

UNIVERSIDAD NACIONAL DEL NORDESTE  
FACULTAD DE MEDICINA



Facultad de Medicina  
Universidad Nacional del Nordeste



Universidad Nacional  
del Nordeste

TÍTULO DE LA TESIS DOCTORAL

**IMPACTO EDUCACIONAL DE LOS SISTEMAS MULTIMÉTODOS  
DE EVALUACION DE COMPETENCIAS QUE SE APLICAN  
EN LA CARRERA DE MEDICINA DE LA UNNE**

Trabajo de investigación que presenta

**Juan José Di Bernardo**

para la obtención del Título de

Doctor de la Universidad Nacional del Nordeste en Medicina

Bajo la dirección de la doctora

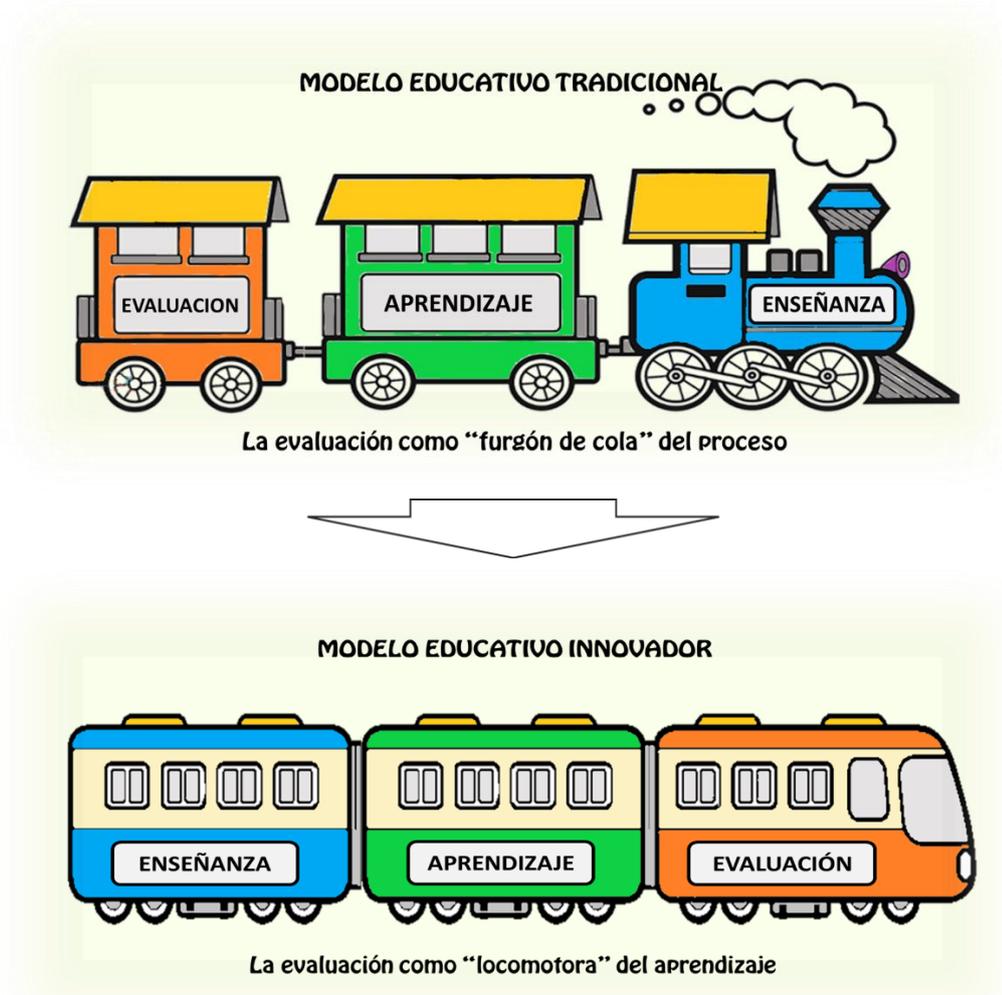
**Patricia Belén Demuth Mercado**

Corrientes, Argentina

**2022**



## UN CAMBIO DE PARADIGMA NECESARIO



**“Todo aquello que el hombre ignora no existe para él.  
Por eso el universo de cada uno se resume al tamaño de su saber “**

Albert Einstein (1930)





## DEDICATORIAS

A BETTINA mi esposa,  
que resiste y acompaña desde hace más de cuatro décadas, todas  
mis “aventuras educativas” con admirable comprensión.

A BODI mi maestro,  
que me inspiró el amor por la clínica, los valores del profesionalismo  
y la enorme “pasión” por la Educación Médica.

## AGRADECIMIENTOS

A PATÍN (Pato) mi directora de tesis por su constante apoyo,  
sus acertadas recomendaciones, el valioso tiempo que me regaló  
y su increíble tolerancia ante mi TOC. Gracias AMIGA

A todos los estudiantes que participaron voluntariamente en esta investigación,  
pues sin sus revelaciones no hubiera escrito una sola línea.

A los profesores de las asignaturas “Medicina, Hombre y Sociedad” y “Medicina III”  
por permitirme acceder libremente a los registros y materiales curriculares.

Al director del Doctorado Prof. Dr. Jorge Gorodner y a su cuerpo académico, por  
brindarme las herramientas necesarias para afrontar con firmeza este desafío.

A las profesoras Milagros Delfino y Mariana Sottile por revisar y corregir mis  
lamentables textos en inglés. Thank you teachers

A las autoridades de la Facultad de Medicina de la UNNE, por dejarme seguir  
aprendiendo de nuestras propias experiencias educativas.

Y un especial agradecimiento a mis nietos: Marcos, Tomi, Aramí, Santiago, María  
Emilia, Juan Cruz, Paula, Liz e Isidro; que fueron y son mi “recreo” más maravilloso.  
(hay un ejemplar para cada uno así no se pelean por leerla)



## INDICE GENERAL

ÍNDICE DE FIGURAS.....	13
ÍNDICE DE GRÁFICOS.....	13
ÍNDICE DE TABLAS.....	14
RESUMEN.....	15
ABSTRACT.....	17
Capítulo 1: INTRODUCCIÓN A LA INVESTIGACIÓN.....	21
1.1. El tema y su problematización.....	21
1.2. Fundamentación del problema de investigación.....	22
1.3. Supuestos básicos de la investigación.....	24
1.4. Supuestos de anticipación de sentido.....	26
1.5. Propósitos y Objetivos de la Investigación.....	27
1.6. Estructura del informe de investigación.....	29
Capítulo 2: MARCO TEÓRICO REFERENCIAL.....	33
2.1. Educación por competencias.....	33
2.2. Evaluación de competencias:.....	36
2.2.1. Instrumentos para evaluar competencias.....	39
2.2.2. Sistemas de evaluación de competencias.....	44
2.2.3. Calidad de las evaluaciones.....	45
2.2.3.1. Calidad de los instrumentos de evaluación.....	45
2.2.3.2. Calidad de los sistemas de evaluación.....	46
2.3. Impacto educativo de la evaluación.....	47
2.4. Principales variables vinculadas al impacto educativo de la evaluación.....	50
2.4.1. Metas de logro.....	50
2.4.2. Enfoques de aprendizaje:.....	52
2.4.3. Comportamientos de estudio de los estudiantes.....	54
2.4.4. Bases cognitivas del aprendizaje.....	56
2.4.4.1. Sistemas de Memoria.....	56
2.4.4.2. Procesamiento cognitivo.....	57
2.4.4.3. Carga cognitiva.....	58

2.4.5. Retroalimentación y autorreflexión de los estudiantes .....	59
2.4.6. Percepciones de los estudiantes en las evaluaciones .....	62
2.5. Estado del Arte.....	65
2.5.1. Sobre los efectos de la evaluación.....	65
2.5.2. Sobre los efectos de la retroalimentación .....	66
2.5.3. Sobre instrumentos de evaluación específicos.....	68
2.5.4. Sobre la evaluación programática .....	70
2.5.5. Investigaciones en Latinoamérica y en la región .....	71
Capítulo 3: MARCO METODOLÓGICO .....	77
3.1. Paradigma de Investigación.....	77
3.2. Diseño Metodológico.....	78
3.2.1. Tipo de Investigación.....	78
3.2.2. Selección de los casos de estudio .....	80
3.2.3. Categorías preliminares de análisis .....	80
3.2.3.1. Entorno de aprendizaje y evaluación en las asignaturas .....	80
3.2.3.2. Aspectos motivacionales de los estudiantes .....	80
3.2.3.3. Enfoques de Aprendizaje de los estudiantes .....	81
3.2.3.4. Comportamientos de estudio de los estudiantes.....	82
3.2.3.5. Experiencias y percepciones de los estudiantes en las evaluaciones .....	82
3.2.4. Técnicas e Instrumentos de recolección de la información .....	82
3.2.4.1. Documentos, registros, y materiales curriculares .....	82
3.2.4.2. Inventarios estandarizados y cuestionario - INSTRUMENTO A.....	83
3.2.4.3. Entrevistas electrónicas con cuestionario abierto - INSTRUMENTO B.....	83
3.2.5. Técnicas de Procesamiento y Análisis de la información .....	84
3.2.5.1. Análisis de la información cualitativa .....	84
3.2.5.2. Análisis de los datos cuantitativos.....	85
3.2.6. Matriz analítica-metodológica: Ejes y categorías de análisis .....	86
Capítulo 4: PRESENTACIÓN DE LOS RESULTADOS.....	91
4.1 Presentación de los casos de estudio .....	91
4.1.1. Contexto institucional de los casos de estudio .....	91

4.1.2. Caso de estudio MHS.....	92
4.1.3. Caso de estudio MED-III .....	94
4.2. Primer Eje de Análisis: Evaluación del aprendizaje .....	96
4.2.1. Entorno de aprendizaje y evaluación .....	96
4.2.1.1. Caso MHS.....	96
4.2.1.2. Caso MED-III .....	99
4.2.2. Estudiantes evaluados.....	103
4.2.2.1. Caso MHS.....	103
4.2.2.2. Caso MED-III .....	105
4.3. Segundo Eje de Análisis: Evaluación para el aprendizaje .....	108
4.3.1. Factores de la evaluación.....	108
4.3.1.1. Caso MHS.....	108
4.3.1.2. Caso MED-III .....	111
4.3.2. Efectos de aprendizaje .....	114
4.3.2.1. Caso MHS.....	114
4.3.2.2. Caso MED-III .....	116
4.4. Tercer Eje de Análisis: Evaluación como aprendizaje .....	118
4.4.1. Antes de la evaluación.....	118
4.4.1.1. Caso MHS.....	118
4.4.1.2. Caso MED-III .....	120
4.4.2. Durante la evaluación.....	122
4.4.2.1. Caso MHS.....	122
4.4.2.2. Caso MED-III .....	123
4.4.3. Después de la evaluación .....	125
4.4.3.1. Caso MHS.....	125
4.4.3.2. Caso MED-III .....	126
4.5. Diferencias entre los Casos de Estudio.....	127
4.5.1. Diferencias en el escenario curricular de la carrera .....	127
4.5.2. Diferencias en los entornos de aprendizaje y evaluación .....	128
4.5.3. Diferencias en los estudiantes evaluados .....	129

4.5.4. Diferencias en los factores de la evaluación promotores de aprendizaje .....	131
4.5.5. Diferencias en los efectos de aprendizaje de las evaluaciones.....	132
4.5.6. Comportamientos, experiencias y percepciones de los estudiantes.....	133
Capítulo 5: DISCUSIÓN DE LOS RESULTADOS .....	137
5.1. Primer Eje de Análisis: Evaluación del aprendizaje .....	137
5.1.1. Entorno de aprendizaje y evaluación .....	137
5.1.1.1. ¿Como se diseñaron y planificaron los sistemas de evaluación multimétodo? .....	137
5.1.1.2. ¿Por qué se eligieron esas herramientas de evaluación?.....	138
5.1.1.3. ¿Como se elaboraron y validaron las herramientas de evaluación? .....	138
5.1.1.4. ¿Cómo funcionaron las herramientas de evaluación? .....	139
5.1.2. Estudiantes evaluados.....	140
5.1.2.1. ¿Cuáles eran las metas de logro de los estudiantes?.....	140
5.1.2.2. ¿Qué enfoques de aprendizaje aplicaban? .....	142
5.1.2.3. ¿Qué resultados obtuvieron en las evaluaciones? .....	143
5.2. Segundo Eje de Análisis: Evaluación para el aprendizaje.....	145
5.2.1. Factores de la evaluación .....	145
5.2.1.1. ¿Qué sabían los estudiantes sobre los objetivos y criterios de las evaluaciones? y ¿Qué interacción tuvieron con los docentes?.....	145
5.2.1.2. ¿Cómo influyeron en el estudio las evaluaciones múltiples y la proximidad temporal a cada evaluación? .....	147
5.2.2. Efectos de aprendizaje .....	148
5.2.2.1. ¿Qué procesos cognitivos y metacognitivos aplicaron en las distintas evaluaciones?.....	148
5.2.2.2. ¿Cómo fueron las devoluciones en las diferentes evaluaciones? .....	151
5.3. Tercer Eje de Análisis: Evaluación como aprendizaje .....	153
5.3.1. Antes de la evaluación.....	153
5.3.1.1. ¿Cómo influyeron las evaluaciones en los comportamientos de estudio de los estudiantes?; ¿Que ajustes hicieron en sus estrategias de estudio?; ¿Por qué lo hicieron?.....	153
5.3.1.2. ¿Qué factores motivaron o desanimaron el estudio, en la fase de preparación para las evaluaciones?.....	158

5.3.2. Durante la evaluación.....	162
5.3.2.1. ¿Qué beneficios aportaron las herramientas de evaluación a la formación y aprendizaje de los estudiantes?.....	162
5.3.2.2. ¿Qué factores o elementos de las evaluaciones facilitaron y cuales obstaculizaron el buen rendimiento?.....	163
5.3.3. Después de la evaluación .....	165
5.3.3.1. ¿Cómo percibieron su capacidad y sus habilidades intelectuales para afrontar las evaluaciones? .....	165
5.3.3.2. ¿Qué consecuencias percibieron como riesgosas en el sistema de evaluación? (Riesgos percibidos).....	167
5.3.3.3. ¿Qué aportó el sistema de evaluación aplicado, al aprendizaje de la materia? .....	169
Capítulo 6: CONCLUSIONES .....	175
6.1. Conclusiones de la investigación.....	175
6.2. Implicaciones de la investigación para la práctica docente.....	183
6.3. Limitaciones de la investigación.....	189
BIBLIOGRAFIA .....	193
ANEXO 1.1: Instrumento A (1).....	213
ANEXO 1.2: Instrumento A (2).....	215
ANEXO 1.3: Instrumento A (3).....	217
ANEXO 2: Instrumento B.....	219



## ÍNDICE DE FIGURAS

Figura 2.1: Evaluación Formativa: Modelo de Clark (2010) .....	38
Figura 2.2: Instrumentos de evaluación y resultados de aprendizaje ensamblados a la Pirámide de Miller .....	42
Figura 2.3: Diferentes estados de evaluación en los programas de educación médica .....	45
Figura 2.4: Modelo de los efectos de aprendizaje pre-evaluación de la evaluación sumativa .....	48
Figura 2.5: Modelo de la memoria humana de Atkinson–Shiffrin .....	56
Figura 3.1: Síntesis del proceso de triangulación, codificación y categorización de los datos cualitativos aplicando la Teoría fundamentada .....	85
Figura 4.1: Estructura de materias en la Carrera de Medicina y asignaturas casos de estudio .....	92
Figura 4.2: Diagrama que muestra la organización y articulación de los contenidos de MHS .....	93
Figura 6.1: Interacción de los factores que tienen efectos sobre el aprendizaje .....	188

## ÍNDICE DE GRÁFICOS

Gráfico 4.1: Instrumento de evaluación que más aportó al aprendizaje de MHS (ranking de las respuestas de los estudiantes) .....	111
Gráfico 4.2: Instrumento de evaluación que más aportó al aprendizaje de MED-III (ranking de las respuestas de los estudiantes) .....	113
Gráfico 4.3: Actividades cognitivas de los estudiantes en las evaluaciones de MHS .....	115
Gráfico 4.4: Actividades cognitivas de los estudiantes en las evaluaciones de MED-III ....	117
Gráfico 4.5: Diferencia entre los puntajes logrados en las pruebas de MHS y MED-III .....	130

## ÍNDICE DE TABLAS

Tabla 3.1: Matriz analítica-metodológica .....	87
Tabla 4.1: Metas de logro de los 322 estudiantes que completaron el cuestionario A-AGQ-R .....	103
Tabla 4.2: Enfoques de aprendizaje de los 322 alumnos que completaron el R-SPQ-2F .....	104
Tabla 4.3: Puntajes de los 462 estudiantes de MHS en cada evaluación .....	105
Tabla 4.4: Metas de logro de los 139 estudiantes que completaron el cuestionario A-AGQ-R .....	106
Tabla 4.5: Enfoques de aprendizaje de los 139 estudiantes que completaron el R-SPQ-2F .....	106
Tabla 4.6: Puntajes de los 149 estudiantes de MED-III en cada evaluación .....	107
Tabla 4.7: Opiniones de los alumnos sobre las oportunidades y beneficios que tuvieron en las evaluaciones de MHS .....	122
Tabla 4.8: Opiniones de los alumnos sobre las oportunidades y beneficios que tuvieron en las evaluaciones de MED-III .....	124
Tabla 4.9: Opiniones de los alumnos sobre las oportunidades y beneficios que tuvieron en las evaluaciones .....	133

## RESUMEN

### **Título: IMPACTO EDUCACIONAL DE LOS SISTEMAS MULTIMÉTODOS DE EVALUACION DE COMPETENCIAS QUE SE APLICAN EN LA CARRERA DE MEDICINA DE LA UNNE**

**INTRODUCCION:** Numerosas evidencias demuestran que la evaluación tiene efectos positivos en el aprendizaje de los estudiantes, pero también reportan efectos menos deseables. Este fenómeno conocido como “impacto educacional” de las evaluaciones, ha sido muy estudiado en diferentes instrumentos de evaluación en forma aislada, pero pocos investigadores lo han explorado en el contexto de sistemas de evaluación que utilizan múltiples herramientas. En tal sentido se planteó como problema de investigación: ¿Qué impacto educacional puede producir en los estudiantes de medicina la aplicación de sistemas multimétodos para evaluar competencias?; y como principales interrogantes que dan sentido a los objetivos específicos de este trabajo: ¿Cómo se integran las evaluaciones al entorno de aprendizaje de los estudiantes?; ¿Como influyen las evaluaciones sobre el aprendizaje de los estudiantes?; ¿Qué puede aportar cada momento de la evaluación al aprendizaje de los estudiantes?

**METODOLOGÍA:** Desde un paradigma “interpretativo-constructivista”, la investigación fue desarrollada con un enfoque metodológico mixto a predominio cualitativo; diseñada como un estudio de caso de tipología múltiple, diseño incrustado con diferentes unidades de análisis; y con alcance interpretativo. Los casos de estudio seleccionados fueron las dos asignaturas que aplicaban sistema de evaluación multimétodo en la Carrera de Medicina: “Medicina, Hombre y Sociedad” (MHS) de 1° año y “Medicina III” (MED-III) de 5° año. Las categorías preliminares de análisis fueron: el entorno de aprendizaje y evaluación en los casos de estudio; los aspectos motivacionales y enfoques de aprendizaje de los estudiantes, sus comportamientos de estudio, y sus experiencias y percepciones en las evaluaciones. Para recabar la información se utilizaron: documentos, registros y materiales curriculares; inventarios estandarizados; y cuestionarios autoadministrados. Para analizar la información cualitativa se aplicó la Teoría Fundamentada constructivista; y para los datos cuantitativos, rúbricas específicas y análisis estadísticos. Durante el trabajo de tesis, se elaboró una matriz analítica-metodológica con tres ejes de análisis: evaluación del aprendizaje, evaluación para el aprendizaje y evaluación como aprendizaje, y diferentes categorías de las que se desprenden las preguntas exploratorias que guiaron la investigación.

**RESULTADOS:** En relación al entorno de aprendizaje-evaluación: los casos de estudio diseñaron sus propios sistemas de evaluación de competencias en forma integrada a los procesos de aprendizaje, con diversos instrumentos adecuados a los objetivos de cada evaluación, acorde al nivel curricular y al contexto de cada materia; con propósitos diferentes, sumativo (MHS) y formativo-sumativo (MED-III). La mayoría de los estudiantes

tenían orientación por metas de aproximación al aprendizaje (predisposición por aprender) y aplicaban enfoques de aprendizaje profundo, lo que se vio reflejado en los resultados que obtuvieron en las evaluaciones. En cuanto a los efectos de las evaluaciones sobre el aprendizaje: Hubo diferencias entre los casos, debidas principalmente al elevado número de cursantes de MHS (1363) que contrastó notoriamente con MED-III (151). En ese contexto “masivo” de cursado, los alumnos principiantes (P): no conocían (antes de las pruebas) los objetivos de aprendizaje y los criterios de evaluación; tuvieron poca interacción personal con los docentes; afrontaron instancias de evaluación muy cercanas en el tiempo; y no recibieron devoluciones oportunas y efectivas. Estas experiencias fueron inversas en los estudiantes avanzados (A). En cuanto a los procesos cognitivos realizados durante las evaluaciones, las diferencias entre (P) y (A) estuvieron relacionadas a su nivel de formación en la carrera, los primeros aplicaron principalmente procesos mentales de “orden inferior” (Taxonomía de Bloom) como recordar, comprender y aplicar; y los (A), operaciones mentales de “orden superior” como analizar, evaluar y crear. Los tres momentos de la evaluación tuvieron efectos de aprendizaje. Antes de la evaluación, los alumnos adecuaron sus comportamientos de estudio para mejorar su preparación, valorando el “aprendizaje grupal colaborativo” y los beneficios de revisar exámenes de otras cohortes; los principales motivadores para el estudio fueron: “las exigencias del sistema de evaluación de MHS” y la posibilidad de “promoción meritosa” en MED-III; y como factores de desánimo, el escaso tiempo entre las pruebas (MHS) y el cursado intenso con evaluaciones simultáneas (MED-III). Durante las evaluaciones, facilitaron el rendimiento: la diversidad de instrumentos aplicados y las cualidades de algunas herramientas; y obstaculizaron el buen rendimiento, fallas en el diseño de los instrumentos y, en la aplicación y corrección de ciertas evaluaciones; algunos alumnos tenían autoconcepto académico “negativo” auto atribuido a la formación inadecuada en el nivel medio (P) y a ciertas limitaciones intelectuales (A); Los principales riesgos percibidos fueron, no ingresar a la carrera de medicina (P) y no poder cursar la POF (A). Entre los beneficios que aportó el sistema de evaluación al aprendizaje de la materia, reconocieron como más significativos: ampliar y profundizar los conocimientos, identificar las necesidades de aprendizaje, y desarrollar habilidades para trabajar en equipo y para reflexionar sobre el aprendizaje.

**CONCLUSIONES:** Esta investigación permitió comprender los factores, mecanismos y efectos de aprendizaje vinculados al impacto educacional de los sistemas multimétodo de evaluación de competencias en las dos materias que aplican dichos sistemas en la Carrera de Medicina de la UNNE. Además, posibilitó agrupar los factores relacionados a los efectos de aprendizaje de las evaluaciones en tres categorías: “factores que promovieron el aprendizaje”; “factores que dificultaron el aprendizaje”; y “factores con efectos ambivalentes”, que son los principales aportes de esta investigación para las prácticas docentes en Educación Médica.

## ABSTRACT

### **Title: EDUCATIONAL IMPACT OF COMPETENCE MULTIMETHOD ASSESSMENT SYSTEMS APPLIED IN THE MEDICINE CAREER AT UNNE**

**INTRODUCTION:** Numerous pieces of evidence show that assessment has positive effects on student learning, but less desirable effects are also reported. This phenomenon known as the "educational impact" of assessments has been extensively studied in different assessment instruments in isolation, but few researchers have explored it in the context of assessment systems that use multiple tools. In this sense, the following research problem was raised: What educational impact can the application of multi-method systems to assess competencies produce in medical students? and as main questions that give meaning to the specific objectives of this work: How are evaluations integrated into the student learning environment? How does assessment influence student learning?; What can each moment of evaluation contribute to student learning?

**METHODOLOGY:** From an "interpretivist/constructivist" paradigm, the research was developed with a mixed methodological approach with a predominance of qualitative; designed as a case study of multiple typology, embedded design with different units of analysis; and with interpretive scope. The case studies selected were the two subjects that applied a multi-method evaluation system in Medicine Career: "Medicine, Man and Society" (MMS) in the 1st year and "Medicine III" (MED-III) in the 5th year. The preliminary categories of analysis were: the learning and evaluation environment in the case studies; the motivational aspects and learning approaches of the students, their study behaviors, and their experiences and perceptions in the evaluations. To collect the information, the following were used: documents, records and curricular materials; standardized inventories; and self-administered questionnaires. To analyze the qualitative information, the constructivist Grounded Theory was applied; and for quantitative data, specific rubrics and statistical analyses. During the thesis work, an analytical-methodological matrix was elaborated with three axes of analysis: evaluation of learning, evaluation for learning and evaluation as learning, and different categories from which the exploratory questions that guided the research emerged.

**RESULTS:** In relation to the learning-assessment environment: the case studies designed their own competency assessment systems in an integrated manner to the learning processes, with various instruments appropriate to the objectives of each assessment, according to the curricular level and to the context of each subject; with different purposes, summative (MMS) and formative-summative (MED-III). Most of the students were oriented by learning-

approach goals (predisposition to learn) and applied deep learning approaches, which was reflected in the results they obtained in the evaluations.

Regarding the effects of the evaluations on learning: There were differences between the cases, mainly due to the high number of MMS students (1363) that contrasted markedly with MED-III (151). In this “massive” course context, the beginning students (P): did not know (before the tests) the learning objectives and the evaluation criteria; they had little personal interaction with teachers; they faced evaluation instances very close in time; and did not receive timely and effective returns. These experiences were reversed in advanced students (A). Regarding the cognitive processes carried out during the evaluations, the differences between (P) and (A) were related to their level of training in the career, the former mainly applied “lower order” mental processes (Bloom’s Taxonomy) such as remember, understand and apply; and the (A), “higher order” mental operations such as analyze, evaluate, and create. The three moments of the evaluation had learning effects. Before the evaluation, the students adapted their study behaviors to improve their preparation, valuing “collaborative group learning” and the benefits of reviewing exams from other cohorts; the main motivators for the study were: “the demands of the MMS evaluation system” and the possibility of “meritorious promotion” in MED-III; and as factors of discouragement, the short time between the tests (MMS) and the intense course with simultaneous evaluations (MED-III). During the evaluations, they facilitated the performance: the diversity of applied instruments and the qualities of some tools; and hindered the good performance, flaws in the design of the instruments and, in the application and correction of certain evaluations; some students had a “negative” academic self-concept attributed to inadequate training at the secondary level (P) and to certain intellectual limitations (A); The main perceived risks were not entering the medical degree (P) and not being able to take the POF (A). Among the benefits that the assessment system brought to the learning of the subject, they recognized as the most significant: broaden and deepen knowledge, identify learning needs, and develop skills for teamwork and reflection on learning.

**CONCLUSIONS:** This research made it possible to understand the factors, mechanisms and learning effects linked to the educational impact of multi-method competency assessment systems in the two subjects that apply these systems in the UNNE Medical School. In addition, it made it possible to group the factors related to the learning effects of the evaluations into three categories: “factors that promoted learning”; “factors that hindered learning”; and “factors with ambivalent effects”, which are the main contributions of this research for teaching practices in Medical Education.

# CAPITULO 1

## INTRODUCCIÓN A LA INVESTIGACIÓN





## Capítulo 1: INTRODUCCIÓN A LA INVESTIGACIÓN

### 1.1. El tema y su problematización

El impacto educacional de la evaluación es desde hace varios años tema importante de investigación y controversias en la Educación Superior y, en particular, en el campo de la Educación Médica.

Numerosos trabajos han demostrado que la evaluación tiene un efecto positivo en el aprendizaje de los estudiantes, pero al mismo tiempo han reportado que puede promover el desarrollo de estrategias de aprendizaje que son menos deseables (Newble & Jaeger, 1983; McManus et al., 1998; McLachlan 2006; Segers & Dochy, 2006; Cilliers et al., 2010; Al-Kadri et al., 2012; Cobb et al., 2013; Heeneman et al., 2015).

Para Monroe (2016), el impacto educacional de la evaluación no se refleja solamente en la cantidad y calidad del aprendizaje construido sino también en los cambios de aptitud de los estudiantes para aprender y en el desarrollo de habilidades para el aprendizaje autónomo. Sin embargo, no hay suficientes evidencias que demuestren que las evaluaciones puedan impactar positivamente sobre el desempeño clínico de los estudiantes y de los graduados (Miller & Archer, 2010).

Para muchos autores “la evaluación guía el aprendizaje” de los estudiantes pues constituye una fuerza motivadora para que aprendan. Para otros, esta afirmación es incorrecta y argumentan que el plan de estudios “es la clave para el aprendizaje” en cualquier curso o carrera. Para dirimir estas diferencias es oportuno recordar la afirmación de George Miller (reconocido como el padre de la pedagogía médica): *“cambiar el sistema de evaluación sin modificar el programa de enseñanza ha tenido una mayor repercusión sobre la naturaleza y la calidad del aprendizaje, que si se hubiera modificado el programa sin tocar los exámenes”* (citado por Nolla-Domenjó, 2009:228)

Por otro lado, los estudios de currícula ocultos han mostrado que la evaluación no impulsa el aprendizaje en todos los contextos, y en algunos casos los autores observaron que la evaluación “solo impulsa el aprendizaje para esa evaluación”, no el aprendizaje per se (Sambell & McDowell, 1998; McLachlan, 2006).

Otras investigaciones estudiaron el impacto educacional de la evaluación diferenciando sus efectos en tres momentos: 1°- antes de la evaluación (reflejado en los comportamientos de estudio de los estudiantes cuando se preparan para la evaluación); 2°-

durante la evaluación (condicionado por el tipo de pruebas y evidenciado por los resultados obtenidos); y 3°- después de la evaluación (manifestada durante la retroalimentación). (Cilliers et al., 2012).

La gran mayoría de estos estudios han explorado aisladamente el impacto educacional de cada uno de los instrumentos que se utilizan para evaluar competencias, y si bien los resultados son de mucha utilidad para elegir una determinada herramienta de evaluación, el impacto educacional no es directamente aplicable a los Sistemas de Evaluación que integran diferentes instrumentos de evaluación, como ocurre en algunas asignaturas de la carrera de medicina de la Universidad Nacional del Nordeste.

Por ello, se planteó como problema de investigación el siguiente interrogante:

**¿Qué impacto educacional puede producir en los estudiantes de medicina la aplicación de sistemas multimétodos para evaluar competencias?**

Para dar respuesta a este interrogante se abordaron como unidad a los Sistemas de Evaluación de competencias, objeto que no ha sido estudiado en los currícula estructurados en asignaturas y que es el enfoque original de esta propuesta.

## **1.2. Fundamentación del problema de investigación**

El impacto educacional de la evaluación no es un tema nuevo. En 1984, el investigador Norman Fredericksen afirmó que *"el método más poderoso para cambiar la manera de estudiar de los estudiantes y la manera en que los profesores enseñan, es cambiar la manera de evaluar"* (citado por Brailovsky, 2001:114). En la misma época David Newble postuló que *"la evaluación impulsa el aprendizaje de los estudiantes"* (citado por Al-Kadri et al., 2013:S31). Estas declaraciones siguen vigentes hasta el presente.

Las evidencias indican que la evaluación es el motor más apropiado para aprovechar el currículum, pues los estudiantes sobrecargados por las actividades de la carrera, responden estudiando sólo las partes del programa que se evalúan. En consecuencia, para promover el aprendizaje, la evaluación debe ser amplia y formativa, los estudiantes deben aprender de las evaluaciones y recibir la retroalimentación adecuada para construir sus conocimientos y habilidades; pero también la evaluación debe tener una función sumativa, para asegurar el desarrollo de las competencias (Wass et al, 2001).

En tal sentido, Epstein (2007) propone para los currícula de grado, sistemas de evaluación longitudinal multimétodo (ELM), ya que posibilitan evaluar una mayor amplitud

de contenidos y aportan el enfoque formativo para que estos estudiantes que están avanzando en un cuerpo relativamente no estructurado de conocimientos tengan los puntos de referencia para orientar su aprendizaje, reforzar su motivación e inspirar la adopción de estándares más estrictos para su formación.

En los últimos años ha surgido, como enfoque integral de la evaluación, la “evaluación programática para el aprendizaje” (PAL) que ya se ha extendido a varios países, reemplazando a los modelos de evaluación tradicional conductista que tienen como principal propósito determinar si un estudiante aprendió lo suficiente como para avanzar en sus estudios. Este nuevo modelo de evaluación aprovecha la relación que existe entre la evaluación y el comportamiento de aprendizaje de los estudiantes, aplicando la evaluación específicamente para dirigir el aprendizaje de los mismos de una manera más significativa, mediante múltiples puntos de evaluación con retroalimentación (cualitativa y cuantitativa) relevante para el aprendizaje, para finalmente tomar las decisiones de progreso cuando se han recogido suficientes datos y la información es clara, completa y objetiva. (Heeneman et al., 2015; Van der Vleuten & Heeneman, 2016; Schuwirth et al., 2017a; Schuwirth et al., 2017b; Schut et al, 2018).

Para Brailovsky (2001), está muy claro que la evaluación dirige el aprendizaje de los estudiantes, y lo hace a través de los contenidos y del formato de las herramientas de evaluación que se aplican, por eso recomienda utilizarla de manera estratégica para reforzar el aprendizaje y no con fines punitivos.

Sin embargo, aunque la mayoría de los autores coincide sobre la importancia de la evaluación en la mejora de los procesos de aprendizaje y en el efectivo desarrollo de las competencias, la realidad muestra que en los currícula organizados en asignaturas las prácticas de evaluación se repiten de forma casi mecánica, y muchas materias no evalúan la calidad de sus métodos de evaluación ni sus efectos, creando un escenario que perpetúa las prácticas evaluativas vigentes.

Esto es lo que ha ocurrido en la Carrera de Medicina de la UNNE que tiene un modelo pedagógico basado en las competencias del “médico general”, pero su diseño curricular está estructurado en asignaturas, lo que dificulta establecer una adecuada articulación e integración curricular que es fundamental para el aprendizaje y evaluación de las competencias, y para asegurar el afectivo desarrollo de las mismas a lo largo de la carrera.

En ese contexto, y a partir del análisis de las planificaciones anuales de las materias, se observó que en las evaluaciones predominaban los instrumentos tradicionales: pruebas escritas “multiple choice” de pocas preguntas, observaciones informales de las prácticas, y exámenes orales abiertos; y en algunos casos no se especificaba con claridad el sistema de evaluación formativa. Por otro lado, en las encuestas que completaron los estudiantes a través del SIU Guaraní al finalizar el cursado de cada asignatura, uno de los aspectos más cuestionados fue la falta de devoluciones en la mayoría de las evaluaciones (Di Bernardo y col, 2014).

Para resolver estas debilidades, algunas asignaturas como Medicina III (5° año) y Medicina Hombre y Sociedad (1° año), reestructuraron sus sistemas de evaluación, articulando diferentes instrumentos para evaluar todo el espectro de las competencias, mejorando además sus mecanismos de retroalimentación para fortalecer la evaluación formativa (Di Bernardo y col, 2009; Di Bernardo y col, 2017).

La mayoría de los instrumentos de evaluación utilizados en dichas materias han sido validados por diferentes autores en otros contextos, que además han reportado las propiedades psicométricas de cada uno. Sin embargo, no se ha estudiado el impacto educacional del conjunto de instrumentos aplicados por estas asignaturas en forma integrada a sus procesos formativos como sistemas de evaluación multimétodo, lo que constituye el propósito principal de este trabajo de investigación.

### **1.3. Supuestos básicos de la investigación**

Estudiar el impacto educacional que reciben los estudiantes de medicina a través de los sistemas multimétodo de evaluación de competencias es un desafío que demanda explorar, en el contexto de la carrera de medicina de la UNNE, las diferentes dimensiones y elementos abordados en el marco teórico referencial y en el marco metodológico de esta tesis. Su caracterización, análisis e interpretación se realizará a partir de la comprensión de las siguientes proposiciones:

- **El impacto educacional de las evaluaciones se refleja principalmente en el aprendizaje y su máxima expresión es “la evaluación como aprendizaje”.**

Se presupone que el concepto “impacto educacional” de las evaluaciones hace referencia esencialmente a los “efectos de aprendizaje” que tienen las evaluaciones (Al-Kadri et al. 2012; Cilliers et al. 2012; Heeneman et al. 2015), y está en consonancia con los efectos

“educativo” y “catalizador” enunciados en la Conferencia de Ottawa 2010 como criterios para una buena evaluación por sus beneficios sobre el aprendizaje (Norcini et al, 2011:211).

Además, dicho significado es concordante con las expresiones de otros autores que sostienen que en los currícula basados en competencias, la evaluación es el verdadero motor del aprendizaje (Schuwirth & Van der Vleuten, 2011b); (Ruiz de Gauna et al., 2015).

En consecuencia, esta investigación está dirigida a explorar e interpretar el impacto educacional de las evaluaciones, poniendo el foco especialmente en sus efectos sobre el aprendizaje de los estudiantes.

En la misma dirección, y desde la perspectiva de “la evaluación como aprendizaje” se buscará reconocer y caracterizar los procesos y escenarios curriculares que facilitan dicha perspectiva en el contexto de la carrera de medicina de la UNNE.

Para este abordaje se aplicará la conceptualización formulada por van der Vleuten et al (2017), de tres perspectivas diferentes en las prácticas de la evaluación de competencias: la evaluación **del** aprendizaje, la evaluación **para** el aprendizaje y la evaluación **como** aprendizaje. La primera está dirigida a verificar si los estudiantes lograron los objetivos de aprendizaje esperados o alcanzaron determinados estándares educativos y tomar decisiones correctas sobre ellos (ingreso, promoción, graduación, etc); la evaluación para el aprendizaje está centrada en los efectos de aprendizaje que tiene la evaluación y que permite guiar el aprendizaje de los estudiantes y hacerlo más eficaz; y la evaluación como aprendizaje que abarca las dos perspectivas anteriores integradas al proceso formativo, de tal manera que los programas de evaluación se planifican, implementan, gestionan, evalúan y adaptan a lo largo de todo el plan de estudios para conducir el aprendizaje de los estudiantes.

- **Los efectos de aprendizaje de las evaluaciones están condicionados por otros factores que son ajenos al instrumento o instancia de evaluación.**

Se presupone que los efectos que tienen las evaluaciones sobre el aprendizaje de los estudiantes no dependen solamente de las características de los instrumentos e instancias de evaluación aplicados, también intervienen otros factores que pueden impulsar u obstaculizar el aprendizaje, tales como: las metas de logros y los enfoques de aprendizaje de los estudiantes (Wilson & Fowler, 2005; Al Kadri et al., 2011; Cobb et al., 2013); los comportamientos (habilidades, hábitos y actitudes) de estudio (Credé & Kuncel, 2008; Dunlosky et al, 2013; Yan et al, 2014; Biwer et al, 2020); las habilidades de procesamiento cognitivo y de regulación metacognitiva (Van Merriënboer & Sweller, 2010; Cilliers et al.,

2012; Evans & Stanovich, 2013; de Bruin et al., 2019); la capacidad reflexiva y de autorregulación de los estudiantes (Sandars, 2009; Pelgrim et al., 2013; Wen et al, 2015; Uygur et al, 2019; Wood, 2019); y las percepciones que tienen los mismos sobre sus capacidades intelectuales y sobre las evaluaciones (Segers et al, 2006; Cilliers et al, 2010; Schut et al, 2018; Watling & Ginsburg, 2019); entre otros.

