueden acceder a las encuestas una vez que han cargado las calificaciones.

De la misma manera, un director parece advertir la potencial limitación que supone utilizar como requisito necesario la encuesta para conocer la calificación que obtuviste en el curso:

“No lo sé, yo supongo que la mayoría...a ver no es un espacio donde los pibes le pongan buena dedicación, me parece que más bien hacen ‘tiki tiki tiki’. Me parece, salvo algunas ocasiones, me parece que todo el mundo se quiere sacar de encima la evaluación para llegar a la nota.” (Director 4, fragmento 19)

Esto se pone en tensión con la hipótesis antedicha, que sugiere que la actitud de los estudiantes se ve comprometida o influenciada por la calificación que esperan recibir. Bajo esta hipótesis, el mecanismo parecería ser acertado, pues no hay forma de que los estudiantes conozcan su nota. Sin embargo, es limitado por eso mismo, pues aún no puede escapar de las actitudes suscitadas por la incertidumbre de la calificación.

Por otra parte, otro director señala como la simultaneidad en la administración de las encuestas de todos los cursos atenta contra la dedicación que un estudiante podría tener para responderla:

“Una cosa es cuando respondes una, y estás re comprometida y tenés el tiempo para hacerla. Pero otra cosa es cuando estás respondiendo 4 o 5 materias, que me imagino que cuando llegas a la quinta no es lo mismo que la primera” (Director 5, fragmento 14)

En este caso, la preocupación pasa por cómo impacta en los estudiantes el desgano producido por la obligación de completar varios cuestionarios en un periodo de tiempo relativamente corto, a la hora de calificar a los docentes.

Estas preocupaciones también parecen estar latentes en estudiantes que, por ejemplo, temen que las encuestas no sean anónimas: “siempre temo que no sean anónimas y que el profesor se entere de los nombres de cada uno de los encuestados” (Estudiante 7, fragmento 3). En relación a este punto, si bien no fue algo cuestionado, resulta interesante exponer la particular postura de un docente que propuso que la encuesta deje de ser anónima con la intención de combatir la cultura de la crítica no fundada, bajo el fundamento de que “es una cuestión de responsabilidad” (Docente 8, fragmento 4). Asimismo, otro docente sugiere que el anonimato de las encuestas habilita el *troleos*, es decir, que los estudiantes adquieran posturas provocadoras u ofensivas:

“Hay una cultura del *troleo* que se ha extendido a todo. Con eso que podemos hacer...Bueno lo mismo pasa en las encuestas, porque sobre todo al no poder ser identificados se sienten con esa libertad de la actitud del troll en las redes, me parece que eso está contaminando un poco a los chicos” (Docente 3, fragmento 7)

<sup>8</sup> Proviene del mundo digital, y refiere a la actitud de intervenir de forma ofensiva o perniciosa, con el ánimo de hacer fracasar algo (Fundeu BBVA, 2015).

En otras palabras, si bien el anonimato es una condición que se recomienda mantener (Centra, 1993) debe reconocerse que en algunos casos puede acabar provocando actitudes indeseadas y ofensivas para los docentes, potenciadas por el sentimiento de impunidad de los estudiantes.



Universidad de  
**San Andrés**

## CONCLUSIONES Y REFLEXIONES FINALES

En este trabajo nos propusimos indagar sobre las concepciones que los estudiantes, los docentes y los directores de carrera de la Universidad de San Andrés, construyen alrededor de la encuesta, como instrumento de evaluación del desempeño docente.

### Conclusiones

Preguntarnos por las concepciones que construyen a alrededor de la encuesta, implicó preguntarnos por el por qué y el para qué de la encuesta. Asimismo, conllevó preguntarse cuán confiables son resultados producidos por los CEDA para evaluar el desempeño docente.

#### *Del propósito...*

A partir del análisis llevado a cabo podemos señalar que para los directores de carrera y docentes el propósito de la evaluación docente a través de los cuestionarios estudiantiles resulta ambiguo. Por un lado, creen que se encuentra marcado por la intención de ser una instancia de control del desempeño docente, es decir, para detectar problemas relativos a su performance. Por el otro, consideran que las encuestas tienen la finalidad de recoger información útil para la mejora de la práctica pedagógica. Asimismo, señalan el potencial límite que supone para su potencialidad formativa, la burocratización de esta instancia.

Con relación a las creencias en torno a la heteroevaluación, podemos argumentar que, en su mayoría, tanto docentes, como directores, creen que la voz de los estudiantes es importante. En concreto, las posturas van desde los que consideran que es sumamente esencial, hasta lo que adoptan posturas más críticas que creen que si bien es un factor importante, debe entenderse que es imperfecto y que es uno de los tantos factores a considerar. A pesar de su inminente imprecisión, todos los entrevistados consideran que su opinión resulta indispensable para evaluar el desempeño de un docente, ya que son ellos los protagonistas de la práctica de enseñanza y aprendizaje.

Por otra parte, a partir de las entrevistas realizadas pudimos constatar que casi ningún actor conoce el uso que la institución hace de los resultados provistos por las encuestas. En los casos de docentes y directores donde se evidencia un ligero

conocimiento de su uso, se hace referencia a cuestiones anecdóticas. Cabe resaltar que algunos estudiantes y docentes expresan cierta preocupación respecto a la falta de contextualización o información provista sobre el propósito de esta instancia.

*Del uso...*

Por lo que refiere al uso de las encuestas, encontramos que la dimensión formativa es mucho más aceptada que la dimensión sumativa. En otras palabras, encontramos que los docentes encuentran que la encuesta les es provechosa para mejorar sus prácticas de enseñanza, pero que no debería orientar acciones *high-stake*.