- **El estudio de casos como estrategia adecuada para abordar en profundidad los efectos que producen los sistemas de evaluación multimétodo sobre el aprendizaje.**

Se presupone que un estudio de casos múltiples con diseño incrustado (Tipo 4) y diferentes unidades de análisis (Yin, 2003) es la elección más apropiada para explorar en profundidad los efectos de aprendizaje de las evaluaciones en las dos asignaturas de la carrera de medicina de la UNNE que aplican sistemas de evaluación multimétodo, ya que permitirá hacer una indagación empírica amplia del fenómeno estudiado dentro de su contexto real (cada materia), teniendo en cuenta que *“los límites entre el fenómeno y el contexto no son claramente evidentes”*. (Yin, 2003:13)

Para el mismo propósito y desde el paradigma interpretativo (Bunniss & Kelly, 2010), se aplicará un enfoque metodológico mixto de complementación cuali-cuatitativa para ampliar el conocimiento de la realidad estudiada, pues *“no se trata de buscar convergencia ni confirmación de resultados, sino de contar simultáneamente con dos imágenes que enriquezcan nuestra comprensión de los hechos...”*. (Bericat, 1998:106)

#### **1.4. Supuestos de anticipación de sentido**

Por ser una investigación exploratoria e interpretativa de dos casos de estudio (asignaturas) de la misma carrera, se espera encontrar que:

- ✓ Los entornos de aprendizaje y evaluación serán diferentes en los casos de estudio por tratarse de asignaturas ubicadas en ambos extremos del plan de estudio (1° y 5° año).
- ✓ Los alumnos aspirantes tendrán metas de logro orientadas al rendimiento y enfoque de aprendizaje más superficial que los estudiantes avanzados, porque su motivación principal es extrínseca (ingresar a la carrera) y temen al fracaso.
- ✓ El nivel de conocimiento de los estudiantes sobre los objetivos y criterios de evaluación aplicados será el adecuado para afrontar las evaluaciones.

- ✓ Los procesos cognitivos que aplican los estudiantes en las distintas evaluaciones abarcarán niveles cognitivos más altos en los alumnos avanzados que en los principiantes, al igual que la calidad de las devoluciones recibidas.
- ✓ Los cambios en los comportamientos de estudio serán más significativos en los alumnos principiantes pues en general no tienen experiencias previas con la evaluación por competencias.
- ✓ A partir de sus experiencias de evaluación, tanto los alumnos principiantes como los estudiantes avanzados percibirán a las evaluaciones formativas como decisivas o punitivas, ya que es el enfoque tradicional en el que se han formado.
- ✓ Los efectos de aprendizaje de los sistemas de evaluación serán reconocidos tanto por los alumnos principiantes como por los avanzados, y la relevancia de los aprendizajes será valorada según el nivel de formación de los estudiantes en la carrera.

## 1.5. Propósitos y Objetivos de la Investigación

### Propósito general de la investigación

Comprender los factores, mecanismos y efectos de aprendizaje vinculados al impacto educacional de los sistemas multimétodo de evaluación de competencias en el contexto de las dos asignaturas que aplican dichos sistemas en la Carrera de Medicina de la UNNE.

Para hacer el desglose del propósito general se presentan a continuación los principales interrogantes que dan sentido a los objetivos específicos de este trabajo:

- **¿Cómo se integran las evaluaciones al entorno de aprendizaje de los estudiantes?**
- **¿Como influyen las evaluaciones sobre el aprendizaje de los estudiantes?**
- **¿Qué puede aportar cada momento de la evaluación al aprendizaje de los estudiantes?**

**Los objetivos específicos son:**

- 1) Caracterizar el entorno de aprendizaje y los sistemas e instrumentos de evaluación aplicados en las asignaturas que constituyen los casos de estudio.**

Este primer objetivo lleva a realizar una descripción detallada de las características del entorno de aprendizaje y de los sistemas de evaluación en cada caso de estudio, sus

fundamentos, propósitos, diseño, planificación y aplicación, en el contexto de cada materia, los instrumentos de evaluación utilizados, sus propiedades psicométricas y el alcance cognitivo de las tareas y consignas planteadas en las pruebas.

**2) Identificar las metas de logro y los enfoques de aprendizaje que tienen los alumnos ingresantes y los estudiantes avanzados de la carrera de medicina.**

Este objetivo está dirigido a los estudiantes evaluados (principiantes y avanzados) y abarca las metas de logro que tienen los mismos, los enfoques y estrategias de aprendizaje que utilizaron en el contexto del sistema de evaluación aplicado, y los resultados obtenidos.

**3) Determinar el nivel de conocimiento y el grado de interacción con los docentes, que tuvieron los estudiantes en cuanto a la planificación y aplicación de las evaluaciones.**

Este objetivo conduce a los factores de la evaluación y comprende los objetivos y criterios de las evaluaciones; el material evaluado; la interacción con los docentes; la proximidad temporal a la evaluación; y la sobrecarga de tareas simultaneas.

**4) Distinguir los procesos cognitivos y metacognitivos que usaron los estudiantes en las distintas evaluaciones y las características de las devoluciones recibidas.**

Este objetivo se enfoca en los efectos de aprendizaje y lleva a explorar la naturaleza de las actividades de procesamiento cognitivo y de regulación metacognitiva en cada evaluación (cantidad y distribución del esfuerzo, elección de recursos y contenidos; estrategias de monitoreo); y los mecanismos de las devoluciones.

**5) Identificar los cambios en los comportamientos de estudio que adoptaron los estudiantes en su preparación para las diferentes evaluaciones.**

Este objetivo conduce a la fase de pre-evaluación y abarca los comportamientos de estudio (hábitos, habilidades y actitudes) que los estudiantes aplicaron en su preparación para las evaluaciones, incluyendo gestión del tiempo, establecimiento de metas, selección de qué, cómo, y dónde estudiar; y los factores que motivaron y desanimaron el estudio.

**6) Caracterizar las experiencias que tuvieron los estudiantes dentro del entorno de aprendizaje y evaluación**

Este objetivo remite al proceso mismo de cada evaluación, y comprende las experiencias de los estudiantes en las evaluaciones, los beneficios obtenidos en su formación y

aprendizaje, y las oportunidades que tuvieron de mostrar su progreso y reflexionar sobre lo aprendido; así como los factores del sistema que facilitaron u obstaculizaron dichas oportunidades.

**7) Especificar las percepciones que tuvieron los estudiantes sobre el entorno de aprendizaje y evaluación, su autoconcepto académico y logros percibidos.**

Este objetivo lleva a indagar las percepciones de los estudiantes dentro del entorno de aprendizaje y evaluación, como perciben su capacidad intelectual para afrontar las evaluaciones y los riesgos percibidos (no regularizar o aprobar, bajo promedio, etc); y cuanto aportaron las evaluaciones al aprendizaje de la materia.

## **1.6. Estructura del informe de investigación**

El informe de este trabajo de tesis está estructurado en seis capítulos que describen todo el proceso de investigación. Luego de este primer capítulo introductorio continúan:

En el **Capítulo 2** se construye el Marco Teórico Referencial que abarca los conceptos medulares de la temática abordada en la investigación, se presenta el modelo de educación médica basado en competencias y los aspectos más relevantes de la evaluación por competencias, sus dimensiones e instrumentos y sistemas aplicados; se profundiza en el concepto de “impacto educativo” de la evaluación y caracterizan las principales variables motivacionales, cognitivas y afectivas, vinculadas al mismo; y se revisan las últimas evidencias internacionales y latinoamericanas relacionadas al tema de investigación.

En el **Capítulo 3** se presenta el Marco Metodológico, partiendo del posicionamiento paradigmático adoptado para esta investigación, se describen y fundamentan los diferentes elementos que componen el diseño metodológico elegido, para finalmente exponer la matriz analítica-metodológica construida, con los ejes, categorías y preguntas que organizaron y guiaron el proceso de investigación.

El **Capítulo 4** comprende la Presentación de los Resultados, comenzando con las características contextuales y particulares de los Casos de estudio, para luego mostrar los resultados obtenidos en cada uno de los tres ejes de análisis explorados en la investigación: “Evaluación del aprendizaje”, “Evaluación para el aprendizaje” y “Evaluación como aprendizaje”; y finalmente las diferencias entre los Casos de estudio.

En el **Capítulo 5** se Discuten los Resultados del estudio, siguiendo las preguntas derivadas en cada uno de los ejes y categorías de análisis que orientaron la investigación: entornos de aprendizaje y evaluación; estudiantes evaluados; factores de la evaluación; efectos de aprendizaje; y los aportes de las evaluaciones en sus tres momentos (antes, durante y después de la evaluación).

En el **Capítulo 6** se presentan las Conclusiones de la investigación, partiendo del diseño y elaboración de la tesis, recuperando e integrando los interrogantes que guiaron el proceso de investigación, y los resultados, interpretaciones y reflexiones que fueron emergiendo en los diferentes momentos del recorrido investigativo; para luego abordar las implicaciones de la investigación para la práctica docente y sus limitaciones.

## CAPITULO 2

MARCO TEÓRICO  
REFERENCIAL





## Capítulo 2: MARCO TEÓRICO REFERENCIAL

### 2.1. Educación por competencias<sup>1</sup>

La educación médica de las últimas cuatro décadas ha venido desarrollando el modelo de “educación basada en los resultados” (outcome-based medical education), ponderando el concepto de que el profesional que esperamos define su proceso de formación (no a la inversa) y esto está determinado por dos referentes externos: las necesidades de la sociedad y de los sistemas de salud (Harden et al., 1999). Al mismo tiempo, surgieron críticas de rechazo a este modelo por parte de algunos autores que afirmaron (citados por Harden et al, 1999:9) *“...hay que valorar la educación por sí misma, no porque conduce a algún resultado”; “...con estos planes, el contenido académico tradicional se omite o se entierra en un pantano de declamaciones pedagógica...”, “...definir la educación con un conjunto de resultados de aprendizaje preestablecidos, atenta contra los impredecibles y maravillosos viajes de exploración que caracterizan el aprendizaje a través del descubrimiento y la investigación”.*

En respuesta a estas críticas, Harden et al (1999) sostiene que estos cuestionamientos y conceptos liberales de la educación, son muy apropiados para las artes, las humanidades y otras disciplinas, pero la educación médica no puede ignorar el perfil del profesional que forma y por ende necesita planes de estudios que especifiquen claramente los resultados del aprendizaje.

En esa dirección, Shumway & Harden (2003) describieron 12 resultados de aprendizaje para asegurar la formación de un médico competente y reflexivo: Conocimientos Biomédicos; Manejo de pacientes; Investigación; Manejo de la información; Promoción de la salud y prevención; Destreza en procedimientos; Habilidades clínicas; Comunicación; Actitudes, Ética y responsabilidad legal; Toma de decisiones y razonamiento clínico; Desempeño profesional; Desarrollo personal.

Los “resultados de aprendizaje” son los organizadores curriculares de la Educación Médica Basada en Competencias (EMBC), modelo que se ha extendido y establecido sólidamente en numerosos países debido a que se asume, permite formar profesionales mejor capacitados para la práctica y solución de los problemas de salud (Morcke et al. 2013);

---

<sup>1</sup> Competencia: *“Uso habitual y juicioso de la comunicación, el conocimiento, las habilidades técnicas, el razonamiento clínico, las emociones, los valores y la reflexión en la práctica diaria, para el beneficio de los individuos y de la comunidad a la que se sirve”* (Epstein & Hundert, 2002:226)

facilitando la evaluación objetiva de las actividades propias del médico y su vinculación con las demandas del medio (Morales Castillo & Varela, 2015).

El concepto de competencia es polisémico y tiene múltiples acepciones porque abarca diferentes niveles del saber (saber hacer, saber ser, saber estar, y saber convivir), y los términos “competencias”, “competencia” y “competente”, muy utilizados en la EMBC, son aplicados con diferentes significados<sup>2</sup> (Frank et al, 2010a). No obstante, todos están referidos a los desempeños, al “saber hacer”, y en este “hacer” se articulan conocimientos, habilidades, actitudes y valores, que deben ser congruentes con el contexto y con la complejidad del problema que se atiende para que el actuar (hacer) sea responsable y efectivo. (Morales Castillo & Varela, 2015)

El concepto de Educación Médica Basada en Competencias más utilizado, fue desarrollado por Frank et al. (2010b-636), en una revisión sistemática de los términos existentes en la literatura de educación médica, como: *“un enfoque para la preparación de los médicos para la práctica, orientada fundamentalmente a lograr un graduado formado en las competencias derivadas de un análisis de las necesidades de la sociedad y de los pacientes; que pone menos énfasis en la formación basada en el tiempo y da mayor responsabilidad y flexibilidad, centrando el aprendizaje en el estudiante”*

El grupo de trabajo “Colaboradores Internacionales en EMBC” (Frank et al, 2010a) integrado por expertos en Educación Médica de USA, Canadá y Reino Unido, justifican y fundamentan la adopción de este modelo en cuatro argumentos centrales:

- ✓ **La formación debe focalizarse en los resultados a alcanzar**, para poder asegurar que todos los graduados sean competentes para el ejercicio profesional, y ser transparente en los objetivos y efectividad, con los alumnos, con los docentes y con el público.
- ✓ **El currículum debe estar centrado en las competencias a desarrollar**, que deben ser observables y estar relacionadas a las necesidades de la sociedad y del sistema de salud.

---

<sup>2</sup> **Competencias:** Son las capacidades observables del profesional de la salud que integran múltiples elementos: conocimientos, habilidades, valores y actitudes; y al ser desempeños observables, pueden evaluarse para asegurar su correcto desarrollo y ejecución. Las competencias se pueden organizar como bloques para facilitar el desarrollo progresivo. **Competencia:** Es el conjunto de habilidades en múltiples dominios o aspectos del desempeño de los médicos en un contexto determinado, por lo que requieren calificadores descriptivos para definir las capacidades pertinentes, el contexto y la etapa de formación. La competencia es multidimensional y dinámica, cambia con el tiempo, la experiencia y el escenario. **Competente:** Que posee las habilidades requeridas en todos los dominios en un contexto determinado en una etapa definida de la educación o de la práctica médica. (Frank et al, 2010a)

- ✓ **La formación (tradicional) estructurada en el tiempo debe ser abandonada**, pues los estudiantes aprenden a ritmos diferentes, y pueden desarrollar las competencias más rápido o más lento que el promedio de sus pares. Por ello hay que poner mayor énfasis en la progresión de las habilidades y en las medidas de rendimiento.
- ✓ **La educación médica debe estar centrada en el estudiante**, promoviendo su participación con metas claras a lo largo de la carrera, para permitirle regular su propio aprendizaje, y desarrollar las competencias esperadas.

Para van der Vleuten et al (2017), la EMBC cambia el modelo curricular tradicional “de entrada”, estructurado en horas de cursado para diferentes disciplinas, por un modelo “de salida” basado en un conjunto definido de resultados o competencias a desarrollar, algunas de las cuales van más allá del dominio del conocimiento y abarcan habilidades profesionales como ser capaz de trabajar en equipo, poder comunicarse, poder escribir académicamente y poder comportarse profesionalmente, que son muy relevantes para el desempeño futuro como profesional. Otro cambio importante de la EMBC es el didáctico, donde la educación cambia del aprendizaje atomista al holístico, del enfoque dominante en la enseñanza a métodos de aprendizaje más activos, del aprendizaje centrado en el docente al aprendizaje centrado y autodirigido por el estudiante.

En tal sentido, la aplicación de la EMBC ha sido para muchas carreras un gran desafío, por la complejidad de su modelo, por las dificultades que presenta su implementación y por el impacto que genera en las culturas y sistemas existentes. Sin embargo, la EMBC sigue creciendo e innovando, y la inclusión de “hitos” y “actividades profesionales confiables”<sup>3</sup> dentro de los marcos de competencias existentes habla de la evolución dinámica y de su potencial transformador de la Educación Médica. (Holmboe et al, 2017; Ferguson et al, 2017)

Por otro lado, la Educación Basada en Competencias también ha sido y es objeto de numerosas críticas, como las planteadas por Barnett (2001) quien considera que establecer las competencias de acuerdo a las necesidades del contexto laboral constituye una visión instrumental y reduccionista de la educación que hace peligrar los aspectos humanistas, y enfatiza que en estos modelos, el conocimiento está cada vez más divorciado de la acción y en consecuencia, tanto el conocimiento como la acción se empobrecen y se achican,

---

<sup>3</sup> **Actividades Profesionales Confiables - EPA** (Entrustable Professional Activity): Son unidades de práctica profesional, tareas o responsabilidades que los estudiantes deben realizar sin supervisión una vez que alcanzaron las competencias específicas suficientes. **Hitos** (Milestone): Son descriptores de comportamientos que marcan el nivel de desempeño que debe demostrar el estudiante en una competencia en un momento del proceso formativo. (Ten Cate, 2013)

convirtiendo el conocimiento en simple información y la acción, en simple trabajo. Por su parte, Mulder et al. (2007) en una revisión crítica sobre el modelo de competencia en la educación y la formación profesional en Inglaterra, Francia, Alemania y Países Bajos señala aspectos tales como la falta de una definición coherente del concepto de competencia; la ausencia de un vínculo directo entre competencia y desempeño; el riesgo de restar atención al conocimiento cuando la tarea está centrada en competencias; la imposibilidad de incluir en los programas de grado competencias que solo pueden aplicarse luego de la graduación; las dificultades e inconsistencias que se plantean en las evaluaciones cuando las competencias están formuladas en términos muy generales; la falta de correspondencia que existe entre las evaluaciones en contextos simulados y el desempeño en la realidad; y otras.

Sin embargo, como lo expresa Sanchez-Mendiola (2015:2), *“muchos de los ensayos y editoriales que cuestionan o atacan este modelo educativo exageran sobre el daño (potencial y real) que el uso de las competencias en educación médica puede causar en los educandos y educadores (...) El concepto original de EBC, que surgió en la industria como un intento de homogeneizar lo que eran capaces de hacer los graduados de carreras técnicas para ser más eficientes en la cadena productiva y ser más competitivos en la economía global, tiene poco que ver con el concepto moderno que se maneja en la literatura de educación médica, (...) en los que se plantean diversos argumentos que, parecen coherentes y cuando menos atendibles e interesantes”*.

Pero más allá de las ventajas que puede ofrecer este modelo, algunos autores se preguntan: ¿en qué medida la EBC garantiza que los estudiantes, al finalizar el cursado, logren las competencias definidas en el programa? Pues si no se dan estas garantías las competencias quedan como simples declaraciones de intenciones de los profesores. Por eso es importante arbitrar todos los medios y estrategias necesarias para lograr que la “intenciones de aprendizaje” se conviertan en “resultados de aprendizaje” y aquí toma fuerte protagonismo la evaluación (Planas-Coll, 2013).

## **2.2. Evaluación de competencias:**

La evaluación es un aspecto clave en todo curriculum por competencias, y aunque ha sido el área más explorada de la Educación Médica en las últimas décadas, y ha crecido y desarrollado mucho más que en cualquier otra disciplina (van der Vleuten, 2016), sigue siendo tema de discusión y reflexión entre los principales autores que se preguntan: *¿Estamos evaluando correctamente las competencias y con los instrumentos adecuados?(...)*

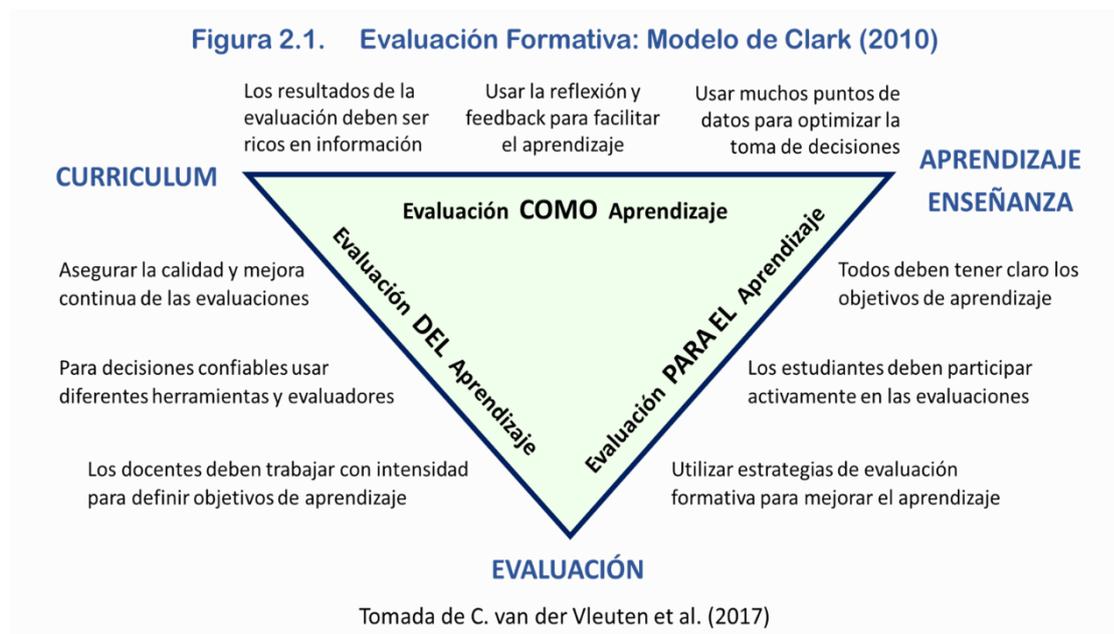
*¿Hemos establecido programas integrados y progresivos de evaluación de competencias o seguimos evaluando desde nuestros compartimentos estancos? (...) ¿Estamos seguros de que una evaluación clínica objetiva estructurada al final de las prácticas tuteladas asegura de forma inequívoca que los graduados han adquirido de manera satisfactoria todas las competencias previstas?” (Palés-Argullós, 2016a:110)*

Para dar respuesta afirmativa a estas preguntas y dado que las competencias tienen un desarrollo progresivo durante el programa de formación, el sistema de evaluación debería estar configurado a lo largo de la carrera para que todo el cuadro de competencias pueda ser evaluado con instrumentos adecuados al nivel de formación de los estudiantes, aplicando criterios claramente establecidos y métodos combinados cuidadosamente elegidos, lo que es válido también para programas educativos tradicionales, que las subdividen en conocimientos, habilidades y actitudes (Schuwirth et al., 2011b).

Al respecto, profesores de cinco Facultades de Medicina de las Universidades públicas de Cataluña consensuaron estas recomendaciones para la evaluación de competencias: evaluar de forma integrada sus componentes (conocimientos, habilidades y actitudes); los instrumentos deben ser congruentes con el diseño curricular, con la metodología docente y con el tipo de competencias a evaluar; utilizar como referente criterios de evaluación claros y acordes al nivel de logro esperado; el estudiante tiene que tener un papel activo en su evaluación (reflexión, autoevaluación y/o co-evaluación); y la evaluación debe ser progresiva a largo de la carrera. (Carreras Barnés et al, 2009)

La secuencia y el momento de aplicación de las evaluaciones requiere una planificación cuidadosa, especialmente cuando se busca reflejar las trayectorias de aprendizaje de los estudiantes individuales en un programa educativo, pues el conocimiento, las habilidades y los comportamientos evolucionan con el tiempo, pero la competencia se puede lograr antes del punto final del programa. Sobre esta cuestión hay dos enfoques diferentes. El primero (más tradicional), consiste en “calibrar” las evaluaciones según los resultados de aprendizaje esperados para cada etapa/fase de un programa, lo que permite seguir el desarrollo evolutivo de la competencia. El segundo enfoque es “calibrar” la evaluación para evaluar los resultados del aprendizaje, con la expectativa que al final del programa el estudiante alcance el nivel más alto de competencia (punto final), lo que asegura la preparación para la práctica independiente. (Norcini et al, 2018)

Desde otra óptica, van der Vleuten et al (2017), describen tres perspectivas en las prácticas de la evaluación de competencias: la evaluación **del** aprendizaje, la evaluación **para** el aprendizaje y la evaluación **como** aprendizaje. Figura 2.1.



- La **evaluación del aprendizaje** está relacionada al termino “evaluación sumativa” y pone la atención en la toma de decisiones adecuadas sobre los estudiantes evaluados (verificar si alcanzaron los estándares educativos y tomar decisiones correctas sobre ellos). Por eso en esta evaluación se aplican varios instrumentos de evaluación, cada uno de los cuales evaluará la competencia con diferentes niveles de autenticidad, cubriendo entre todos, las cuestiones psicométricas necesarias (confiabilidad, validez, objetividad, etc.)
- La **evaluación para el aprendizaje**, también llamada “evaluación formativa”, está centrada en los efectos de la evaluación sobre el aprendizaje (proporcionar retroalimentación oportuna y significativa, promover estrategias de aprendizaje profundo, apoyar el aprendizaje continuo) lo que permite guiar el aprendizaje de los estudiantes y hacerlo más eficaz.
- La **evaluación como aprendizaje**, es la que fundamenta a la “evaluación programática” donde las dos perspectivas mencionadas anteriormente se integran al proceso formativo, de tal forma que los programas de evaluación se planifican, implementan, gestionan, evalúan y adaptan a lo largo de todo el plan de estudios para conducir el aprendizaje de los estudiantes. Esto permite además superar las limitaciones individuales de cada método de evaluación.

### 2.2.1. Instrumentos para evaluar competencias

En relación a este tema, todos los autores coinciden que para evaluar adecuadamente las competencias es necesario articular diferentes instrumentos de evaluación, para cubrir todos los aspectos de la actividad educativa; que estén integrados en un programa; y sean congruentes con las tareas programadas. Cada instrumento debe ser pertinente con el objeto a evaluar y los criterios de evaluación deben ser anticipados y compartidos con los estudiantes. En consecuencia, la elección del instrumento de evaluación depende en primer lugar de lo que se pretende evaluar.

Para evaluar **conocimientos declarativos** se utilizan mucho las pruebas estructuradas con *preguntas de opción múltiple* (POM) ricas en contexto<sup>4</sup>, basadas en viñetas o problemas clínicos, que permiten evaluar comprensión y aplicación de conocimientos, reconocimiento de patrones, y razonamiento clínico. Este tipo de prueba, por su alta confiabilidad, es preferible a las preguntas de ensayo u otro formato abierto como las de respuesta corta, salvo en situaciones en las que la generación espontánea de la respuesta es un aspecto esencial de la pregunta. (Schuwirth & Van der Vleuten, 2004), (Palés-Argullós, 2010).

Para evaluar **razonamiento clínico y habilidades para resolver problemas**, además de las POM, se aplican también otros instrumentos como: *“Problemas de manejo de pacientes”* (PMP)<sup>5</sup>, *“Problemas Basados en Puntos Clave”* (KFP)<sup>6</sup> y *“Preguntas de Emparejamiento Ampliado”* (EMQ), estas herramientas tienen en común que describen casos cortos, e indagan sobre aspectos esenciales, características claves o decisiones importantes

---

<sup>4</sup> Las preguntas de opción múltiple “ricas en contexto”, son preguntas que están directamente relacionada con un caso o escenario clínico y permiten evaluar más efectivamente el razonamiento clínico y las competencias del nivel “sabe cómo”; a diferencia de las preguntas “pobres en contexto”, que evalúan solo conocimiento de tipo factual (nivel “sabe”) más que el razonamiento, aunque pueden ser valiosas para la comprensión de los conceptos. (Schuwirth & Van der Vleuten, 2004)

<sup>5</sup> Problemas de manejo de pacientes - PMP (Patient Management Problem) Es un método escrito que permite evaluar la capacidad de resolución de casos clínicos, dificultades diagnósticas y/o decisiones terapéuticas. Pero dado que tiene el inconveniente de requerir mucho tiempo para cada problema, se le han hecho algunas modificaciones para mejorar su validez. (Farmer & Page, 2005)

<sup>6</sup> Problemas Basados en “Puntos Clave” - KFP (Key Features Problem): Es una modificación del PMP donde se presenta un problema clínico con información para interpretar y se pide a los estudiantes identificar características claves o decisiones esenciales (no rutinarias). Las respuestas pueden tener diferentes formatos (respuesta corta, POM, EMQ), y en algunos diseños las elecciones erróneas restan puntaje. La principal ventaja de esta innovación es que permite presentar muchos más casos clínicos en un tiempo menor, en comparación con el PMP convencional. (Farmer & Page, 2005)

(Farmer & Page, 2005). Otro instrumento de mucha utilidad para estudiantes avanzados y graduados, es el “*Test de Concordancia Script*” (SCT) <sup>7</sup> cuyo objetivo principal no es medir conocimientos, sino la organización de ese conocimiento clínico en redes conceptuales (scripts), que se revelan en la acción y en la toma de decisiones ante una situación clínica específica (Fournier et al, 2008).

Para evaluar **habilidades clínicas y comunicacionales** se utilizan evaluaciones basadas en la observación con instrumentos acordes al contexto, así, en los escenarios clínicos reales y lugares de trabajo asistencial, además del tradicional “*Caso largo*” <sup>8</sup> (Ponnamperuma, 2009), los más utilizados son el “*Mini-ejercicio de Evaluación Clínica*” (Mini-CEX)<sup>9</sup> (Norcini, 2005) y la “*Observación directa de habilidades procedimentales*” (DOPS)<sup>10</sup> (Norcini & Burch, 2007); por otro lado, en los escenarios estandarizados, las prácticas con pacientes simulados o con simuladores tecnológicos se evalúan a través del “*Examen Clínico Objetivo Estructurado*” o ECOE <sup>11</sup> (Newble, 2004). Todas estas pruebas se realizan aplicando

---

<sup>7</sup> **Test de Concordancia Script** - SCT (Script Concordance Test): La prueba comprende viñetas clínicas que describen situaciones problemáticas de la práctica; hipótesis diagnósticas o terapéuticas concordantes con cada situación planteada; y aporta nueva información relevante para cada hipótesis. El examinando debe responder, cuál es el efecto de esta nueva información sobre cada una de las hipótesis (si la descarta, no la modifica o la confirma). (Fournier et al, 2008)

<sup>8</sup> El **Caso Largo** es un método tradicional de evaluación con numerosas variantes actuales, que consiste básicamente en asignar al estudiante un paciente ambulatorio o internado en un centro asistencial, para que en un tiempo determinado realice una entrevista y/o examen físico y luego presente sus hallazgos y planes a los evaluadores, quienes preguntan sobre el paciente y temas relacionados, para juzgar la calidad del desempeño clínico del estudiante. (Ponnamperuma, 2009)

<sup>9</sup> **Mini-ejercicio de Evaluación Clínica** - mini CEX (Clinical Evaluation Exercise) En esta prueba, los estudiantes son observados en acción con un paciente, con el observador presente, aunque también se pueden usar cámaras en directo o el video de la entrevista. Permite evaluar las habilidades clínicas para la entrevista, la exploración física y las habilidades de comunicación. Por su variabilidad intra e interobservador, se aplican formularios específicos para mejorar su confiabilidad. (Norcini, 2005)

<sup>10</sup> **Observación directa de habilidades procedimentales** - DOPS (Direct observation of procedural skills): Se enfoca en evaluar las habilidades de los estudiantes observándolos durante la realización de procedimientos médicos en escenarios clínicos reales y aplicando un formulario específico para valorar el desempeño. Abarca un tiempo de observación aproximado de 15 minutos y luego 5 minutos para la retroalimentación. (Norcini & Burch, 2007)

<sup>11</sup> El **Examen Clínico Objetivo Estructurado** (ECOE) es un método de evaluación estandarizado y en contexto simulado, que comprende diferentes estaciones (mínimo 10) donde los estudiante deben interactuar con simuladores tecnológicos, entrevistar a pacientes estandarizados, interpretar estudios complementarios u otras situaciones de la práctica clínica, mientras son observados en forma directa o con cámara gesell por los evaluadores que utilizan listas de cotejo o formularios globales para evaluar el desempeño clínico de los mismos. (Newble, 2004).

instrumentos de observación estructurados como rúbricas<sup>12</sup> y listas de cotejo<sup>13</sup> (López Cabrera y Silva López, 2017) y para que tengan impacto formativo, deben complementarse con una oportuna y adecuada “*devolución constructiva*”, definida como el acto de dar información a un estudiante describiendo su desempeño en una situación clínica observada y conforme a un patrón o estándar pre-establecido. Esto requiere que el estudiante esté dispuesto a exponer sus debilidades y errores y sea permeable a una crítica constructiva. (Alves de Lima, 2008; Lefroy et al, 2015)

Para evaluar las **actitudes y conductas profesionales** se aplica la observación de las conductas de los estudiantes en diferentes ámbitos y a lo largo de la carrera, aplicando tanto instrumentos formales como no normalizados. Pero también son muy valiosos los “*portfolios*” que son herramientas muy utilizadas en la educación de profesionales de la salud con fines de aprendizaje, evaluación o ambos. Pueden ser en papel o digitales y su contenido varía con el propósito del mismo. Hoy se prefieren los portfolios electrónicos (e-portfolios) que permiten insertar hipervínculos para facilitar la lectura y búsqueda de los materiales, y para establecer conexiones entre las evidencias, las reflexiones y las imágenes que pueden almacenarse en otro lugar. (Van Tartwijk & Driessen, 2009)

Cuando se trata de **evaluar una práctica profesional**, Ruiz de Gauna y col (2015) proponen abordarla desde dos enfoques complementarios: 1) durante la acción, por observación directa de uno o más observadores, que incluye el preguntar al actor para obtener una mayor comprensión del aprendizaje; y 2) después de la acción (resultados), que puede ser efectuada por el mismo actor (autoevaluación reflexiva), por los informes de los tutores o docentes y/o a través de registros de la práctica clínica.

Todos los instrumentos de evaluación mencionados anteriormente tienen aplicación en uno o más de los diferentes niveles de desarrollo de las competencias. Para explicarlo mejor vamos a utilizar la clásica Pirámide propuesta por George Miller (1990) para graficar el desarrollo creciente de las competencias profesionales, que permite comprender mejor el proceso de formación en competencias y seleccionar las herramientas más apropiadas para evaluar cada nivel.

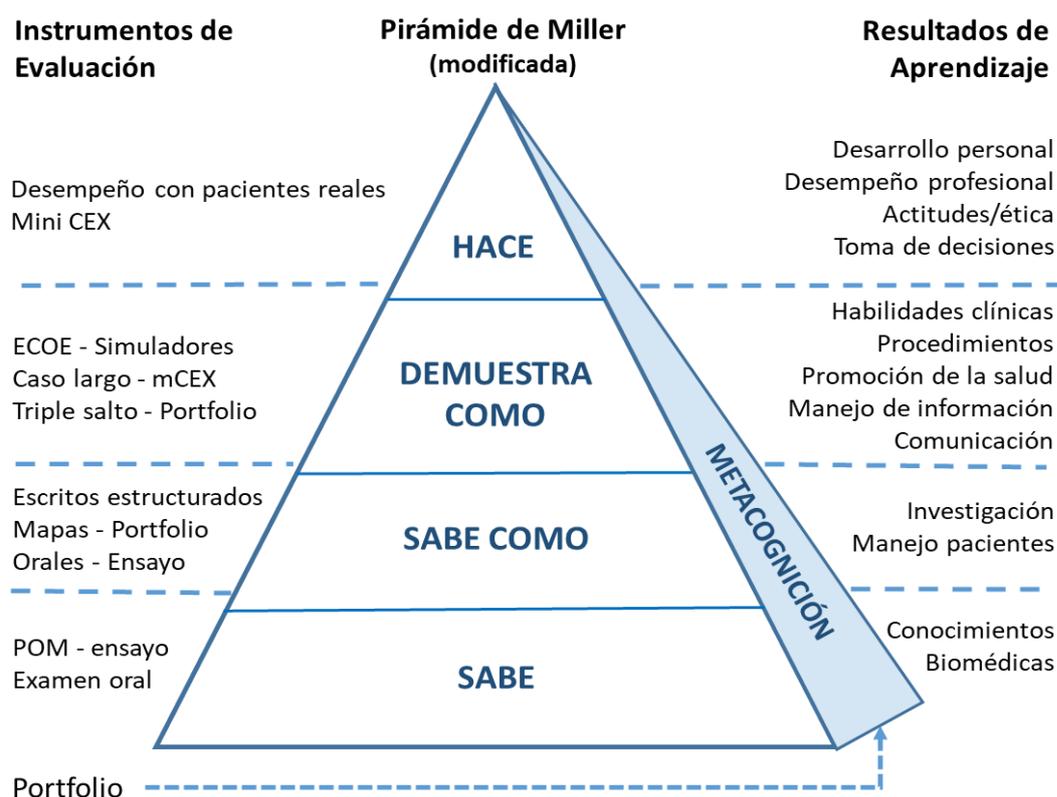
---

<sup>12</sup> **Rúbricas:** guías precisas que valoran el aprendizaje y los productos realizados por medio de tablas que desglosan los niveles de desempeño de los estudiantes en un aspecto determinado, empleando criterios específicos sobre rendimiento.

<sup>13</sup> **Lista de cotejo:** herramienta que permite realizar observaciones del desempeño de un estudiante mediante parámetros donde una actividad ha sido realizado o no.

Esta gráfica (Figura 2.2.) tiene como base los aspectos cognitivos del estudiante y en el vértice el desempeño en la práctica profesional. En los dos primeros niveles (sabe y sabe cómo) se evalúan habilidades cognitivas y en las dos capas superiores (demuestra cómo y hace) se evalúa el desempeño conductual.

**Figura 2.2.:** Instrumentos de evaluación y resultados de aprendizaje ensamblados a la Pirámide de Miller (traducido y modificado de Shumway & Harden 2003).



La base de la pirámide representa el nivel "sabe" donde la evaluación está centrada en explorar los conocimientos biomédicos declarativos a través de pruebas escritas u orales en sus diferentes formatos. El segundo nivel es el denominado "sabe cómo", donde se evalúa la capacidad del estudiante para utilizar el conocimiento, aplicarlo a casos clínicos, razonar y/o resolver problemas, aquí los métodos de evaluación pueden ser similares al primer nivel en cuanto a formato pero deben ser ricos en contexto (casos, problemas, escenarios) pues se evalúan habilidades cognitivas de orden superior (análisis, síntesis y evaluación), ya que se explora principalmente el razonamiento y por eso se aplican también mapas conceptuales, informes y ensayos. En el siguiente nivel, "demuestra cómo", las conductas y habilidades se evalúan por observación en escenarios estandarizados, como el ECOE donde los estudiantes deben interactuar con simuladores tecnológicos, entrevistar pacientes estandarizados,

interpretar estudios complementarios u otras situaciones de la práctica clínica, mientras son observados en forma directa o con cámara gesell por los evaluadores que utilizan listas de cotejo o formularios globales para evaluar el desempeño clínico de los mismos. En el último nivel de competencia, el "hace", se evalúa el desempeño conductual en entornos asistenciales y laborales de la vida real, y esto puede realizarse de tres maneras: en forma directa "juzgando el desempeño observado" a través de observaciones no normalizadas o aplicando instrumentos como miniCEX y DOPS; en forma indirecta "evaluando informes, registros y productos del trabajo"; y combinando múltiples fuentes o diferentes evaluadores como la "evaluación 360 grados" que recopila datos de desempeño del estudiante aportados por diferentes actores del contexto o la "evaluación mini-pares" (mini-PAT) realizada por varios evaluadores seleccionados por los estudiantes. (Norcini & Burch, 2007)

A la gráfica original de la Pirámide de Miller algunos autores le agregaron otro plano, el de las "habilidades metacognitivas", que para Durante (2005-21) *"define la percepción que el individuo tiene de por qué sabe lo que sabe"* y sin las cuales se vuelve dificultoso seguir aprendiendo. Para evaluar esta dimensión el instrumento más utilizado es el Portfolio donde se pueden incluir, las evidencias del trabajo realizado por el estudiante, sus reflexiones sobre su aprendizaje, la retroalimentación recibida, los avances realizados y los planes para mejorar sus competencias; lo que le otorga valor tanto para el aprendizaje como para la evaluación. (Van Tartwijk & Driessen, 2009)

En relación al gráfico de Miller, Van der Vleuten et al (2017) remarcan que los instrumentos de evaluación que se utilizan en los tres primeros niveles de la pirámide son mayoritariamente pruebas estandarizadas y/o estructuradas, mientras que en el último nivel la evaluación es no estandarizada. Esta diferencia es importante pues en las evaluaciones estandarizadas, todas las condiciones de evaluación son las mismas (o muy similares) para todos los estudiantes examinados, incluso en las evaluaciones más auténticas del tercer nivel, esto permite que los resultados sean más confiables y fácilmente comparables. Por el contrario, cuando se evalúa en la práctica real, la estandarización es imposible y las evaluaciones se basan en juicios más subjetivos, a menudo más holísticos, con lo que se enriquece mucho el proceso de evaluación.

Otro aspecto relevante en cuanto a los instrumentos aplicados, es el desarrollo que ha tenido la simulación clínica en los últimos años, pues se ha instalado y expandido como herramienta didáctica para la educación de los profesionales de la salud, en respuesta a los cambios producidos en los escenarios de práctica clínica. El creciente número de estudiantes

en los centros asistenciales, la estadía cada vez más breve de los pacientes internados, la mayor severidad clínica de los enfermos hospitalizados, los aspectos asistenciales ético-legales y las normas para proteger la seguridad del paciente, entre otras, hacen cada vez más difícil que los estudiantes puedan aprender y formarse exclusivamente en escenarios clínicos reales. Por otro lado, estos espacios de práctica constituyen una tensión para el docente clínico pues deben centrar su atención en el desempeño de los estudiantes y al mismo tiempo cuidar que no se perjudique o incomode a los pacientes. (Motola et al, 2013)

La simulación clínica permite representar experiencias reales de pacientes a través de escenarios adecuadamente diseñados, controlados y estandarizados. Los formatos más utilizados son: la simulación tecnológica (máquinas y maniqués de baja y alta fidelidad) y la simulación humana (paciente simulado o estandarizado) que es una persona que no padece una enfermedad, sino que la simula o actúa para fines docentes, siendo capacitado para aportar a la docencia o evaluación desde el rol de paciente. (Moore y col, 2016)

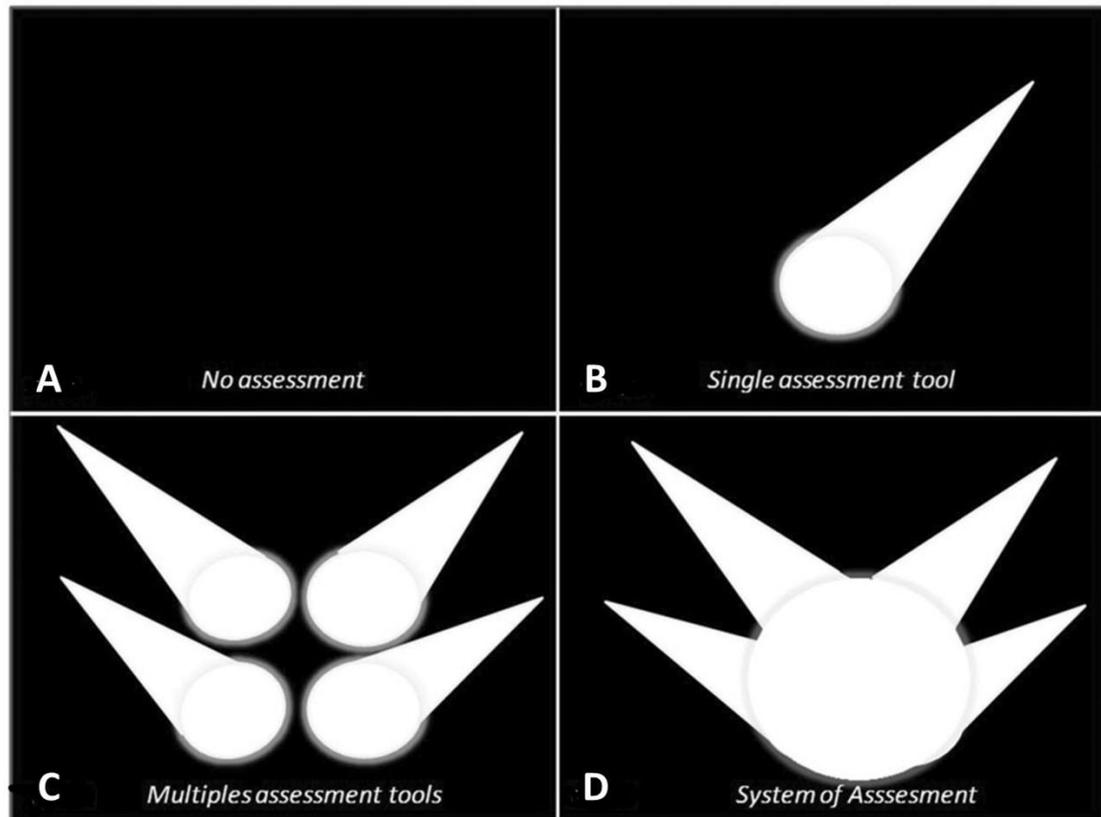
### **2.2.2. Sistemas de evaluación de competencias**

La educación y la práctica de los profesionales de la salud requieren múltiples habilidades cognitivas, psicomotoras, actitudinales y relacionales que los métodos únicos de evaluación no pueden capturar en su totalidad, y por lo tanto requieren la aplicación de múltiples evaluaciones complementarias. Sin embargo, estas evaluaciones a menudo se aplican de forma aislada o de manera no coordinada para llegar a una decisión global basada en ponderaciones dictadas por la tradición. En contraste con este último formato surgieron los “sistemas de evaluación”, que combinan explícitamente evaluaciones individuales ensambladas para lograr los diferentes propósitos y para abordar las múltiples necesidades de las partes interesadas. (Norcini et al, 2018)

Las modalidades de evaluación que se aplican en todo el mundo fueron definidas e ilustradas en la Conferencia de Ottawa 2018 como se muestran en la Figura 2.3.

El cuadro A representa a los programas educativos que por diferentes razones no cuentan con evaluación. El cuadro B muestra una situación muy común, donde por facilidad o practicidad solo se evalúa un aspecto de la competencia. El cuadro C muestra lo que ocurre en muchas instituciones que aplican una serie de evaluaciones dirigidas a diferentes dimensiones de las competencias, pero con limitada integración que conduce a lagunas en lo que cubren. El cuadro D se acerca más a un sistema de evaluación que funciona bien, pues ofrece una mejor cobertura del universo de contenidos y el uso más eficiente de los recursos.

**Figura 2.3:** Diferentes estados de evaluación en los programas de educación médica



Fuente: tomado de Norcini et al (2018)

### 2.2.3. Calidad de las evaluaciones

#### 2.2.3.1. Calidad de los instrumentos de evaluación

Los criterios para una buena evaluación, aplicables a una sola evaluación o proceso de evaluación de un solo propósito, fueron consensuados en la Conferencia de Ottawa 2010 (Norcini et al, 2011), y más adelante, revisados en la Conferencia de Ottawa 2018 (Norcini et al, 2018) y reformulados por el mismo grupo de trabajo como “marco de una buena evaluación” con los siguientes elementos:

- **Validez (coherencia):** los resultados de la evaluación son apropiados para su propósito. El instrumento mide lo que se supone que mide, en una situación específica y con un grupo de estudiantes específico.
- **Confiabilidad (consistencia):** la evaluación arroja un resultado consistente (reproducible) cuando se repite en circunstancias similares. Es un concepto estadístico, que indica el grado en el cual las puntuaciones de los estudiantes serían similares si fueran examinados de nuevo.

- **Equivalencia:** la misma evaluación produce puntuaciones o decisiones equivalentes cuando se administran en diferentes instituciones o ciclos de pruebas.
- **Factibilidad:** la evaluación es práctica, realista, sensata y acorde a las circunstancias y al contexto.
- **Efecto (impacto) educativo:** la evaluación motiva a quienes la toman a prepararse de una manera que tenga un beneficio en el aprendizaje. Por ello las tareas solicitadas en la misma deben ser lo más cercanas posibles a las que los estudiantes realizarán en la práctica real.
- **Efecto catalizador:** La evaluación proporciona resultados y comentarios (retroalimentación) de una manera que motiva a las partes interesadas a crear, mejorar y apoyar la formación; impulsa el aprendizaje futuro y mejora la calidad general del programa.
- **Aceptabilidad:** las partes interesadas encuentran que el proceso de evaluación y los resultados son creíbles.

Estos elementos del marco de referencia no se aplican de manera universal e igualitaria a todas las evaluaciones, ya que la importancia de cada elemento depende principalmente del contexto, de los actores y del propósito la prueba. Cuando se trata de una **evaluación formativa**, que es de bajo riesgo porque su objetivo es estimular el aprendizaje, los elementos más importantes son, el “efecto educativo” y el “efecto catalizador”, y funciona mejor cuando está integrada en el proceso formativo, brinda retroalimentación específica y procesable, es continua y es oportuna; la validez también es importante, pero los otros elementos no son esenciales. Por el contrario, cuando es una **evaluación sumativa**, que suele ser de riesgo medio o alto porque está destinada a tomar decisiones importantes (selección de postulantes, ingreso, promoción, graduación), los elementos esenciales son: validez, confiabilidad y equivalencia, pues requiere material de prueba coherente y de alta calidad, un proceso sistemático basado en estándares y una administración segura; los otros elementos son relativamente menos importantes. (Brailovsky, 2001; Norcini et al, 2018)

### 2.2.3.2. Calidad de los sistemas de evaluación

El marco para una buena evaluación aplicable a un sistema de evaluación, fue consensado en la Conferencia de Ottawa 2018 y abarca los siguientes elementos:

- **Coherente:** El sistema se compone de múltiples evaluaciones individuales coordinadas y actuaciones independientes ordenadas y alineadas en torno a los mismos propósitos.

- **Continuo:** El sistema de evaluación es continuo y los resultados individuales contribuyen acumulativamente a los propósitos del sistema.
- **Integral:** El sistema de evaluación es inclusivo y efectivo, consta de componentes que son formativos, diagnósticos y/o sumativos según corresponda a sus propósitos. Algunos o todos los componentes son auténticos e integradores.
- **Factible:** El sistema de evaluación y sus componentes son prácticos, realistas, eficientes y sensatos, dados los propósitos, las partes interesadas y el contexto.
- **Orientado a propósitos:** El sistema de evaluación respalda los propósitos para los que fue creado.
- **Aceptable:** Las partes interesadas en el sistema encuentran que el proceso de evaluación y los resultados son creíbles y están basados en evidencia.
- **Transparente y libre de prejuicios:** las partes interesadas entienden el funcionamiento del sistema y se minimizan sus consecuencias no deseadas. Las decisiones son justas y equitativas.

La calidad de los sistemas de evaluación depende en gran medida de la calidad de los componentes individuales, pero también depende de las propiedades emergentes del sistema que surgen de la planificación temporal de los componentes individuales (cantidad y relación entre los mismos) y de su impacto sobre el proceso formativo y los resultados de aprendizaje. (Norcini et al, 2018)

Los sistemas de evaluación son mucho más que combinar puntajes a lo largo del tiempo para tomar una decisión de promoción o graduación; los propósitos del sistema deben ser claros y coherentes con la visión/misión del programa al que sirve, debe estar alineado con el plan de estudio y con los resultados de aprendizaje esperados. (Moore et al, 2009; Nolla-Domenjó, 2009; Schuwirth et al., 2011b; Norcini et al, 2018)

### 2.3. Impacto educativo de la evaluación

El impacto educativo o educacional de una evaluación o sistema de evaluación, hace referencia a los efectos positivos o negativos (estimulo o inhibición) que puede tener una evaluación sobre el aprendizaje y la formación de los estudiantes (Newble & Jaeger, 1983; McManus et al., 1998; McLachlan 2006; Segers & Dochy, 2006; Cilliers et al., 2010; Al-Kadri et al., 2012; Cobb et al., 2013; Heeneman et al., 2015)

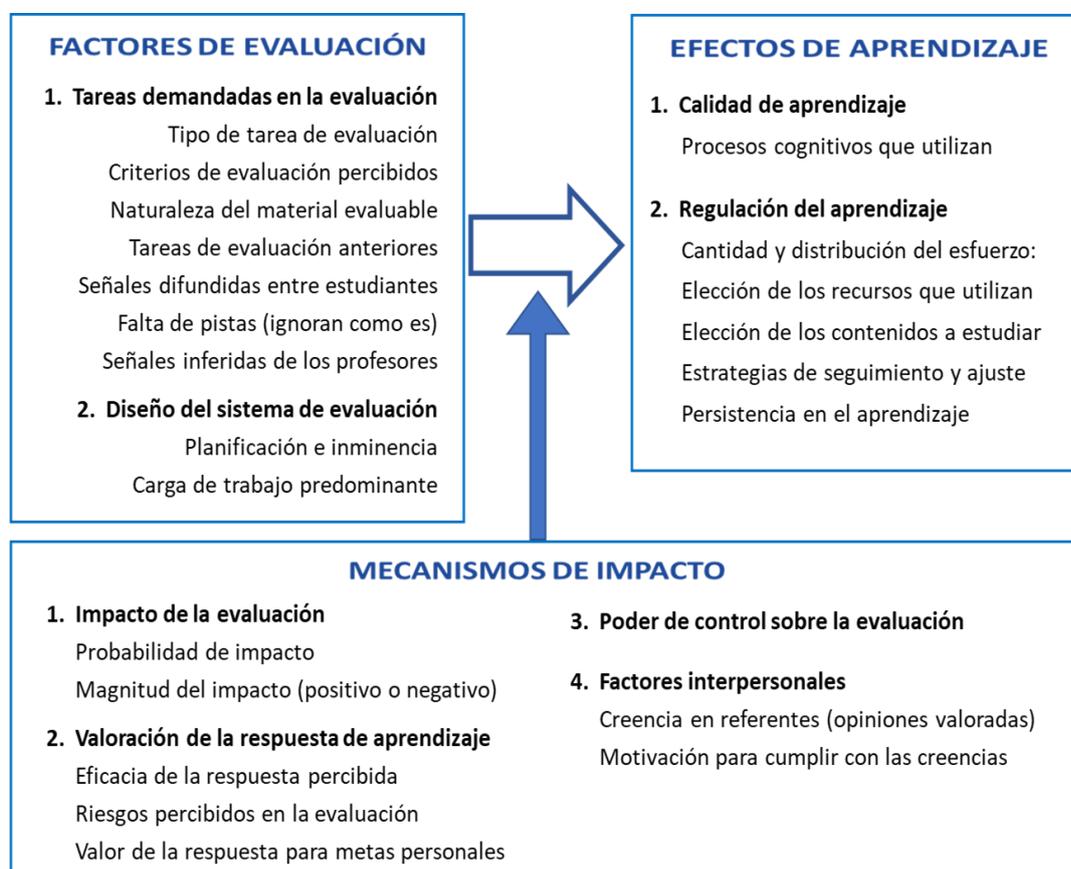
Estos efectos de la evaluación no se reflejan solamente en la cantidad y/o calidad del aprendizaje sino también en los cambios de aptitud de los estudiantes y en el desarrollo de

habilidades para el aprendizaje autodirigido (Monroe, 2016). Aunque no hay evidencias concluyentes de que las evaluaciones puedan impactar positivamente sobre el desempeño clínico de estudiantes y graduados (Miller & Archer, 2010).

Los investigadores que estudiaron el impacto educacional de la evaluación han diferenciado estos efectos en tres momentos: 1°- antes de la evaluación, que se refleja en el comportamiento y los hábitos de estudio que aplican los estudiantes cuando se preparan para un evento de evaluación; 2°- durante la evaluación, que está condicionado por las características de la prueba y evidenciado por los resultados de la misma y/o en el contenido de los portfolios; y 3°- después de la evaluación, manifestada durante el feedback y en los logros de aprendizaje. (Cilliers et al., 2012a).

Estudios conjuntos realizados por las Universidades de Maastricht (Holanda) y Stellenbosch (Sud Africa) posibilitaron el desarrollo de un modelo teórico para explicar los efectos de aprendizaje previos a la evaluación sumativa, y sus mecanismos. Analizaron los “factores de evaluación”, los “efectos de aprendizaje” y los “mecanismos de impacto”, estos componentes son explicitados en la Figura 2.4. (Cilliers et al., 2012b).

**Figura 2.4:** Modelo de los efectos de aprendizaje pre-evaluación de la evaluación sumativa



Fuente: Traducido de Cilliers et al (2012b)

Para medir el impacto educacional de la evaluación, varios autores analizaron los enfoques de aprendizaje (profundo, superficial y de logro), cada uno de los cuales está fuertemente asociado con una forma de motivación y una estrategia de aprendizaje. (Pintrich, 2004; Heijne-Penninga et al, 2010; Stegers-Jager et al, 2012; Cobb et al., 2013).

Por otro lado, el impacto de cualquier sistema de evaluación sobre el aprendizaje está mediado por las percepciones de cada estudiante, por ello es fundamental conocer estas percepciones para comprender qué significado tienen para los estudiantes, y cómo y por qué estas evaluaciones permiten o no, que los mismos puedan optimizar y autorregular su aprendizaje a partir de las evaluaciones (Eva et al, 2012; Bowen et al, 2017; Schut et al, 2018). En los currícula basados en competencias, el impacto educacional de la evaluación es un aspecto esencial, pues en este modelo la evaluación es principalmente formativa y constituye el verdadero motor del aprendizaje. Por ello los autores recomiendan que en este contexto debe evaluarse no solo “EL” aprendizaje sino también “PARA” el aprendizaje, desde el inicio y a lo largo de todo el proceso formativo. Esto facilitará que el estudiante sea consciente de su nivel de desarrollo de las competencias y de su progreso, y le permitirá tomar las medidas necesarias para mejorar su formación (Schuwirth & van der Vleuten, 2011b; Dannefer, 2013; Ruiz de Gauna et al., 2015).

Este enfoque de evaluación “para” el aprendizaje es la base conceptual de la “evaluación programática”, una visión integral de la evaluación que considera como objetivo el programa completo de evaluación y donde cada instancia evaluativa es solo un punto de información, que si bien es insuficiente para decidir si un estudiante aprueba o no, es muy valioso para proporcionar feedback (cuantitativo o cualitativo) relevante para el aprendizaje. Las decisiones sobre el progreso del estudiante se hacen cuando se han recogido suficientes datos y la información es clara y completa. La toma de decisiones se realiza a través de comités y se toman las medidas necesarias para que estas decisiones sean sólidas. (Heeneman et al., 2015; van der Vleuten & Heeneman, 2016; Schuwirth et al., 2017a)

La “evaluación programática” es modular (en bloques) y también longitudinal. La evaluación de cada bloque abarca distintos momentos y utiliza múltiples herramientas. La evaluación longitudinal es cognitiva y conductual, y el portfolio es la principal herramienta. Cada estudiante tiene un “mentor” (tutor) que lo guía durante toda la carrera y tiene acceso a su portfolio, los tutores son entrenados para su rol y no tienen voz en las decisiones del comité, esto permite proteger y sincerar la relación estudiante-mentor. El comité utiliza criterios narrativos para juzgar el portfolio y fundamenta todas sus decisiones. Para que la toma de

decisiones sea sólida, el análisis debe ser objetivo, para ello en lugar de intentar que cada punto de datos sea objetivo, el conjunto de datos debe ser objetivo. Las decisiones no son una sorpresa para los estudiantes pues ellos ya han tenido ciclos de retroalimentación individual durante todo el proceso. (van der Vleuten & Heeneman, 2016; Schuwirth et al., 2017b; Schut et al, 2018)

## **2.4. Principales variables vinculadas al impacto educativo de la evaluación**

### **2.4.1. Metas de logro**

Se sabe que la motivación puesta en juego a la hora de aprender está relacionada con las metas que tienen los estudiantes y evocan diferentes escenarios mentales en los mismos haciendo que tengan actitudes positivas o negativas hacia el estudio.

Tapia Castillo, F. (2017) en su trabajo de tesis, marca la diferencia entre los conceptos de “metas” y “metas de logro”; las metas son representaciones cognitivas de aquellas cosas que se desean cumplir, son como guías que proveen dirección y llevan a la acción; mientras que, las metas de logro se relacionan con la búsqueda de competencia y representan diferentes orientaciones motivacionales de los estudiantes en situaciones específicas como un determinado curso. Por lo tanto, las metas de logro tienen un rol muy relevante en contexto de la evaluación, pues condicionan la motivación de los estudiantes, influyendo en sus hábitos de estudio, en el desempeño y los resultados en las pruebas, y en los efectos de aprendizaje de las evaluaciones.

El constructo de las metas de logro, fue definido como los propósitos u objetivos que los alumnos persiguen en ámbitos educacionales y que dan lugar a diferentes resultados de desempeño. (Elliot, 2005; citado en Elliot & Murayama, 2008)

La teoría de las metas de logro nace en los años de 1970, diferenciando al comienzo dos dimensiones: las metas de aprendizaje o dominio y las metas de rendimiento. Las primeras tenían como propósito incrementar la propia competencia y dominar una tarea nueva; en cambio las metas de rendimiento estaban dirigidas a obtener juicios positivos de su competencia o evitar juicios negativos del mismo. A este primer componente de dominio y rendimiento se le agregó más adelante un segundo componente que indica si el foco está en aproximarse a la competencia o en evadir la incompetencia. Así surge el modelo 2 x 2 (Elliot & Murayama, 2008) que abarca cuatro metas de logro: meta de aproximación al dominio (aprender más), meta de evitación al dominio (evitar aprender menos), meta de

aproximación al rendimiento (rendir más que otros) y meta de evitación al rendimiento (evitar el fracaso). (Tapia Castillo, 2017)

Para Kool et al (2016), las metas de dominio han sido relacionadas con el interés continuo en el aprendizaje, la persistencia frente a obstáculos del cursado, el aprendizaje autorregulado, el comportamiento de buscar ayuda y el uso de estrategias de procesamiento cognitivo profundo. En cambio, la orientación hacia el rendimiento se ha vinculado con el uso de estrategias de procesamiento cognitivo más superficiales, pero también con un mayor rendimiento académico. Ambas orientaciones de metas tienen una función importante en la formación de las profesiones de la salud, pues se espera que los profesionales se concentren en el desempeño por el bien de sus pacientes, y al mismo tiempo necesitan aprender continuamente nuevas habilidades.

Según Moller y Elliot (2006), en una revisión sobre la investigación empírica del modelo 2x2, las metas de aproximación al dominio y de aproximación al rendimiento estuvieron relacionadas con resultados académicos positivos. Mientras, las metas de evitación al dominio y de evitación al rendimiento fueron predictores negativos del rendimiento académico.

En el trabajo de Kool et al (2016), se reconocen tres constructos predominantes, relacionados con las diferencias en la orientación de metas logro de los estudiantes: el género, el rendimiento académico y la autoeficacia (entendida como el grado de creencia de los alumnos en su propia capacidad para completar ciertas tareas). Los varones, por lo general, están más orientados al desempeño, mientras que las mujeres están más orientadas al dominio; los estudiantes con promedio alto de calificaciones, muestran niveles más altos de dominio y orientación al desempeño; y los estudiantes con una alta autoeficacia (que esperan un buen desempeño), tienen niveles más altos de dominio y orientación al desempeño comparados con los de autoeficacia más baja.

Otro aspecto importante en relación a las metas de logro, tiene que ver con los cambios y fluctuaciones que pueden tener los estudiantes en sus orientaciones a determinadas metas a lo largo de un curso o carrera. Este fenómeno ha sido poco estudiado y los resultados no son concluyentes (Kool et al, 2016), sin embargo, los trabajos realizados en estudiantes de medicina por Artino et al (2012) mostraron fluctuaciones en la orientación al dominio entre los alumnos de los tres primeros años, pero no hubo cambios en la orientación al desempeño. Estas fluctuaciones en las orientaciones de las metas de los estudiantes a lo largo del tiempo pueden estar relacionadas con cambios en el rendimiento

académico y en la autoeficacia del alumno en alguna etapa del cursado. Por ejemplo, el nivel de autoeficacia de un estudiante puede cambiar durante el año académico debido a la reprobación de un examen, provocando fluctuaciones en los niveles de dominio y orientación al desempeño o viceversa. (Kool et al, 2016)

#### **2.4.2. Enfoques de aprendizaje:**

La investigación de los enfoques de aprendizaje, como lo explican Freiberg Hoffmann & Fernández Liporace (2016), se inicia con Marton y Säljö (1976), quienes, a partir de un estudio fenomenográfico sobre casos de estudiantes universitarios, describen dos tipos de enfoque de aprendizaje cualitativamente dispares, que denominan enfoque profundo (procesamiento dirigido a comprender el significado de lo que debe aprenderse) y enfoque superficial (procesamiento orientado al aprendizaje del texto, mediante estrategias memorísticas). A estos conceptos Biggs (1988) le añade aspectos motivacionales y estratégicos, como cualidades constitutivas de cada tipo de enfoque. Los primeros se vinculan con los motivos y metas que el sujeto se plantea, guiando su interés por asimilar determinada información. Los segundos, aluden a las estrategias que puede emplear para incorporar la información requerida.

De esta manera, poniendo de relieve los motivos y las estrategias, surge el modelo de Biggs (1988) con un **enfoque profundo**, que corresponde a los estudiantes que manifiestan una elevada motivación intrínseca y un alto nivel de implicación con lo que están aprendiendo; un **enfoque superficial**, que describe a los estudiantes que tienden a cumplir con los requisitos mínimos para aprobar las instancias de evaluación, con mínimo esfuerzo y aplicando estrategias dirigidas a aprender en forma repetitiva y memorística; y un **enfoque de logro** (o estratégico), que implica a los estudiantes que quieren obtener el máximo rendimiento posible y mejores calificaciones que sus compañeros, aplicando estrategias de planificación y organización orientadas a ese objetivo. La preferencia por uno de los enfoques deriva de la interacción entre ciertas características personales (motivación y estrategias) y de las diferentes actividades que la situación de aprendizaje plantea, pudiendo en ocasiones, utilizar un enfoque diferente del que le es habitual.

Al respecto, Al Kadri et al (2011) explica que los estudiantes de grado, adoptan estrategias de aprendizaje de acuerdo con sus interpretaciones de los requisitos de las tareas asignadas dentro de un entorno de aprendizaje específico. Tales interpretaciones varían según la percepción de los alumnos sobre la calidad académica de un curso, la naturaleza del plan de estudios, las teorías implícitas de aprendizaje que tienen al ingresar a la educación

superior y el entorno de aprendizaje al que están expuestos. Estas percepciones que tienen los estudiantes sobre el entorno en el que aprenden y son evaluados, pueden modificar en cualquier momento su elección por un enfoque de aprendizaje, por eso es difícil predecir qué elecciones harán.

Existe un consenso general, que en la educación superior es deseable un enfoque profundo del aprendizaje, pues hay evidencia de investigación que respaldan una asociación entre enfoques profundos y mejores resultados de aprendizaje. Los estudiantes que usan un enfoque profundo parecen más capaces de desarrollar nuevos conceptos, demostrar su comprensión y demostrar un mayor desarrollo de habilidades genéricas; pero no hay suficiente evidencia de que obtengan calificaciones más altas debido al efecto del tipo de evaluación utilizado. (Wilson & Fowler, 2005; Heijne-Penninga et al, 2010)

Los enfoques de aprendizaje que los estudiantes usan en la práctica, según Wilson & Fowler (2005), están influenciados por una variedad de factores personales y situacionales, al respecto, las investigaciones sugieren que, si bien los estudiantes pueden tener una predisposición general para un enfoque de aprendizaje profundo o superficial, esto se ve influido por sus percepciones de una situación de aprendizaje específica. Por ejemplo, las percepciones de los estudiantes que tienen objetivos claros y están recibiendo una buena formación promueve un enfoque profundo; por el contrario, la percepción de una gran carga de trabajo se asocia con un enfoque superficial. Pero dado que los alumnos pueden alternar entre distintas estrategias en respuesta a diferentes factores ambientales y contextuales, es posible que un estudiante con enfoque profundo, ocasionalmente pueda usar estrategias que implican "memorizar sin comprender" y "conocimiento fragmentado" según las circunstancias (Al Kadri et al, 2011).

Lo mismo ocurre según Cilliers et al (2012) con las demandas de procesamiento cognitivo en las evaluaciones (habilidades requeridas para procesar el conocimiento y generar la respuesta solicitada), que influyen en el enfoque de aprendizaje de los estudiantes a través de señales inferidas del evaluador y de la tarea de evaluación. Es más probable que las preguntas de respuesta abierta y las evaluaciones abiertas generen un enfoque profundo del aprendizaje; mientras que las preguntas de opción múltiple y las evaluaciones cerradas, estimulan un enfoque superficial.

Estos ajustes en los enfoques de aprendizaje están relacionados al tipo de prueba y a su alcance cognitivo, y mediados por las percepciones de los estudiantes. Como lo explica Cilliers et al (2012), si los alumnos percibieron que las preguntas abiertas estaban dirigidas a

evaluar niveles cognitivos altos, es probable que adopten enfoques profundos, aunque las preguntas estén dirigidas a evaluar solo conocimientos o comprensión. De igual forma, cuando los estudiantes que estaban intrínsecamente motivados para comprender el contenido percibían que una evaluación requería memorización, solo memorizaban los hechos; y cuando percibían que las preguntas de redacción requerían simplemente extraer información de los libros, adoptaban un enfoque de aprendizaje superficial.

### 2.4.3. Comportamientos de estudio de los estudiantes

Los comportamientos de estudio, según Credé & Kuncel (2008), engloban a las habilidades, los hábitos y las actitudes de estudio de los estudiantes. Las **habilidades** de estudio se refieren al conocimiento del estudiante de las estrategias y métodos de estudio apropiados y la capacidad de administrar el tiempo y otros recursos (organización y técnicas aplicadas) para cumplir con las demandas de las tareas académicas. Los **hábitos** de estudio denotan el grado en que el estudiante se involucra en actos regulares de estudio que se caracterizan por rutinas de estudio apropiadas (frecuencia y tiempo de las sesiones, revisión de los materiales, concentración puesta en la tarea, autoevaluación) que ocurren en un entorno propicio para el estudio. Las **actitudes** de estudio que se refieren a la actitud positiva de un estudiante hacia el acto específico de estudiar (voluntad, ansiedad, motivación, interés por el estudio y deseo de aprender) e incluyen la aceptación y aprobación del estudiante de las metas más amplias de una educación universitaria.

Más allá de estas categorías de comportamientos de estudio que pueden ser muy heterogéneas en diferentes carreras y asignaturas, la cuestión más relevante es como los estudiantes gestionan su propio aprendizaje, ya que el aprendizaje autorregulado<sup>14</sup> es esencial para su formación y su desarrollo profesional continuo a lo largo de la vida.

Uno de los mayores desafíos que enfrentan los estudiantes al ingresar a la universidad es la autorregulación de su aprendizaje, pues a diferencia de la educación secundaria, los profesores universitarios ofrecen una orientación muy limitada sobre cómo, cuándo y qué aprender, mientras que los estudiantes, por lo general, carecen de conocimientos sobre estrategias de aprendizaje y confían en las intuiciones y rutinas que

---

<sup>14</sup> El término "aprendizaje autorregulado" se utiliza para describir la forma en que los estudiantes controlan su comportamiento de aprendizaje estableciendo y logrando metas, administrando recursos y adaptándose a la retroalimentación externa. (Wood, 2019)

desarrollaron en la escuela secundaria y que no siempre son efectivas en la universidad. (Biwer et al, 2020)

Las estrategias de estudio y su efectividad sobre el aprendizaje han sido investigadas por décadas en el ámbito universitario (Dunlosky et al, 2013; Yan et al, 2014; Biwer et al, 2020), y estos trabajos han revelado que las estrategias que parecen más laboriosas para los estudiantes, son las más efectivas para su aprendizaje. Al respecto, Bjork (1994) citado por Yan et al, 2014, acuñó el término "dificultades deseables" para referirse al conjunto de condiciones y estrategias de estudio que parecen ralentizar la comprensión de los materiales para ser aprendidos y hacen que el proceso de aprendizaje parezca más laborioso, pero al mismo tiempo mejoran la retención y la comprensión a largo plazo.

Las estrategias de estudio que pueden crear tales "dificultades deseables" son la práctica de recuperación, la práctica distribuida y la práctica intercalada. La práctica de recuperación, se refiere al acto de recuperar información de la memoria de forma activa respondiendo preguntas de prueba o recordando libremente (en lugar de leer o releer la información). La práctica distribuida implica espaciar las sesiones de estudio a lo largo del tiempo que, en comparación con la práctica masiva, aumenta la dificultad de recuperación, pero mejora la retención a largo plazo. La práctica intercalada se refiere a la combinación de diferentes temas durante una sesión de estudio, que contrasta con la práctica bloqueada, en la que los estudiantes estudian un tema hasta terminarlo antes de pasar al siguiente tema. (Biwer et al, 2020)

Sin embargo, trabajos basados en encuestas y experimentales, han revelado que los estudiantes siguen aplicando técnicas de estudio poco efectivas, y no logran comprender el efecto de prueba (recupera información de la memoria respondiendo preguntas) y el efecto de espaciamento (espaciar las sesiones de estudio repetidas), que son los dos efectos más sólidos demostrados en psicología cognitiva para mejorar la retención y el aprendizaje a largo plazo. (Yan et al, 2014)

En consonancia con lo anterior, Hartwig & Dunlosky (2012), observaron que los estudiantes subestiman los beneficios de aprendizaje de las pruebas, y prefieren seguir aplicando la relectura que es la estrategia de estudio más habitual. Por otro lado, aunque comprenden los beneficios del estudio espaciado, no lo aplican en sus sesiones de estudio pues generalmente están más condicionados por los plazos inminentes de las evaluaciones, esto último fue más notorio en los estudiantes que tenían bajo promedio general de calificaciones que en aquellos con mejor rendimiento académico.

Si bien la evidencia de la investigación sobre las estrategias de aprendizaje efectivas es clara y conocida por los psicólogos cognitivos, la información que ha llegado a los centros académicos de formación y a sus actores (docentes y estudiantes) es muy limitada. Los estudiantes universitarios apenas reciben instrucción sobre cómo estudiar de manera efectiva, y las intervenciones educativas específicas centradas en crear “dificultades deseables” durante el aprendizaje son muy escasas. (Biwer et al, 2020)

#### 2.4.4. Bases cognitivas del aprendizaje

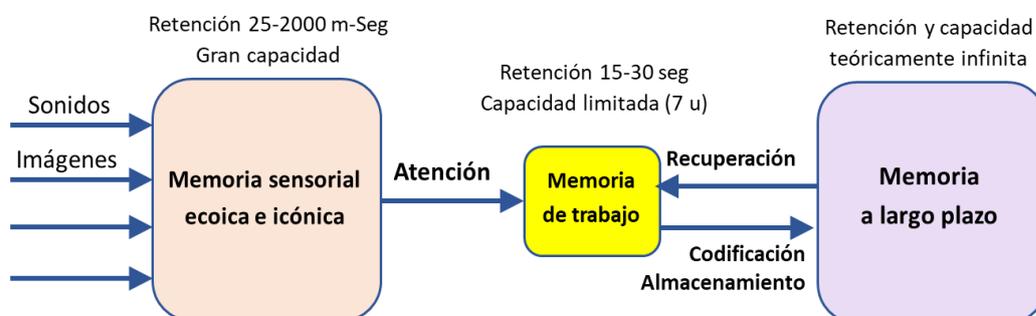
Abordar los efectos de aprendizaje de la evaluación, implica necesariamente considerar algunos aspectos inherentes a los sistemas de memoria, el procesamiento cognitivo y la carga cognitiva.

##### 2.4.4.1. Sistemas de Memoria

El sistema cognitivo humano, utiliza funcionalmente, tres tipos de memoria: memoria sensorial, memoria de trabajo y memoria a largo plazo. (Figura 2.5.)

La memoria sensorial tiene una capacidad enorme, los sistemas visual, auditivo, táctil y olfatorio, perciben una gran cantidad de información entrante, pero solo pueden ser retenidas un tiempo muy breve (menos de 2 segundos). En la memoria sensorial comienza el proceso de aprendizaje, las palabras e imágenes impresas percibidas (textos, gráficos, expresión de un paciente) son retenidas en la memoria icónica (“boceto visuoespacial”); las palabras habladas y otros sonidos (latido cardíaco, respuestas del paciente) son retenidas en la memoria ecoica (bucle fonológico). La mayor parte de esta información percibida no llega a la conciencia, pero cuando el estudiante presta “atención” a la misma en la memoria sensorial, la información se traslada a la memoria de trabajo. (Young et al, 2014)

**Figura 2.5:** Modelo de la memoria humana de Aktinson–Shriffin



Fuente: traducido de Young et al (2014)

La memoria de trabajo tiene una capacidad y duración muy limitadas, solo puede contener siete ( $\pm 2$ ) unidades de información y procesar, de manera intencional o consciente,

no más de cuatro simultáneamente; casi toda la información se pierde después de unos 20 segundos, sino se actualiza durante el proceso. Estas limitaciones solo se aplican a la “información nueva” obtenida a través de la memoria sensorial, pues cuando la información es recuperada de la memoria a largo plazo, la memoria de trabajo no tiene limitaciones. (Van Merriënboer & Sweller, 2010)

La memoria a largo plazo comprende el almacenamiento masivo de las experiencias de cada persona a largo de toda su vida, aunque, la capacidad para recuperar información específica depende en gran medida, de cuán significativa sea para la persona, con qué frecuencia la utilice y la fuerza de las asociaciones con las señales que recibe, pues son las señales del entorno las que activan automáticamente la memoria a largo plazo y en consecuencia cuando las señales y las asociaciones son fuertes, los recuerdos tienen mayor claridad. (de Bruin et al 2019)

#### **2.4.4.2. Procesamiento cognitivo**

Se describen dos tipos principales de procesamiento cognitivo: automático y controlado. Los procesos automáticos utilizan directamente a la memoria a largo plazo sin involucrar la memoria de trabajo, son rápidos y se reclutan inconscientemente; mientras los procesos controlados se basan en la memoria de trabajo, son relativamente más lentos y se reclutan conscientemente. Si bien ambos procesos contrastan entre sí, pueden ocurrir en paralelo y complementarse mutuamente, o aplicarse en forma secuencial, como se ve con frecuencia en la progresión del aprendizaje de los estudiantes que comienzan aplicando procesos controlados y con el tiempo y la práctica confían más en los procesos automáticos. (Evans & Stanovich, 2013)

En el campo de la Educación Médica estos procesos cognitivos se han integrado en la teoría del procesamiento dual, que abarca dos tipos o sistemas de razonamiento: el Tipo 1 o Sistema 1, automático, subconsciente, más rápido y de menor esfuerzo, que se basa en las experiencias almacenadas en la memoria a largo plazo (patrones o esquemas construidos); y el Tipo 2 o Sistema 2, que es analítico, más lento y de mayor esfuerzo, que combina funciones ejecutivas intensivas en recursos, incluida la memoria de trabajo y la atención. (Young et al, 2014)

Entre los docentes de medicina, prevalece la creencia de que el procesamiento de tipo 2 es superior para la tarea de diagnóstico médico, y por ende recomiendan la aplicación de pautas, algoritmos y eliminar los sesgos de la medicina basada en la evidencia; cuando en realidad todos deberían comprender que los procesos automáticos y controlados son

importantes y se reclutan según sea necesario para cumplir con las demandas de la tarea. (de Bruin et al 2019)

En relación al proceso de aprendizaje, Van Merriënboer & Sweller (2010) explican que la memoria de trabajo inevitablemente debe tener una capacidad limitada cuando la información es novedosa porque, a medida que la cantidad de elementos que deben organizarse aumenta linealmente, la cantidad de posibles combinaciones de elementos aumenta exponencialmente, y esto solo puede resolverse limitando severamente el número de unidades de información que se procesan simultáneamente. En consecuencia, el proceso de aprendizaje, se apoya en la memoria a largo plazo con la construcción y automatización de los esquemas. Los esquemas se pueden construir durante el proceso de aprendizaje reuniendo elementos e incorporando nuevos elementos en esquemas ya disponibles en la memoria a largo plazo o, más comúnmente, obteniendo información ya esquematizada de otras personas. Entonces, los esquemas pueden tratarse como un solo elemento en la memoria de trabajo y esto disminuye considerablemente la carga cognitiva asociada con el desempeño de tareas posteriores ya que la construcción y automatización de esquemas puede liberar la capacidad de la memoria de trabajo para otras actividades.

#### **2.4.4.3. Carga cognitiva**

La teoría de la carga cognitiva, descrita por John Sweller en 1988 (Young et al, 2014), es una teoría del aprendizaje que integra tres componentes clave de la arquitectura cognitiva: los sistemas de memoria (sensorial, de trabajo y a largo plazo) y procesos de aprendizaje (comentados anteriormente) y los tipos de carga cognitiva impuesta sobre la memoria de trabajo. Esta teoría ha tomado mucha relevancia en la educación médica debido a que las tareas y actividades profesionales que deben aprender los estudiantes para desarrollarse hacia la experiencia requieren la integración simultánea de múltiples y variados conjuntos de conocimientos, habilidades y comportamientos, en un momento y lugar específicos (competencias); y en consecuencia, cuando la carga cognitiva asociada con una tarea (aprendizaje o evaluación) excede la capacidad de la memoria de trabajo del alumno, el rendimiento y el aprendizaje se ven afectados.