Como hemos mencionado, los docentes señalan que las encuestas son un instrumento de gran utilidad para mejorar su desempeño, en tanto les permiten identificar que cosas les gustaron o no les gustaron, les sirvieron o no le sirvieron a la totalidad de los estudiantes y, consiguientemente, tomar medidas que resulten pertinentes. De hecho, varios docentes manifestaron realizar cambios a partir de la devolución de sus alumnos, que van desde la dinámica de la clase, hasta el tiempo dedicado a cada tema, entre otros. Sin embargo, el criterio para determinar la introducción de una modificación –además de la recurrencia, es que dicha modificación se encuentre alineada con su propia convicción acerca de lo que es (o no) beneficioso para el aprendizaje.

Las creencias de los estudiantes respecto a la utilidad formativa de la encuesta, distan de lo que expresan docentes y directores. A partir de lo relevado en el análisis, podemos concluir que los estudiantes tienen cierta desconfianza hacia los profesores, en la medida que no logran percibir variaciones en el dictado de los cursos a partir de sus devoluciones. En este punto, señalan casos de docentes a los que se les ha criticado sistemáticamente lo mismo a lo largo de los años y que no han cambiado. Además, piensan que, de cambiar algo, los docentes no modifican cuestiones estructurales sino más bien secundarias. Por último, los estudiantes creen que a lo que los docentes le prestan más atención es a los comentarios. Esto parece ser efectivamente así, pues los profesores se han mostrado particularmente positivos frente a la utilidad de los comentarios o el espacio de sugerencias, especialmente debido a su especificidad.

También vale la pena recuperar aquellas características de los CEDA y su administración que limitan su oportunidad formativa. Primero, los docentes consideran que el momento de administración del cuestionario es problemático, pues no admite la posibilidad de cambio o de respuesta. Segundo, han señalado cuestiones ligadas al

cuestionario en sí, como el contenido de las preguntas, que no permiten recoger información del todo relevante para mejorar su práctica.

Por otra parte, como hemos anticipado, los docentes ofrecen resistencias frente a la utilización de sus resultados para orientar acciones ligadas al salario, la promoción, la permanencia y el reconocimiento de los docentes. A pesar de que no existe un conocimiento certero de la utilización del instrumento con fines sumativos en la Universidad, la postura que ha prevalecido en su mayoría ha sido la de aceptar su uso, pero bajo la condición de que no sean el factor único y decisivo. Además, hay un grupo de docentes cuya postura es más taxativa, que consideran que bajo ningún punto de vista los CEDA son un criterio relevante para orientar decisiones *high-stake*. En cuanto a las creencias de los directores de carrera, se encuentran más bien alineadas con las del primer grupo de docentes, es decir, consideran que no debería ser el único, sino un indicador más a la hora de evaluar el desempeño docente. Si bien consideran que es valioso conocer que opinan los estudiantes, sostienen que es un diagnóstico incompleto signado por las limitaciones del instrumento.

Por último, respecto a las concepciones de los estudiantes, han sido variadas. Si bien hay quienes manifiestan que las encuestas deben ser totalmente tenidas en cuenta para orientar acciones de alto impacto, lo que predomina son posturas más conciliadoras, que entienden que la encuesta es un factor a tener en cuenta siempre y cuando se instale una cultura de la evaluación en la Universidad y se haga una lectura adecuada de los datos. Ahora bien, pese a los ideales que los estudiantes pudieran tener respecto al uso de las encuestas, lo que la experiencia les ha demostrado es que las encuestas en la Universidad tienen impacto dependiendo de quién sea el profesor en cuestión. En otras palabras, los estudiantes creen que, a pesar de los fuertes llamados de atención plasmados en las encuestas, cuando se trata de profesores de alta jerarquía (“eminencias”), permanecen impunes.

Otro tema que hemos indagado relacionado a los potenciales usos de la encuesta, ha sido las consecuencias que se desprenden de utilizarla como una instancia decisiva y de responsabilización individual de los docentes. Aunque la encuesta no parece utilizarse de esta manera en la Universidad, los docentes y directores expresan cierta preocupación frente a los desenlaces que esto podría tener. Dichos desenlaces refieren principalmente a los incentivos que podrían tener los docentes, de bajar los estándares de evaluación y de adoptar un comportamiento demagógico respecto a la demanda de los estudiantes, si su éxito o fracaso dependiera únicamente de la calificación obtenida en la encuesta.

*De la confiabilidad...*

La tercera cuestión que indagamos son las concepciones que los sujetos construyen acerca de la confiabilidad del instrumento para evaluar el ejercicio de los docentes. De esta forma, el primer tópico que emerge abarca las actitudes que adoptan los estudiantes a la hora de evaluar a sus profesores, el impacto que tiene la calificación recibida sobre la evaluación del docente y del curso y las creencias alrededor de su capacidad para evaluar la enseñanza.

Conforme a las entrevistas que realizamos, podemos decir que las creencias de docentes y directores en relación a la seriedad y el compromiso con la los estudiantes asumen la evaluación han sido de lo más variadas. Por un lado, podemos encontrar un grupo que percibe que los estudiantes son honestos y ponen mucho empeño cuando realizan las encuestas. De manera totalmente opuesta, existe otro grupo que encuentra que los estudiantes son deshonestos, es decir, que mienten y ponen cualquier cosa en las encuestas. En medio, el grueso de los entrevistados no se encuentra identificado con ninguna de las posiciones anteriores y considera que en general los alumnos se toman las encuestas a conciencia, aunque creen que existe un grupo que de seguro no le pone tanta dedicación. Las actitudes de los estudiantes se corresponden más con las creencias de este último grupo. En relación a esto, un hallazgo interesante es que la actitud con la que los estudiantes asumen la evaluación depende de el carácter extremista de la experiencia, es decir, se toman la encuesta con más compromiso y mayor grado de rigurosidad cuando el docente fue sobresaliente o, por el contrario, fue pésimo. Los directores y docentes advierten lo anterior, pues creen que quienes se toman el tiempo de completarla con detenimiento son aquellos estudiantes que se encuentra enojados.