La teoría de la carga cognitiva sostiene que la instrucción puede imponer tres tipos de carga en el sistema cognitivo de un alumno: la complejidad de la tarea y el conocimiento previo del estudiante determinan la “carga intrínseca”; las características de la instrucción que no son beneficiosas para el aprendizaje contribuyen a la “carga externa”; y las características de instrucción que son beneficiosas para el aprendizaje (formación de

esquemas y automatización) contribuyen a la “carga pertinente”. Por lo tanto, según esta teoría, para el diseño instruccional deberían seleccionarse tareas de aprendizaje que coincidan con el conocimiento previo de los alumnos y así optimizar la carga intrínseca; reducir la carga ineficaz para minimizar la carga externa; y posibilitar a los estudiantes participar en actividades que impongan una mayor carga pertinente para reducir la carga intrínseca. (Leppink & Duvivier, 2016)

La carga pertinente es regulada por cada alumno, pues representa el nivel de concentración del mismo dedicado al aprendizaje. Cuando las cargas, externa y/o intrínseca, son demasiado altas y se acercan o superan los límites de la memoria de trabajo del estudiante, no habrá suficientes recursos disponibles en dicha memoria para la carga pertinente necesaria para el aprendizaje, tales como combinar los nuevos elementos de información con esquemas ya existentes en la memoria a largo plazo. (Young et al, 2014; Cook et al, 2017)

Esta interacción de cargas cognitivas permite explicar porque los alumnos novatos o con menos conocimientos pueden aprender mejor, optimizando la carga intrínseca, con ejemplos resueltos o completando un problema parcialmente resuelto, que trabajando en forma autónoma en la resolución de problemas; mientras que los estudiantes avanzados o con más conocimientos se benefician mucho más con la resolución autónoma de problemas, pues cuentan con esquemas cognitivos suficientes para resolver el problema sin orientación, y si tuvieran que trabajar con ejemplos resueltos podrían experimentar un efecto de redundancia que aumenta la carga cognitiva externa. En resumen, cuando la carga cognitiva intrínseca es óptima y la carga extrínseca es baja, los estudiantes pueden participar en procesos de construcción de conocimientos como autoexplicación y argumentación, que aportan carga cognitiva pertinente y facilitan el aprendizaje. (Leppink et al, 2013)

#### **2.4.5. Retroalimentación y autorreflexión de los estudiantes**

Una instancia fundamental en la evaluación formativa de las competencias es la “retroalimentación” o “devolución constructiva”, definida como el acto de dar información a un estudiante describiendo su desempeño en una situación clínica observada y conforme a un patrón o estándar pre-establecido. Esto requiere que el estudiante esté dispuesto a exponer sus debilidades y errores y sea permeable a una crítica constructiva. El mayor impacto se obtiene cuando el estudiante compara la devolución del docente con su propio desempeño, pues la disonancia entre el desempeño deseado y el realizado, es un potente generador de motivación y de aprendizaje profundo. Es muy importante en la devolución dar

recomendaciones, provocar reacciones y auto reflexión para finalmente definir un plan de acción para mejorar. (Alves de Lima, 2008; Eva et al, 2012; Lefroy et al, 2015)

En las últimas décadas, la retroalimentación ha cobrado especial importancia y atención en la formación médica, y ha sido tema de estudio muy explorado por diferentes autores. Tuma & Nassar (2022) en su revisión, registraron que en marzo 2019 había más de 10000 publicaciones en PubMed centradas en la retroalimentación en la educación médica (810 eran artículos de revisión), lo que indica su relevancia en el proceso educativo y en la evaluación formativa.

Para Wood (2019), la retroalimentación proporciona la ruta por la cual la evaluación se convierte en una herramienta para el aprendizaje, y es esencial para la evaluación formativa. La retroalimentación posterior a la evaluación alienta al estudiante a trabajar junto al profesor, para mejorar su aprendizaje. El profesor muestra que está interesado en las opiniones de los estudiantes y lo anima a abordar su aprendizaje de una manera diferente. La retroalimentación proporcionada sin prejuicios y de manera abierta permite que el estudiante se sienta más seguro para discutir sus dificultades y planificar su aprendizaje.

Sin embargo, Watling & Ginsburg (2019) plantean que, combinar evaluación y retroalimentación puede ser problemático, ya que sus propósitos pueden parecer contrarios. Mientras que la evaluación enfatiza el juicio y la toma de decisiones, la retroalimentación enfatiza el desarrollo y el aprendizaje. En consecuencia, mezclar esos ingredientes distintos crea una receta arriesgada para aprender, y la gestión de ese riesgo requiere equilibrar las prioridades y tensiones en competencias relacionadas con la evaluación y la retroalimentación.

Por lo tanto, para que la retroalimentación sea efectiva, debe centrarse en la tarea y no en el individuo; debe ser específica, concreta y de alto rendimiento; debe focalizarse en el análisis de lo observado y no en la transmisión de juicios o inferencias; debe permitir a los estudiantes reflexionar e identificar sus fortalezas y debilidades; no debe socavar la autoestima pero tampoco limitarse a alabanzas; es más efectiva cuando se realiza en forma inmediata que tardíamente, y el lugar debe ser el más apropiado. (Archer, 2010)

Estas recomendaciones son muy valiosas para administrar una retroalimentación, pues las percepciones de los docentes y los estudiantes sobre las devoluciones entregadas y recibidas pueden ser muy diferentes. Esta situación se puso en evidencia en el estudio de Perera et al (2008), donde el 75% de los docentes manifestó entregar regularmente retroalimentación, pero sólo el 55% de los estudiantes coincidieron con esto, y la mayoría

manifestó que esperaban recibir más explicaciones sobre las calificaciones obtenidas, tener oportunidad para discutir los resultados y recibir preguntas reflexivas para comprender mejor.

Las diferentes técnicas y estrategias para administrar una adecuada retroalimentación, por su extensión, no van a ser abordadas en este capítulo, en cambio, es interesante hacer la distinción entre la retroalimentación “creíble” y la “constructiva”. La retroalimentación “creíble” se basa en la observación directa del desempeño del estudiante, originada en una fuente confiable. La retroalimentación “constructiva” contiene un plan de acción que ofrece dirección y orientación al alumno sobre cómo mejorar. (Watling & Ginsburg, 2019)

Un aspecto importante sobre el que hay gran consenso en la literatura, es que los eventos de retroalimentación únicos no son suficientes, sino que se requiere una “cultura” de retroalimentación dentro de la cual se da y recibe retroalimentación en forma regular y repetida; y para lograr dicha “cultura” es fundamental capacitar tanto a los profesores como a los estudiantes en la retroalimentación. (Lefroy et al, 2015; Thrien et al, 2020)

La reflexión en el marco de la educación médica fue definida por John Sandars (2009:685) como *“un proceso metacognitivo<sup>15</sup> que ocurre antes, durante y después de las situaciones con el propósito de desarrollar una mayor comprensión tanto de uno mismo como de la situación para que los encuentros futuros con la situación se basen en encuentros previos”*. Este autor, obtuvo evidencias de que la reflexión está asociada con un enfoque más profundo del aprendizaje y que permite integrar el nuevo aprendizaje con el conocimiento y las habilidades existentes.

En consecuencia, un practicante reflexivo y una racionalidad reflexiva, permiten acortar la brecha entre la teoría y la práctica, con una actitud de trabajar desde la experiencia narrando, repensando y reformulando esa brecha. La reflexión brinda a los profesionales y estudiantes la oportunidad de pensar amplia y profundamente sobre su práctica, por ello la capacidad reflexiva se ha incorporado a una de las competencias profesionales de un médico eficaz. (Telio et al, 2015; Wen et al., 2015)

---

<sup>15</sup> La metacognición es un proceso de autorregulación que selecciona, monitorea y evalúa un proceso cognitivo, que en este caso, es el acercamiento a la reflexión. Un proceso que se puede controlar y que permite desarrollar diversas estrategias de formación para que se potencie la reflexión. (Sandars, 2009)

Con respecto al uso de la reflexión en la educación médica, los autores refieren tres métodos de aplicación: a) utilizar un modelo reflexivo (como el ciclo reflexivo de Gibbs), para construir preguntas reflexivas que faciliten reflexiones amplias y profundas sobre una experiencia realizada; b) hacer una reflexión (diario o portfolio) sobre la acción después de que el evento haya ocurrido en la práctica médica; y c) la reflexión guiada por un mentor (docente o tutor) que alienta al estudiante a compartir sus pensamientos, sentimientos o reacciones a sus experiencias y promueve múltiples perspectivas y acciones alternativas. Esta última es la que se aplica en el acto de la retroalimentación. (Sandars, 2009; Wen et al, 2015; Uygur et al, 2019)

Analizando la relación entre el acto de retroalimentación, la reflexión del estudiante y el uso de la retroalimentación en el desarrollo de un plan de acción para el aprendizaje futuro, Pelgrim et al (2013), observaron que solo los comentarios específicos de los docentes son seguidos por la reflexión del estudiante, y esta reflexión a su vez promueve el uso de la retroalimentación para el desarrollo de planes de acción.

Esta respuesta reflexiva a la retroalimentación también está relacionada a ciertos rasgos de los estudiantes. Según Wood (2019), los estudiantes “autorregulados” generan su propia retroalimentación interna que los ayuda a evaluar su progreso hacia las metas y adaptar sus procesos de aprendizaje ante obstáculos o cambios en su motivación; son conscientes de sus propios conocimientos, creencias y habilidades cognitivas, y los utilizan para interpretar la retroalimentación externa de manera efectiva.

#### **2.4.6. Percepciones de los estudiantes en las evaluaciones**

Según diferentes autores (Segers et al, 2006; Cilliers et al, 2010; Schut et al, 2018), el impacto de cualquier sistema de evaluación sobre el aprendizaje está mediado por las percepciones de los estudiantes. Por lo tanto, es fundamental conocer estas percepciones para comprender lo que significan las evaluaciones para los estudiantes en cuanto a riesgos y consecuencias, y cómo y por qué las mismas permiten, o no, a los alumnos optimizar y autorregular su aprendizaje utilizando estas evaluaciones,

Los riesgos y consecuencias de una evaluación individual, constituyen “lo que está en juego” para el estudiante en esa evaluación, que puede ser muy poco (bajo riesgo y sin consecuencias) como ocurre en las evaluaciones formativas cuyo principal propósito es favorecer el aprendizaje; pero también puede ser de “alto riesgo” por sus consecuencias (selección, promoción o graduación) como ocurre con algunas evaluaciones sumativas que

pueden llevar a los estudiantes a enfocar la evaluación como un obstáculo (hay mucho en juego) y no como una oportunidad para el aprendizaje.

Los estudiantes interactúan con la evaluación basándose en como perciben su propósito, y estas percepciones no siempre coinciden con las intenciones de los profesores que diseñaron los instrumentos. Los estudiantes de medicina a menudo perciben a la evaluación formativa como si el propósito fuera sumativo. Esa percepción puede interferir con el impacto educativo pretendido disminuyendo sus efectos sobre el aprendizaje. (Harrison et al, 2017)

Según Schut et al (2018), la conceptualización de los estudiantes sobre “lo que está en juego” en la evaluación está fuertemente relacionada con su capacidad percibida para actuar, controlar y tomar decisiones dentro del entorno de aprendizaje y evaluación, y esto depende en gran medida de las oportunidades que les ofrece el programa de evaluación, de la interacción entre sus experiencias y su confianza, y de la relación con los profesores.

Al respecto, las pruebas de conocimiento estandarizadas (como el “multiple choice”) donde los estudiantes experimentan el éxito cuando son capaces de "encontrar la respuesta correcta", pueden producir en los mismos una sensación de dependencia al contenido y la calidad de la prueba, y una percepción de poco o ningún control sobre la evaluación y los resultados, especialmente cuando este tipo de evaluación adjudica calificaciones. Por el contrario, cuando tienen la oportunidad de interactuar con el evaluador (examen oral o durante una observación directa) y mostrar su progreso y habilidades, pueden percibir un mayor control sobre el proceso y el resultado de la evaluación, lo que reduce los riesgos percibidos. (Schut et al, 2018)

Sin embargo, los estudiantes pueden tener cierta ambivalencia sobre la observación directa, pues si bien reconocen su valor formativo, también identifican cierta incomodidad al ser observados y sienten que son evaluados. Esta situación puede alterar sus rutinas y hacerlos interactuar de manera diferente con el paciente, adoptando un enfoque tipo lista de cotejo para tomar los antecedentes y el examen físico como si estuvieran actuando en un entorno de examen de alto riesgo, donde el propósito es obtener buenos resultados en la evaluación en lugar de servir como base para la retroalimentación y el aprendizaje. (Watling & Ginsburg, 2019)

Una forma de reducir los riesgos percibidos por los estudiantes en las pruebas individuales y brindar más oportunidades para que demuestren su progreso y mejora, es aplicar programas de evaluación con múltiples evaluaciones complementarias, tales como la

evaluación programática y la evaluación longitudinal multimétodo, que dan a los alumnos una mayor sensación de tener el control, debido a las múltiples ocasiones que tienen de mostrar su desempeño y mejorar su aprendizaje, especialmente cuando son acompañadas con mensajes de retroalimentación recurrentes. Sin embargo, aunque el número de evaluaciones complementarias puede influir positivamente en la percepción de “lo que está en juego” en la evaluación, también puede ser percibido por los estudiantes como una “exageración”, y en tal caso la evaluación deja de ser percibida como formativa y se convierte en una “cacería de evaluaciones” que hay que atravesar para cumplir con los requisitos del programa. (Schut et al, 2018)

Un fenómeno similar al anterior se ha descrito en la evaluación programática cuando todos los puntos de datos (parciales) contribuyen a una decisión sumativa final, los estudiantes tienden a percibir todos los puntos de datos como sumativos, cuando en realidad tienen un propósito formativo. Esa percepción puede ser problemática, pues los estudiantes pueden buscar evaluadores más indulgentes o casos más sencillos en respuesta al conocimiento de que los momentos formativos contribuyen a los juicios acumulativos. Por otro lado, esta distorsión en las percepciones de los estudiantes puede obstaculizar la aceptación de la evaluación como verdaderamente formativa y en consecuencia no cumplir con su propósito. (Heeneman et al, 2015; Watling & Ginsburg, 2019)

Otro aspecto relacionado a las percepciones de los estudiantes, muy importante para esta investigación, es como ellos auto-perciben su capacidad y sus habilidades intelectuales para afrontar las evaluaciones en la asignatura que cursan. Esta percepción puede vincularse desde ópticas diferentes, a tres constructos, que para la Psicología Educativa tienen distintos significados: “autoconcepto académico”<sup>16</sup>; “autopercepción de competencia para aprender”<sup>17</sup>; y “autoeficacia percibida”<sup>18</sup>; aunque para algunos autores como Yeung et al, (2014) los tres constructos están abarcados en el de “autoconcepto académico”.

---

<sup>16</sup> El “autoconcepto académico” designa el conjunto de percepciones y creencias que la persona tiene de sí misma sobre determinadas características, habilidades y motivos, para afrontar el aprendizaje en un contexto instruccional. (Valle Arias y col, 1998)

<sup>17</sup> La “autopercepción de competencia para aprender” es la apreciación que el estudiante tiene de sí mismo respecto de su mayor o menor facilidad para aprender, en el contexto de la carrera que cursa (Morchio & Difabio, 2019).

<sup>18</sup> La “autoeficacia percibida” alude a las creencias en las propias capacidades para organizar y ejecutar las acciones requeridas para producir determinados logros futuros. (Bandura, 1997, citado por Morchio & Difabio, 2019)

De acuerdo a como los estudiantes perciben su “autoconcepto académico”, como auto-perciben su “competencia para aprender”, y cuáles son sus creencias frente a una situación futura (“autoeficacia percibida”), gravitan sobre sus pensamientos, emociones, decisiones, interacción social, motivación, esfuerzo, actitud ante los obstáculos y, en su conjunto, sobre sus resultados académicos; y la influencia de estos aspectos no se explica por los eventos en sí, sino en función de cómo son procesados subjetivamente por los estudiantes. (Morchio & Difabio, 2019)

Sin embargo, desde el enfoque abarcativo de Yeung et al, (2014), el “Autoconcepto Académico” puede explicar los efectos sobre el aprendizaje y el rendimiento de los estudiantes de medicina, ya que si el autoconcepto es “positivo” las energías del estudiante quedan potenciadas en la dirección del aprendizaje, pero si es “negativo” esas mismas energías se reducen, disminuyendo su capacidad objetiva para las tareas. Por esta razón, el rendimiento académico no guarda relación directa con la capacidad objetiva del sujeto en sus tareas, sino que esa capacidad objetiva esta mediada por el concepto que el sujeto tiene de su propia capacidad.

## **2.5. Estado del Arte**

La mayoría de las investigaciones relacionadas al “impacto educacional de las evaluaciones” se han enfocado en diferentes momentos y escenarios del proceso evaluativo y en la aplicación de herramientas de evaluación específicas, pero muy pocos autores han abordado los sistemas de evaluación multimétodo, salvo en la “evaluación programática” (Schuwirth & Van der Vleuten, 2011b), que está tomando cada día más protagonismo.

### **2.5.1. Sobre los efectos de la evaluación**

Con el propósito de desarrollar un modelo para explicar el impacto educativo de las evaluaciones en el aprendizaje de los estudiantes, antes, durante y después de la prueba, Kordestani Moghaddam et al (2019) realizaron un estudio cualitativo que incluyó a ocho estudiantes de medicina y siete profesores de la Universidad de Ciencias Médicas de Teherán, a través de entrevistas semiestructuradas, grupos focales y recolección de notas de campo, aplicando para su análisis el enfoque de la teoría fundamentada. Los conceptos extraídos se clasificaron en cuatro categorías: elementos de la evaluación que afectan el aprendizaje; mecanismo por el cual ejercen sus efectos; factores contextuales; e impacto sobre el aprendizaje. Los elementos y mecanismos identificados fueron similares a los

publicados en otros trabajos. Las consecuencias educativas observables derivadas de la evaluación fueron cambios en el contenido del aprendizaje, la profundidad del aprendizaje, la durabilidad del aprendizaje y, en última instancia, la implicación emocional de los estudiantes.

Lynam & Cachia (2017) realizaron un estudio cualitativo con tres grupos focales donde participaron 23 estudiantes de psicología de la Facultad de Ciencias Humanas y Sociales de la Universidad de West London (Reino Unido). Los datos fueron analizados mediante análisis temático inductivo experiencial. Se identificaron dos temas (y subtemas) que afectaron el aprendizaje de los estudiantes: los factores de enseñanza (oportunidad y tipo de evaluación) y los factores estudiantiles (madurez académica y emociones), ambos temas estaban mediados por la relación entre tutores y estudiantes. Pudieron comprobar el papel que tienen las cualidades personales y los factores emocionales en el compromiso del alumno, y observar que el compromiso de los estudiantes con la evaluación se alineó positivamente con el aumento de la madurez académica, como reflejo de la creciente autenticidad de la evaluación a medida que avanza el plan de estudios y las actividades y evaluaciones se enfocan en tareas y prácticas clínicas, en contraste con las pruebas del conocimiento científico y el desarrollo de habilidades genéricas de los primeros años. Las autoras concluyeron que aplicar estrategias para promover la madurez académica, reducir el estrés y el miedo en los estudiantes podrían fomentar un enfoque más constructivo del aprendizaje.

### **2.5.2. Sobre los efectos de la retroalimentación**

Preston et al (2020) estudiaron una muestra de estudiantes (de los seis años) de la carrera de medicina de la Universidad “James Cook” (Australia) con el objetivo de explorar la percepción que tenían sobre el valor de la evaluación y la retroalimentación en su aprendizaje, y cómo se relacionaba con su desempeño en los exámenes. El diseño de la investigación fue mixto, a través de cuestionarios autoadministrados y discusiones de grupos focales. Los exámenes escritos en sus diferentes formatos fueron percibidos como las herramientas que mejor reflejan el esfuerzo realizado en el aprendizaje, mientras que la escritura reflexiva, el ensayo y el mini-CEX fueron calificados como menos precisos. Las percepciones sobre el nivel de dificultad de las evaluaciones fueron mejorando a medida que los estudiantes avanzaban en la carrera. La mayoría reconoció que la retroalimentación es importante para mejorar y debería ser más oportuna, de mejor calidad y personalizada. En

general, los estudiantes expresaron su satisfacción con las habilidades clínicas aprendidas y sintieron que el curso los preparó bien para la práctica profesional, pero observaron que la evaluación y la carga de trabajo del curso fueron extensas y las identificaron como fuentes comunes de estrés y ansiedad para los exámenes. La mayoría de los estudiantes no apreciaron el valor de la reflexión y el vínculo con la práctica profesional, por lo que los autores consideran que los estudiantes pueden "jugar el juego" siguiendo una "receta" para aprobar, pero tienen poca comprensión de los motivos de la evaluación.

Harrison et al (2016) diseñaron un estudio con el objetivo de explorar los factores que ayudan o dificultan la receptividad a la retroalimentación. Realizaron seis grupos focales en tres Facultades de Medicina de tres países (EE.UU., Reino Unido y Países Bajos) que tenían diferentes enfoques de evaluación, que iba desde una evaluación sumativa tradicional a un sistema de evaluación programática. Observaron que la receptividad a la retroalimentación se vio reforzada por las culturas de evaluación que tenían los estudiantes; por la provisión de una evaluación auténtica y relevante; y por un andamiaje apropiado para ayudar a interpretar la retroalimentación. La adjudicación de calificaciones proporcionó una referencia externa útil, pero pareció obstaculizar la promoción de la excelencia. Los estudiantes percibieron como muy beneficioso para su aprendizaje y formación, poder tener un mayor control sobre los procesos de evaluación y retroalimentación, que las evaluaciones sean lo más auténticas posible, y contar con tutorías efectivas a largo plazo.

Con el propósito de incorporar dentro de los programas de evaluación sumativa a la retroalimentación sistemática como estrategia de aprendizaje, Harrison et al (2017) realizaron un estudio en la Carrera de Medicina de la Universidad de Keele (Inglaterra) que tiene un sistema de promoción y graduación basado en evaluaciones sumativas con instancias de retroalimentación electivas. El trabajo se efectuó con estudiantes de 3° a 5° año de medicina y docentes de diferentes niveles, a través de entrevistas individuales y grupales que se analizaron con un enfoque socio-cultural. La gran mayoría de los participantes (en particular los docentes) que solo habían experimentado una cultura de evaluación sumativa, tuvieron mucha dificultad para conceptualizar un cambio radical en la cultura de evaluación. Muchos compartían suposiciones y creencias comunes sobre la importancia del paradigma de evaluación sumativa y no consideraron viable la aplicación de elementos de una cultura de "evaluación para el aprendizaje", como la tutoría a largo plazo y la provisión de comentarios sin calificaciones. Todas las sugerencias de cambio se alinearon predominantemente con el paradigma de la evaluación sumativa y los participantes dieron

por sentado que la evaluación sumativa es la forma en que se debe evaluar a los estudiantes de medicina.

### **2.5.3. Sobre instrumentos de evaluación específicos**

Para explorar los factores que influyen sobre el impacto educativo de las evaluaciones por observación (Mini-CEX y DOPS), Lörwald et al (2017) realizaron una revisión sistemática de la literatura y seleccionaron 12 artículos que revelaron como factores influyentes cuatro temas (contexto, usuarios, implementación y resultado) y nueve subtemas (tiempo; aplicación; habilidad y actitud del observador; conocimientos, experiencia y percepción del estudiante; calidad y evaluación de la retroalimentación). Los autores concluyen que el impacto educativo no depende solo del instrumento utilizado sino de su adecuada aplicación y de la relación de confianza entre estudiante y evaluador (alianza educativa), para lograr una retroalimentación efectiva.

Coob et al (2013) evaluaron el impacto educativo de dos instrumentos de evaluación diferentes, DOPS (observación directa de habilidades de procedimiento) y MCQ (examen de opción múltiple) analizando los enfoques de aprendizaje de los estudiantes y el rendimiento en cada examen. Se aplicó el Cuestionario del Proceso de Estudio (SPQ) a 87 estudiantes del último año de veterinaria y se realizaron entrevistas semiestructuradas a 16 para explorar sus percepciones sobre las evaluaciones. Los resultados mostraron que los estudiantes adoptaron un enfoque de aprendizaje más profundo con el DOPS y más superficial con MCQ, observándose una correlación positiva entre el enfoque de logro y el rendimiento en los exámenes. El análisis cualitativo reveló siete influencias principales en sus enfoques de aprendizaje: motivación, propósito, consecuencia, aceptabilidad, retroalimentación, presión de tiempo y diferencias individuales.

Raupach et al (2016) realizaron un ensayo cruzado aleatorizado con 124 estudiantes de cuarto año de la Universidad de Medicina de Göttingen (Alemania) con el objetivo de probar la hipótesis de que las pruebas repetidas con “preguntas de puntos clave” son más efectivas que las sesiones repetidas de “aprendizaje basado en casos” para fomentar el razonamiento clínico. Los estudiantes participantes, en la condición de “control” estudiaron las narrativas de los casos; y en la condición de “intervención” trabajaron los mismos contenidos, con el agregado de preguntas de puntos clave. El desempeño de los mismos se evaluó a las 13 semanas (examen de egreso) y a los nueve meses (prueba de retención). Se comprobó que las pruebas repetidas fueron más efectivas que el aprendizaje basado en casos

solo, demostrando así el efecto del “aprendizaje mejorado por pruebas” en el razonamiento clínico evaluado con preguntas de puntos clave. Los autores concluyeron que la implementación curricular de este tipo de pruebas en forma longitudinal puede mejorar considerablemente los resultados de aprendizaje de los estudiantes en aspectos relevantes de la medicina clínica.

En la misma dirección, Schuelper et al (2019) realizaron un ensayo prospectivo, no aleatorizado, con 75 estudiantes de cuarto año de la Universidad de Medicina de Göttingen (Alemania) con el objetivo de probar la hipótesis de que la exposición repetida a preguntas de puntos clave basadas en videos mejora el razonamiento clínico en mayor medida que la exposición repetida a preguntas de puntos clave basadas en texto. Los estudiantes asistieron a 12 seminarios de casos basados en computadora donde trabajaron en historias de casos disponibles en formato de texto y video que contenían preguntas de puntos clave, cada estudiante eligió el formato de su preferencia. El desempeño de los mismos fue evaluado con exámenes formativos de entrada, salida y retención con la misma herramienta. Observaron que la mayoría (70%) eligió el formato basado en texto y no encontraron diferencias significativas en las puntuaciones del examen de salida entre los grupos de texto y video. Si en cambio hubo diferencias en los resultados de las pruebas de retención (a los nueve meses) a favor de los estudiantes que trabajaron con casos en video y que los autores atribuyen al mayor tiempo de exposición que demanda la tarea de visualización de los videos.

Para evaluar el impacto que puede tener sobre el rendimiento académico de los estudiantes eliminar las calificaciones de las evaluaciones, Pippitt et al (2022) realizaron un estudio en la Facultad de Medicina de la Universidad de Utah (EE.UU.) con estudiantes de medicina de 1° y 2° año que completaron evaluaciones con preguntas de opción múltiple y un examen final que hasta 2020 adjudicaban calificaciones y desde ese año no cuentan con calificación. Se compararon las puntuaciones obtenidas entre cohortes de pruebas calificadas y pruebas sin calificar, y se recogieron comentarios de estudiantes y profesores. Observaron que, aunque no hubo diferencias significativas en las calificaciones de ambos grupos, las cohortes “sin calificar” obtuvieron puntajes por encima del promedio nacional en comparación con las cohortes “calificadas”, que obtuvieron puntajes por debajo del mismo. Los comentarios de los estudiantes fueron positivos y los profesores sintieron que mejoraron su relación con los mismos. Los autores concluyen que una combinación de evaluaciones sin calificar y calificadas puede promover de manera efectiva la evaluación para el aprendizaje.

#### 2.5.4. Sobre la evaluación programática

Schut et al (2018), entrevistaron 26 estudiantes de tres países (Canadá, EE.UU. y Holanda) y cinco programas de grado y posgrado. Exploraron las experiencias y percepciones de los estudiantes en el contexto de la evaluación programática, aplicando un enfoque abierto y cualitativo. En todos los programas, los estudiantes conceptualizaron la evaluación como un continuo fuertemente vinculado con la capacidad percibida de actuar, controlar y elegir, lo que les permite utilizar la evaluación para apoyar sus aprendizajes. Reconocieron que las evaluaciones estandarizadas brindan pocas oportunidades para aprender y destacaron la importancia de la autonomía y la confianza como facilitadores del aprendizaje.

Heeneman et al. (2015) estudiaron los elementos de la evaluación que los estudiantes percibían como apoyo o inhibición del aprendizaje, y los factores que influyeron en sus aprendizajes. Aplicando una metodología cualitativa, entrevistaron 17 estudiantes de la etapa preclínica de la carrera de Medicina de la Universidad de Maastricht (Países Bajos), sobre sus percepciones de la evaluación y los enfoques de aprendizaje. Los resultados mostraron que los estudiantes percibieron la evaluación formativa como sumativa y como un importante disparador para el aprendizaje; los exámenes orales, el portfolio y el sistema de calificaciones tuvieron efectos positivos para el aprendizaje; sobre los otros instrumentos algunas percepciones fueron de apoyo y otras de inhibición para el aprendizaje, diferencia que no dependía tanto de las cualidades del instrumento como de la percepción individual de cada estudiante.

Con el propósito de optimizar las prácticas de evaluación en un sistema de evaluación programática de reciente implementación, Roberts et al (2022) realizaron un estudio cualitativo con 14 grupos focales donde participaron 112 estudiantes de la Facultad de Medicina y Salud de la Universidad de Sydney (Australia), aplicando un enfoque realista crítico que en el análisis incluyó inducción (identificación de patrones), abducción (interpretación teórica) y retroducción (explicación causal). Observaron que las estructuras de evaluación y los sistemas culturales dentro de la evaluación programática proporcionaron condiciones (restricciones y habilitaciones) y condicionamientos (aceptación o rechazo de nuevos procesos 'no tradicionales') para que los estudiantes puedan ejecutar sus opciones de aprendizaje. El mecanismo subyacente que más influyó en la experiencia de evaluación programática fue el de equilibrar las complejas relaciones entre la gestión del alumno, las estructuras de evaluación y el sistema cultural. Los autores concluyeron que es necesario

aplicar estrategias para lograr cambios sostenidos en la práctica de la evaluación, tales como promover la reflexividad colectiva, mejorar la integración teoría-práctica, cambiar la cultura de aprendizaje, y generar aceptación y confianza con el nuevas normas, creencias y comportamientos en la evaluación para y del aprendizaje.

#### **2.5.5. Investigaciones en Latinoamérica y en la región**

Los estudios relacionados al impacto educativo de la evaluación o a sus efectos sobre el aprendizaje de los estudiantes son muy limitados en América Latina, la mayoría de las publicaciones son artículos de revisión (no sistemáticas), informes de experiencias, editoriales y resúmenes presentados en congresos o reuniones científicas, pero hay muy pocos artículos originales de investigación sobre este tema y ninguno referido a los sistemas de evaluación multimétodo, por lo que se puede inferir que la temática abordada en esta tesis es un área de vacancia en la región.

Para evaluar el impacto del cambio curricular, de un enfoque de evaluación del aprendizaje a uno de evaluación para el aprendizaje, Pérez et al (2015) realizaron un estudio en un curso de la Escuela de Medicina de la Pontificia Universidad Católica de Chile, comparando las actividades, las percepciones de los estudiantes y sus calificaciones finales, antes y después del cambio curricular que consistió en eliminar las clases magistrales e incorporar evaluaciones intermedias seguidas de una sesión de retroalimentación. Los resultados mostraron que con el cambio curricular se redujo el tiempo de clases (19.5%), aumentó el tiempo de evaluación (8.5%) y la retroalimentación ocupó el 7% de la carga horaria. Hubo mejoras significativas en las percepciones de los estudiantes en todas las áreas evaluadas por la encuesta, principalmente las relativas a las evaluaciones y a la retroalimentación. Las calificaciones de los estudiantes también mejoraron de un promedio de 5.8 a 6.4 (escala 1-7). Los autores concluyeron que la evaluación para el aprendizaje logró el impacto educativo deseado sin incrementar el tiempo curricular asignado.

En la misma dirección, con el objetivo de evaluar cómo los estudiantes percibían el proceso educativo y el sistema de evaluaciones que se realizaban diariamente, Ferreira (2021) realizó un estudio en la Facultad de Medicina de Botucatu de la Universidad Estadual Paulista (Brasil), donde participaron todos los estudiantes de 3° año del curso de Enfermería. Se aplicó un cuestionario con preguntas abiertas y cerradas, y la recolección de datos se realizó al final de cada actividad diaria y al final del curso. La mayoría de los estudiantes calificó como excelentes los recursos didácticos, las actividades desarrolladas y la

comprensión de los contenidos. Las evaluaciones diarias los mantuvo motivados y el apoyo del tutor fue un facilitador del aprendizaje por la objetividad y la discusión crítica. Los objetivos y criterios de cada evaluación permitían a los estudiantes evaluar su progreso, posibilitando adaptar la enseñanza a sus necesidades, y el docente podía evaluar la efectividad de su desempeño. La adecuada duración de cada bloque, la baja relación docente-alumno y la evaluación diaria, fueron factores que favorecieron el aprendizaje.

Con el propósito de evaluar la validez aparente del de examen clínico objetivo estructurado (ECO) como aporte al estudio de su factibilidad local, Ferrando-Castagnetto et al (2019) realizaron un estudio al final del ciclo propedéutico “Introducción a la Medicina Integral”, con una subcohorta de estudiantes que cursaban en el Hospital Universitario de la Facultad de Medicina de la Universidad de la República (Uruguay). Se aplicó un cuestionario de 28 ítems dirigido a explorar las percepciones sobre las propiedades del ECO en los potenciales factores diseño y validez aparente, y sus efectos formativos. Los resultados mostraron un alto grado de aceptación en cuanto a la autenticidad de los estímulos, dinámica, pertinencia y justicia del examen. Los estudiantes percibieron obstáculos organizativos durante la prueba, manifestando desacuerdo en incluir alguna estación dirigida a evaluar solamente habilidades para la comunicación y requiriendo instancias de devolución personalizada. Hubo acuerdo en que el examen evalúa claramente los objetivos de aprendizaje específicos del curso, resulta una evaluación justa, exige un intenso entrenamiento clínico que va más allá de la carga horaria curricular y permite reconsiderar la pertinencia del curso.

Para investigar el valor del Portfolio Reflexivo Electrónico en la práctica reflexiva y en el proceso de aprendizaje, Maia & Struchiner (2016) realizaron un estudio en el contexto del internado (electivo) en Anestesiología en la Escuela de Medicina de la Universidad Federal de Rio de Janeiro (Brasil). En 10 estudiantes se analizó: el contenido de los portfolios, el proceso de elaboración de los mismos y las percepciones que tenían sobre la realización y utilidad del portfolio. Se observó que la mayoría de los estudiantes reflexionaron al preparar su portfolio y documentaron la reflexión en sus narrativas. A través de sus narraciones, demostraron su proceso de aprendizaje, haciendo conexiones entre los hechos prácticos y los conocimientos teóricos involucrados y discutiendo otros temas, como aspectos éticos y de trabajo en equipo, contribuyendo a una evaluación formativa individualizada. Las autoras concluyen que el Portfolio Reflexivo es una herramienta que estimula la práctica reflexiva y contribuye a la construcción del conocimiento de manera significativa.

Extendiendo el contexto de investigación a las residencias médicas y con el objetivo de determinar el impacto educativo de la evaluación basada en el desempeño de los residentes, Díaz-Plasencia et al (2016) realizaron un estudio descriptivo a 93 residentes y 34 tutores del Programa de Segunda Especialización de la Universidad Privada Antenor Orrego (Perú), aplicando una encuesta mixta e indicadores de salud y evaluado el impacto educativo según los cuatro niveles del modelo de Kirkpatrick. Los resultados mostraron impacto positivo en los niveles 1 a 3 de Kirkpatrick, que los residentes adjudicaron a la discusión de casos clínicos e incidentes críticos y a la retroalimentación del portfolio; mientras los tutores lo vincularon a las pruebas de progreso, portfolio, mini-CEX y observación directa de actividades procedimentales. Las evaluaciones también tuvieron impacto en el nivel 4 de Kirkpatrick al mejorar la mortalidad neta e infecciones intrahospitalarias en los hospitales académicos.



## CAPITULO 3

**MARCO  
METODOLÓGICO**





## Capítulo 3: MARCO METODOLÓGICO

### 3.1. Paradigma de Investigación

Esta investigación está enmarcada en los postulados de un paradigma<sup>19</sup> “interpretativo-constructivista”, a partir del cual se comprende que la realidad social no sólo se construye por condiciones objetivadas sino por la forma particular con que los actores sociales perciben, interpretan y construyen esta trama, tan múltiple como compleja. La realidad social (educativa), en consecuencia, no se trata de una cosa dada, sino de una pluralidad de construcciones, mentalmente intangibles, basadas en la experiencia social de los actores, de naturaleza local y específica. Estas construcciones no son ni más ni menos verdaderas, sino simplemente más o menos estructuradas y/o sofisticadas, dependiendo en la forma y contenido que adoptan de las personas individuales o los grupos que las sostienen (Guba & Lincoln, 2015).

Dentro del paradigma “interpretativo”, la generación del conocimiento ocurre cuando los conocimientos relevantes emergen naturalmente a través del discurso del investigador de su interacción investigador-participante en el entorno natural (Bunniss & Kelly, 2010). De igual manera, el paradigma “constructivista” ve el conocimiento como activamente construido y cocreado como el producto de las interacciones y relaciones humanas; y por ende, los datos y el análisis se crean a partir de “experiencias y relaciones compartidas con los participantes y otras fuentes de datos” (Charmaz, 2006 citado por Watling & Lingard, 2012).

Por lo tanto, ambos paradigmas adoptan:

- Una **ontología** (naturaleza de la realidad) relativista, la realidad es subjetiva y cambiante, no hay una verdad última.
- Una **epistemología** (naturaleza del conocimiento) transaccional, el conocimiento es subjetivo, hay múltiples y diversas interpretaciones de la realidad, no existe una forma definitiva o "correcta" de saber.
- Una **metodología** (naturaleza de la investigación) hermenéutica y dialéctica, centrada en la comprensión y el razonamiento inductivo, donde el significado se construye en la

---

<sup>19</sup> Los paradigmas son conjuntos de creencias y prácticas, compartidas por comunidades de investigadores, que regulan la investigación dentro de las disciplinas. Los diversos paradigmas se caracterizan por diferencias ontológicas, epistemológicas y metodológicas en sus enfoques para conceptualizar y realizar investigaciones, y en su contribución a la construcción del conocimiento disciplinario. (Weaver & Olson, 2006 citado por Bunniss & Kelly, 2010)

interacción investigador-participante en el medio natural, y puede reunir diversas interpretaciones como teoría fundamentada y etnografía.

El paradigma “interpretativo”, asume que un estudio nunca puede estar libre de sesgos; por lo tanto, eliminar el sesgo no es una intención de la investigación. Por el contrario, uno de los objetivos de la investigación interpretativa es atender a cómo los pensamientos, sentimientos, opiniones y experiencias del investigador pueden influir en lo que observa y registra. Esta reflexividad sobre cómo el investigador construye el conocimiento en conjunto con los participantes es crucial para la interpretación de los datos. (Bunniss & Kelly, 2010)

Por otro lado, la aplicación de la “Teoría fundamentada constructivista” como método de análisis para la información cualitativa (que se explica más adelante), es pertinente en este estudio, pues apunta a la comprensión interpretativa y el conocimiento situado, enfatizando la reflexividad, reconociendo los roles del investigador, los participantes de la investigación, y la situación y el proceso de investigación en la construcción del conocimiento. (Watling & Lingard, 2012; Charmaz, 2014)

## 3.2. Diseño Metodológico

### 3.2.1. Tipo de Investigación

Esta investigación fue desarrollada con un enfoque metodológico mixto<sup>20</sup>, cualitativo y cuantitativo; diseñada como un estudio de caso<sup>21</sup> de tipología múltiple con diferentes unidades de análisis y con un alcance exploratorio-interpretativo.

Como **enfoque metodológico** mixto, de las tres estrategias básicas de integración cuali-cuantitativa (complementación, combinación y triangulación) planteadas por Bericat (1998), se aplicó la estrategia de complementación, que tiene una finalidad meramente aditiva, porque se basa en el deseo de contar con dos imágenes distintas de la realidad social investigada, en el que cada método revela aspectos diferentes, permitiendo ampliar nuestro

---

<sup>20</sup> “Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada (metainferencias) y lograr un mayor entendimiento del fenómeno bajo estudio” (Hernández-Sampieri et al, 2014:534)

<sup>21</sup> “Un estudio de caso es una indagación empírica que investiga un fenómeno contemporáneo dentro de su contexto de la vida real, especialmente cuando los límites entre el fenómeno y el contexto no son claramente evidentes”. (Yin, 2003:13)

conocimiento de la realidad estudiada. La complementación es la estrategia de integración más adecuada para este trabajo, pues como lo explica Bericat, 1998:106, *“no se trata de buscar convergencia ni confirmación de resultados, sino de contar simultáneamente con dos imágenes que enriquezcan nuestra comprensión de los hechos..... y que, en su máximo nivel de integración, puede dar lugar a una síntesis interpretativa que integra los resultados procedentes de cada método”*.