Otra cuestión que se abordó son las creencias en torno a la relación que existe entre la encuesta y la calificación que obtuvieron los alumnos. Para la gran mayoría de los docentes y directores, los estudiantes se encuentran muy condicionados por la calificación que recibir o esperan recibir a la hora de realizar el cuestionario. En otras palabras, creen que, si a los estudiantes les fue bien, entonces van a ser positivos en sus evaluaciones. Si, por el contrario, les fue mal, van a adoptar una actitud vengativa en la encuesta. Por su parte, las actitudes de los estudiantes son variadas. Aunque en su gran mayoría reconocen esta relación y la dificultad que entraña diferencias sus sentimientos de la evaluación del docente, muchos niegan comportarse así. Quienes no lo niegan, al menos manifiestan intentar ser objetivos y que la calificación que obtuvieron

no nuble su juicio. Una cuestión interesante es que generalmente los estudiantes entrevistados atribuyeron este comportamiento vengativo a sus compañeros, es decir que creen que el resto sí se comporta así.

Por último, en relación a las características propias de los estudiantes, se indagó sobre la capacidad que tienen, o no, de evaluar la enseñanza. Los estudiantes consideran que sí se encuentran capacitados, en tanto entienden que pueden mostrarse objetivos a la hora de evaluar el docente. Es decir, que pueden separar el interés propio y la calificación obtenida en el curso, del desempeño del profesor propiamente dicho. Sin embargo, al mismo tiempo han señalado que su capacidad de evaluar es también deficiente, ya que no pueden terminar de identificar y evaluar cualidades que podrían ser fundamentales para una buena docencia. Por su parte, las creencias de los docentes se encuentran muy ligadas a ese último argumento, pues consideran que la capacidad de los estudiantes es parcial, pues su carácter de aprendices les impide darse cuenta qué estrategias son apropiadas o no, que temas son necesarios o no, entre otros.

El segundo tópico tiene que ver con las creencias alrededor de los factores que los estudiantes tienen en cuenta a la hora de evaluar, más precisamente, cuáles son los atributos que tienen en cuenta a la hora de calificar un profesor positivamente. Uno de los aspectos que ha sido valorado fuertemente por los estudiantes es el compromiso que los docentes manifiestan para con el aprendizaje de sus alumnos. Es decir, hasta qué punto se esfuerzan por hacer que entiendan y que la materia se vuelva interesante. Muchos docentes y directores perciben esto que los estudiantes manifiestan, pues consideran que los estudiantes valoran la voluntad del docente por hacer que sus alumnos aprendan. Otro aspecto que los estudiantes valoran son las estrategias de enseñanza que utiliza el docente para hacer que ellos aprendan y para que la clase sea más dinámica. Sin embargo, mientras que los estudiantes asocian las estrategias de enseñanza y el dinamismo a cuestiones que favorecen su aprendizaje, los docentes y directores asocian esta valoración a cualidades personales de los docentes como *entertainers*. Como resultado, algunos docentes creen que los estudiantes valoran la personalidad por sobre la capacidad de sus profesores de hacerlos aprender, lo que produce en algún punto comienzan a desconfiar de la encuesta. Además, otro aspecto que ha sido valorado es el trato de los docentes. Es decir, cuando los alumnos perciben que el profesor tiene un trato ameno con los alumnos, lo suelen calificar positivamente. De lo contrario, el costo es muy alto. Las concepciones de los docentes no se alejan de esto, pues varios piensan que los alumnos tienen en cuenta este factor.

El tercer tópico se relaciona con creencias alrededor de las características asociadas a la naturaleza del curso y que pueden impactar en la evaluación del docente. La disconformidad surge en tanto son cosas que exceden su desempeño, pero que aún así lo acaban perjudicando. A partir de las entrevistas realizadas podemos decir que son dos las características que docentes y directores encuentran desventajosas, a saber: el tamaño y la electividad del curso. En relación a lo primero, docentes y directores creen que quienes dictan en clases numerosas se encuentran en desventaja frente a los cursos chicos, porque no la misma atención o la posibilidad de generar *rappport* que se habilita en clases pequeñas que en clases grandes. Respecto a la electividad del curso, es decir, a si los alumnos han elegido el curso o lo tienen que hacer por obligación, varios docentes y directores creen que tener alumnos que están ahí obligatoriamente es desventajoso porque resulta mucho más difícil hacer que la temática sea de su interés. Con todo, algunos docentes se han mostrado contrarios a este último argumento, pues creen que es una cuestión ligada a las expectativas. En otras palabras, no tiene que ver si la eligió o no, sino de las expectativas que el estudiante había formado antes de empezar el curso. En este escenario, un estudiante que eligió una materia puede verse fuertemente frustrado y calificar negativamente y un estudiante que curso lo materia obligatoriamente puede verse positivamente sorprendido, y otorgar una muy buena calificación.

Para finalizar, se abordaron cuestiones relacionadas a la administración del cuestionario. Una de las preocupaciones principales está asociada al mecanismo que se utiliza como incentivo para que sea respondida en tiempo y forma y al momento en que las encuestas son administradas. De acuerdo con lo recogido en las entrevistas, los estudiantes creen que tener que completar los cuestionarios como requisito excluyente para poder acceder a sus calificaciones, atenta contra la dedicación que esta instancia requiere. Es decir que, en muchos casos, la ansiedad por ver la calificación que obtuvieron en el curso nubla la intención de tomarse con cierto grado de compromiso la encuesta. Por otra parte, en relación al momento de administración del cuestionario, algunos directores creen que administrar todas las encuestas de manera simultánea, produce cierto desgano en los estudiantes, lo que trae aparejado una distorsión en los criterios y la calidad de la evaluación. Por último, algunos estudiantes expresan cierta desconfianza frente al anonimato de la encuesta. En este punto, si bien no fue una postura general, algunos docentes creen que la condición de anonimidad incita a los estudiantes a comportarse de manera ofensiva, marcada por un cierto grado de impunidad que no resulta beneficioso esta instancia evaluativa.

## Reflexiones finales

Habiendo expuesto ya las principales conclusiones a las que hemos arribado a partir del trabajo de campo, creemos resulta importante compartir las múltiples reflexiones que la elaboración de esta tesina, desde su inicio hasta su culminación, a suscitado. Es así que, para dar cierre a este proceso, plasmamos a continuación algunas de ellas.

La primera, refiere a la dificultad que entraña evaluar una práctica tan compleja como es la docencia (Murphy, MacLaren y Flynn en Surgenour 2013). Como bien nos advertían Rueda y Cohen (2010; en Surgenour, 2013), no existe un consenso de lo que es una buena enseñanza, ni mucho menos esta definidas cuales son las cualidades que un “buen docente” debe tener. Entonces ¿qué se evalúa cuando se evalúa el desempeño a través de las encuestas? Como hemos evidenciado, las cualidades que valoran los estudiantes, muchas veces distan de aquellas cosas que consideran valiosas los docentes. Incluso, quienes toman las decisiones podrían valorar cosas distintas y esto supondría un problema. Si bien los ítems, es decir, las preguntas de los cuestionarios, establecen un criterio y dan un mensaje sobre aquello que es valorado, sostenemos que para que la evaluación sea lo más justa posible, todos los actores involucrados deben conocer los criterios que se han de considerar para evaluar y ser evaluados, según corresponda.

La segunda, íntimamente relacionada a la anterior, se desprende el sentido mismo de la evaluación docente. La evaluación es un juicio de valor y como tal, la práctica evaluativa resulta indisociable de un propósito o una finalidad (Vain, 1998). Como hemos podido argumentar, para ninguno de los sujetos entrevistados resulta clara cuál es la finalidad de la encuesta en la Universidad. Más aún, pudimos constatar que la carencia de sentido en la práctica resulta perjudicial para la práctica misma, pues dar lugar a la desconfianza y especulaciones. Al mismo tiempo, genera una suerte de desgano, propio de una actividad cuyo impacto se encuentra signado por la incertidumbre: ¿por qué un estudiante se tomaría el tiempo de hacer una evaluación de manera comprometida si cree que el profesor no va a realizar ningún cambio al respecto?

Actualmente, los estudiantes facilitan retroalimentación significativa solo en los casos donde desean manifestar lo excelente o lo pésimo que fue un curso o un docente, pero en el resto de los casos es una instancia que se toman “así nomás”. Es en este punto donde resulta elemental dar a conocer para qué se evalúa a los docentes y es la institución la que resulta responsable de llevar a cabo las acciones y las políticas necesarias para

lograr este cometido. Asimismo, debe esforzarse por instaurar, al menos paulatinamente, una cultura de la evaluación que reconozca su dimensión formativa, para que de esta forma las encuestas se vuelvan una oportunidad para promover la mejora al interior de la Universidad (Fernández Lamarra, 2010).

La tercera reflexión, refiere a las tensiones y limitaciones que pueden surgir cuando un mismo instrumento es utilizado con finalidades disímiles. A partir de las entrevistas realizadas, pudimos dar cuenta que la encuesta en la Universidad tiene funciones tanto sumativas como formativas, no obstante, no termina por cumplir ninguna de las dos funciones satisfactoriamente. Desde nuestro punto de vista esto se debe, principalmente, a que la disconformidad que tienen los docentes frente al uso del instrumento con fines sumativos, termina impactando en la percepción que tienen la instancia en su totalidad, produciendo que su dimensión formativa sea desatendida. A modo de ejemplo, sostenemos que un docente podría ser receptivo a las sugerencias de los alumnos, de no ser una instancia sumativa (con la que no esta de acuerdo) el contexto en el que tiene lugar esa sugerencia.

Por otra parte, la información provista por la encuesta resulta insatisfactoria para orientar decisiones *high-stake*, en tanto ha quedado evidenciado que resulta ser un indicador un tanto imperfecto como para orientar acciones de reprimenda o reconocimiento de los docentes de la Universidad. Valerse solamente de este indicador sin contemplar las múltiples variables que hemos mencionado en el desarrollo de la investigación, creemos, sería algo imprudente. Además, se correría el riesgo de producir incentivos algo perversos para los docentes, que distorsionarían su comportamiento con tal de conseguir mejor calificación en las encuestas: ¿Qué debería hacer un docente si su trabajo dependiera de la calificación de los estudiantes? ¿ser demagógico? ¿bajar los estándares? ¿ser fiel a sus convicciones pedagógicas y poner en riesgo su trabajo?

Con esto, no queremos decir que los CEDA deberían ser eliminados, sino que deben reconocerse sus limitaciones y actuar en consecuencia. La evaluación docente es un fenómeno complejo que requiere no de una, sino un conjunto de estrategias que permitan abordar a los diferentes integrantes del sistema educativo, es decir, pares de los docentes, directivos, incluso a los docentes mismo. La selección de un actor determinado supone asumir ciertos aportes y limitaciones (Vain, 1998), de ahí que el modelo de evaluación docente desde la perspectiva de los estudiantes resulta necesario, pero insuficiente, debiendo integrarse con otras variables e indicadores. Las sugerencias de Chan *et al* (2014) resultan valiosas este punto, pues proponen que los CEDA deben ser

parte de un plan estratégico que triangule evidencia desde diversas perspectivas para la mejora de los procesos de enseñanza y aprendizaje.