El **diseño metodológico** de esta investigación es un estudio de casos de tipología múltiple y diseño incrustado, con múltiples unidades de análisis (Tipo 4)<sup>22</sup>, y se basa en una indagación empírica que investiga los casos abordados aplicando las preguntas de "cómo" o "por qué" con respecto al fenómeno de interés (Yazan, 2015). Sin embargo, para Stake (2005) citado por Van Wynsberghe & Khan, (2007) el estudio de caso no es una elección metodológica sino una elección de lo que se va a estudiar; y desde esta óptica, un “caso” es un problema a estudiar, que revelará una comprensión profunda del mismo o sistema acotado, lo que implica la comprensión de un evento, actividad, o proceso. En pocas palabras, “el sistema se convierte en el caso, y el investigador elige un evento, actividad o proceso dentro de este sistema para iluminarlo” (Van Wynsberghe et al, 2007:82).

Cabe aclarar que los casos seleccionados para este trabajo no constituyen una muestra estadística representativa de una población donde se infieren generalizaciones, por el contrario, se buscó estudiar y comprender la especificidad de cada uno para luego hacer una lectura del conjunto, detectando las características que tienen en común y aquellas que hacen a sus diferencias, aunque la intención de estudiar de forma comparada los casos no se limitó a simplemente detectar sus similitudes y desviaciones sino, a partir de ellas, poder reconocer un arco de propiedades en torno a cada categoría que pueda servir para interpretar el fenómeno estudiado.

Esta investigación tiene esencialmente un **alcance interpretativo**, que en el marco del paradigma interpretativo, Bunniss & Kelly (2010) destacan dos elementos centrales: utiliza una epistemología subjetiva que anticipa múltiples y diversas interpretaciones de la realidad en lugar de buscar revelar una "verdad" general; y se asocia con un esfuerzo

---

<sup>22</sup> Robert Yin (2003) diferencia cuatro tipos de diseño para los estudios de caso, considerando cuántos casos se decide estudiar y si se distinguen o no en el interior de éstos unidades de análisis: Tipo 1: caso único, diseño holístico (una unidad de análisis); Tipo 2: caso único, diseño “incrustado” (varias unidades de análisis); Tipo 3: múltiples casos, diseño holístico; Tipo 4: múltiples casos, diseño “incrustado”. Yin, R. (2003:39)

interpretativo para recopilar una variedad de elementos con el objetivo de construir una imagen detallada de cómo un fenómeno particular es entendido por aquellos que tienen una experiencia personal de él.

### 3.2.2. Selección de los casos de estudio

Para la elección de los casos de estudio se tuvo en cuenta que los mismos cumplieran con las características planteadas en el problema de investigación y que permitieran realizar las experiencias empíricas necesarias para alcanzar los objetivos propuestos.

En esa dirección y dado que el objetivo principal de esta investigación es “analizar el impacto educacional que tienen los sistemas de evaluación multimétodo en los estudiantes de la carrera de medicina”, se eligieron como casos de estudio dos (2) materias que aplican dicho sistema de evaluación en la Carrera de Medicina de la Universidad Nacional del Nordeste (UNNE): la asignatura **Medicina Hombre y Sociedad (MHS)**, primera materia de la carrera; y la asignatura **Medicina III (MED-III)** que se cursa en quinto año de la misma carrera y son las únicas que (hasta el momento) aplican sistemas de evaluación multimétodo.

### 3.2.3. Categorías preliminares de análisis

Para explorar el impacto educacional de la evaluación multimétodo se definieron como categorías preliminares de análisis: el entorno de aprendizaje y evaluación en los casos de estudio; los aspectos motivacionales y enfoques de aprendizaje de los estudiantes, sus comportamientos de estudio, y sus experiencias y percepciones en las evaluaciones.

#### 3.2.3.1. Entorno de aprendizaje y evaluación en las asignaturas

Esta categoría comprende las características del entorno de aprendizaje y de los sistemas de evaluación aplicados en los casos de estudio, sus fundamentos, propósitos, diseño, planificación y aplicación, en el contexto y nivel formativo de cada materia. Los tipos y calidad de los instrumentos de evaluación utilizados, sus propiedades psicométricas, los resultados obtenidos por los estudiantes en las diferentes evaluaciones y los efectos catalíticos desarrollados después de su aplicación. Además, se consideró el alcance cognitivo de los procesos mentales que los estudiantes debían realizar en las tareas y consignas planteadas en las evaluaciones, aplicando la nueva taxonomía de Bloom que abarca seis niveles crecientes: recordar, comprender, aplicar, analizar, evaluar y crear. (Adams, 2015)

#### 3.2.3.2. Aspectos motivacionales de los estudiantes

Para esta categoría se consideraron las **metas de logro** pues constituyen una las dimensiones más relevantes vinculadas a la motivación. Se aplicó el modelo de Elliot &

Murayama (2008), que utiliza un "esquema 2 x 2 de metas de logro" aplicado a las metas de aprendizaje<sup>23</sup> y de rendimiento, donde la "aproximación" implica moverse hacia un objeto valorado positivamente o mantenerse en él; mientras que la "evitación" supone alejarse de un objeto valorado negativamente o permanecer distanciado de él. Este modelo comprende cuatro tipos de metas de logro: la meta de **aprendizaje-aproximación**, cuyo objetivo es el de comprender/aprender lo más posible; la meta de **aprendizaje-evitación**, que trata de eludir la incompetencia, el no aprender todo lo posible; la meta de **rendimiento-aproximación**, que se centra en la capacidad relativa del sujeto comparándose con los compañeros y buscando superarlos; y la meta de **rendimiento-evitación**, que intenta escapar al fracaso y evitar juicios negativos de otros.

### 3.2.3.3. Enfoques de Aprendizaje de los estudiantes

Para esta categoría se aplicó el modelo de Biggs (1988) que clasifica los enfoques de aprendizaje en profundo, superficial y de logro, y según este autor: el **enfoque profundo** está relacionado con alta motivación intrínseca, con estudiantes comprometidos con lo que están aprendiendo, que demuestran gran interés por la materia y encuentran al aprendizaje emocionalmente satisfactorio, por lo que utilizan estrategias dirigidas a descubrir el significado de lo que van a aprender estableciendo relaciones con sus conocimientos previos relevantes. Al **enfoque superficial**, lo relaciona con la motivación extrínseca, donde los estudiantes tienen miedo al fracaso y muestran preocupación por el tiempo que invierten en las tareas, intentan cumplir solo con los requerimientos y consignas, con un mínimo de esfuerzo y poco compromiso, por lo que aplican estrategias dirigidas a aprender mecánica y repetitivamente la información para reproducirla en el momento oportuno. Finalmente, al **enfoque de logro**, lo relaciona con la competitividad del estudiante por obtener las mejores calificaciones y satisfacer todos los requisitos formales, intentando economizar esfuerzos, por eso también lo denomina enfoque "estratégico" e implica a los estudiantes con intención claramente definida de obtener el máximo rendimiento posible a través de una adecuada planificación de sus tareas, del esfuerzo y del tiempo disponible, por lo que aplican estrategias de planificación y organización de sus actividades con el objetivo prioritario de obtener logros académicos lo más altos posibles. Este último enfoque no es explorado con el cuestionario R-SPQ-2F que se aplica en este estudio (se aclara más adelante).

---

<sup>23</sup> Es importante aclarar que se utiliza el término "Aprendizaje" por el significado que tiene para este trabajo, aunque en el modelo de Elliot & Murayama (2008) se aplica el término "Maestría", y que otros autores lo traducen como "Dominio".

#### 3.2.3.4. Comportamientos de estudio de los estudiantes

Esta categoría comprende a los hábitos, habilidades y actitudes de estudio, que los estudiantes aplican para adquirir, organizar, sintetizar, evaluar, recordar y utilizar la información. Según Credé & Kuncel (2008), los **hábitos de estudio** abarcan principalmente la frecuencia y el tiempo de las sesiones de estudio, el entorno utilizado, la revisión de los materiales, la concentración puesta en la tarea, el ensayo del material aprendido, y la autoevaluación; las **habilidades de estudio** comprenden, la capacidad para administrar el tiempo y asignar otros recursos de acuerdo con las demandas de las tareas académicas, capacidad para organizar, resumir e integrar los materiales de estudio y las técnicas aplicadas; y las **actitudes para el estudio** incluyen principalmente la voluntad, la ansiedad y la motivación, el interés por el estudio y el deseo de aprender.

#### 3.2.3.5. Experiencias y percepciones de los estudiantes en las evaluaciones

Esta categoría abarca las oportunidades que les brindó a los estudiantes el sistema de evaluación para: mostrar su progreso, evaluar su desarrollo de conocimientos, reflexionar sobre su aprendizaje, ejercer control dentro del programa, influir en los resultados, interactuar con los docentes, y mejorar su rendimiento. Además, se incluyeron los riesgos percibidos y la capacidad percibida de actuar, controlar y hacer elecciones dentro del entorno de aprendizaje y evaluación, siguiendo las pautas propuestas por Schut, et al (2018). Entre las **experiencias** se indagaron los factores del sistema de evaluación que facilitaron u obstaculizaron las oportunidades para que los estudiantes puedan orientar y desarrollar sus aprendizajes; y entre las **percepciones**: la capacidad percibida, como percibió cada estudiante su capacidad intelectual para afrontar las evaluaciones (autoconcepto académico) y los riesgos percibidos, como el riesgo de no regularizar o aprobar la materia, de fracaso académico, de bajo promedio, de perder prestigio, etc.

### 3.2.4. Técnicas e Instrumentos de recolección de la información

#### 3.2.4.1. Documentos, registros, y materiales curriculares

Con el apoyo de los profesores responsables de las asignaturas MHS y MED-III, se obtuvieron los programas y las planificaciones de cursado, las guías didácticas e instructivos para los docentes, los instrumentos de evaluación aplicados y su contenido, las tablas de especificaciones de elaboración, los criterios de calificación, las rúbricas y pruebas patrón, los registros del lector óptico, los análisis psicométricos efectuados a los instrumentos de evaluación, las planillas de calificaciones, las formas de información/devolución de los

resultados a los estudiantes; y se recuperaron del sistema SIU Guaraní las encuestas de opinión de los estudiantes sobre el desempeño de los docentes y el cursado de las materias.

#### **3.2.4.2. Inventarios estandarizados y cuestionario - INSTRUMENTO A**

Este instrumento integra tres herramientas con propósitos diferentes.

Para indagar los cuatro tipos de metas de logro que abarca el modelo de Elliot & Murayama (2008), se utilizó el Cuestionario de Metas de Logro (A-AGQ-R) validado para estudiantes universitarios argentinos por Sánchez Rosas (2015). (ANEXO 1.1.)

Para explorar los Enfoques de Aprendizaje se aplicó el Cuestionario Revisado de Procesos de Estudio, Dos Factores (R-SPQ-2F), validado y adaptado para estudiantes universitarios argentinos por Freiberg Hoffmann & Fernández Liporace (2016). (ANEXO 1.2.)

Para indagar los efectos que tuvieron los sistemas de evaluación de MHS y MED-III sobre el aprendizaje de los estudiantes, se utilizó un cuestionario de elaboración propia, con preguntas precodificadas y abiertas. (ANEXO 1.3.)

Las tres herramientas se administraron en forma anónima en el mismo formulario de Google Forms®, identificado como INSTRUMENTO A, y fue aplicado en la primera etapa del proyecto, al finalizar el cursado de las dos asignaturas.

#### **3.2.4.3. Entrevistas electrónicas con cuestionario abierto<sup>24</sup> - INSTRUMENTO B**

Para indagar los efectos de las evaluaciones sobre el aprendizaje de los estudiantes, sus hábitos, habilidades y actitudes de estudio, los cambios que adoptaron para afrontar los sistemas de evaluación, los factores que motivaron y desanimaron el estudio, y sus experiencias y percepciones en las evaluaciones aplicadas en los casos analizados, se utilizaron preguntas abiertas.

Las preguntas fueron elaboradas a partir de los primeros resultados obtenidos en el análisis de los materiales curriculares y con el Instrumento A, con el propósito de recabar información más detallada, triangular temporalmente los datos e interpretar con mayor profundidad las unidades y categorías emergentes de los datos recopilados previamente.

Este cuestionario (ANEXO 2) se aplicó en formato online en la última etapa del proceso de investigación.

---

<sup>24</sup> Se optó por efectuar “entrevistas electrónicas” (Fontana & Frey, 2015:186) ante la imposibilidad de realizar los grupos focales previstos en el proyecto de investigación, por las medidas de aislamiento social obligatorio establecidas en el contexto de la pandemia COVID-19.

### 3.2.5. Técnicas de Procesamiento y Análisis de la información

#### 3.2.5.1. Análisis de la información cualitativa

Para analizar los documentos, registros y materiales curriculares, que Peräkylä (2015) denomina materiales empíricos de ocurrencia natural, se adoptó el “enfoque informal” que propone la misma autora, en este enfoque se busca establecer los temas claves y trazar una imagen de los presupuestos y significados que constituyen el material textual, sin ajustarse a ningún protocolo de análisis predefinido.

Para valorar el alcance cognitivo de las evaluaciones se estimaron los procesos mentales que demandaba responder las preguntas y consignas de las pruebas, y se categorizaron según la nueva taxonomía de Bloom. (Adams, 2015)

La información cualitativa de los cuestionarios se analizó con el método iterativo de la “Teoría Fundamentada constructivista”, que comienza codificando las respuestas abiertas para organizarlas en áreas o temas conceptuales clave, haciendo una comparación constante con otras respuestas en las categorías preliminares y emergentes, para luego llevar el análisis de lo categórico a lo conceptual y así poder generar teoría. (Watling & Lingard, 2012)

Para la codificación inicial (abierta) de los datos se utilizó la técnica de “unidades constantes” eligiendo como unidad de análisis, el “párrafo” (en el lenguaje que usaron los estudiantes), cada párrafo fue codificado (según su significado) con números, letras o colores, para facilitar su identificación, agrupamiento y categorización. Los datos codificados en las tres fuentes de información (documentos, instrumento A y B) fueron comparados por triangulación y agrupados, de manera inductiva, en categorías conceptuales (algunas fueron subcategorías) según su contenido y significado. (Figura 3.1.)

En el segundo nivel de codificación, se compararon estas categorías emergentes, buscando similitudes y diferencias (conceptuales) entre sus unidades de análisis y posibles vínculos entre las categorías; de esa forma, se identificaron las subcategorías y se integraron en categorías generales de mayor amplitud conceptual (codificación axial), con lo cual se redujo el número inicial de categorías emergentes. Finalmente, en la “codificación enfocada” (Charmaz, 2014), se interpretaron y analizaron las relaciones entre las categorías emergentes, y con las categorías preliminares y teóricas desprendidas del marco teórico<sup>25</sup>,

---

<sup>25</sup> La revisión bibliográfica en la Teoría Fundamentada Constructivista da la oportunidad de aclarar ideas y hacer comparaciones profundas, ampliando y desafiando otras evidencias, sin agotar la creatividad del investigador ni limitar la teoría (Charmaz, 2014).

para construir las categorías centrales que constituyen la matriz analítico-metodológica de este trabajo. (Figura 3.1.)

Este proceso recursivo entre datos, categorías e interpretaciones permitió también ajustar y reformular las preguntas de investigación y, con ello, la reorganización y resignificación de categorías y conceptos, ampliando los alcances del marco analítico-metodológico preliminar.

**Figura 3.1.** Síntesis del proceso de triangulación, codificación y categorización de los datos cualitativos aplicando la Teoría fundamentada.



- 1: Codificación abierta de los datos triangulados y agrupamiento en subcategorías conceptuales.
- 2: Codificación enfocada, interpretación, comparación y construcción de categorías centrales.
- 3: Agrupamiento de las categorías en ejes analíticos según los propósitos de la investigación.

### 3.2.5.2. Análisis de los datos cuantitativos

Para valorar la calidad de diseño y construcción de los instrumentos de evaluación aplicados se utilizaron rubricas específicas para cada uno, y la Escala de Galofré en las pruebas con preguntas de opción múltiple.

Para interpretar las variables psicométricas que los profesores realizaron a los exámenes escritos se tomó como referencia la Teoría Clásica de los Test, en cuanto a la Confiabilidad (Alfa de Crombach) de las pruebas, y los Índices de Dificultad y de Discriminación de cada una de las preguntas o ítems.

Las respuestas a los cuestionarios R-SPQ-2F y A-AGQ-R se procesaron siguiendo las pautas de los autores Biggs (1988) y Elliot & Murayama (2008) respectivamente.

Los datos cuantitativos se expresan en valores absolutos, porcentajes y proporciones según corresponda. Como medidas de resumen se calcularon los valores de tendencia central y de dispersión. Para la presentación de los datos resumidos se utilizan tablas de columnas y para la representación gráfica de las variables, diagramas de caja (box plots). Los análisis estadísticos (descripción y visualización de los datos, pruebas paramétricas y pruebas de asociación) se realizaron con el software XLSTAT (Addinsoft, 2021).

### **3.2.6. Matriz analítica-metodológica: Ejes y categorías de análisis**

Finalmente, se presenta el instrumento analítico-metodológico (Tabla 3.1.) que permite estructurar y organizar el análisis de los elementos empíricos de la investigación, que surgió como resultado de un proceso recursivo entre el encuadre teórico referencial y las indagaciones realizadas sobre los casos de estudio, que permitió ajustar, modificar y proponer nuevas categorías explicativas para el fenómeno estudiado.

No es un constructo desarrollado a priori que busca verificación en el apartado analítico, sino una matriz elaborada durante el trabajo de tesis, conformada por distintos ejes y categorías que permiten ordenar (registros-descripciones) y procesar (interpretaciones-reflexiones) los datos recogidos de las diferentes fuentes de información.

El instrumento está compuesto por tres ejes de análisis, a partir de los cuales se agrupan el conjunto de categorías de las que se desprenden la preguntas que organizan la investigación. La definición de dichos ejes se vincula a las tres perspectivas enunciadas por van der Vleuten et al (2017), en las prácticas de evaluación de competencias: la evaluación del aprendizaje, la evaluación para el aprendizaje y la evaluación como aprendizaje.

**Tabla 3.1: Matriz analítica-metodológica**

<b>EJE 1: Evaluación del aprendizaje</b>		
¿Cómo se integran las evaluaciones al entorno de aprendizaje de los estudiantes?		
<b>Categorías de Análisis</b>	<b>Breve Descripción</b>	<b>Preguntas Derivadas</b>
Entorno de aprendizaje y evaluación	Comprende las características del entorno de aprendizaje y de los sistemas de evaluación, sus fundamentos, propósitos, diseño, planificación y aplicación, en el contexto de cada materia, los instrumentos de evaluación utilizados, sus propiedades psicométricas y el alcance cognitivo de las tareas y consignas planteadas en las pruebas.	<p>¿Cómo se diseñaron y planificaron los sistemas de evaluación multimétodo?</p> <p>¿Por qué se eligieron esas herramientas de evaluación?</p> <p>¿Cómo se elaboraron y validaron las herramientas de evaluación?</p> <p>¿Cómo funcionaron las herramientas de evaluación?</p>
Estudiantes evaluados	Abarca las metas de logro que tienen los estudiantes evaluados, los enfoques y estrategias de aprendizaje que utilizaron en el contexto del sistema de evaluación aplicado, y los resultados obtenidos.	<p>¿Cuáles eran las metas de logro de los estudiantes?</p> <p>¿Qué enfoques de aprendizaje aplicaban?</p> <p>¿Qué resultados obtuvieron en las evaluaciones?</p>
<b>EJE 2: Evaluación para el aprendizaje</b>		
¿Cómo influyen las evaluaciones sobre el aprendizaje de los estudiantes?		
<b>Categorías de Análisis</b>	<b>Breve Descripción</b>	<b>Preguntas Derivadas</b>
Factores de la evaluación	Comprende los objetivos y criterios de las evaluaciones; el material evaluado; la interacción con los docentes; la proximidad temporal a la evaluación; y la sobrecarga de tareas simultáneas.	<p>¿Qué sabían los estudiantes sobre los objetivos y criterios de las evaluaciones?</p> <p>¿Qué interacción tuvieron con los docentes?</p> <p>¿Cómo influyeron en el estudio las evaluaciones múltiples y la proximidad temporal a cada evaluación?</p>
Efectos de aprendizaje	Explora la naturaleza de las actividades de procesamiento cognitivo y de regulación metacognitiva en cada evaluación (cantidad y distribución del esfuerzo, elección de recursos y contenidos; estrategias de monitoreo); y mecanismos de las devoluciones.	<p>¿Qué procesos cognitivos y metacognitivos aplicaron en las distintas evaluaciones?</p> <p>¿Cómo fueron las devoluciones en las diferentes evaluaciones?</p>

### EJE 3: Evaluación como aprendizaje

¿Qué puede aportar cada momento de la evaluación al aprendizaje de los estudiantes?

Categorías de Análisis	Breve Descripción	Preguntas Derivadas
Antes de la evaluación	Abarca los comportamientos de estudio (hábitos, habilidades y actitudes) que los estudiantes aplicaron en su preparación para las evaluaciones, incluyendo gestión del tiempo, establecimiento de metas, selección de qué, cómo, y dónde estudiar; y factores que motivaron y desanimaron el estudio.	<p>¿Cómo influyeron las evaluaciones en los comportamientos de estudio de los estudiantes?</p> <p>¿Qué ajustes hicieron en sus estrategias de estudio?; ¿Por qué lo hicieron?</p> <p>¿Qué factores motivaron o desanimaron el estudio, en la fase de preparación para las evaluaciones?</p>
Durante la evaluación	Abarca las experiencias de los estudiantes en las evaluaciones, los beneficios obtenidos en su formación y aprendizaje, y las oportunidades que tuvieron de mostrar su progreso y reflexionar sobre lo aprendido; así como los factores del sistema que facilitaron u obstaculizaron dichas oportunidades.	<p>¿Qué beneficios aportaron las herramientas de evaluación a la formación y aprendizaje de los estudiantes?</p> <p>¿Qué factores o elementos de las evaluaciones facilitaron y cuales obstaculizaron el buen rendimiento?</p>
Después de la evaluación	Abarca las percepciones de los estudiantes dentro del entorno de aprendizaje y evaluación, como perciben su capacidad intelectual para afrontar las evaluaciones y los riesgos percibidos (no regularizar o aprobar, bajo promedio, etc); y cuanto aportaron las evaluaciones al aprendizaje de la materia.	<p>¿Cómo percibieron su capacidad y sus habilidades intelectuales para afrontar las evaluaciones?</p> <p>¿Qué consecuencias percibieron como riesgosas en el sistema de evaluación?</p> <p>¿Qué aportó el sistema de evaluación aplicado, al aprendizaje de la materia?</p>

# CAPITULO 4

## PRESENTACIÓN DE LOS RESULTADOS





## Capítulo 4: PRESENTACIÓN DE LOS RESULTADOS

### 4.1 Presentación de los casos de estudio

#### 4.1.1. Contexto institucional de los casos de estudio

Los casos de estudio que se analizan en este trabajo son unidades curriculares de la Carrera de Medicina de la Universidad Nacional del Nordeste (UNNE), que se desarrolla en el ámbito de la Facultad de Medicina de la UNNE, donde además se dictan otras dos carreras de grado, la Licenciatura en Enfermería y la Licenciatura en Kinesiología y Fisiatría; y diferentes actividades de posgrado y de educación continua.

La Carrera de Medicina<sup>26</sup>, fue creada en 1953, y hasta 1999 aplicó un plan de estudio de siete años de duración, con un diseño curricular estructurado en cátedras (especialidades) y un perfil de graduación de “Médico Cirujano”. A partir del año 2000, la carrera adoptó un nuevo plan de estudio<sup>27</sup> (Plan 2000) de seis años de duración, estructurado en asignaturas, con un modelo curricular orientado en competencias y un perfil de graduación de “Médico General”.

El Plan 2000<sup>28</sup> de la carrera de medicina, con 5600 horas de carga horaria total, contiene 25 asignaturas trocales (obligatorias) y 19 materias optativas (electivas); y los programas vigentes de todas las asignaturas tienen sus objetivos, contenidos y actividades (formativas y de evaluación), adecuados al modelo de competencias, establecido en 2007 por el Ministerio de Educación<sup>29</sup> para la acreditación de las carreras de medicina.

En consecuencia, el sistema de evaluación que utiliza cada materia, es esencial para controlar y asegurar el desarrollo efectivo de las competencias, pues la estructura curricular está “fragmentada en asignaturas” y las competencias se desarrollan a lo largo de la carrera.

Sin embargo, en el análisis realizado en 2012 a las planificaciones anuales de las materias (Di Bernardo y col, 2014), se observó que en las evaluaciones predominaban instrumentos tradicionales como pruebas escritas “multiple choice” de pocas preguntas,

---

<sup>26</sup> Resolución N°749/10-CONEAU: Acreditación de la carrera de Medicina de la Facultad de Medicina de la Universidad Nacional del Nordeste por un período de seis años.

<sup>27</sup> Resolución N°349/00-CS: Aprobación del nuevo plan de estudio de la Carrera de Medicina de la Universidad Nacional del Nordeste

<sup>28</sup> Resolución N°2400/02-CD: Reglamentación del Plan 2000 de la Carrera de Medicina de la Universidad Nacional del Nordeste.

<sup>29</sup> Resolución N°1314/07-MECyT: Contenidos curriculares básicos, carga horaria mínima, criterios de intensidad de la formación práctica y estándares para la acreditación de las carreras de Medicina.

observaciones informales de las prácticas, y exámenes orales abiertos; y en algunos casos no se especificaba el sistema de evaluación formativa. Por otro lado, en las encuestas que completaron los estudiantes a través del SIU Guaraní al finalizar el cursado de cada asignatura, uno de los aspectos más cuestionados fue la falta de devoluciones en la mayoría de las evaluaciones. Para resolver estas debilidades, algunas materias como Medicina III (Di Bernardo y col, 2009) y Medicina Hombre y Sociedad (Di Bernardo y col, 2017), reestructuraron sus sistemas de evaluación, articulando diferentes instrumentos para evaluar todo el espectro de las competencias, y mejoraron sus mecanismos de retroalimentación para fortalecer la evaluación formativa.

En este contexto, y dado que el objetivo principal de este trabajo de investigación es “analizar el impacto educacional que tienen los sistemas de evaluación multimétodo en los estudiantes de la carrera de medicina”, los casos de estudio analizados son las únicas dos asignaturas que aplican dicho sistema de evaluación en la Carrera de Medicina de la UNNE: la asignatura Medicina Hombre y Sociedad (MHS), primera materia de la carrera; y la asignatura Medicina III (MED-III) que se cursa en quinto año de la misma. Figura 4.1.

**Figura 4.1:** Estructura de materias en la Carrera de Medicina y asignaturas casos de estudio (recuadros amarillos)

PLAN DE ESTUDIO 2000	Primer Año		Segundo Año		Tercer Año		Cuarto Año		Quinto Año		Sexto Año
	1º C	2º C	1º C	2º C	1º C	2º C	1º C	2º C	1º C	2º C	1º y 2º C
<b>MHS</b>		ANA	FIS	AF-P	MED-I	EME	CIR-I	MED-II	<b>MED-III</b>	CIR-II	PFO
		HIS	MIC	APS		S-PUB					
		BIO	ING	FAR		D-IMG					
			OPT 1	OPT 2		S-MEN					
						GIN					
						OPT 3	PED-I	PED-II	OBS		
					OPT 4	OPT 5	M LEG				

**MHS:** Medicina, hombre, sociedad; **ANA:** Anatomía; **HIS:** Histología-Embriología; **BIO:** Bioquímica; **FIS:** Fisiología; **MIC:** Micro-Parasito-Inmunología; **ING:** Inglés; **AF-P:** Anatomía y Fisiología Patológicas; **APS:** Atención Primaria de la Salud; **FAR:** Farmacología; **MED-I:** Medicina I; **EME:** Emergentología; **S-PUB:** Salud Pública; **D-IMG:** Diagnóstico por Imágenes; **S-MEN:** Salud Mental; **CIR-I:** Cirugía I; **GIN:** Ginecología; **MED-II:** Medicina II; **PED-I:** Pediatría I; **MED-III:** Medicina III; **PED-II:** Pediatría II; **CIR-II:** Cirugía II; **OBS:** Obstetricia; **M-LEG:** Medicina Legal; **PFO:** Práctica Final Oblig; **OPT:** Optativas

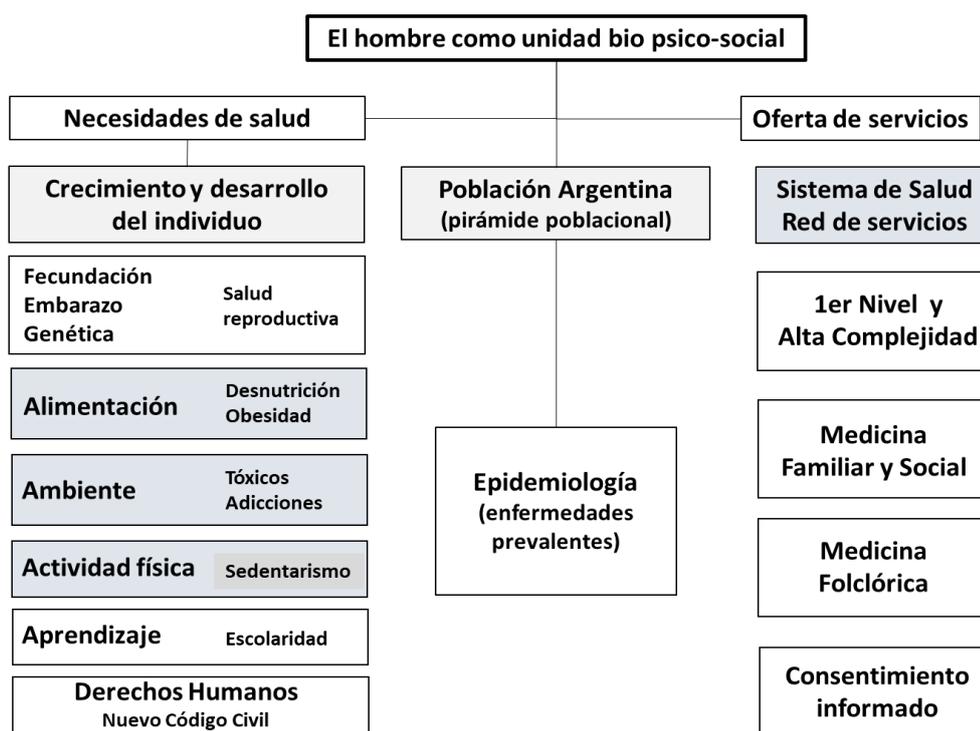
Fuente: elaboración propia

#### 4.1.2. Caso de estudio MHS

La asignatura Medicina Hombre y Sociedad es la primera materia de la Carrera de Medicina de la UNNE que los alumnos inscriptos deben regularizar y aprobar para ingresar como estudiantes regulares a dicha carrera. El programa de MHS está orientado en

competencias y enriquecido con contenidos de Medicina Familiar y Comunitaria, paradigma en el que está centrado el perfil de graduación de la carrera. El cursado, con 400 horas de carga horaria, está estructurado en cuatro ejes temáticos: “Alimentación”, “Actividad Física”, “Ambiente” y “Sistemas de Salud”; y un eje de “Contenidos Transversales” (antropología médica, epistemología, metodología de estudio, informática e inglés) dirigido a la formación humanística, científica y metodológica. Figura 4.2.

**Figura 4.2:** Diagrama que muestra la organización y articulación de los contenidos de MHS



Fuente: tomado del programa de la asignatura MHS

Según la planificación 2019 (año del estudio)<sup>30</sup>, las actividades de aprendizaje se desarrollaron en tres escenarios: **áulicas presenciales**, comisiones de 80 estudiantes subdivididos en pequeños grupos coordinados por docentes-tutores, donde trabajaron en formatos de aula inversa, talleres y cine debate; **aula virtual** donde se desarrollaron los contenidos transversales a través de diferentes actividades y recursos, en forma integrada con los otros ejes longitudinales de la materia; y **trabajos en terreno**, en grupos de ocho estudiantes con apoyo tutorial, que consistieron en observaciones sistemáticas, realización de entrevistas a equipos de salud, aplicación de encuestas, y análisis de los datos epidemiológicos publicados.

<sup>30</sup> Resolución N°4268/17-CD: Aprobación del Programa de la asignatura Medicina Hombre y Sociedad vigente a partir del ciclo lectivo 2018.

El cursado se desarrolló en dos tramos con cinco instancias de evaluación al final de cada tramo (dos grupales y tres individuales), que los estudiantes debían aprobar al menos tres de las cinco para mantener su regularidad en la asignatura.

Las instancias grupales abarcaron, la evaluación de los **informes de los trabajos en terreno** (ITT) que los estudiantes realizaron durante el cursado de cada tramo y una **evaluación de contenidos transversales** (ECT) a través de diferentes actividades sincrónicas en el aula virtual más un cuestionario de lecto-comprensión de inglés.

Las evaluaciones individuales de cada tramo fueron pruebas escritas, **dos exámenes integradores de ejes**: “Alimentación con Actividad física” (ALAF) y “Sistemas de salud con Ambiente” (SISAM); y **un examen integrador de tramo** (EIT). Los dos primeros con preguntas abiertas de respuesta corta y el último con preguntas de estímulo formato opción múltiple.

**Cuerpo docente:** En el cursado 2019 participaron 37 docentes: un (1) profesor titular, tres (3) profesores adjuntos, 21 jefes de trabajos prácticos, 10 docentes adscriptos y dos (2) ayudantes alumno.

**Alumnos cursantes:** Para el ciclo lectivo 2019, se inscribieron 1.363 alumnos, de los cuales 462 completaron el cursado y rindieron las evaluaciones del segundo tramo

#### 4.1.3. Caso de estudio MED-III

La asignatura Medicina III es una materia de 240 horas de carga horaria que se dicta en quinto año de la Carrera de Medicina de la UNNE, su programa está orientado en competencias relacionadas al manejo clínico de las enfermedades hematológicas, reumáticas, digestivas, hepáticas, endocrinas, metabólicas, nutricionales e infecciosas.

Según la planificación 2019 (año del estudio)<sup>31</sup>, las actividades de aprendizaje se desarrollaron en diferentes escenarios como: **aulas invertidas** que abarcaron una fase de aprendizaje autónomo de los alumnos en aula virtual (casos clínicos o problemas, videoclases y material bibliográfico) guiados por consignas, y luego de algunos días aplicaron ese conocimiento trabajando en pequeños grupos en aula presencial con apoyo tutorial docente; **talleres** subdivididos en pequeños grupos donde a partir de un texto, imagen o grabación de una situación problemática real desarrollaron un trabajo sistemático (análisis, investigación y discusión) que culminó con la entrega de una producción grupal (mapa conceptual, video, narración o poster); **trabajos prácticos con pacientes** que realizaron en pequeños grupos

---

<sup>31</sup> Resolución N°613/18-CD: Aprobación del Programa de la asignatura Medicina III vigente a partir del ciclo lectivo 2018.

supervisados por un JTP, en sectores de internación de los hospitales y en centros del primer nivel de atención; **actividades en aula virtual** como foros de discusión de casos clínicos, cápsulas de microaprendizaje, cuestionarios, textos en línea, encuestas y un e-portfolio individual; y otras **actividades electivas** como ateneos anatómico-clínicos hospitalarios, prácticas en consultorios externos de clínica y medicina familiar, mini-guardias en servicios de urgencia de los hospitales, que los estudiantes pudieron realizar como actividades complementarias según sus preferencias.

Las evaluaciones aplicadas en 2019, se integraron en un sistema de evaluación longitudinal multimétodo con doble propósito, formativo y sumativo.

Para la evaluación formativa, se utilizaron cuatro tipos de herramientas en forma articulada: **pruebas con preguntas de opción múltiple**, para evaluar el conocimiento declarativo y la resolución de problemas; **mini-ejercicios de evaluación clínica (mini-CEX)**, para evaluar por observación habilidades, conductas y actitudes clínicas; **elaboración de una historia clínica completa y búsqueda de evidencias** relacionadas, para evaluar razonamiento clínico, toma de decisiones y habilidades para la medicina basada en la evidencia; y **tareas en el aula virtual** (cafecito infectológico y portfolio) para evaluar la capacidad de resolución de problemas, el razonamiento clínico y aspectos metacognitivos relacionados con el aprendizaje. Todas estas instancias se acompañaron de una fase de **devolución constructiva** para que los estudiantes tomen conciencia de sus avances, identifiquen sus dificultades, y puedan hacer los ajustes necesarios para mejorar y reforzar su aprendizaje.

Para la evaluación sumativa se aplicó la sumatoria de los puntos obtenidos en cada instancia de evaluación formativa (escala 0 a 100 cada una) y la sumatoria de los puntos obtenidos por la participación y desempeño de los estudiantes en las actividades de aprendizaje desarrolladas en el cursado: aulas invertidas, talleres, trabajos prácticos y foros virtuales (escala 0 a 50 cada una), con lo cual los máximos puntajes posibles fueron 800 puntos en evaluaciones y 200 puntos en participación y desempeño, máximo total: 1000 puntos.

**Cuerpo docente:** En el cursado 2019 participaron 35 docentes: tres (3) profesores titulares, seis (6) profesores adjuntos, 18 jefes de trabajos prácticos, seis (6) docentes adscriptos y dos (2) ayudantes alumno.

**Alumnos cursantes:** Para el ciclo lectivo 2019, se inscribieron a la asignatura 151 estudiantes de los cuales 149 completaron el cursado y regularizaron la materia.

## 4.2. Primer Eje de Análisis: Evaluación del aprendizaje

### 4.2.1. Entorno de aprendizaje y evaluación

Comprende las características del entorno de aprendizaje y de los sistemas de evaluación, sus fundamentos, propósitos, diseño, planificación y aplicación, en el contexto de cada materia, los instrumentos de evaluación utilizados, sus propiedades psicométricas y el alcance cognitivo de las tareas y consignas planteadas en las pruebas.

#### 4.2.1.1. Caso MHS

Como se desprende del programa, el entorno de aprendizaje de los estudiantes en esta asignatura está compuesto principalmente por actividades grupales, centradas en tareas específicas (discusión de un problema, análisis de un caso, discusión de un video, observaciones en terreno, análisis de datos epidemiológicos y otras) y en producciones conjuntas (informes, posters, videos, y otros productos); todas con apoyo tutorial de los docentes y con materiales didácticos de elaboración propia, que se canalizan a través del campus virtual de la Facultad de Medicina.

En este escenario, el sistema de evaluación está integrado a los procesos de enseñanza y aprendizaje, y cobra valor en tanto aporta información que enriquece y fortalece el desarrollo curricular de la propuesta, pues intenta influir en las formas de aprendizaje de los estudiantes y en los obstáculos que encuentran en dicho proceso. Todas las evaluaciones están enmarcadas en el principio de “evaluación auténtica” y hacen foco en el estudiante y su desempeño, ubicándolo en su propio contexto y frente a situaciones de aprendizaje significativas y complejas, tanto a nivel individual como grupal.

De acuerdo a la planificación 2019, la asignatura se desarrolló en dos tramos de cursado con cinco instancias de evaluación al final de cada tramo (tres individuales y dos grupales), que los alumnos debían aprobar (al menos tres de las cinco) para mantener su regularidad.

Para las evaluaciones individuales se aplicaron dos exámenes integradores de ejes: ALAF (Alimentación con Actividad física) y SISAM (Sistemas de salud con Ambiente); y un examen integrador de tramo (EIT). Cada integrador de ejes incluyó cinco ítems con cinco consignas o preguntas de respuesta abierta (PRA) por ítem, que se evaluaron por comparación con “pruebas patrón” (respuestas correctas esperadas) y rúbricas específicas que indicaban los criterios de evaluación y de asignación de puntajes, estas herramientas fueron elaboradas por los docentes que diseñaron las pruebas. Para el EIT se aplicó un cuestionario con 80 preguntas de estímulo, de opción múltiple (POM), ricas en contexto, que se calificaron con lector óptico.

En las instancias grupales se evaluaron los informes de los trabajos en terreno (ITT) que los estudiantes realizaron durante el cursado de cada tramo y una evaluación de contenidos transversales (ECT) a través de diferentes actividades del aula virtual y un cuestionario de lecto-comprensión de inglés.