La cuarta reflexión parte de la preocupación que suscita la decisión de instaurar un instrumento de evaluación en una organización sin ningún tipo de noción, al menos aparente, de las opiniones y creencias de los evaluadores y los evaluados. Como hemos visto, las concepciones que un individuo pudiera tener, determinan en gran medida su forma de actuar y la intencionalidad de dichos actos (Hidalgo y Murillo, 2017). En este caso, la desconfianza los CEDA puede limitar fuertemente su utilidad, mismo el desconocimiento del propósito producir fuertes desencuentros al interior de la organización. De ahí la importancia de pensar que:

los procesos de evaluación no se llevan a cabo en el marco de un vacío de subjetividades, sino que se insertan en un escenario con actores que manifiestan (activa o pasivamente) enfoques diferentes y muchas veces contradictorios en torno a los procesos evaluatorios (Claverie, González y Pérez, 2008: 162).

Por otro lado, a pesar de que ha sido ampliamente debatido en el campo de educación superior, creemos que las polémicas que lo rodean han hecho que se vuelva un tema tabú en la Universidad. Sin embargo, a partir de las impresiones que recogimos en nuestro trabajo de campo, nos atrevemos a decir que los sujetos entrevistados no solo estaban dispuestos a hablar del tema, sino que deseaban hacerlo. Habilitar canales o espacios de discusión podría ser una alternativa valiosa para captar los diversos enfoques acerca de las encuestas. Conocer las concepciones de los distintos actores, insistimos, podría ser un factor decisivo para la potencial mejora de la Universidad.

Por último, la quinta reflexión parte de la preocupación que despierta el escaso desarrollo de investigaciones científicas que aborden esta temática en la región –y en Argentina– a pesar de que los CEDA se han vuelto la estrategia más predominante a la hora de evaluar la docencia en las instituciones de educación superior. Las investigaciones que dieron pie a este trabajo se han desarrollado mayoritariamente en países e instituciones donde los sistemas de evaluación docente ya se encuentran altamente instalados. Por lo cual, si bien hemos encontrado puntos de acuerdo entre sus hallazgos y los nuestros, creemos es necesario llevar adelante nuevas investigaciones que tengan en consideración las particularidades propias del contexto nacional.

## BIBLIOGRAFÍA

Abedin, N. F. Z., Taib, J. M., & Jamil, H. M. T. (2014). Comparative study on course evaluation process: Students' and lecturers' perceptions. *Procedia: Social and Behavioral Sciences*, 123 (4), 380–388.

Aleamoni, L. M. (1981) Student ratings of instruction, in: Millman, J. (Ed.) *Handbook of Teacher Evaluation*, 110-145 (Beverly Hills, Sage).

Alexander, F. (2000). The Changing Face of Accountability: Monitoring and Assessing Institutional Performance in Higher Education. *The Journal of Higher Education*, 71(4), 411-431.

Algozzine, B., Gretes, J., Flowers, C., Howley, L., Beattie, J., Spooner, F., Monhanty, M., & Bray, M. (2004). Student Evaluation Of College Teaching: A Practice In Search Of Principles. *College Teaching*, 52(4), 134-141.

Allen, A. (2012) Cultivating the myopic learner: the shared project of high-stakes and low-stakes assessment. *British Journal of Sociology of Education*, 33 (5), 641-659.

Amarante, A.M. y Daura, F.T. (2012). La evaluación de la cátedra como recurso regulador de las prácticas docentes. *Revista de Docencia Universitaria*, 10 (3), 379-898.

Asassfeh, S., Al-Ebous, H., Khwaileh, F., & Al-Zoubi, Z. (2014). Student faculty evaluation (SFE) at Jordanian universities: A student perspective. *Educational Studies*, 40(2), 121-143.

Balam, E., & Shannon, D. (2010). Student ratings of college teaching: A comparison of faculty and their students. *Assessment & Evaluation in Higher Education*, 35(2), 209-221.

Ball, S. (2013). Performatividad y fabricaciones en la economía educacional: rumbo a una sociedad performativa. *Pedagogía y Saberes*, (38), 103-113

Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18 (2), 215-228.

Ball, S. J. (2014). Globalización, mercantilización y privatización: tendencias internacionales en educación y política educativa. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, (22), 1-13.

Braskamp, L.A. & Ory, J.C. (1994). *Assessing Faculty Work: Enhancing Individual and Institutional Performance. Jossey-Bass Higher and Adult Education Series*. San Francisco: Jossey-Bass Inc.

Camilloni, A. (s/f). "Las funciones de la evaluación", ficha de cátedra de Didáctica I, Depto. de Ciencias de la Educación, Facultad de Filosofía y Letras, UBA, Mimeo.

Campbell, J., & Bozeman, W. (2007). The Value of Student Ratings: Perceptions of Students, Teachers, and Administrators. *Community College Journal of Research and Practice*, 32(1), 13-24.

Capa-Aydin, Y., Cobanoglu, R., & Gulmez-Dag, G. (2017) Student Course Evaluation in Higher Education: a qualitative Inquiry into Faculty perceptions. Paper presented at European Conferencia on Educational Research, August 2017.

Carlino, F. (2003). *La evaluación educacional. Historia, problemas y propuestas*. Buenos Aires, Larousse.

Centra, J. A. (1979) *Determining Faculty Effectiveness* (San Francisco, Jossey-Bass).

Centra, J. A. (1993). *Reflective Faculty Evaluation: enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.

Chan, C., Luk, L., & Zeng, M. (2014). Teachers' perceptions of student evaluations of teaching. *Educational Research and Evaluation*, 20(4), 275-289.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies, *Review of Educational Research*, (51), 281-309.

CONEAU (2019) *Comisión Nacional de Evaluación y Acreditación Universitaria*. Universidad de San Andrés. Jorge Lafforgue (Ed). Ciudad Autónoma de Buenos Aires: CONEAU-Comisión nacional de Evaluación y Acreditación Universitaria.

Corbetta, P. (2003) *Metodología y técnicas de investigación social*. Madrid: Mc Graw Hill.

De Ketele, J. y Roegiers, X. (1995) *Metodología para la recogida de información*, Madrid: La Muralla.

Douglas, J. & Douglas, A. (2006). Evaluating Teaching Quality. *Quality in Higher Education*, 12 (1), 3-13.

Elola, N. et. al. (2010). La evaluación educativa. Fundamentos teóricos y orientaciones prácticas. Buenos Aires: Aique Grupo Editor.

Feldman, K. A. (1976). Grades and college students' evaluations of their courses and teachers. *Research in Higher Education*, (4), 69-111.

Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis, *Research in Higher Education*. (6), 223-274.

Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't. *Research in Higher Education*, (9), 199-242.

Feldman, K. A. (1979). The significance of circumstances for college students' ratings of their teachers and courses. *Research in Higher Education*, (10), 149-172.

Feldman, K. A. (1989). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. *Research in Higher Education*, (30), 137-189.

Feldman, K. A. (1992). College students' views of male and female college teachers: Part I—Evidence from the social laboratory and experiments. *Research in Higher Education*, (33), pp. 317-375.

Feldman, K. A. (1993) College students' views of male and female college teachers: Part II—Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, (34), 151.

Fernández Lamarra, N. y Coppola, N. (2009). La Evaluación de la Docencia Universitaria desde un abordaje institucional. *Revista Iberoamericana de RAES*, 6 (9)

Fernández Lamarra, N., y Aiello, M. (2013). “Evaluación para la mejora. El impacto de las evaluaciones institucionales en la planificación estratégica y en la gestión de las universidades argentinas”, en Almuiñas Rivero, J.L. (Compilador) y Fernández Lamarra, N. (Introducción) *La Planificación Estratégica en las Instituciones de Educación Superior*, Montevideo, Universidad de la República, Facultad de Veterinaria. Red de Dirección Estratégica de la Educación Superior (RED-DEES).

Fernández Lamarra, N., y Coppola, N. (2008a). La Evaluación de la Docencia Universitaria en Argentina. Situación, problemas y perspectivas. *Revista Iberoamericana de Evaluación Educativa*, 1 (3), 97-123.

Fernández Lamarra, N., y Coppola, N. (2008b). Aproximaciones a la evaluación de la docencia universitaria en algunos países iberoamericanos. Una perspectiva comparada entre similitudes, diferencias y convergencias. *Perspectivas em políticas públicas*, 1 (2), 131-162.

Franklin J. & Theall, M. (1989). Who reads ratings: Knowledge, attitude and practice of users of student ratings of instruction. Paper presented at the Annual Meeting of the American Education Research Association, San Francisco.

Frick, Theodore W. (2010). Improving course evaluations to improve instruction and complex learning in higher education. *Educational Technology Research & Development*, 58(2), 115-137

Fundéu BBVA (2015). Troleo y troleo, con una sola ele, palabras correctas. Madrid, España: Creative Commons. Recuperado de: <https://www.fundeu.es/recomendacion/troleo-troleo/>

García Garduño, J. M. (2000). Las dimensiones de la efectividad docente, validez y confiabilidad de los cuestionarios de evaluación de la docencia: síntesis de investigación internacional. En Mario Rueda y Frida Díaz-Barriga (Eds.), *Evaluación de la docencia*. México. Paidós, 41-62.

García Garduño, J.M., y Medécigo Shej, A. (2014) Los criterios que emplean los estudiantes universitarios para evaluar la in-eficacia docente de sus profesores. *Perfiles educativos*, 26 (143), 124-139.

Haskell, R. E. (1997). Academic freedom, tenure, and student evaluations of faculty: Galloping polls in the 21st century. *Education Policy Analysis Archives*, 5 (6), 1-32.

Hidalgo, N. y Murillo, F. J. (2017.) Las concepciones sobre el Proceso de Evaluación del Aprendizaje de los Estudiantes. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 15 (1), 107-128.

Locke, K. (2015). Performativity, Performance and Education. *Educational Philosophy and Theory*, 47 (3), 247-259. DOI: 10.1080/00131857.2013.857287

Luna Serrano, E. Y Torquemada, A. D. (2008) Los cuestionarios de evaluación de la docencia por los alumnos: balance y perspectivas para su agenda. *Revista Electrónica de Investigación Educativa, Especial*.

Lyotard, J. F (1993) *La condición postmoderna: informe sobre el saber*. Madrid: Planeta-Agostini.

Marlin, J. (1987). Student Perception of End-of-Course Evaluations. *The Journal of Higher Education*, 58(6), 704-716.

Marques, T. E., Lane, D. M. y Dorfman, P. W. (1979). Toward the development of a system for instructional evaluation: Is there consensus regarding what constitutes effective teaching? *Journal of Educational Psychology*, 71, 840-849.

Marsh, H. W. (1987). Students' evaluation of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, (11), 253-388.

Marsh, H. W. & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective, in: J. C. SMART (Ed.) *Higher Education: Handbook of Theory and Research*, (8), 143-233. New York: Agathon Press.

Martínez Carazo, P. C. (2006). El método de estudio de caso: estrategia metodológica de la investigación científica. *Pensamiento y gestión*, 20, 165-193.

McKeachie, W. J. (1979). Student ratings of faculty: A reprise, *Academe*. (65), 384-397.

Moore, S. & Kuol, N. (2005). Students evaluating teachers: exploring the importance of faculty reaction to feedback on teaching, *Teaching in Higher Education*, 10 (1), 57-73.