Los instrumentos de evaluación seleccionados permiten abordar: los dominios declarativos a través de pruebas escritas estructuradas y abiertas (formatos de estímulo y de respuesta); los dominios procedimentales a partir de los informes de los trabajos en terreno, el análisis de los datos y los diagnósticos de situación; y los dominios orientados a la metacognición con preguntas reflexivas sobre el aprendizaje en la evaluación de contenidos transversales.

Para este trabajo se analizaron solo los instrumentos de evaluación aplicados en el segundo tramo que incluyó a los 462 estudiantes que finalizaron el cursado de la materia.

En los instrumentos de evaluación individual se consideraron las propiedades psicométricas y el alcance cognitivo de las preguntas. Las pruebas integradoras de ejes (ALAF y SISAM) se analizaron en forma conjunta, reuniendo así diez ítems con cinco consignas cada uno, que totalizaron 50 PRA, mientras que en el EIT se analizaron 80 POM con tres opciones.

En los instrumentos de evaluación grupal (ITT y ECT) solo se exploró el alcance cognitivo, pues las características heterogéneas de su diseño, no permitió el análisis psicométrico.

#### • **Variables psicométricas de los cuestionarios**

En los análisis psicométricos realizados por los profesores responsables de las evaluaciones luego de su aplicación a los estudiantes de MHS se obtuvieron los siguientes resultados:

Los exámenes integradores de eje (ALAF y SISAM) analizados en forma conjunta, tuvieron un coeficiente de confiabilidad “aceptable”<sup>32</sup> (Alfa de Crombach = 0,74); y las 50 preguntas que los componen, mostraron en promedio ( $\pm$  DE), un Índice de Dificultad “intermedia”<sup>33</sup> ( $iP = 0,49 \pm 0,08$ ) y un Índice de Discriminación “aceptable”<sup>34</sup> ( $iD = 0,36 \pm 0,10$ ).

---

<sup>32</sup> La confiabilidad de los instrumentos de evaluación se reconoce como aceptable con valores de Alfa de Crombach  $\geq 0.70$  (Santos, 1999)

<sup>33</sup> Los valores se interpretaron según los niveles propuestos por Haladyna et al (2002) en Muy fácil ( $iP = 0,92-1,00$ ); Medianamente fácil ( $iP = 0,76-0,91$ ); Dificultad intermedia ( $iP = 0,45-0,75$ ); Medianamente difícil ( $iP = 0,25-0,44$ ); y Muy difícil ( $iP = 0,00-0,24$ ).

<sup>34</sup> Los valores se interpretaron según la clasificación de Ebel (1965) en Discriminación: Alta ( $iD \geq 0,40$ ); Aceptable ( $iD = 0,30 - 0,39$ ); Baja ( $iD = 0,20 - 0,29$ ); Mala ( $iD = 0,00 - 0,19$ ); e Inaceptable ( $iD < 0,00$ )

El examen integrador de tramo, tuvo un coeficiente de confiabilidad “aceptable” (Alfa de Crombach = 0,71); y las 80 preguntas que integran la prueba, mostraron en promedio ( $\pm$  DE), un Índice de Dificultad “intermedia” ( $iP = 0,51 \pm 0,19$ ) y un Índice de Discriminación “baja” ( $iD = 0,28 \pm 0,18$ ).

- **Alcance cognitivo de las preguntas**

Los ítems de los **exámenes integradores de ejes** utilizaron como disparadores, en el ALAF: etiquetas nutricionales, relato que introduce al metabolismo hidroelectrolítico, listado de nutrientes, dibujo esquemático de la fibra muscular, gráfico de producción de energía en ejercicio; y en el SISAM: familigramas, problema relacionado a determinantes de salud, relato que introduce a la división y ciclo celular, casos de salud en diferentes etapas de la vida, problemas del sistema de salud. Cada uno de estos ítems comprenden cinco PRA con consignas que llevó a los estudiantes a realizar algunos de estos procesos: interpretar, identificar, ordenar, comparar, seleccionar, combinar, esquematizar, graficar, integrar, valorar, justificar. La aplicación de estos procesos dio a las PRA un alcance cognitivo de cinco niveles en la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar.

Analizando las 80 POM del **EIT** se observó que todas incluyeron un breve enunciado relacionado a los contenidos de la asignatura, redactados como: problema social, situación sanitaria, caso clínico, experiencia de laboratorio, proceso biológico, mecanismo bioquímico, fenómeno de la naturaleza, teoría científica, suceso histórico y referencia antropológica. Debajo de cada enunciado se planteó una pregunta con tres opciones de respuesta (una correcta y dos distractores), que exigió a los estudiantes realizar algunos de estos procesos: Recordar (conceptos, definiciones, propiedades, modelos); Identificar (elementos, compuestos, estructuras, funciones); Interpretar (datos, formulas, tablas, gráficos, mecanismos); Reconocer (características, patrones, factores, categorías); Ordenar (elementos, números, etapas). En consecuencia, el alcance cognitivo de las POM abarcó los tres primeros niveles de la Taxonomía de Bloom: Recordar, Comprender y Aplicar.

Para el **Trabajo en Terreno**, los grupos de estudiantes, en el entorno de su núcleo de convivencia y en el ámbito barrial, tuvieron que recolectar y registrar datos relacionados a los ejes de la materia (alimentación, actividad física y ambiente) y desarrollar como tareas: reconocer en la alimentación familiar, hábitos saludables y poco saludables; caracterizar sus actividades físicas y deportivas; identificar conductas barriales que podían impactar sobre el ambiente; y elaborar e interpretar un familigrama estructural y funcional. En el informe registraron valores y unidades, calcularon porcentajes, proporciones y medidas de tendencia

central, elaboraron tablas y gráficos, y los comentarios y conclusiones correspondientes. Con lo cual, el alcance cognitivo de este ejercicio abarcó los cinco niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar.

En la **Evaluación de Contenidos Transversales** los estudiantes en pequeños grupos, tuvieron que: reconocer el paradigma de investigación de un estudio y recordar sus características metodológicas; identificar el tipo de información de un texto y explicar su uso, comparar la cita de un artículo con las normas APA; identificar los criterios de validez de un artículo; analizar y valorar el diseño de un mapa conceptual; recordar y ordenar los pasos a seguir para buscar un artículo en internet. Además, respondieron un cuestionario de lecto-comprensión de inglés. Con lo cual, el alcance cognitivo de este ejercicio abarcó los cinco niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar.

#### **4.2.1.2. Caso MED-III**

Según lo enunciado en el programa de la asignatura, el aprendizaje de los estudiantes se desarrolla en pequeños grupos, con apoyo tutorial y/o supervisión docente; y está diversificado en actividades de integración teórico-prácticas (aula inversa y talleres), trabajos prácticos con pacientes (hospitales y centros de salud), actividades en aula virtual (foros de casos clínicos, mini-test, cápsulas de microaprendizaje, y otras), a las que suman otras actividades electivas (ateneos hospitalarios, prácticas en consultorio, mini-guardias, y otras), que los estudiantes realizan según sus preferencias.

El sistema de evaluación longitudinal multimétodo de MED-III, fue diseñado para recoger información más amplia y completa sobre el desarrollo de las competencias, y para superar las limitaciones que tienen los instrumentos de evaluación aplicados en forma aislada. Con esa intención, se eligieron herramienta de evaluación que exploran diferentes dimensiones de las competencias: para evaluar conocimientos declarativos y resolución de problemas se aplicaron exámenes parciales "multiple choice"; para evaluar habilidades procedimentales y actitudes se utilizó mini-CEX; para evaluar razonamiento clínico, toma de decisiones y habilidades para la búsqueda de información, se solicitaron historias clínicas con evidencias científicas; y para evaluar la capacidad de resolución de problemas, razonamiento clínico y aspectos metacognitivos se plantearon trabajos en el aula virtual.

Los **exámenes parciales** fueron dos pruebas escritas, con 80 preguntas de base clínica multi-items, formato A de 3 opciones<sup>35</sup>, complementados con una fase de retroalimentación luego de su corrección con lector óptico.

Los **mini-CEX**, se aplicaron en los trabajos prácticos con pacientes, para evaluar el desempeño de los estudiantes en una o más de estas dimensiones: habilidades para conducir la entrevista médica; habilidades para conducir el examen físico; cualidades humanísticas y profesionalismo; criterio clínico; habilidades para asesor al paciente; organización y eficiencia; y competencia clínica global. Durante el curso, se realizaron dos evaluaciones a cada estudiante acompañadas de una devolución constructiva.

El **Trabajo Final** (individual) consistió en elaborar: a) una base de datos de Historia Clínica; b) listar y jerarquizar los problemas del paciente; c) redactar una discusión, planteando y fundamentando las hipótesis diagnósticas; d) proponer un plan de estudio y e) un plan terapéutico; y f) efectuar una búsqueda bibliográfica a partir de uno de los problemas identificados. Cada estudiante eligió “su paciente” en uno de sus prácticos, y desarrollo el trabajo durante el cursado con apoyo de los JTP y tutorías en pequeños grupos. La evaluación de estos trabajos finales fue realizada por los docentes a través de una lista de cotejo, para evaluar el cumplimiento de todas las etapas del trabajo, y de rubricas, para determinar con mayor objetividad la calidad del mismo.

El “**Cafecito infectológico**” se implementó como un foro virtual donde los estudiantes en grupos de 20, acompañados por un tutor-facilitador, discutieron casos clínicos de infectología y fueron evaluados con Rúbricas específicas que consideraron: la pertinencia y relevancia de las participaciones; la capacidad de análisis; la calidad de los aportes; el dominio del lenguaje; y la frecuencia de las participaciones.

El **Portfolio** consistió en un archivo de texto donde los estudiantes registraron sus experiencias de aprendizaje a través de la escritura reflexiva y aportaron evidencias de lo aprendido. En el cursado 2019 se les solicitó: a) la evidencia y reflexión de su aprendizaje en un Ateneo Clínico (de su elección); b) el relato de una situación vinculada al Profesionalismo (médico o estudiantil) o a la Seguridad del Paciente (positiva o negativa), observada durante

---

<sup>35</sup> Las preguntas “múltiple choice” son clasificadas por la NBME (National Board of Medical Examiners) en los siguientes formatos, Selección de una sola respuesta: Formato A únicas o multi-ítem (3 o más opciones); Formato B emparejamiento (4-5 opciones en 2-5 preguntas); Formato F-G (agrupadas por temas específicos); Formato R emparejamiento ampliado (5 a 20 preguntas). Selección de todas las opciones verdaderas: Formato C (A/B/Ambas/Ninguna); Formato K (complejas de verdadero/falso); Formato X (simples de verdadero/falso). (Paniagua & Swygert, 2016)

el año, con una reflexión de lo aprendido; y c) una reflexión global sobre lo que aportó el portfolio a su aprendizaje.

Estos instrumentos se aplicaron como evaluación formativa durante el cursado, realizando una fase de retroalimentación después de cada evaluación; y con la sumatoria de los resultados obtenidos se construyó el puntaje final de cursado (evaluación sumativa).

- **Variables psicométricas de los cuestionarios**

En los análisis psicométricos que realizaron los profesores responsables de las evaluaciones luego de su aplicación a los estudiantes de MED-III se obtuvieron los siguientes resultados:

El primer examen parcial, tuvo un coeficiente de confiabilidad “aceptable” (Alfa de Crombach = 0,76); y las 80 preguntas de opción múltiple que los componen, mostraron en promedio ( $\pm$  DE), un Índice de Dificultad “intermedia” ( $iP = 0,59 \pm 0,23$ ) y un Índice de Discriminación “baja” ( $iD = 0,21 \pm 0,12$ ).

El segundo examen parcial, tuvo un coeficiente de confiabilidad “aceptable” (Alfa de Crombach = 0,72); y las 80 preguntas que integran la prueba, mostraron en promedio ( $\pm$  DE), un Índice de Dificultad “intermedia” ( $iP = 0,57 \pm 0,24$ ) y un Índice de Discriminación “baja” ( $iD = 0,20 \pm 0,13$ ).

- **Alcance cognitivo de las evaluaciones**

Analizando las POM de los dos **exámenes parciales** aplicados, se observó que todas incluyeron como enunciado una breve viñeta clínica con los datos esenciales (clínicos y complementarios) necesarios para elegir la mejor respuesta. Debajo de cada enunciado se plantearon dos o más preguntas con tres opciones de respuesta (una correcta y dos distractores), que exigió a los estudiantes realizar algunos de estos procesos: Recordar (etiología más probable, entidad de mayor prevalencia, objetivos terapéuticos, tiempo de tratamiento, método para confirmar un diagnóstico); Identificar (dato anormal, estudio más sensible, efecto adverso, antibiótico de elección); Interpretar (signo o síntoma, resultado de laboratorio, imagen o estudio); Reconocer (patrones o síndromes, signos de gravedad, determinación que falta, resultado esperado); Comparar (estudios o tratamientos indicados); Tomar decisiones (hipótesis diagnóstica más probable, tratamiento de elección o más efectivo, medida de urgencia recomendada, conducta más correcta, suspensión del tratamiento); Justificar (preguntas para orientar el diagnóstico, elementos que fundamentan una conducta); Planificar (secuencia de pruebas diagnósticas, estudios para evaluar la

evolución, plan de inmunización); Sintetizar (diagnóstico que explica la mayoría de las manifestaciones); Valorar (razonamiento más apropiado, compromiso más relevante, pronóstico más probable, lugar más apropiado para internar). En consecuencia, el alcance cognitivo de las POM abarcó cinco niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar.

Con los **mini-CEX** se examinaron las habilidades de los estudiantes para conducir la entrevista clínica y el examen físico de los pacientes, y se evaluó la capacidad para: identificar los problemas activos, interpretar las manifestaciones integrándolas en síndromes o asociaciones patogénicas, proponer hipótesis diagnósticas fundadas, plantear y justificar la secuencia de estudios diagnósticos, establecer un plan terapéutico adecuado y pautas para la educación del paciente, priorizando los estudios y estrategias de intervención, en relación al contexto, al análisis costo/beneficio y a la oportunidad. Con lo cual el alcance cognitivo de los mini-CEX abarcó cinco niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar.

Respecto al **Trabajo Final**, los estudiantes tuvieron que realizar las siguientes tareas: redactar con claridad y precisión una base de datos para la Historia Clínica; identificar y jerarquizar los problemas que presenta el paciente; elaborar con los datos disponibles, una discusión diagnóstica coherente; formular y fundamentar las hipótesis diagnósticas; diseñar y justificar un plan de estudios pertinente; establecer un plan terapéutico adecuado; y confeccionar un resumen claro y preciso de la Historia Clínica. Además, para el trabajo bibliográfico debían: seleccionar un problema del paciente; convertir el problema en una pregunta estructurada (formato PICO); realizar la búsqueda bibliográfica; confeccionar la ficha bibliográfica; y evaluar la validez externa del artículo. Con lo cual el alcance cognitivo de este ejercicio abarcó seis niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar, Evaluar y Crear.

En el **“Cafecito infectológico”** los estudiantes discutieron casos clínicos de infectología para lo cual tuvieron que identificar y jerarquizar problemas, describir e interpretar imágenes radiológicas, proponer y justificar diagnósticos diferenciales y alternativas terapéuticas, y aportar evidencias científicas para fundamentar sus comentarios. En consecuencia, el alcance cognitivo de esta actividad abarcó cinco niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar.

A través del **Portfolio** se buscó que los estudiantes desarrollaran habilidades metacognitivas, analizando y reflexionando sobre los procesos cognitivos que aplicaron para

su aprendizaje en otras actividades realizadas durante el cursado, y de esa forma tomaran conciencia de sus propios recursos cognitivos y consideraran posibles estrategias para lograr un aprendizaje autónomo más efectivo.

#### 4.2.2. Estudiantes evaluados

Abarca las metas de logro que tienen los estudiantes evaluados, los enfoques y estrategias de aprendizaje que utilizaron en el contexto del sistema de evaluación aplicado, y los resultados obtenidos.

##### 4.2.2.1. Caso MHS

De los 462 alumnos que finalizaron en 2019 el cursado de la materia, 322 estudiantes (70%) completaron los cuestionarios de Metas de Logro (A-AGQ-R) y de Procesos de Estudio - Dos Factores (R-SPQ-2F), que se administraron en el mismo formulario online en forma voluntaria y anónima.

En relación a las **Metas de Logro** y desde el esquema 2 x 2 de Elliot & Murayama (2008), los estudiantes de MHS, que fueron los alumnos ingresantes a la carrera de medicina en 2019, mostraron mayor inclinación por las metas de aproximación, tanto al aprendizaje como al rendimiento. Tabla 4.1.

**Tabla 4.1:** Metas de logro de los 322 estudiantes que completaron el cuestionario A-AGQ-R

Metas de logro	n	Media	DE	Mediana	Rango
Aprendizaje - Aproximación	322	4.4	0.6	4.7	1 - 5
Aprendizaje - Evitación	322	3.5	1.2	3.7	1 - 5
Rendimiento - Aproximación	322	4.2	1.1	5.0	1 - 5
Rendimiento - Evitación	322	3.4	1.3	3.7	1 - 5

Fuente: propia

En el cuestionario A-AGQ-R, las metas de “aproximación al aprendizaje” implican llegar a dominar los materiales de estudio, comprender todos los contenidos de la asignatura y aprender tanto como sea posible. Mientras que las metas de “aproximación al rendimiento” conllevan desempeñarse bien y obtener los mejores resultados con respecto a los otros compañeros del curso. Este resultado puede explicarse por la exigencia que tiene este grupo de alumnos de regularizar y aprobar la primera asignatura del plan de estudios para poder ingresar a la carrera de medicina.

En cuanto a los **Enfoques de Aprendizaje** y según los criterios establecidos por Biggs (1988), los estudiantes de MHS, en su mayoría, adoptaron un enfoque profundo con intensidad media, mientras un grupo minoritario mostró un enfoque superficial. Tabla 4.2.

**Tabla 4.2:** Enfoques de aprendizaje de los 322 alumnos que completaron el R-SPQ-2F

Enfoque	Intensidad	N	%
Profundo	Alta	61	0.19
Profundo	Media	170	0.53
Profundo	Baja	52	0.16
Neutro		1	0.00
Superficial	Baja	38	0.12

Fuente: propia

Esta mayoritaria presencia de enfoques profundos entre los alumnos “ingresantes” puede explicarse por la fuerte motivación intrínseca que genera la “nueva experiencia” (ingresar a la universidad y a la carrera de medicina) y el interés por aprender contenidos “nuevos”.

Pero también es comprensible que en este grupo de alumnos (ingresantes) haya algunos con enfoque de aprendizaje superficial, pues este enfoque está vinculado a la motivación extrínseca, donde el estudiante tiene miedo al fracaso (a no poder ingresar a la carrera) y se preocupa por cumplir con todas las obligaciones y consignas que exige la asignatura para su regularidad y aprobación.

En el cuestionario R-SPQ-2F un alto porcentaje de alumnos estuvo “**bastante**” o “**muy de acuerdo**” con estos enunciados: “Me esfuerzo por leer la mayor parte de los materiales que me recomiendan en clase” (91%); “Me parecen muy interesantes todos los temas cuando los empiezo a estudiar (83%); “Trabajo duro en mis estudios cuando creo que el material o el contenido son interesantes” (91%).

En cambio, estuvieron “**poco**” o “**nada de acuerdo**” con las siguientes proposiciones: “Solo estudio lo que se ve en clase” (79%); “La mejor forma de aprobar los exámenes es memorizar las respuestas de las preguntas más probables” (88%); “Puedo aprobar los exámenes memorizando partes claves .... sin intentar comprenderlas” (84%)

En cuanto a los **resultados** que obtuvieron los estudiantes en las evaluaciones del segundo tramo 2019 de MHS, se resumen en la Tabla 4.3., y están expresados como “puntajes” (escala 0 a 100) pues son los valores que arroja la corrección y que luego son

convertidos en letras (A, B o C) para informar los resultados a los estudiantes y decidir sobre la aprobación o no de cada instancia.

**Tabla 4.3:** Puntajes de los 462 estudiantes de MHS en cada evaluación (Escala 0-100)

Instancias de evaluación	Media	DE	Rango*
Examen Integrador de ejes (ALAF)	61.2	22.6	10 - 100
Examen Integrador de ejes (SISAM)	64.3	21.0	10 - 100
Examen Integrador de tramo (EIT)	65.3	11.0	32 - 91
Evaluación de Trabajos en terreno	66.3	18.3	20 - 100
Evaluación Contenidos transversales	81.6	14.5	30 - 100

\* El rango es el observado en los registros (puntaje mínimo y máximo obtenidos)

En la Tabla se observa que los mayores puntajes corresponden a las evaluaciones de los trabajos en terreno y de los contenidos transversales, que son las dos evaluaciones que se realizan en forma grupal, pues las otras (ALAF, SISAM y EIT) son pruebas individuales. La diferencia de puntajes queda más clara cuando se compara el promedio de los tres exámenes individuales ( $63.6 \pm 18.2$ ) con el promedio de los dos grupales ( $73.9 \pm 16.4$ ), que muestra una diferencia de medias muy significativa: 10.3 con un IC (95%) 8.06 – 12.53 ( $p= 0.0000$ ).

Esta disparidad en los resultados puede explicarse no solo por la forma en como los estudiantes respondieron a la prueba (en grupos e individualmente), también pudo influir el tiempo de desarrollo y discusión grupal, que demandaron los trabajos evaluados.

Por otro lado, el amplio rango que se observa en los resultados indica un alto poder de “discriminación” de los instrumentos de evaluación, que permiten diferenciar a los alumnos “bien preparados” para la prueba de aquellos que no estaban suficientemente preparados para rendir.

#### 4.2.2.2. Caso MED-III

De los 149 estudiantes que finalizaron en 2019 el cursado de la materia, 139 (93%) completaron los cuestionarios de Metas de Logro (A-AGQ-R) y de Procesos de Estudio - Dos Factores (R-SPQ-2F), que integraban el mismo formulario on line.

Como **Metas de Logro** y en el esquema 2 x 2 de Elliot & Murayama (2008), los estudiantes de MED-III, que estaban en el último año de cursado antes de realizar la PFO, mostraron una fuerte preferencia por las metas de aproximación al aprendizaje. Tabla 4.4.

**Tabla 4.4:** Metas de logro de los 139 estudiantes que completaron el cuestionario A-AGQ-R

Metas de logro	n	Media	DE	Mediana	Rango
Aprendizaje - Aproximación	139	4.5	0.6	4.7	1 - 5
Aprendizaje - Evitación	139	3.4	1.2	3.7	1 - 5
Rendimiento - Aproximación	139	2.9	1.3	3.0	1 - 5
Rendimiento - Evitación	139	3.0	1.3	3.0	1 - 5

Fuente: propia

Este resultado es el esperado en estudiantes avanzados de la carrera de medicina, que están dispuestos a esforzarse para avanzar en su aprendizaje, y desarrollar las habilidades y competencias que demanda su formación profesional.

En relación a los **Enfoques de Aprendizaje** y siguiendo los criterios de Biggs (1988), los estudiantes en MED-III, casi en su totalidad, tuvieron un enfoque profundo que aplicaron con intensidad media y alta, mayoritariamente. Tabla 4.5.

Tabla 4.5. Enfoques de aprendizaje de los 139 estudiantes que completaron el R-SPQ-2F

Enfoque	Intensidad	N	%
Profundo	Alta	49	0.35
Profundo	Media	69	0.50
Profundo	Baja	13	0.09
Neutro		3	0.02
Superficial	Baja	5	0.04

Fuente: propia

La mayor intensidad del enfoque profundo en este grupo de estudiantes puede estar relacionada con alta motivación intrínseca, ya que se trata de estudiantes avanzados que están llegando al final de la carrera y tienen un mayor compromiso con lo que están aprendiendo, pues encuentran al aprendizaje emocionalmente satisfactorio y esencial para su futura (y cercana) vida profesional.

En el cuestionario R-SPQ-2F un alto porcentaje de alumnos estuvo **“bastante”** o **“muy de acuerdo”** con estas proposiciones: “Tengo momentos en los que estudiar me produce gran satisfacción” (88%); “Debo estudiar bastante un tema para poder formar mis propias

conclusiones y así quedar conforme” (76%); “Cuando estudio temas importantes me autoevalúo tanto como sea necesario hasta lograr comprenderlos por completo” (85%).

En cambio, estuvieron “poco” o “nada de acuerdo” con estas proposiciones: “Generalmente me limito a estudiar solo lo que se me pide, porque creo innecesario hacer cosas extra” (84%); “Creo que no es conveniente estudiar los temas a fondo. Eso produce confusión y pérdida de tiempo...” (87%); “Mi objetivo es aprobar la materia haciendo el menor trabajo posible” (83%).

Los **resultados de las evaluaciones** que obtuvieron los estudiantes durante el cursado 2019 de MED-III, se resumen en la Tabla 4.6., y se expresan como “puntajes” (escala 0 a 100) pues son los valores que arroja la corrección, y luego son convertidos en Notas (0-10) para informar los resultados a los estudiantes y decidir sobre su regularidad en la materia.

**Tabla 4.6:** Puntajes de los 149 estudiantes de MED-III en cada evaluación (Escala 0-100)

Instancias de evaluación	Media	DE	Rango*
Examen parcial (primero)	61.7	7.9	45 - 91
Examen parcial (segundo)	59.2	7.8	40 - 81
Mini CEX (primero)	79.6	11.4	50 - 100
Mini CEX (segundo)	77.7	12.2	50 - 100
Historia Clínica (completa)	83.4	12.4	16 - 100
Resumen HC y evidencia	85.5	14.0	20 - 100
Cafecito infectológico	77.2	15.4	50 - 100
Portfolio personal	84.8	15.7	30 - 100

\* El rango es el observado en los registros (puntaje mínimo y máximo obtenidos)

Llama la atención la similitud de puntajes que se observa en los resultados que provienen del mismo de tipo de instrumento de evaluación; los puntajes más bajos (< 62) corresponden a los exámenes “multiple choice”; los valores intermedios (77 a 80), a las evaluaciones del desempeño clínico y discusión de casos clínicos; y los puntajes más altos (> 83), a las pruebas de ensayo como la Historia Clínica y el Portfolio.

Estas diferencias, observadas en el mismo grupo de estudiantes (donde uno podría esperar resultados más homogéneos), pueden explicarse en los objetivos que tienen las distintas evaluaciones, pues cada instrumento evalúa diferentes componentes de las competencias, y puede ocurrir que un alumno que no tenga un conocimiento teórico

suficiente, demuestre un buen desempeño clínico, y tenga mucha capacidad para redactar una Historia Clínica con claridad y precisión.

Por otro lado, las diferencias observadas en la amplitud de los rangos de puntaje entre los distintos instrumentos, indica que no todos tuvieron un buen poder de “discriminación” pero aplicados en conjunto permitieron diferenciar a los alumnos “bien preparados” para las pruebas de aquellos que no estaban suficientemente preparados para rendir.

### **4.3. Segundo Eje de Análisis: Evaluación para el aprendizaje**

#### **4.3.1. Factores de la evaluación**

Comprende los objetivos y criterios de las evaluaciones; el material evaluado; la interacción con los docentes; la proximidad temporal a la evaluación; y la sobrecarga de tareas simultaneas.

##### **4.3.1.1. Caso MHS**

En la documentación aportada por la asignatura, se observa que las guías didácticas y los instructivos de los diferentes instrumentos de evaluación (excepto el “multiple choice” integrador de tramo), tenían explicitados los objetivos de aprendizaje explorados y los criterios generales de evaluación y calificación de cada prueba. Pero se aclara que los estudiantes no tuvieron acceso antes de rendir, a las “pruebas patrón” (posibles respuestas correctas de las preguntas abiertas) ni a las rubricas de evaluación que fueron utilizadas por los docentes para la corrección y calificación de los exámenes.

Sin embargo, en el cuestionario online, la mayoría de los estudiantes (68%) manifestó no tener conocimiento previo, de los objetivos de aprendizaje explorados y de los criterios de evaluación aplicados; y solo el 14% indicó tener un amplio conocimiento de estos aspectos antes de rendir cada evaluación.

Dos semanas antes de las evaluaciones, la asignatura habilitó un foro online de consultas sobre las pruebas y una solicitud de tutorías a demanda con los docentes, estas actividades (según el informe) tuvieron muy baja participación de los estudiantes, menos del 15% de los cursantes utilizaron las mismas. No obstante, la mayoría de los alumnos (82%) reconoce haber recibido de los docentes responsables de su comisión, apoyo y orientación para la preparación de los exámenes.

Las cinco evaluaciones del segundo tramo fueron aplicadas en un período de dos semanas con suspensión de todas las actividades de cursado y con un intervalo de 1 a 2 días

entre cada examen según el ordenamiento de las comisiones. Esta planificación de evaluaciones muy cercanas en el tiempo fue cuestionada por algunos estudiantes con expresiones como: *“No somos genios...es imposible rendir cinco exámenes en dos semanas”*; *“El estrés sometido a los alumnos es exuberante...”*; *“...el lunes a la tarde rendimos el ALAF y el miércoles a la mañana tuvimos el SISAM, una locura”*; *“...al estar dos semanas presionada con 5 exámenes llegue el ultimo día agotada”*; *“Las evaluaciones no me dejaban dormir”*; *“Si no quieren que ingresemos, que nos tomen un solo examen eliminatorio así no pasamos por tanto estrés”*

Otros estuvieron más conformes: *“Las evaluaciones están bien organizadas y estructuradas”*; *“...evalúan el contenido que se fue desarrollando a lo largo del tramo, si estudias para las clases no es muy complicado preparar los exámenes”*; *“...me gusta que las evaluaciones dejen día de por medio para mejorar los temas siguientes”*

La estructura de las evaluaciones y las tareas solicitadas dependieron de los instrumentos utilizados; en los exámenes integradores de ejes (ALAF y SISAM) fueron preguntas abiertas de respuesta corta, y en el examen integrador de tramo, preguntas de opción múltiple; aunque ambos formatos tuvieron en común, enunciados previos en cada pregunta, para darle un contexto y funcionar como disparadores. Estos enunciados, en varias preguntas, eran casos clínicos o problemas de salud, redactados en forma muy breves, que los estudiantes percibieron como muy valiosos para el aprendizaje: *“que hayan aplicado casos fue lo que más me gustó ya que tuve una visión de lo que es la profesión”*; *“me pareció muy bueno que nos presenten un caso y nosotros debamos razonar y resolverlo con lo aprendido”*; *“...las evaluaciones son muy buenas ya que presentaron casos clínicos que deberíamos saber responder para luego actuar en el día que ya seamos médicos”*.

Sin embargo, hubo algunos cuestionamientos a la redacción y/o calidad de las preguntas y a las opciones de respuesta: *“algunas consignas son poco claras y ambiguas”*; *“deben ser más claras, porque cuando no entendemos que es exactamente lo que nos piden... capaz sabíamos la respuesta, pero no entendimos la consigna”*; *“los interejos evaluaron la memoria del alumno, ya que varias consignas solicitaban valores o datos puntuales de los materiales de estudio, antes que la comprensión de un tema”*; *“El múltiple choice es más objetivo y seguramente más fácil de corregir, pero las respuestas distractoras a veces pueden confundir y más cuando uno está nervioso rindiendo”*; *“...muchas de las preguntas eran de carácter eliminatorio y no nivelatorio como se supone que debería ser”*

Un aspecto particular de MHS es la utilización de evaluaciones grupales, como la del Trabajo en terreno y la de Contenidos transversales, que se aplican en forma combinada con instrumentos de evaluación individuales (ALAF, SISAM, EIT).

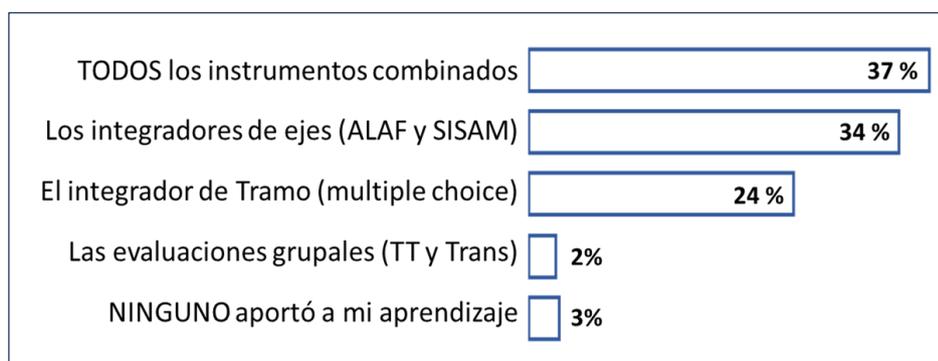
Estas evaluaciones grupales, donde las consignas de la prueba se discuten y se resuelven en grupos, ha surgido como un elemento controversial, pues para algunos estudiantes no es un método apropiado: *“...las personas del grupo nunca tienen la misma predisposición o compromiso y el trabajo siempre recae en unos pocos”*; *“...no me parece bien la evaluación grupal dado que es diferente el nivel de conocimiento entre los integrantes”*; *“En mi grupo, el TT presencial lo realizamos 3 integrantes y el resto no aportó nada”*; *“Las evaluaciones grupales no miden el desempeño individual”*; *“...no me parece justo que las evaluaciones sean grupales, ya que muchos aprueban a costa del esfuerzo de otros”*.

Sin embargo, otros se muestran complacidos: *“Me gustan las evaluaciones grupales ya que el hecho de tener que trabajar en conjunto con otras personas con diferentes perspectivas, a fin de superar un obstáculo es hasta divertido, ... me siento a gusto”*; *“...la posibilidad de compartir opiniones diferentes y respetarnos me pareció lo más valioso del trabajo grupal”*; *“me gustan..., pero todo depende de la capacidad para ponernos de acuerdo en plasmar una idea ya que todos tenemos varias maneras de abordar una misma consigna”*; *“...fue una experiencia muy interesante trabajar juntos por y para un objetivo común”*.

Sobre el sistema de evaluación multimétodo hubo también opiniones favorables: *“Me encuentro muy satisfecha con los métodos de evaluación, ya que siento que se evalúa más que nada el desempeño de cada persona de una manera más detallada y específica”*; *“...en relación al año pasado las evaluaciones están mejor organizadas, estructuradas, y las rubricas fueron más claras”*; *“Las evaluaciones me parecieron muy prácticas, ...no fue lo que esperaba, creí que iban a ser más rigurosas. La metodología me pareció muy favorable”*; *“En mi opinión, las evaluaciones de MHS fueron muy abarcativas lo cual me resultó positivo ya que integran la mayoría de los contenidos que damos en clase y en el CV”*. En cambio, otras voces expresaron su disconformidad: *“...el sistema de corrección tiene muchas imperfecciones que perjudican a los alumnos capaces de aprobar la materia”*; *“No evalúan la capacidad que tiene cada uno de nosotros. Son exámenes eliminatorios”*; *“... funciona como un gran examen de ingreso encubierto que solo pretende limitar el ingreso...”*.

Finalmente, cuando se les pregunto cuál de los instrumentos de evaluación había aportado más a su aprendizaje de la materia, las respuestas se resumen en el Gráfico 4.1.

**Gráfico 4.1:** Instrumento de evaluación que más aportó al aprendizaje de MHS  
(ranking de las respuestas de los estudiantes)



#### 4.3.1.2. Caso MED-III

En el programa de la asignatura (disponible en el aula virtual), están especificados los objetivos de aprendizaje explorados y los criterios generales de evaluación de los diferentes instrumentos que integran el sistema de evaluación longitudinal multimétodo. Dichos objetivos y criterios, también están explicitados en los formularios de mini-CEX (que conocen los estudiantes); en la Guía para elaborar el Trabajo Final (Historia Clínica y evidencias); en la introducción al Cafecito Infectológico; y en la Guía para el Portfolio. Para los exámenes “multiple choice” se informó solamente los contenidos que abordaba cada uno. Además, en el aula virtual de la asignatura, hay una presentación que explica las características del sistema y de los instrumentos de evaluación aplicados, y los criterios de calificación y promoción. Estas diversas maneras de llevar la información a los cursantes, puede explicar el resultado del cuestionario online, donde la mayoría de los estudiantes (74%) manifestó tener un amplio conocimiento previo, de los objetivos de aprendizaje explorados y de los criterios de evaluación aplicados.

La agenda con las fechas y los horarios de las evaluaciones estuvo disponible en el aula virtual desde la primera semana de cursado. Las instancias de evaluación se planificaron en forma secuencial a lo largo de todo el cuatrimestre, con intervalos de dos o más semanas entre cada examen, y sin interrumpir las otras actividades del cursado.

Este escenario de evaluaciones y actividades teórico-prácticas concomitantes, generó en algunos estudiantes una limitación del tiempo disponible para estudiar: “...la forma de evaluar es correcta, pero hay que tener en cuenta el tiempo que toma al alumno cumplir con todas las actividades solicitadas y esto quita horas que se dedicarían al estudio”; “...nos llenan de cosas para hacer, y por más que uno quiera y se esfuerce por estudiar bien

*cada tema. NO LE DA EL TIEMPO!"; "...el tiempo del cursado es muy corto y con tantas evaluaciones no nos permite estudiar los temas dados en las aulas en profundidad"; "...estar pendiente todo el tiempo de tantos exámenes, perdemos tiempo que podríamos utilizar para estudiar y comprender mejor los temas básicos de la carrera".*

Para otros, la dificultad estuvo en las exigencias del cursado: *"...no digo que esté mal, al contrario, pienso que está bien la exigencia, pero el stress que a veces causa la sobrecarga de tareas... te hace querer volver con mamá y papá"; "El estrés que estaba manejando fue impresionante, hasta que me superó... y deje de leer, ... yo decidí hacerlo así para no volver a ese grado de estrés"; "El desgaste mental, físico y social que tiene un alumno en esta materia no se compara con ninguna otra".*

La estructura de las evaluaciones y las tareas solicitadas dependieron de los instrumentos utilizados.