Muñoz, J., y Alonso, F. (2015). Un análisis crítico del Espacio Europeo de Educación Superior como reforma de la enseñanza universitaria: Los programas Verifica y Docencia en el contexto español. *Educar Em Revista*, (57), 81-97.

Nasser, F., & Fresko, B. (2002). Faculty Views of Student Evaluation of College Teaching. *Assessment & Evaluation in Higher Education*, 27(2), 187-198.

Neave, G. (1988). On the cultivation of quality, efficiency and enterprise: An overview of recent trends in higher education in Western Europe 1986-1988, *European Journal of Education*, 23.

Ornstein, A. C. (1990) A look at teacher effectiveness research—theory and practice, *National Association of Secondary School Principals Bulletin*, (74), 78-88.

Perrenoud, P. (2010). La evaluación de los alumnos. De la producción de la excelencia a la regulación de los aprendizajes. Buenos Aires: Colihue.

Powell, R. (1978) Faculty rating scale validity: The selling of a myth, *College English*, (39), 616-629.

Remmers, H. H. & Branderburg, G. C. (1927) Experimental data on the Purdue ratings scale for instructors, *Educational Administration and Supervision*, (13), 519-527.

Ricco, G. (2001). Plan Nacional de Evaluación de la calidad de las Universidades. Buenos Aires: Consejo de Universidades de Argentina, CIN.

Rueda Beltrán, M. (2006). Evaluación de la labor docente en el aula universitaria. México DF: Red de Investigadores de Evaluación de la Docencia (RIED)/ Centro de Estudios sobre la Universidad-UNAM.

Rueda, M., Elizalde, L. y Torquemada, A. (2003). La evaluación de la docencia en las universidades mexicanas. *Revista de la Educación Superior*, 127, 71-77.

Rueda, M.B, Luna Serrano, E., y Loredó Enríquez, J. (2010) La Evaluación de la Docencia en las Universidades Públicas Mexicanas: un diagnóstico para su comprensión y mejora. *Revista Iberoamericana de Evaluación Educativa*, 3 (1), 78-92.

Ryan, P. (2015). Quality assurance in higher education: A review of literature. *Higher Learning Research Communications*, 5(4). <http://dx.doi.org/10.18870/hlrc.v5i4.257>

Sautú, R. (2005) Todo es teoría: objetivos y métodos de investigación. Buenos Aires: Lumiere.

Schettini, P., y Cortazzo, I. (2015). Análisis de datos cualitativos en la investigación social: Procedimientos y herramientas para la interpretación de información cualitativa. Buenos Aires: Editorial de la Universidad de La Plata (EDULP).

Scriven, M. (1967). The methodology of evaluation. En Tyler, R. W., Gagné, R. M. y Scriven, M. (Eds.), *Perspectives of curriculum evaluation*, 39-83. Chicago, IL: Rand McNally.

Scriven, M. (1981) Summative teacher evaluation, in: Milman, J. (Ed.), *Handbook of Teacher Evaluation*, 244-271 (Beverly Hills, Sage).

Seldin, P. (1993) The use and abuse of student ratings of professors, *Chronicle of Higher Education*, 39(46), 40.

Slocombe, T., Miller, D., & Hite, N. (2011). A Survey of Student Perspectives Toward Faculty Evaluations. *American Journal of Business Education*, 4(7), 51-57.

Sojka, J., Gupta, A., & Deeter-Schmelz, D. (2002). Student and Faculty Perceptions of Student Evaluations of Teaching: A Study of Similarities and Differences. *College Teaching*, 50(2), 44-49.

Stark, P. B. & Freishtat, R. (2014). An evaluation of course evaluations. *ScienceOpen Research*, 1, 1-7.

Tejedor Tejedor, F. J., y García-Valcárcel, A. (2010). Evaluación del desempeño docente. *Revista Española De Pedagogía*, 68 (247), 439-459.

Theall, M. & Franklin, J. L. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? In M. Theall, P. C. Abrami, & L. A. Mets (Eds.) *The student ratings debate: are they valid? How can we best use them? (New directions for Institutional Research*, 109, 45-56). San Francisco: Jossey-Bass.

Tomlinson, M. (2018). Conceptions of the value of higher education in a measured market. *Higher Education*, 75(4), 711-727.

Trowler, P. (2001) The Higher Education policy context of evaluative practices. En Saunders, M., Trowler, P., & Bamber, V. (Eds.) *Reconceptualising Evaluation in Higher Education: the practice turn*, 18-31. New York: McGraw Hill.

Vain, P. (1998). La evaluación de la docencia universitaria: un problema complejo. Documentos de trabajo. Buenos Aires: CONEAU.

Wachtel, H. K. (1998) Student Evaluation of College Teaching Effectiveness: a brief review. *Assessment & Evaluation in Higher Education*, 23 (2), 191-212. <http://dx.doi.org/10.1080/0260293980230207>.

Yorke, J. & Vidovich, L. (2016). Learning Standards and the Assessment of Quality in Higher Education: Contested Policy Trajectories. Doi:10.1007/978-3-319-32924-6

Zoller, U. (1992). Faculty teaching performance evaluation in higher science education: Issues and implications (a “cross-cultural” case study). *Science Education*, 76 (6), 673–84.

## ANEXOS

### Anexo 1. Guías de entrevista.

#### Estudiantes

##### A. *Introducción*

- Presentación personal y de los objetivos de la investigación.
- Garantizar el anonimato de las respuestas.
- Pedir permiso para grabación.

##### B. *Caracterización del entrevistado*

- Sexo
- Carrera que está cursando dentro de la universidad
- Año de cursada

##### C. *Preguntas Orientadoras*

¿Cuál crees que es el fin, en general, de las encuestas como instrumento de evaluación docente por parte de los estudiantes?

¿Consideras que las calificaciones de las encuestas constituyen una buena medida para evaluar el desempeño docente?

¿Crees que es importante que se incluya a los estudiantes como evaluadores de los docentes? por qué?

¿Cuál es tu percepción respecto a tu rol como evaluador de los docentes?

¿Crees que los docentes miran/prestan atención a los resultados de las encuestas?

¿Crees que los comentarios son tenidos en cuenta por los docentes para realizar cambios?

¿Consideras que es un instrumento adecuado para proveer feedback a los docentes?

¿Crees que alguna persona además del docente ve/accede a los resultados?

¿Conoces los usos de los resultados de las encuestas?

¿Crees que las encuestas son un criterio importante para tomar decisiones importantes? (permanencia o no permanencia en la institución, cambio en el salario, etc.)?

¿Cuál crees que es el impacto de utilizar las encuestas con estos fines?

¿Crees que los docentes ajustan su comportamiento para recibir mejores calificaciones en las encuestas?

¿Cuáles son los factores o aspectos que tenés en cuenta a la hora de evaluar a un docente? (amable, si sabe mucho, si es entretenido, etc.)

¿Te tomas las encuestas seriamente?

¿Crees que lo que pones en las encuestas tiene que ver, o este asociado a la calificación que recibiste o esperas recibir en la materia? ¿alguna vez tomaste una postura vengativa por una mala calificación?

A grandes rasgos, ¿crees que la encuesta es un instrumento confiable para evaluar el desempeño de los docentes?

#### *D. Cierre*

- Agradecimiento

### Docentes

#### *A. Introducción*

- Presentación personal y de los objetivos de la investigación.
- Garantizar el anonimato de las respuestas.
- Pedir permiso para grabación.

#### *B. Caracterización del entrevistado*

- Sexo
- Campo disciplinar
- Grado de formación alcanzado
- Dedicación
- Tendencia en las encuestas

#### *C. Cuestionario*

¿Cuál crees que es el fin, en general, de las encuestas como instrumento de evaluación docente por parte de los estudiantes? ¿Cuál consideras que es su principal propósito?

¿Cómo percibís el hecho de que tus cursos sean evaluados por tus estudiantes?

¿Consideras que las calificaciones de las encuestas constituyen una buena medida del desempeño docente?

¿Mirás los resultados de las encuestas?

La información recogida ¿es significativa para mejorar tu desempeño como docente?  
¿qué parte? (información cuantitativa o cualitativa)

¿Has hecho cambios a partir de las encuestas?

¿Podrías mencionar que tipos de cambios, qué cosas cambiaste particularmente?

¿Conoces los usos de los resultados de las encuestas?

¿Cuál crees que es el rol que deben tener las evaluaciones en la toma de decisiones?

¿Considerás que las calificaciones provistas por las encuestas constituyen un criterio relevante para la toma de decisiones? ¿y en los casos de permanencia/no permanencia en la institución? ¿y en los casos de promoción?

¿Cuál crees que es el impacto de utilizar las encuestas con estos fines?

¿Crees que la calificación recibida en la encuesta refleja la calidad/que tan bueno sos como docente?

¿Cuáles pensás que son los aspectos que los estudiantes tienen en cuenta para evaluarte?

¿Con qué grado de seriedad/compromiso considerás que los estudiantes se toman las encuestas?

¿Crees que los estudiantes se encuentran capacitados como para evaluar el curso?

¿Considerás que hay variables externas a la enseñanza que podrían influenciar o impactar en los resultados? ¿Cuáles? (tamaño de la clase, obligatoriedad, etc.)

#### *D. Cierre*

- Agradecimiento

#### Directores de carrera

##### *A. Introducción*

- Presentación personal y de los objetivos de la investigación.
- Garantizar el anonimato de las respuestas.
- Pedir permiso para grabación.

##### *B. Caracterización del entrevistado*

- Sexo
- Carrera de la cual es director
- Año
- Grado de formación alcanzado

- Dedicación
- Tendencia en las encuestas

### *C. Cuestionario*

¿Cuál crees que es el fin, en general, de las encuestas como instrumento de evaluación docente por parte de los estudiantes? ¿Cuál consideras que es su principal propósito?

¿Cómo percibís el hecho de que los cursos sean evaluados por estudiantes?

¿Consideras que las calificaciones de las encuestas constituyen una buena medida del desempeño docente?

¿Mirás los resultados de las encuestas de tus docentes?

La información recogida ¿crees que es significativa para mejorar su desempeño como docentes? ¿Alguna parte de la encuesta en particular? (información cuantitativa o cualitativa)

¿Crees que los docentes hacen cambios a partir de las encuestas?

¿Podrías mencionar que tipos de cambios?

¿Conoces los usos de los resultados de las encuestas?

¿Cuál crees que es el rol que deben tener las evaluaciones en la toma de decisiones?

¿Considerás que las calificaciones provistas por las encuestas constituyen un criterio relevante para tomar decisiones? ¿y en los casos de permanencia/no permanencia en la institución? ¿y en los casos de promoción?

¿Cuál crees que es el impacto de utilizar las encuestas con estos fines?

¿Crees que la calificación recibida en la encuesta refleja la calidad los docentes?

¿Cuáles piensas que son los aspectos que los estudiantes tienen en cuenta para evaluar a los docentes?

¿Con qué grado de seriedad/compromiso considerás que los estudiantes se toman las encuestas?

¿Crees que los estudiantes se encuentran capacitados como para evaluar el curso?

¿Considerás que hay variables externas a la enseñanza que podrían influenciar o impactar en los resultados? ¿Cuáles? (tamaño de la clase, obligatoriedad, etc.)

### *D. Cierre*

- Agradecimiento.