Para los exámenes parciales se utilizaron preguntas de opción múltiple que tenían como enunciado una viñeta clínica con los datos necesarios para elegir la mejor respuesta. Este formato, fue muy bien valorado por algunos estudiantes: *"...me encantó el desafío de resolver casos clínicos"; "...la formulación de las preguntas fueron muy buenas"; "...para mí fue la mejor evaluación"; "...aprendí mucho más estudiando para los parciales que para las clases".* Aunque para otros, fue objeto de varias críticas: *"...los enunciados no son claros, nos enseñan conceptos en clase que después los toman como incorrectos, y no es solo percepción mía sino consultando a varios doctores han llegado a la misma conclusión..."; "... deberían dar más tiempo para resolver las consignas, ya que, a mi parecer el objetivo del parcial es evaluar el razonamiento y el reconocimiento de patrones, y dado que recién estamos empezando a implementar dichas aptitudes, creo que no es importante resolver un caso en 1 minuto, sino poder resolverlo a fin de cuentas"; "...el parcial me pareció una injusticia con el tema del tiempo, no puede ser que nos digan 1 min por pregunta, y la grilla??? que onda no somos maquinas archimega rápidas... cuando hay consignas que llevan más de un minuto solo leerlas".*

Los miniCEX que se aplicaron en los trabajos prácticos con pacientes, fueron para algunos estudiantes instrumentos muy valiosos: *"A esta altura de la carrera, para mí, los mini-cex son los que mejor evalúan el aprendizaje y nos exige mucho más que cualquiera de las otras"; "...son los que más aportan en vincular conocimientos a la práctica clínica"; "...me permitió mejorar mucho mi habilidad para la práctica del examen físico"; "...con los minicex... siento que aprendí mucho más incluso que en otros años"; "...fueron los más útiles para mí".*

No obstante, algunos cuestionaron su aplicación: *"Mi única crítica para el Mini-cex, no es el método en sí que me parece genial, sino lo sesgado que está la aplicación del mismo si se compara un docente y otro"; "siento que a veces algunas evaluaciones pueden caer en la subjetividad del evaluador"; "el mini cex me sirvió mucho, pero los resultados varían mucho entre los distintos JTP... lo ideal sería que todo se corrija por el mismo docente"*

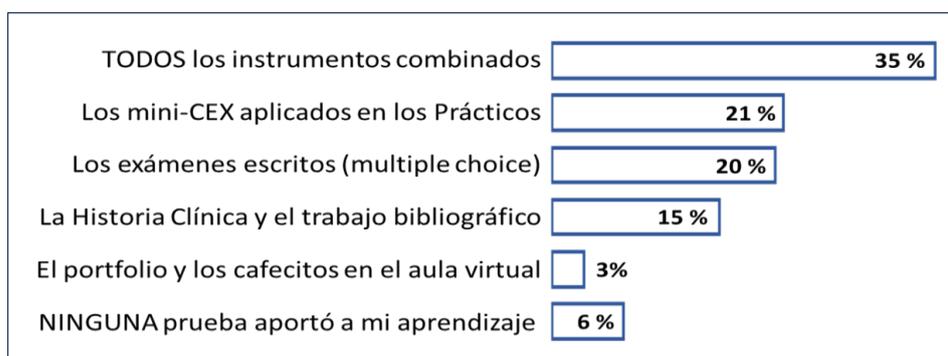
En relación a los otros instrumentos de evaluación no hubo tantas opiniones encontradas, y la observación más frecuente tuvo que ver con el tiempo de realización: *"...la historia clínica, la ficha y el cafecito, no me parece que sean malos instrumentos, pero el tiempo que te lleva hacerlas, te hace perder el objeto de estudio, y la mayoría de nosotros se pasó el equivalente a un mes para realizarlas"; "... la HC, el portfolio y el cafecito demandan mucho tiempo que no tenemos si queremos dedicar a estudiar para los niveles de conocimientos q exige esta materia"*

Sobre el sistema de Evaluación Longitudinal Multimétodo hubo opiniones favorables: *"Me parece un sistema muy acertado. Constantemente somos evaluados en muchos aspectos"; "El planteamiento de las evaluaciones está muy bien hecho, solamente hace falta más tiempo, estaría bueno que sea una materia anual..."; "Está bueno que hayan diferentes instrumentos y no solo una evaluación que nos defina"; "Es una forma diferente y nueva y como todo lo nuevo cuesta, me gusta, creo que nunca leí tanto para una materia como para esta (bueno para anatomía si jajajaja) pero de clínica, no"*

Otras voces plantearon críticas y hasta propuestas de cambio: *"...al final la forma de evaluarnos es bastante subjetiva y sujeta al ojo de quien corrige, y no me parece q esta sumatoria de puntos califique realmente al que estudio del que no"; "No me gusta en absoluto el sistema de evaluación de esta materia. No estamos acostumbrados a este sistema que nos imponen de golpe luego de 5 años de otra cosa"; "Si de verdad quieren saber cuanto sabe un alumno, tómenle un oral, denle tiempo para responder y ayúdenlo a rendir (no dándole la respuesta, sino incentivando a que razone y busque la respuesta en lo estudiado, generen un verdadero dialogo de confianza)".*

Finalmente, cuando se les pregunto cuál de los instrumentos de evaluación había aportado más a su aprendizaje de la materia, las respuestas se resumen en el Gráfico 4.2.

**Gráfico 4.2:** Instrumento de evaluación que más aportó al aprendizaje de MED-III (ranking de las respuestas de los estudiantes)



Fuente: propia

### 4.3.2. Efectos de aprendizaje

Explora la naturaleza de las actividades de procesamiento cognitivo y de regulación metacognitiva en cada evaluación (cantidad y distribución del esfuerzo, elección de recursos y contenidos; estrategias de monitoreo); y mecanismos de las devoluciones.

#### 4.3.2.1. Caso MHS

Analizando las actividades de procesamiento cognitivo que realizaron los estudiantes en cada uno de los instrumentos de evaluación que completaron al finalizar el segundo tramo de cursado, se observó una gran diversidad de acciones, operaciones y procesos mentales que estuvieron relacionados con las características y formatos de las diferentes pruebas.

Los **exámenes integradores de ejes**, por su formato con preguntas de respuesta abierta referidas a problemas, casos, tablas, imágenes y gráficos, tenían un alcance cognitivo que abarcaba cinco niveles de la Taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar. Por lo tanto, en estas pruebas los estudiantes realizaron operaciones mentales como: identificar, ordenar, interpretar, comparar, esquematizar, seleccionar, combinar, graficar, integrar, valorar, y justificar.

En el **examen integrador de tramo**, diseñado con preguntas de opción múltiple vinculadas a un enunciado (problema, caso clínico, proceso biológico, fenómeno natural, relato, etc) y con un alcance cognitivo que abarcaba los tres primeros niveles taxonómicos: Recordar, Comprender y Aplicar; los alumnos realizaron operaciones como: recordar, identificar, interpretar, reconocer, ordenar y jerarquizar.

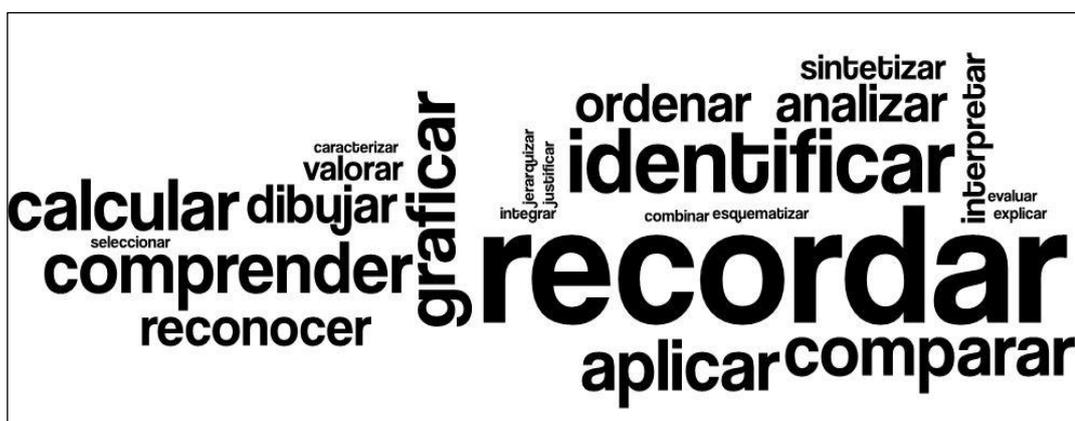
Para el **Trabajo en Terreno**, los alumnos en grupos y trabajando en su entorno familiar y barrial, recolectaron y registraron datos específicos con los que realizaron tareas como: reconocer hábitos alimentarios, caracterizar actividades físicas y deportivas,

identificar conductas de impacto ambiental y elaborar familigramas, que presentaron y discutieron en un informe final, aplicando herramientas estadísticas básicas. Por ende, para este ejercicio aplicaron cinco niveles cognitivos: Recordar, Comprender, Aplicar, Analizar y Evaluar.

En la **Evaluación de Contenidos Transversales** también aplicaron estos cinco niveles taxonómicos pues tuvieron que: reconocer paradigmas de investigación, recordar aspectos metodológicos; identificar la información de un texto y explicar su uso, aplicar las normas APA, identificar criterios de validez de un artículo, valorar mapas conceptuales, ordenar los pasos para buscar un artículo en internet y demostrar lecto-comprensión de inglés.

La frecuencia con la que los estudiantes utilizaron cada uno de estos procesos cognitivos en las evaluaciones, se reflejan en la nube de palabra del Grafico 4.3., donde se observa un predominio de acciones y procesos que corresponden al nivel cognitivo de Orden Inferior en la Taxonomía de Bloom (recordar, comprender y aplicar).

**Gráfico 4.3:** Actividades cognitivas de los estudiantes en las evaluaciones de MHS



Fuente: propia

El contenido de este gráfico, es coincidente con las opiniones vertidas por los estudiantes en el cuestionario aplicado, donde en la pregunta: ¿Qué tipo de proceso mental utilizaste con mayor frecuencia para rendir las pruebas de MHS?, respondieron “Solo reconocer o recordar conceptos teóricos” (48%); “Identificar o explicar mecanismos y procesos” (34%); y “Razonar y aplicar lo aprendido a casos o problemas” (18%).

Otro importante efecto de aprendizaje que debe ser considerado, es el “efecto catalítico” que puede tener una evaluación según el tipo de retroalimentación o devolución que ofrece a los estudiantes.

Al respecto y según el procedimiento aplicado en MHS, la devolución que recibieron los alumnos en todas las evaluaciones fue diferida, y consistió básicamente en informar las calificaciones que obtuvieron en cada prueba luego de completar todas las instancias de evaluación. A partir de la publicación de los resultados, se habilitó un mail para consultas y reclamos de los estudiantes, y por ese medio recibieron información individual y detallada sobre su desempeño en las pruebas.

Este procedimiento que utiliza MHS se ve reflejado en las opiniones de los estudiantes en el cuestionario, donde en la pregunta: ¿Qué tipo de devolución recibiste de los docentes después de rendir las pruebas?, respondieron “Solo la calificación obtenida” (85%); “Pude ver las respuestas correctas” (14%); “Me explicaron los aciertos y errores” (1%).

#### 4.3.2.2. Caso MED-III

Analizando los procesos mentales realizados por los estudiantes en cada uno de los instrumentos de evaluación que integraban el sistema de evaluación longitudinal multimétodo, se observó diferentes acciones, operaciones y procesos mentales relacionados con las características y formatos de las distintas pruebas.

En los **exámenes parciales**, que incluían preguntas de opción múltiple referidas a viñetas de casos clínicos y que tenían un alcance cognitivo de cinco niveles de la taxonomía de Bloom: Recordar, Comprender, Aplicar, Analizar y Evaluar; los estudiantes realizaron operaciones como: recordar, identificar, interpretar, reconocer, ordenar, comparar, tomar decisiones, justificar, planificar, sintetizar y valorar.

En los **mini-CEX**, los estudiantes fueron evaluados por observación, en sus habilidades para conducir la entrevista clínica y el examen físico de los pacientes, abarcando un alcance cognitivo de cinco niveles taxonómicos: Recordar, Comprender, Aplicar, Analizar y Evaluar, donde los estudiantes realizaron operaciones como: identificar problemas, interpretar e integrar manifestaciones, proponer y fundamentar hipótesis diagnósticas, plantear y justificar estudios diagnósticos y tratamientos.

En el **Trabajo Final**, los estudiantes utilizaron cinco niveles taxonómicos: Recordar, Comprender, Aplicar, Analizar y Evaluar, realizando tareas como: redactar una base de datos de Historia Clínica, identificar y jerarquizar problemas del paciente; elaborar una discusión diagnóstica, formular y fundamentar hipótesis diagnósticas, diseñar y justificar planes de estudios y de tratamiento, y buscar evidencias científicas relacionadas al paciente.

En el **Cafecito infectológico**, los estudiantes aplicaron también cinco niveles taxonómicos: Recordar, Comprender, Aplicar, Analizar y Evaluar, discutiendo casos clínicos de infectología donde tuvieron que identificar y jerarquizar problemas, interpretar imágenes, proponer y justificar diagnósticos diferenciales y tratamientos, y aportar evidencias científicas.

En el **Portfolio**, los estudiantes aplicaron habilidades metacognitivas, pues tuvieron que analizar y reflexionar sobre los procesos cognitivos que utilizaron para su aprendizaje en otras actividades de la asignatura.

La frecuencia con la que los estudiantes utilizaron cada uno de estos procesos cognitivos en las evaluaciones, se reflejan en la nube de palabra del Grafico 4.4., donde se observa un predominio de acciones y procesos que corresponden al nivel cognitivo de Orden Superior en la Taxonomía de Bloom (analizar, evaluar y crear).

El contenido de este gráfico, coincide con las opiniones aportadas por los estudiantes en el cuestionario aplicado, donde en la pregunta: ¿Qué tipo de proceso mental utilizaste con mayor frecuencia para rendir las pruebas de MED-III?, respondieron “Solo reconocer o recordar conceptos teóricos” (3%); “Identificar o explicar mecanismos y procesos” (13%); y “Razonar y aplicar lo aprendido a casos o problemas” (84%).

**Gráfico 4.4:** Actividades cognitivas de los estudiantes en las evaluaciones de MED-III



Fuente: propia

Respecto a los mecanismos de retroalimentación utilizados en las evaluaciones de MED-III, el sistema de evaluación longitudinal multimétodo, con propósito formativo, ofreció a los estudiantes devoluciones en todas las evaluaciones, con formatos diferentes según el instrumento aplicado. En los exámenes parciales (multiple choice), se presentaron y discutieron en clases plenas, las preguntas en las que mostraron mayor grado de dificultad;

en los mini CEX se efectuaron devoluciones constructivas inmediatas a cada estudiante sobre su desempeño clínico; para el Trabajo Final se realizaron tutorías (presenciales y online) para asesorar y facilitar la producción de las historias clínicas y el trabajo bibliográfico; en el Cafecito infectológico se aplicaron también tutorías online; y en el Portfolio se realizaron devoluciones constructivas individuales después de su evaluación.

Estos mecanismos de retroalimentación quedaron plasmados en las opiniones aportadas por los estudiantes en el cuestionario, donde en la pregunta: ¿Qué tipo de devolución recibiste de los docentes después de rendir las pruebas?, respondieron “Solo la calificación obtenida” (5%); “Pude ver las respuestas correctas” (17%); “Me explicaron los aciertos y errores” (78%)

#### 4.4. Tercer Eje de Análisis: Evaluación como aprendizaje

##### 4.4.1. Antes de la evaluación

Abarca los comportamientos de estudio (hábitos, habilidades y actitudes) que los estudiantes aplicaron en su preparación para las evaluaciones, incluyendo gestión del tiempo, establecimiento de metas, selección de qué, cómo, y dónde estudiar; y factores que motivaron y desanimaron el estudio.

##### 4.4.1.1. Caso MHS

Indagando posibles cambios adoptados por los estudiantes en sus comportamientos de estudio, para afrontar las evaluaciones de MHS, los comentarios fueron diversos.

Algunos modificaron sus técnicas de estudio: *“era todo muy nuevo para mi...tenía que leer en voz alta varias veces los apuntes de la cátedra para poder comprender...”*; *“...demasiados temas...a medida que iba leyendo iba escribiendo lo más importante y así me quedaba un resumen que me servía para repasar en los exámenes”*; *“para entender esta materia tuve que hacer muchos esquemas y dibujos...”*; *“...analizaba y trataba de entender lo que estaba leyendo porque me cuesta mucho memorizar”*; *“... con mis compañeras nos íbamos haciendo preguntas... y discutíamos las respuestas...”*

Otros se refirieron a los recursos que tuvieron disponibles: *“...las guías que teníamos en el campus nos orientaron bastante, pero muchos temas no estaban en los materiales... y después te los preguntaban...”*; *“...los ejemplos que pusieron de ejercicios resueltos fueron muy útiles...”*; *“...no pudimos leer tantos libros... usamos mucho los resúmenes de temas que hicieron los docentes”*; *“...no teníamos libro de referencia...tuve que buscar muchas cosas en internet y eso lleva tiempo”*; *“...nos dijeron que todo lo necesario estaba en el aula virtual... no fue así”*

Otros plantearon sus estrategias para gestionar el tiempo: *"... estudiamos para todas las clases... para no atrasarnos, porque en las semanas de examen había muy poco tiempo"; "...fue una locura, nunca estudié tanto en mi vida, ...no tuve ni un minuto libre"; "...con mi compañera estuvimos a full... se terminaron las salidas..."*;

Finalmente, otros destacaron sus experiencias de estudio en las Academias<sup>36</sup>: *"...estudie mejor con mis compañeros en la academia..., nos daban preguntas de años anteriores y ejercicios para responder..."; "me gustó estudiar en la academia porque nos enseñaban a preparar los exámenes y como rendir..."; "...ver las respuestas de exámenes de otros años nos ayudó a comprender mejor las preguntas..."; "en casa no me podía concentrar, siempre me estaban preguntando sobre el cursado..., estudiaba mucho mejor con mis compañeros en la academia"*.

En relación a los factores que motivaron o desanimaron el estudio en la fase de preparación para las evaluaciones, algunos estudiantes expresaron que las mismas exigencias del sistema de evaluación fueron una motivación para el estudio previo: *"...me parecieron bastante complicadas, porque requerían un estudio minucioso y profundo de todos los materiales... me gusta leer todo, pero no siempre hago una lectura significativa del todo, a veces para las clases dejo de enfocarme en algunas cosas que considero no son importantes de acuerdo a los objetivos propuestos"; "El sistema... es muy exigente y estoy de acuerdo con eso porque me ayuda a estudiar a fondo los temas para ir a las evaluaciones segura de mí misma"*;

Otros identificaron factores que desanimaron el estudio: *"Las evaluaciones grupales no te motivan para nada, porque el resultado depende del compromiso que tenga el grupo para estudiar durante el cursado"; "Las evaluaciones no siempre se enfocan en lo que uno considera importante, es decir, en algunas ocasiones hacen hincapié en temas que algunos consideramos de poca importancia al momento de formarnos. Tal vez, eso sea algo que nos juega en contra al momento de estudiar y prepararnos para rendir los exámenes, ya que solemos pasar por alto algunas de esas cosas"; "...para que voy a romperme estudiando si en los exámenes todo es cuestión de suerte.."*

---

<sup>36</sup> Las "Academias" son institutos privados, sitios en Corrientes y Resistencia, que ofrecen a los alumnos ingresantes, preparación pre-universitaria y en algunas materias básicas de varias carreras de grado.

#### 4.4.1.2. Caso MED-III

En relación a los posibles cambios adoptados por los estudiantes en sus comportamientos de estudio, para afrontar las evaluaciones de MED-III, algunos modificaron sus técnicas de estudio: *"...me costó mucho aprender el razonamiento clínico,... tuve que buscar artículos y ver videos para poder hacerlo,... es lo que más se necesita para las evaluaciones..."; "Hacer la H. Clínica fue la mejor experiencia porque me hizo invertir el orden del estudio, primero tomé los datos del paciente y después fui a buscar bibliografía para explicarlos..."; "...hacer esquemas y mapas conceptuales me ayudó mucho a comprender..."; "...con mi compañero de estudio simulábamos un ateneo, uno presentaba el caso y el otro hacía las preguntas..."; "Con mi compañera conseguimos preguntas de exámenes anteriores y así pudimos aprender muchos temas de la materia porque todos eran casos clínicos"; "...lo que más me costó fue identificar los temas importantes... son los que más toman en los exámenes"; "Nunca había hecho una historia clínica completa, tuve que leer varias de otros años para poder hacerla..."; "...engancharme en las discusiones del cafecito me sirvió mucho,...te obliga a buscar evidencias y a redactar con claridad..."; "...fuimos haciendo resúmenes de los temas más importantes y eso nos facilitó mucho la preparación para los parciales"; "Haber cursado basada (Medicina Basada en la evidencia) me sirvió mucho para buscar y seleccionar artículos para fundamentar mis comentarios...."*

Otros manifestaron inconvenientes con la disponibilidad y gestión del tiempo: *"...no estamos acostumbrados a tantos trabajos y actividades...y no supimos, creo yo, organizarnos con los tiempos y encontrar buenos momentos para estudiar"; "Haber puesto las evaluaciones sin suspender el cursado de esa semana nos alteró muchísimo los tiempos de estudio"; "...el cursado me limitó mucho el tiempo de estudio, por lo que no pude estudiar y comprender los temas como hubiera querido, no pude dedicarle el tiempo suficiente ya que tuve que hacer otros tipos de actividades..."; "Las distintas formas de evaluación y cursado, dispersaron mi objeto de estudio"*

También hubo comentarios resaltando el valor que tuvieron las aulas inversas para su aprendizaje y para afrontar las evaluaciones de la asignatura: *"...para mí la herramienta que más me aportó e inspiró en esta materia fueron las aulas invertidas, ... es un método espectacular para aprender, se aprende haciendo las aulas y escuchando a los demás compañeros, sobre todo cuando el docente que evalúa aporta datos desde su experiencia o nos recomienda tratamientos, estudios..."; "Creo que las aulas invertidas constituyen el*

*recurso evaluativo más valioso que tiene la cátedra y el que más beneficios pedagógicos supone para los alumnos...”; “...lo que más aportó a mi construcción de conocimientos en el cursado son las aulas invertidas me parece muy interesante poder exponer lo que sabemos sobre un tema y hasta dónde llega la calidad de la información que tenemos...”; “Las aulas invertidas fueron las que más contribuyeron con mi aprendizaje... y me sirvieron mucho para prepararme para los exámenes”*

Respecto a los factores que motivaron o desanimaron el estudio en la fase de preparación para las evaluaciones, algunos estudiantes destacaron el efecto motivador que tuvo el mecanismo de “promoción meritoria”<sup>37</sup> utilizado por la asignatura: *“Creo que lo que más te motiva a estudiar para los exámenes es la posibilidad que nos da la cátedra de aprobarla sin rendir el final cuando tenés un buen puntaje”; “... todos queríamos llegar a ser un estudiante destacado y eso nos motivó a estudiar mucho para cada examen...”*

Por otro lado, hubo manifestaciones de desánimo por el estrés que generó el cursado y las evaluaciones: *“El estrés que estaba manejando fue impresionante, hasta que me superó..., fui sin leer temas por elección, no por falta de tiempo, yo decidí hacerlo así para no volver a ese grado de estrés... lo único que puedo afirmar es que no sabemos manejar las situaciones estresantes y eso resta mucho”; “Tantos exámenes me ponían muy nervioso... y cuando llegaba el momento, mi cabeza se ponía en blanco y no podía repetir lo que estudié”.*

Otros expresaron su desánimo por los instrumentos de evaluación utilizados: *“Desde el ingreso me sucede esto, la verdad siento que me iría mucho mejor o creo yo, que aportaría más en mí rendir exámenes escritos, creo que eso motiva mucho más a un alumno a estudiar y a entender lo que estudiar”; “...soy una persona que estudie muchísimo o poco siempre obtiene los mismos resultados en exámenes múltiple choice, por eso es que los choices no me motivan tanto”.*

---

<sup>37</sup> La “promoción meritoria” es un reconocimiento que reciben los estudiantes destacados, y consiste en otorgarles su promoción sin rendir el examen final de la materia. Para ser un estudiante destacado hay que cumplir los siguientes criterios: Acreditar un 90% o más de asistencia a todas las actividades del cursado; haber rendido todas las evaluaciones; y lograr un puntaje total de 800 puntos o más (equivalente a nota 8).

#### 4.4.2. Durante la evaluación

Abarca las experiencias de los estudiantes en las evaluaciones, los beneficios obtenidos en su formación y aprendizaje, y las oportunidades que tuvieron de mostrar su progreso y reflexionar sobre lo aprendido; así como los factores del sistema que facilitaron u obstaculizaron dichas oportunidades.

##### 4.4.2.1. Caso MHS

Indagando en las experiencias de los estudiantes, las oportunidades y beneficios que tuvieron en las evaluaciones para mejorar su formación, se observó que una alta proporción de los mismos manifestó efectos positivos sobre sus aprendizajes. Tabla 4.7.

**Tabla 4.7:** Opiniones de los alumnos sobre las oportunidades y beneficios que tuvieron en las evaluaciones de MHS

Proposiciones	De acuerdo y muy de acuerdo
Para los exámenes de MHS tuve que estudiar varios materiales que no había leído para las clases.	143 44%
Preparando y rindiendo los exámenes siento que he aprendido más que cursando la materia.	202 63%
Las evaluaciones me permitieron comprender conceptos que no entendí durante el cursado.	158 49%
Las evaluaciones NO me motivaron a estudiar más de lo que estudio habitualmente.	27 8%
Gracias a las evaluaciones pude relacionar y articular mejor los diferentes contenidos de MHS.	203 63%
En los exámenes pude aplicar mejor los conocimientos de la materia a situaciones prácticas.	221 69%
A través de las evaluaciones NO he aprendido más de lo que aprendo en las clases.	47 15%
A través de las evaluaciones pude identificar mejor lo que todavía me falta aprender.	272 84%

Fuente: propia

De la tabla se desprende que la mayoría de los estudiantes concuerda en que a través de las evaluaciones enriquecieron su aprendizaje y tuvieron las oportunidades para integrar y aplicar los diferentes contenidos de la materia, y principalmente con la posibilidad que les dio el sistema de evaluación de identificar sus necesidades de aprendizaje.

Indagando los factores del sistema de evaluación que facilitaron u obstaculizaron el buen rendimiento; para algunos alumnos la diversidad y calidad de los instrumentos de

evaluación facilitó un buen rendimiento: *“Creo que lo mejor de las evaluaciones es la relación que tenemos que realizar durante el examen, capaz cuando los estudiamos por separado no tiene todo el sentido pero cuando lo podemos unir a otra área resulta más fácil de comprender”*; *“...sentí que al dar exámenes tan distintos entre un horario y otro, el aprobar se vuelve una cuestión de suerte, pero puedo decir que me sirvieron de ayuda para estudiar y afianzar contenidos”*; *“...ya que los integradores de ejes fueron individuales y no grupales como el año pasado, mi desempeño académico y calificación final no se vieron afectadas por errores o falta de estudio ajenas”*; *“son mucho más prácticas que las pruebas comunes, no se pueden estudiar de memoria, te hacen pensar y relacionar conceptos para aplicarlos a la práctica, y si estudiaste así los resultados van a ser buenos”*

Para otros, las fallas en el diseño de algunos instrumentos fue un factor que obstaculizó el buen rendimiento en las pruebas: *“...algunas consignas no son completamente claras y generan confusión, sobre todo las del parcial integrador”*; *“...cuando no entendemos que es exactamente lo que piden y preguntamos a algún profesor este no nos resuelve las dudas tampoco. Y capaz sabíamos la respuesta, pero no entendimos la consigna”*; *“Podrían esclarecer un poco las preguntas, algunas son difíciles de comprender lo que facilita que se responda mal”*; *“Las pruebas integradoras son muy buenas para relacionar los contenidos entre sí, sin embargo, en vez de preguntar sobre temas importantes y centrales, algunos preguntan detalles que se encuentran en la última frase de un apunte, y eso puede hacerte salir mal”*; *“Considero que los trabajos transversales y de terreno suelen ser confusos, y que al momento de realizarlos no siempre se está en lo correcto, y por ende no se logra una buena calificación”*.

#### **4.4.2.2. Caso MED-III**

Analizando las opiniones de los estudiantes, en cuanto a las oportunidades y beneficios que tuvieron en las evaluaciones para mejorar su formación, se observó que un alto número de ellos reconocen varios efectos positivos sobre sus aprendizajes. Tabla 4.8.

En la tabla se observa que, mediante las evaluaciones, un alto porcentaje de estudiantes mejoró su aprendizaje, pudo integrar sus conocimientos básicos con la clínica y aplicar lo aprendido a situaciones prácticas; y para la gran mayoría, las evaluaciones los llevaron a ampliar sus fuentes de información y les dio la oportunidad de identificar sus necesidades de aprendizaje.

**Tabla 4.8:** Opiniones de los alumnos sobre las oportunidades y beneficios que tuvieron en las evaluaciones de MED-III

Proposiciones	De acuerdo y muy de acuerdo	
Para los exámenes de MED-III tuve que estudiar varios materiales que no había leído para las clases.	128	92%
Preparando y rindiendo los exámenes siento que he aprendido más que cursando la materia.	66	47%
Las evaluaciones me permitieron comprender conceptos que no entendí durante el cursado.	52	37%
Las evaluaciones NO me motivaron a estudiar más de lo que estudio habitualmente.	25	18%
Gracias a las evaluaciones pude relacionar mejor mis conocimientos de ciencias básicas con la clínica	73	53%
En los exámenes pude aplicar mejor los conocimientos de la materia a situaciones de la práctica.	92	66%
A través de las evaluaciones NO he aprendido más de lo que aprendo en las clases.	28	20%
A través de las evaluaciones pude identificar mejor lo que todavía me falta aprender.	130	94%
Las múltiples evaluaciones a lo largo del cursado orientaron más mi aprendizaje que los exámenes aislados	79	57%

Fuente: propia

Con respecto a los factores del sistema de evaluación que facilitaron u obstaculizaron el buen rendimiento; algunos estudiantes manifestaron que la asociación de diferentes instrumentos de evaluación facilitó un buen rendimiento: *“...rendir diferentes exámenes me abrió la cabeza... ya no pienso solo en la respuesta correcta, aprendí por ejemplo a ver y encarar un caso clínico y como discernir entre lo importante y lo que no lo es, ...creo que por eso saque buenos puntajes”*; *“...las dos que me parecen más útiles son: los parciales que ayudan hacer un balance de lo que me cuesta y lo que no. Y la elaboración de la historia clínica como ejercicio para el futuro”*; *“Considero que el minicex es una buena herramienta que demuestra nuestro desempeño frente al paciente y el utilizar los casos clínicos para la evaluación parcial refleja nuestra capacidad de comprender los temas. Ya que no podríamos aplicar a un contexto lo aprendido a menos que lo hayamos comprendido”*

Otras opiniones revelaron fallas en la aplicación y corrección de algunos instrumentos que obstaculizaron el buen rendimiento: *“El tiempo que dieron para los*

*parciales me pareció muy poco, ya que necesitaba razonar más los encabezados, además para pasar a la grilla también te lleva mucho tiempo”; “...el mecanismo de evaluación utilizado en el choice es muy complejo y nunca nos hemos enfrenado a ese estilo de examen por lo cual no se puede pretender aprobar con buena nota, lo cual no indica que el alumno no haya estudiado o este mal preparado”; “...presentaron contenidos que debían ser respondidos en base a la experiencia, y eso fue muy difícil ya que no contamos con ella...”; “...el sistema de corrección no es del todo objetivo... y a veces confuso, me encontré con devoluciones que no coinciden o se contraponen en algunos casos con las notas expuestas”.*

#### **4.4.3. Después de la evaluación**

Abarca las percepciones de los estudiantes dentro del entorno de aprendizaje y evaluación, como perciben su capacidad intelectual para afrontar las evaluaciones y los riesgos percibidos (no regularizar o promocionar, bajo promedio, etc); y cuanto aportaron las evaluaciones al aprendizaje de la materia.

##### **4.4.3.1. Caso MHS**

Indagando las percepciones que tenían los estudiantes sobre su capacidad y sus habilidades intelectuales para afrontar las evaluaciones, las expresiones revelaron algunas limitaciones: *“...fue muy fuerte el cambio, en el colegio nunca rendimos este tipo de exámenes..., y la verdad es que me sentí muy inseguro...”; “...las actividades y los exámenes fueron demasiados para el tiempo que tuvimos... ninguno de nosotros viene preparado para un ritmo de estudio tan intenso”; “...comencé en la academia... tres meses antes del cursado y por eso pude aprobar, si lo hubiera intentado solo no habría llegado”; “...considerando que la gran mayoría no tenemos una formación básica en estos temas, no pueden pretender que en 4 meses pensemos y razonemos como médicos”; “No me gusta la matemática... no entiendo por qué tengo que hacer cálculos y gráficos para ingresar a la carrera de medicina... no es ingeniería”*

En relación a las posibles consecuencias de riesgo que los alumnos de MHS percibieron en el sistema de evaluación, casi todos coincidieron en que el riesgo más importante era “desaprobar los exámenes del segundo tramo y perder la regularidad en la asignatura” pues en la condición de “libres” no podían presentarse al Examen Final de la materia, que debían aprobar para ingresar como alumnos regulares a la Carrera de Medicina.

Respecto a los efectos que tuvo el sistema de evaluación de MHS sobre el aprendizaje de los estudiantes, hubo percepciones positivas: *“Creo que las evaluaciones de MHS me*

*permitieron comprender mucho más algunos temas que durante el cursado no los supe interpretar o poner en práctica”; “...hicieron que comenzara a prestar más atención a temas que muchas veces uno no le da tanta importancia”; “La metodología de los exámenes me pareció bastante buena... ya que me permitió desarrollar el pensamiento analítico”; “...me permitieron darme cuenta los temas que me faltan ampliar y mi capacidad para estudiar, los cambios en la metodología de estudio fueron efectivas”; “Me permitieron afianzar los conocimientos y además incentivaron el desarrollo óptimo del trabajo en equipo y la importancia de éste”; “si bien algunas me resultaron difíciles, fueron muy útiles para saber que conocimientos tengo que reforzar”.*

Otras voces fueron más pesimistas: *“...las evaluaciones no te enseñan nada, las clases deberían ser donde se aprende”; “...lo único que me dejaron los exámenes de MHS fue un fuerte dolor de cabeza y ...”*

#### **4.4.3.2. Caso MED-III**

En relación a las percepciones que los estudiantes tenían en cuanto a su capacidad y sus habilidades intelectuales para enfrentar las evaluaciones, algunas voces manifestaron cierta incertidumbre: *“Me gustan las evaluaciones, creo que son muy útiles, pero creo que algo falla en mí que siempre me termino desilusionando de mí misma por más que me esfuerce”; “Me apasiona la carrera que elegí y trato de aprender cosas nuevas todos los días, pero me cuesta muchísimo demostrar en los exámenes todo lo que se... tal vez si fueran orales sería mejor para mí”; “Realmente no sé qué me pasó y estoy MUY decepcionada de mi desempeño. Mi sacrificio fue en vano. Mas allá de eso, el sentimiento de miedo y culpa que me generó me ayudó a cambiar la modalidad de estudio”*

Indagando las posibles consecuencias de riesgo que percibieron los estudiantes en el sistema de evaluación de MED-III, algunos expresaron su preocupación por el riesgo de no lograr la “promoción meritosa” que ofrecía la asignatura: *“...no había un puntaje establecido para aprobar los exámenes, ...teníamos que sacar la mayor cantidad de puntos posibles para no rendir el (examen) final”; “...el ultimo parcial me estreso muchísimo, tenía que sacar 90 puntos para ser estudiante destacado, ...estudí como loco pero lamentablemente no lo logré”; “...en las últimas semanas de cursado mi ansiedad fue enorme, ...todavía no tenía el resultado de la historia clínica y de ese puntaje dependía rendir o no el examen final”; “Estar tan pendiente de un puntaje te motiva a estudiar mucho, pero también te carcome los nervios..., no sé cuál es el beneficio”*

A otros les preocupaba no alcanzar la regularidad: *“Yo sabía que promocionarla era difícil, ...solo me interesaba regularizarla, porque rendirla libre es mucho más difícil...”*; *“El principal riesgo es quedar libre, ...el final es muy complicado y si no lo aprobas no puedes hacer la PFO al año siguiente”*.

Sobre los aportes que hizo el sistema de evaluación de MED-III al aprendizaje de los estudiantes, algunos lo percibieron como beneficioso: *“Las formas de evaluación son todas muy productivas. La combinación del foro, realización de portfolio, trabajos prácticos, minicex y examen parcial ayudan a profundizar más en todos los temas y no solo lo que se debería estudiar para un solo examen aislado”*; *“Creo que todos los métodos sirven, porque te hacen dudar de todo, y a partir de ahí nace el hecho de buscar más información y así adquirir más conocimiento”*; *“...la mayoría de nosotros necesitamos algo como este cursado para valorar nuestro esfuerzo y valorar hasta donde podemos llegar, porque es el esfuerzo lo que le da valor a las cosas y las facilidades se lo quitan”*; *“Me pareció una forma novedosa de evaluar..., hay que estudiar muchísimo, ...pero estoy muy contento, aprendí y conocí enfermedades que jamás había visto”*; *“...sirvió para plantearme muchas cosas primero la forma de estudiar, pero también demostró mis falencias como estudiante, sin embargo, está haciendo que me esfuerce un poco más”*

Finalmente, hay estudiantes que perciben en el sistema de evaluación una intencionalidad encubierta: *“En fin creo que esta cátedra tiene como objetivo que nos vaya mal, tienen maldad, nos odian, no quieren que nos recibamos y disfrutan que nos vaya mal. Nunca me sentí tan basureada y pisoteada como en esta materia. El sistema de evaluación no es bueno. Todas las evaluaciones que se realizan obviamente son útiles y nos obligan a ir a estudiar. ¿Pero es necesaria tanta maldad? ... somos personas no somos máquinas y queremos estudiar pero no es necesario que nos hagan sentir tan mal”*

## **4.5. Diferencias entre los Casos de Estudio**

### **4.5.1. Diferencias en el escenario curricular de la carrera**

La asignatura MHS, con 400 horas de carga horaria, es la primera materia de la carrera de medicina que los alumnos inscriptos deben regularizar y aprobar para ingresar como estudiantes regulares a dicha carrera. Los cursantes, en su gran mayoría, provienen del nivel medio sin experiencia previa en estudios universitarios, y por consiguiente, son los alumnos “principiantes” de la carrera. En el otro extremo, se ubica MED-III, materia de 240 horas de carga horaria que se dicta en quinto año de la carrera antes de ingresar a la Practica

Obligatoria Final con la que culminan su formación de grado, por lo que los cursantes son los estudiantes “avanzados” de la carrera.

Ambas materias (MHS y MED-III) con similar cantidad de docente (37 y 35 respectivamente) recibieron en 2019 un número muy distinto de cursantes (1363 vs 151), que determinó una diferencia muy significativa en la relación docente/alumno (1/37 vs 1/4) y escenarios de aprendizaje y evaluación muy diferentes entre las dos asignaturas. En consecuencia, las actividades de MHS estuvieron condicionadas por el alto número de cursantes, con grandes comisiones de alumnos y amplia utilización de la virtualidad, mientras que en MED-III, se pudo trabajar en pequeños grupos de estudiantes y con mayor supervisión de los docentes.

#### **4.5.2. Diferencias en los entornos de aprendizaje y evaluación**

En MHS, las actividades áulicas de aprendizaje se desarrollaron en comisiones de 80 alumnos y a través de actividades virtuales y trabajos grupales tutorizados, siguiendo un cronograma estricto de actividades, sin posibilidad de flexibilización. En MED-III, el aprendizaje de los estudiantes se desarrolló en pequeños grupos, con apoyo tutorial y/o supervisión docente, abarcando aulas inversas, talleres, trabajos prácticos con pacientes, tareas en aula virtual, y otras actividades electivas que los estudiantes realizaron según sus preferencias.

Para las evaluaciones, MHS aplicó cinco instancias de evaluación al final de cada tramo (dos grupales y tres individuales), con propósito sumativo y de alto impacto, pues los alumnos debían aprobar al menos tres de las cinco para mantener su regularidad en la asignatura. En cambio, MED-III aplicó durante el cursado un sistema de evaluación longitudinal multi-método, con cinco instrumentos diferentes y doble propósito (formativo y sumativo). Como evaluación formativa, no hubo puntaje mínimo de aprobación para las pruebas y se ofreció retroalimentación luego de cada una, y para la sumativa se utilizó la sumatoria de los puntos obtenidos en cada instancia de evaluación formativa.

Con respecto a los instrumentos de evaluación utilizados, las diferencias entre las dos asignaturas están fundamentadas en los objetivos de aprendizaje evaluados. MHS como primera materia de la carrera, explora principalmente contenidos biomédicos básicos, metodológicos antropológicos y sociales, y por eso aplica cuestionarios con preguntas abiertas y estructuradas, informes de trabajos en terreno (comunitarios) y ejercicios en el aula virtual; mientras MED-III como asignatura del ciclo clínico, evalúa el desarrollo de los

conocimientos y de las competencias profesionales, para lo que aplica, cuestionarios con viñetas clínicas, observación del desempeño clínico (miniCEX), redacción de historias clínicas y búsqueda de evidencias, foros de discusión de casos clínicos y un portfolio reflexivo.

En cuanto a las herramientas de evaluación grupal que se utilizan en MHS marcan también otra diferencia con MED-III pues en esta asignatura no se aplican.

#### 4.5.3. Diferencias en los estudiantes evaluados

En relación a las **Metas de Logro** exploradas en los dos grupos de estudiantes, solo hubo diferencias significativas en las metas de Rendimiento-Aproximación entre los alumnos de MHS y MED-III con medias ( $\pm$ DE) de  $4.2 \pm 1.1$  vs  $2.9 \pm 1.3$  respectivamente, diferencia de medias: 1.3 con un IC (95%): 1.051 - 1.549 ( $p= 0.0000$ ).

Dado que estas metas de “aproximación al rendimiento” implican desempeñarse bien y obtener los mejores resultados con respecto a los otros compañeros del curso, la diferencia observada puede explicarse por la exigencia que tienen los alumnos de MHS de regularizar y aprobar esta asignatura para poder ingresar a la carrera de medicina.

En cuanto a los **Enfoques de Aprendizaje**, solo hubo diferencias significativas entre MHS y MED-III en las proporciones de estudiantes con “Enfoque profundo de intensidad alta”: 0.19 vs 0.35 respectivamente, diferencia de proporciones con IC (95%): -0.258 -0.068 ( $p= 0.0003$ ); y alumnos con “Enfoque superficial”: 0.12 vs 0.04 respectivamente, diferencia de proporciones con IC (95%): 0.030 - 0.134 ( $p= 0.0092$ ).

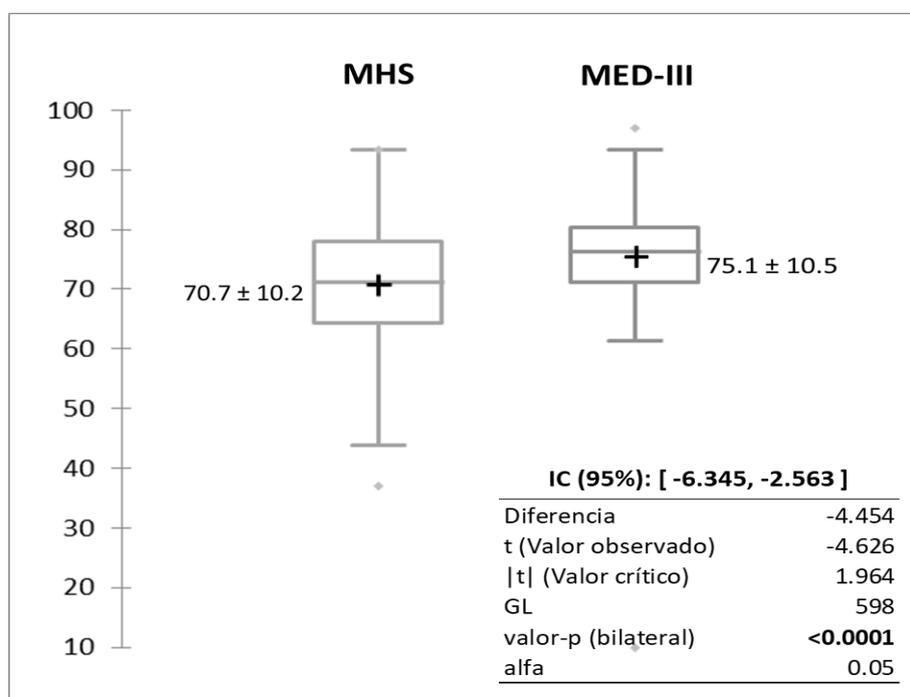
La mayor intensidad del enfoque profundo está vinculada a la alta motivación intrínseca y es esperable que la proporción de alumnos con este enfoque sea más elevada en MED-III pues son estudiantes avanzados que están llegando al final de la carrera y tienen un mayor compromiso con lo que están aprendiendo, pues encuentran al aprendizaje esencial para su futura vida profesional.

Por otro lado, es también comprensible que haya más alumnos de MHS con enfoque de aprendizaje superficial, pues este enfoque está relacionado a la motivación extrínseca, donde el estudiante tiene miedo al fracaso (a no poder ingresar a la carrera) y se preocupa por cumplir con todos las obligaciones y consignas que exige la asignatura para regularizar.

Con respecto a los **resultados** obtenidos por los estudiantes en las evaluaciones, se encontró diferencia significativa entre las medias ( $\pm$ DE) de la sumatoria de los puntajes logrados en las diferentes pruebas de MHS y MED-III (Grafico 4.5.), lo que indica un mejor

rendimiento en las evaluaciones, en el grupo de estudiantes avanzados comparados con los principiantes.

**Gráfico 4.5:** Diferencia entre los puntajes logrados en las pruebas de MHS y MED-III



No fue posible comparar los resultados obtenidos en cada instrumento, porque tenían diseños y tipologías muy diferentes en MHS y en MED-III, salvo los cuestionarios (multiple choice) que tenían similar formato y donde también se encontró diferencia entre las medias de los puntajes:  $65.3 \pm 11.0$  vs  $60.4 \pm 7.9$ ; diferencia de medias: 4.9 con IC (95%): 3.277 – 6.523 ( $p= 0.0000$ ), en este instrumento de evaluación los alumnos principiantes lograron mejores resultados que los avanzados, y eso tal vez podría explicarse por las diferencias en el alcance cognitivo de la preguntas, que fue más elevado en MED-III.

Otra diferencia fue en los rangos de puntaje observados, que fue mayor en MHS (entre 20 y 100 = 80) que en MED-III (entre 38 y 100 = 62), de lo que se puede inferir un mayor poder de “discriminación” global<sup>38</sup> de los instrumentos que utilizó MHS y que permitieron diferenciar mejor a los alumnos “bien preparados” para las pruebas de aquellos que no estaban suficientemente preparados para rendir.

<sup>38</sup> El índice o poder de discriminación global de una prueba se puede calcular dividiendo el rango de puntajes obtenidos (puntaje mayor – puntaje menor) dividida por el puntaje máximo posible (en este caso 100). Este sencillo calculo puede ser útil para para comparar el poder de discriminación de tests con distinto número de ítems o con diferentes formatos de preguntas. (Morales, 2012)

#### 4.5.4. Diferencias en los factores de la evaluación promotores de aprendizaje

En relación al grado de conocimiento que tenían los estudiantes de MHS y MED-III, sobre los objetivos de aprendizaje explorados y los criterios de evaluación y calificación aplicados en los exámenes, ambas asignaturas tenían explícita esa información en las guías didácticas e instructivos que ofrecían a los alumnos al inicio del cursado. Sin embargo, la mayoría de los estudiantes (68%) de MHS manifestó no tener conocimiento previo de dicha información, mientras que la mayoría de los estudiantes (74%) de MED-III señaló tener un amplio conocimiento de los objetivos de aprendizaje evaluados y de los criterios de evaluación, antes de afrontar las pruebas. Esta divergencia entre los alumnos principiantes y avanzados puede explicarse, por las diferentes estrategias de comunicación que utilizaron los docentes en MHS y MED-III, pues en la primera la información fue transmitida “a demanda” por la gran cantidad de cursantes; mientras que en MED-III, la información mencionada fue explicada en la clase inaugural y tenía un espacio exclusivo en el aula virtual de la materia que los estudiantes debían visitar.

Otra diferencia entre los casos de estudio se observó en la planificación de las instancias de evaluación. En MHS las cinco evaluaciones del segundo tramo fueron aplicadas en un período de dos semanas con suspensión de todas las actividades de cursado y con un intervalo de uno a dos días entre cada examen; mientras en MED-III, las instancias de evaluación se planificaron en forma secuencial a lo largo de todo el cuatrimestre, con intervalos de dos o más semanas entre cada examen, y sin interrumpir las otras actividades del cursado.

Estos dos escenarios de evaluación, muy diferentes entre sí, motivaron cuestionamientos por parte de los estudiantes, en MHS las críticas estuvieron enfocadas en la proximidad temporal entre las evaluaciones, ya que estaban muy cercanas en el tiempo y dificultaron la adecuada preparación para los exámenes; y a MED-III le cuestionaron la concomitancia de los exámenes con las actividades teórico-prácticas, pues para las evaluaciones no se suspendieron las actividades de cursado y eso generó, en algunos alumnos, una sobrecarga de tareas simultáneas y limitación del tiempo disponible para estudiar, debido a las exigencias propias del cursado.

Finalmente, en relación a los factores de la evaluación promotores de aprendizaje, una diferencia importante entre MHS y MED-III se observa en los formatos y diseños disímiles de los instrumentos de evaluación que aplica cada asignatura, que como se explicó más

arriba, está fundamentada en los objetivos de aprendizaje evaluados (muy diferentes entre los estudiantes principiantes y los avanzados).

#### 4.5.5. Diferencias en los efectos de aprendizaje de las evaluaciones

Comparando, en los instrumentos de evaluación aplicados, las actividades de procesamiento cognitivo que realizaron los estudiantes durante las evaluaciones, se observó que los alumnos de MHS habían realizado principalmente acciones y procesos que corresponden al nivel cognitivo de Orden Inferior en la Taxonomía de Bloom (recordar, comprender y aplicar); mientras los estudiantes de MED-III, operaciones y procesos mentales que corresponden al nivel cognitivo de Orden Superior (analizar, evaluar y crear) de dicha taxonomía.

Estas diferencias en el alcance cognitivo de las evaluaciones fue confirmada por los mismos alumnos en el cuestionario, cuando se les preguntó sobre qué tipo de proceso mental habían utilizado con mayor frecuencia para rendir las pruebas, los estudiantes de MHS y MED-III respondieron (respectivamente): “Solo reconocer o recordar conceptos teóricos”: 0.48 vs 0.03, diferencia de proporciones con IC (95%): 0.352 - 0.564 ( $p= 0.0000$ ); “Identificar o explicar mecanismos y procesos”: 0.34 vs 0.13, diferencia de proporciones con IC (95%): 0.083 - 0.329 ( $p= 0.0021$ ); y “Razonar y aplicar lo aprendido a casos o problemas”: 0.18 vs 0.84, diferencia de proporciones con IC (95%): -0.780 -0.547 ( $p= 0.0000$ ).

Otra diferencia importante entre los casos de estudio, fue la modalidad de retroalimentación o devolución ofrecida a los estudiantes después de las evaluaciones, que en MHS consistió esencialmente en informar las calificaciones que obtuvieron en cada examen, en forma diferida luego de completar todas las instancias de evaluación y habilitar un mail de consultas y reclamos para informar a cada estudiante sobre su desempeño en las pruebas. En cambio, MED-III, que en su sistema de evaluación longitudinal multimétodo tenía un propósito formativo, ofreció a los estudiantes devoluciones en todas las evaluaciones, con formatos diferentes según el instrumento aplicado.

Estas diferencias en las formas de retroalimentación también se reflejaron en el cuestionario, cuando se les preguntó sobre el tipo de devolución que recibieron de los docentes después de rendir las evaluaciones, los estudiantes de MHS y MED-III respondieron (respectivamente): “Solo la calificación obtenida”: 0.85 vs 0.05, diferencia de proporciones con IC (95%): 0.710 - 0.891 ( $p= 0.0000$ ); “Pude ver las respuestas correctas”: 0.14 vs 0.17, diferencia de proporciones con IC (95%): -0.144 a 0.085 ( $p= 0.7167$ ); y “Me explicaron los

aciertos y errores”: 0.01 vs 0.78, diferencia de proporciones con IC (95%): -0.876 a -0.666 (p= 0.0000).

#### 4.5.6. Comportamientos, experiencias y percepciones de los estudiantes

Tanto los estudiantes principiantes como los avanzados, manifestaron haber realizado algunos cambios en sus estrategias/técnicas de estudio y en la gestión del tiempo, para afrontar las evaluaciones de MHS y MED-III respectivamente, pero no hubo diferencias llamativas entre los casos de estudio en estos ajuste diversos y personales que hicieron los estudiantes, salvo la estrategia que adoptaron varios alumnos de MHS que destacaron sus experiencias de estudio en las “Academias” (Institutos privados de preparación pre-universitaria) por el apoyo y la orientación que les brindaron para preparar los exámenes.

En relación a las experiencias de los estudiantes en los exámenes de MHS y MED-III, la mayoría de las opiniones, en ambos grupos (Tabla 5.9.), se expresaron “de acuerdo” y “muy de acuerdo” con las oportunidades y beneficios que tuvieron en las evaluaciones para mejorar su formación.

**Tabla 4.9:** Opiniones de los alumnos sobre las oportunidades y beneficios que tuvieron en las evaluaciones. Respuestas "de acuerdo" y "muy de acuerdo"

Proposiciones	MHS (n:322)		valor p	MED-III (n: 139)	
	n	%		n	%
Para los exámenes tuve que estudiar varios materiales que no había leído para las clases.	143	0.44	0.0000	128	0.92
Preparando y rindiendo los exámenes siento que he aprendido más que cursando la materia.	202	0.63	0.0032	66	0.47
Las evaluaciones me permitieron comprender conceptos que no entendí durante el cursado.	158	0.49	0.0275	52	0.37
Gracias a las evaluaciones pude relacionar y articular mejor los diferentes contenidos.	203	0.63	0.0442	73	0.53
En los exámenes pude aplicar mejor los conocimientos a situaciones prácticas.	221	0.69	0.6835	92	0.66
A través de las evaluaciones pude identificar mejor lo que todavía me falta aprender.	272	0.84	0.0118	130	0.94

Comparando en la tabla las opiniones de los estudiantes principiantes y avanzados se observan diferencias significativas en casi todas las proposiciones.

Los alumnos de MHS (en mayor proporción que MED-III) acordaron que preparando y rindiendo los exámenes habían aprendido más que cursando la materia, habían comprendido conceptos que no entendieron durante el cursado y lograron relacionar y articular mejor los diferentes contenidos de la materia.

Por otro lado, una proporción mucho más alta de los estudiantes de MED-III concordaron en que las evaluaciones los llevó a buscar y ampliar sus fuentes de información más allá de lo trabajado en el cursado de la materia, y les dio la oportunidad de identificar sus necesidades de aprendizaje. Que son dos habilidades esenciales para el desarrollo continuo y autónomo del aprendizaje.

Otra diferencia entre los estudiantes principiantes y avanzados estuvo en los “riesgos percibidos”, que para los alumnos de MHS el riesgo más importante fue “desaprobar los exámenes del segundo tramo y perder la regularidad en la asignatura” pues en la condición de alumnos “libres” no podían presentarse al examen final de la materia, que debían aprobar para ingresar como alumnos regulares a la Carrera de Medicina. En cambio, para los estudiantes de MED-III el principal riesgo percibido fue no lograr la “promoción meritoria” que ofrecía la asignatura a los estudiantes destacados, exceptuándolos de rendir el examen final de la materia.

# CAPITULO 5

## DISCUSIÓN DE LOS RESULTADOS





## Capítulo 5: DISCUSIÓN DE LOS RESULTADOS

Aplicando la matriz analítica-metodológica presentada en el **Capítulo 3**, los resultados del estudio se discuten, siguiendo las preguntas derivadas en cada uno de los ejes y categorías de análisis que orientaron la investigación.

### 5.1. Primer Eje de Análisis: Evaluación del aprendizaje

#### 5.1.1. Entorno de aprendizaje y evaluación

Comprende las características del entorno de aprendizaje y de los sistemas de evaluación, sus fundamentos, propósitos, diseño, planificación y aplicación, en el contexto de cada materia, los instrumentos de evaluación utilizados, sus propiedades psicométricas y el alcance cognitivo de las tareas y consignas planteadas en las pruebas.

##### 5.1.1.1. ¿Como se diseñaron y planificaron los sistemas de evaluación multimétodo?

Para diseñar un sistema de evaluación en carreras y asignaturas de medicina que tienen su curriculum orientado en competencias, diferentes autores (Nolla-Domenjó, 2009; Moore et al, 2009; Schuwirth & Van der Vleuten, 2011b) recomiendan configurar el sistema de evaluación a lo largo de la carrera para que todo el cuadro de competencias pueda ser evaluado con instrumentos adecuados al nivel de formación de los estudiantes, aplicando criterios claramente establecidos y métodos combinados cuidadosamente elegidos.

Esta recomendación es necesariamente aplicable en carreras con curriculum por competencias y tiempo de formación variable, donde el aprendizaje de cada uno de los estudiantes varía en duración y objetivos, y requiere evaluar y juzgar el desarrollo de las competencias para decidir cuándo un alumno está preparado para pasar a la siguiente fase de su formación (Gruppen et al. 2018). Pero en carreras estructuradas (fragmentadas) en asignaturas con carga horaria fija, como la carrera de medicina de la UNNE, resulta muy complejo y dificultoso diseñar un sistema de evaluación integrado que abarque todo el trayecto formativo.

En consecuencia, los casos de estudio (MHS y MED-III) como asignaturas del mismo Plan (2000), diseñaron y planificaron sus sistemas de evaluación en forma integrada a los procesos de aprendizaje, articulando instrumentos diferentes para evaluar todo el espectro de las competencias abordadas en cada materia y orientadas en el perfil de graduación, y de esa manera superar también las limitaciones que tienen los instrumentos de evaluación aplicados en forma aislada.

### **5.1.1.2. ¿Por qué se eligieron esas herramientas de evaluación?**

Los dos casos de estudio, construyeron sus sistemas de evaluación multi-método con instrumentos diversos y específicos, para poder evaluar objetivos y resultados de aprendizaje diferentes (dominios declarativos, procedimentales y metacognitivos) y así obtener información más amplia y completa sobre el desarrollo de las competencias en el mismo escenario del proceso formativo que transitaban los estudiantes en cada asignatura.

Las razones enunciadas en la elección de herramientas, concuerdan con las recomendaciones para la evaluación de competencias elaboradas por Carreras Barnés et al (2009), que proponen: evaluar de forma integrada sus componentes (conocimientos, habilidades y actitudes); usar instrumentos congruentes con el diseño curricular, con la metodología docente y con el tipo de competencias a evaluar; y utilizar como referente criterios de evaluación claros y acordes al nivel de logro esperado.

En el caso de MED-III, la elección de las herramientas utilizadas en el sistema de “evaluación longitudinal multi-método”, se corresponden con varias de las propuestas formuladas por Epstein (2007) para planificar un sistema de evaluación de competencias en una asignatura del ciclo clínico, donde recomienda incluir, entre otras actividades, la observación directa del estudiante interactuando con pacientes en diferentes situaciones; pruebas de “multiple choice” para evaluar conocimientos; problemas y preguntas para evaluar el razonamiento clínico; y un trabajo de ensayo que demande búsquedas bibliográficas y síntesis de la literatura sobre ciencias básicas y aspectos clínicos de algunas enfermedades observadas por el estudiante. Dichas herramientas integran el sistema de evaluación de MED-III.

### **5.1.1.3. ¿Como se elaboraron y validaron las herramientas de evaluación?**

Tanto en MHS como en MED-III, las preguntas y consignas para los cuestionarios de examen fueron elaboradas por los mismos docentes que desarrollaron los contenidos del programa, siguiendo guías, instructivos y/o recomendaciones que los profesores prepararon para cada formato de instrumento; y posteriormente, corregidas y validadas, por los profesores responsables de cada evaluación antes de su aplicación. Los cuestionarios también incluyeron, preguntas utilizadas en cohortes anteriores que estaban en un banco de preguntas. Los instrumentos estandarizados (miniCEX, Historia Clínica y rúbricas para evaluar ensayos) habían sido elaborados y validados en años anteriores.

La validación (de constructo y contenido) de todas las herramientas fue realizada, en ambas asignaturas, por los profesores y algunos JTP (expertos en contenido); y luego

presentadas y discutidas con los docentes evaluadores para consensuar las pautas de aplicación. Este intercambio de opiniones entre los docentes es muy importante en el proceso de validación de los instrumentos de evaluación, pues como lo señala van der Vleuten (2016:887) *"la validez reside más en los usuarios de los instrumentos que en los instrumentos que se utilizan"*, lo que indica la necesidad de preparar cuidadosamente a los evaluadores y a los estudiantes para su papel en la evaluación.

#### **5.1.1.4. ¿Cómo funcionaron las herramientas de evaluación?**

Para valorar el funcionamiento de los instrumentos de evaluación luego de su aplicación, en ambas asignaturas, los profesores responsables de las evaluaciones analizaron los resultados y el desempeño de los alumnos en cada prueba; y para los exámenes escritos utilizaron como variables psicométricas, la Confiabilidad (Alfa de Crombach) y los Índices de Dificultad y de Discriminación de cada una de las preguntas o ítems.

En el análisis de los resultados de los estudiantes en las pruebas, se observó en los rangos de puntaje obtenidos, diferencias entre MHS y MED-III (80 vs 62 respectivamente), que permite inferir, según Morales (2021), un mayor poder de "discriminación" en los instrumentos de evaluación aplicados por MHS que posibilitaron diferenciar mejor a los alumnos "bien preparados" para las pruebas de aquellos que no estaban suficientemente preparados para rendir.

En los análisis psicométricos se observó: valores Alfa de Crombach  $> 0.70$  en todos los cuestionarios, lo que indica "aceptable" confiabilidad según los criterios propuestos por Santos (1999); los índices de Dificultad de las preguntas, en todas las pruebas, estuvieron en el rango de 0,45 a 0,75, que indica dificultad "intermedia" según los criterios de Haladyna et al (2002); y los índices de Discriminación variaron según el formato de la prueba, en las preguntas "multiple choice" fueron "bajos" ( $iD = 0,20 - 0,29$ ) y en las preguntas de respuesta corta fueron "aceptables" ( $iD = 0,30 - 0,39$ ), según la clasificación de Ebel (1982).

El índice de confiabilidad, indica el grado en el cual las puntuaciones de los estudiantes serían similares si fueran examinados de nuevo, y un coeficiente "aceptable" (como en estos casos) está indicando que cada estudiante obtuvo (probablemente) el puntaje merecido, y se podría inferir que si se aplicara un examen parecido (con otras preguntas del mismo estilo) los alumnos quedarían ordenados de manera semejante en sus puntajes. (Morales, 2012)

Las preguntas con índices de Dificultad en rango “intermedio”, predominantes en ambos casos, son para varios autores (Haladyna et al, 2002; Morales, 2012; Bonillo, 2013; Paniagua et al, 2016) las más adecuadas para los exámenes porque diferencian mejor a los alumnos que “saben” y “no saben”; pues cuando las preguntas son “muy fáciles” las responden todos y si son “muy difíciles” las responden muy pocos o por azar, por ende, ninguna de estas preguntas “muy” fáciles o difíciles son de utilidad para diferenciar a los estudiantes por su preparación.

Los índices de Discriminación fueron “aceptables” en las preguntas abiertas y “bajos” en las preguntas de opción múltiple. Esta diferencia puede explicarse por el formato de la pregunta, pues las abiertas requieren “elaborar” la respuesta, mientras que en las de opción múltiple hay que “elegir” una respuesta, y esto puede hacerse también por azar. En consecuencia, las abiertas permitieron discriminar (diferenciar) mejor que las últimas, a los estudiantes bien preparados (sabían más) de aquellos poco preparados (sabían menos), que es uno de los propósitos de la evaluación sumativa. Un índice “bajo” de discriminación también puede indicar que esa pregunta tiene un alcance cognitivo distinto que la mayoría de los ítems, por ejemplo, una pregunta de comprensión o aplicación puede tener un iD bajo si la mayoría de los ítems son de recordar (Morales, 2012).

Sin embargo, hay que tener presente que el poder de discriminación depende del grado de dificultad de la pregunta, específicamente de su varianza (Rodríguez, 2005; Bonillo, 2013) y las preguntas más discriminantes no son por lo general las más fáciles ni las más difíciles, sino las de dificultad intermedia. (Morales, 2012)

Estos índices, como lo explica Morales (2012:13), “describen cómo ha funcionado una pregunta en una situación dada; no hay que asociar necesariamente juicios de valor sobre la calidad de la pregunta al valor de estos índices (por eso decimos en primer lugar que estos índices describen qué ha sucedido; luego vendrá nuestra valoración)”.

### **5.1.2. Estudiantes evaluados**

Abarca las metas de logro que tienen los estudiantes evaluados, los enfoques y estrategias de aprendizaje que utilizaron en el contexto del sistema de evaluación aplicado, y los resultados obtenidos.

#### **5.1.2.1. ¿Cuáles eran las metas de logro de los estudiantes?**

Los investigadores educativos se refieren a las metas de logro como la razón por la cual los estudiantes se involucran en una tarea. Por ello, la teoría de las metas de logro es

uno de los marcos más importantes para analizar la motivación de los estudiantes y estudiar sus efectos en el aprendizaje y el rendimiento, pues el tipo de meta que persiguen los alumnos influye en sus comportamientos y resultados de aprendizaje. (Stegers-Jager et al, 2012; Kool et al, 2016)

En este trabajo, aplicando el cuestionario de Metas de Logro (A-AGQ-R) se pudo observar que tanto los alumnos principiantes como los estudiantes avanzados tenían una fuerte orientación por las metas de “aproximación al aprendizaje” (ambos grupos con medianas de 4.7 en un rango 1 a 5), lo que según Elliot & Murayama (2008) indica que los estudiantes están enfocados en desarrollar sus propias habilidades y avanzar en su aprendizaje, aplicando un procesamiento cognitivo profundo para llegar a dominar los materiales de estudio, comprender todos los contenidos de la asignatura y aprender tanto como sea posible.

Esta fuerte orientación por las metas de “aproximación al aprendizaje” de los alumnos principiantes puede tener al menos dos explicaciones, la primera está relacionada al brusco cambio de escenario educativo que experimentaron los alumnos al pasar del nivel medio a la universidad, al impacto producido por la “nueva experiencia” de ingresar a la carrera de medicina, que produce una mayor motivación intrínseca y gran interés en aprender contenidos “nuevos”; un fenómeno similar fue observado por Artino et al (2012) en los estudiantes de primer año de medicina. La otra explicación posible podría estar en un aumento de la “autoeficacia” de los alumnos luego de haber superado exitosamente las evaluaciones del primer tramo de MHS, pues según Kool et al (2016) un buen resultado en un examen decisivo, puede aumentar la autoeficacia de los estudiantes y esto conduce a una mayor orientación hacia las metas de “aproximación al aprendizaje” y de “aproximación al rendimiento”, que es otra característica observada en los alumnos principiantes.

En relación a la fuerte orientación que mostraron los estudiantes avanzados por las metas de “aproximación al aprendizaje”, es el resultado esperado, porque son estudiantes del último año de cursado antes de realizar la POF, que según Kool et al (2016) suelen tener autoeficacia alta por sus logros académicos y están dispuestos a esforzarse para desarrollar las habilidades y competencias que demanda su formación profesional.

Con respecto a las metas de “Aproximación al Rendimiento”, la orientación fue más intensa en los alumnos principiantes que en los estudiantes avanzados (medianas 5 y 3 respectivamente en un rango de 1 a 5), esta diferencia puede explicarse por la exigencia que tienen los alumnos de MHS de regularizar y aprobar esa asignatura para poder ingresar a la

carrera de medicina, y es interesante, porque de acuerdo a la revisión del modelo 2 x 2 realizada por Moller & Elliot (2006), las metas de aproximación al rendimiento parecen ser los predictores más fuertes y consistentes de las notas de un examen.

#### **5.1.2.2. ¿Qué enfoques de aprendizaje aplicaban?**

Existe consenso entre los investigadores que los estudiantes pueden abordar el aprendizaje con un enfoque "profundo" (luchando por el significado y la comprensión) o "superficial" (instrumental y reproductiva). Estos enfoques de aprendizaje, según el modelo de Bigg (1988), son una combinación dual de un motivo y una estrategia complementaria. Los motivos, son las razones o los objetivos que tienen los estudiantes para aprender. Las estrategias, son los métodos utilizados por los estudiantes para aprender el material.

Por consiguiente, los estudiantes con una profunda motivación por el aprendizaje buscarán satisfacer un interés intrínseco en el curso y es probable que empleen estrategias profundas (interrelacionar ideas, leer mucho) para lograrlo; mientras que aquellos con una motivación superficial para aprender, buscarán reproducir el material del curso de la manera requerida con mínima comprensión y empleando estrategias superficiales (reproducción de los conceptos y recuerdo textual). (Wilson & Fowler, 2005)

En este trabajo, aplicando el cuestionario revisado de Procesos de Estudio-Dos Factores (R-SPQ-2F), se observó tanto en los alumnos principiantes como en los estudiantes avanzados, un amplio predominio de enfoques de aprendizaje profundo (proporciones: 0.88 vs 0.94). Solo hubo diferencias significativas entre MHS y MED-III en las proporciones de estudiantes con "Enfoque profundo de intensidad alta": 0.19 vs 0.35 respectivamente ( $p=0.0003$ ); y en aquellos con "Enfoque superficial": 0.12 vs 0.04 respectivamente ( $p=0.0092$ ).

La mayor intensidad del enfoque profundo en los estudiantes avanzados, puede explicarse por el entorno clínico de aprendizaje y evaluación que les ofrece MED-III, pues según Al Kadri et al (2011), cuando los estudiantes comienzan a adquirir habilidades y experiencia para la atención de pacientes, aplican métodos de resolución de casos y problemas clínicos, cuentan con supervisión docente, son evaluados en escenarios clínicos y pueden reflexionar sobre sus prácticas, perciben factores importantes para el aprendizaje que pueden aumentar su motivación intrínseca y esto puede llevarlos a adoptar enfoques de aprendizaje profundo.

Además, la mayor intensidad del enfoque profundo en este grupo de alumnos puede estar relacionada con su "madurez académica", que para Lynam & Cachia (2018), incrementa

el compromiso de los estudiantes con la evaluación y esto se desarrolla a medida que el plan de estudios avanza en el ciclo clínico, donde pueden realizar tareas más prácticas y de aplicación clínica, en contraste con las pruebas de conocimiento científico y desarrollo de habilidades genéricas que caracterizan a los primeros años de la carrera. Esto coincide con los resultados de Cobb et al (2013), que observaron que a medida que los estudiantes avanzaban hacia la pasantía estudiantil, su preocupación por la atención al paciente se convirtió en un factor más prominente en su aprendizaje, mientras que al principio del curso estaban preparados para sacrificar su aprendizaje vocacionalmente motivado para reducir los niveles de estrés y aprobar exámenes.

Por otro lado, es también comprensible que haya más alumnos principiantes con enfoque de aprendizaje superficial, pues la percepción de una gran carga de trabajo (como tiene MHS) y el temor al fracaso académico (no poder ingresar a la carrera) son una fuerte motivación extrínseca que se asocia con un enfoque superficial (Wilson & Fowle, 2005). Otra explicación puede estar en las experiencias previas de evaluación que tuvieron los alumnos de MHS en el primer tramo y que superaron exitosamente, pues para Cilliers et al (2012), el grado de familiaridad que tienen los estudiantes con un método de evaluación puede influir en su enfoque de aprendizaje, y es probable que, si un alumno tuvo éxito en un examen anterior usando estrategias superficiales, tenga la percepción de que las estrategias de aprendizaje de nivel profundo no son necesarias para esa prueba, y en consecuencia sigue aplicando un enfoque superficial.

### **5.1.2.3. ¿Qué resultados obtuvieron en las evaluaciones?**

A partir de los resultados obtenidos por los estudiantes en las diferentes evaluaciones, se observó, tanto en los alumnos principiantes como en los avanzados, una importante diversidad de puntajes según el instrumento de evaluación aplicado. Esta observación es la esperada cuando se analiza cualquier sistema de evaluación de competencias multimétodo, pues como lo explica van der Vleuten et al (2017), independientemente de lo que se mida y del método con el que se evalúe, la competencia y el desempeño resultante están ligados al contexto, en consecuencia, el resultado obtenido en un elemento de desempeño (pregunta, situación, estación, escenario), no es predictivo del desempeño en otro elemento. Por eso es necesario muestrear muchos elementos para poder hacer una inferencia sobre la capacidad de alguien independientemente de la muestra utilizada en la evaluación, pues la competencia es específica y no genérica.

Analizando en forma general los resultados, se encontró una diferencia significativa entre las medias ( $\pm$ DE) de la sumatoria de los puntajes logrados en las diferentes pruebas de MHS y MED-III:  $70.7 \pm 10.2$  vs  $75.1 \pm 10.5$ ; diferencia de medias: -4.45 con IC (95%): -6.345, -2.563 ( $p < 0.0001$ ), lo que indica un mejor rendimiento en las evaluaciones, en el grupo de estudiantes avanzados comparados con los principiantes.

En los alumnos de MHS, los mayores puntajes corresponden a las dos evaluaciones grupales (trabajos en terreno y contenidos transversales), que integradas mostraron una media ( $\pm$ DE) de  $73.9 \pm 16.4$  puntos, y comparado con el promedio ( $63.6 \pm 18.2$ ) de las tres pruebas individuales (ALAF, SISAM y EIT) fue significativamente más alto.

Este mejor rendimiento que demostraron los alumnos principiantes en las evaluaciones en grupos, contrasta con las opiniones y percepciones que los mismos dieron en el cuestionario, donde muchos manifestaron su disconformidad y rechazo por los exámenes grupales y al mismo tiempo expresaron su satisfacción con las pruebas individuales. Una posible explicación para esta discrepancia puede estar en los argumentos que dieron algunos estudiantes, cuestionando las diferencias en el grado de compromiso y predisposición de los integrantes del grupo que hacía recaer todo el trabajo en unos pocos. Esta disfunción grupal probablemente pudo afectar el “proceso” de la evaluación, pero por los resultados, no siempre se reflejó en el “producto”.

Un hallazgo interesante en los estudiantes de MED-III, fue la similitud de puntajes que obtuvieron cuando los instrumentos de evaluación eran del mismo formato; los puntajes más bajos ( $< 62$ ) correspondieron a las preguntas de opción múltiple; los valores intermedios (77 a 80 puntos), a los mini CEX y discusiones de casos clínicos; y los puntajes más altos ( $> 83$ ), a la Historia Clínica, búsqueda de evidencias y al Portfolio.

La explicación para estos resultados puede estar en el mismo escenario “clínico” de aprendizaje y evaluación que están transitando, y que es muy similar al descrito por Al Kandri (2011): en el ciclo clínico, además del conocimiento, los estudiantes son evaluados en su capacidad para realizar habilidades específicas en la práctica clínica, organizar pensamientos, vincular el conocimiento teórico al manejo de los pacientes, comunicarse con los pacientes y el entorno, redactar informes y registros de los pacientes, y reflexionar sobre sus prácticas. En este contexto, las evaluaciones del desempeño clínico, los trabajos de ensayo y los ejercicios reflexivos, estimulan a los estudiantes a adoptar enfoques más profundos de aprendizaje, que puede explicar su mejor desempeño con estos instrumentos.

## 5.2. Segundo Eje de Análisis: Evaluación para el aprendizaje

### 5.2.1. Factores de la evaluación

Comprende los objetivos y criterios de las evaluaciones; el material evaluado; la interacción con los docentes; la proximidad temporal a la evaluación; y la sobrecarga de tareas simultáneas.

#### 5.2.1.1. ¿Qué sabían los estudiantes sobre los objetivos y criterios de las evaluaciones? y ¿Qué interacción tuvieron con los docentes?

Según Black & Wiliam (2009), una de los criterios clave para mejorar los efectos de aprendizaje de las evaluaciones es: clarificar y compartir las intenciones de aprendizaje y los criterios para el éxito, con los estudiantes. Esta recomendación, que también comparten otros autores (Wiliam & Thompson, 2007; Yin et al, 2008; Dannefer, 2013; Ruiz de Gauna et al, 2015), es esencial en todo programa formativo cuando el propósito de las evaluaciones está enfocado en “evaluar para el aprendizaje”.

Los casos de estudio abordados en este trabajo, tenían explícitos en las guías didácticas e instructivos que ofrecían a los alumnos al inicio del cursado, los objetivos de aprendizaje explorados y los criterios de evaluación y calificación que se aplicaba en los exámenes. Sin embargo, la mayoría (68%) de los alumnos principiantes manifestó no tener conocimiento previo de dicha información, a diferencia del 74% de los estudiantes avanzados que expresó tener un amplio conocimiento de los objetivos de aprendizaje y de los criterios de evaluación, antes de afrontar las pruebas.

Esta divergencia entre los alumnos principiantes y avanzados puede tener su explicación en las diferentes estrategias de comunicación que utilizaron los docentes en MHS y MED-III. En la primera, por el alto número de cursantes, la información fue “a demanda” a través de un foro habilitado para consultas al que solo ingresaron menos del 15% de los cursantes, aunque la información también estuvo disponible en el aula virtual; en cambio, en MED-III la información fue explicada en detalle en la clase inaugural y durante todo el cuatrimestre tuvo un espacio exclusivo en el aula virtual con presentaciones, guías e instructivos, que los estudiantes debían visitar como un recurso del cursado.

Respecto a la interacción que tuvieron con los docentes, las experiencias y percepciones de los estudiantes fueron distintas entre los principiantes y los avanzados, pues los escenarios de aprendizaje y evaluación fueron muy diferentes en ambas asignaturas en cuanto al número de cursantes y relación docente/alumno.

MHS recibió 1.363 cursantes con un cuerpo académico de 37 docentes, lo que determinó una relación docente/alumno: 1/37; en consecuencia, las actividades estuvieron condicionadas por el alto número de cursantes, con grandes comisiones de alumnos (80 por comisión), que dificultaron mucho la comunicación y el acompañamiento presencial entre docentes y estudiantes, y actividades complementarias en el aula virtual apoyadas en tutorías a demanda, este escenario motivó varias críticas y opiniones negativas de los cursantes. No obstante, la mayoría de los alumnos principiantes (82%) reconoce haber recibido de los docentes responsables de su comisión, apoyo y orientación para la preparación de los exámenes.

Entre las consecuencias de la masificación, Preston et al (2020) comprobaron que los alumnos son conscientes que con un gran número de estudiantes es muy difícil para los docentes hacer seguimiento y devoluciones, y los mismos alumnos recomendaron la provisión de retroalimentación grupal generalizada, indicando que esto sería mejor que no tener retroalimentación.

En MED-III el contexto fue muy distinto, con 151 estudiantes y 35 docentes se estableció una relación docente/alumno: 1/4; con lo cual, se pudo trabajar en pequeños grupos de estudiantes y con mayor supervisión docente en todas las actividades de aprendizaje (aulas invertidas, talleres y trabajos prácticos con pacientes) y aplicar instrumentos de evaluación con retroalimentación inmediata. Este escenario generó un espacio de comunicación más fluido y efectivo entre docentes y estudiantes que no tuvo percepciones negativas.

La necesidad de retroalimentación inmediata fue validada en una revisión realizada por Aguis & Wilkinson (2014), quienes comprobaron que los estudiantes esperan esto, ya que les brinda la oportunidad de mejorar en las tareas de evaluación posteriores y también como una forma de tranquilidad. Las estrategias sugeridas incluyen proporcionar ejemplos, comentarios y apoyo emocional adicional para mejorar las expectativas de los estudiantes.

Por otro lado, también es importante, la oportunidad de interactuar con el evaluador durante un examen o cuando el evaluador interactúa con el alumno durante la observación directa, pues los estudiantes lo perciben como una influencia potencial en el resultado de la evaluación ya que la interacción con el evaluador les brindaba más oportunidades para mostrar su progreso y habilidades, y les hace sentir que tienen más control sobre el proceso y el resultado de la evaluación. (Schut et al, 2018)

### **5.2.1.2. ¿Cómo influyeron en el estudio las evaluaciones múltiples y la proximidad temporal a cada evaluación?**

La planificación de las instancias de evaluación durante el cursado fue muy diferente en los dos casos de estudio.

En MHS las cinco evaluaciones del segundo tramo se aplicaron en un período de dos semanas con un intervalo de uno a dos días entre cada examen, y durante ese período se suspendieron todas las actividades académicas del cursado con el propósito de facilitar a los alumnos la preparación para las pruebas. Esta planificación fue cuestionada por muchos estudiantes con críticas dirigidas principalmente a la proximidad temporal entre cada evaluación, objetando que estaban muy cercanas en el tiempo y eso dificultó la preparación para los exámenes.

La percepción de “falta de tiempo para estudiar” (cuando está próximo el día del examen), es un fenómeno muy común entre los estudiantes, y ha sido explorado entre los efectos que tiene la inminencia temporal de una evaluación sobre el aprendizaje; al respecto, Cilliers et al (2012) en el “modelo de efectos de aprendizaje de la evaluación sumativa”, observaron que a medida que la evaluación se vuelve más inminente, la probabilidad de impacto es inevitable, los estudiantes perciben que no saben lo suficiente e intentan memorizar tantos datos como pueden para aprobar el examen y eso implica muy poca comprensión y análisis; en consecuencia, la gravedad del impacto (desempeño en la prueba) depende de las actividades de procesamiento cognitivo que cada estudiante adopta.

Por otro lado, la cantidad de tiempo que los estudiantes dedican a estudiar aumenta, hasta cierto punto, a medida que aumenta el volumen de material y/o el grado de dificultad del material a estudiar. Las altas cargas de trabajo también llevan a los estudiantes a ser más selectivos sobre qué contenido estudiar y suelen adoptar estrategias de procesamiento cognitivo de bajo nivel. (Cilliers et al, 2012)

Estas consideraciones podrían ser de utilidad para los docentes interesados en planificar las evaluaciones con tiempos y tareas que estimulen a los alumnos la adopción de enfoques más profundos de aprendizaje. Aunque para Wilson & Fowler (2005), los educadores saben mucho sobre los factores que contribuyen a malos resultados del aprendizaje y mucho menos sobre las condiciones ambientales en las que se pueden facilitar los enfoques profundos.

En MED-III, las instancias de evaluación se planificaron en forma secuencial a lo largo de todo el cuatrimestre, con intervalos de dos o más semanas entre cada examen, y sin

interrupción de las otras actividades del cursado. Este diseño también fue cuestionado por los estudiantes, principalmente por la concomitancia de los exámenes con las actividades teórico-prácticas, ya que para las evaluaciones no se suspendieron las actividades de cursado y eso generó, en algunos alumnos, una sobrecarga de tareas simultáneas y limitación del tiempo disponible para estudiar, debido a las exigencias propias del cursado.

Estas críticas coinciden con las observaciones de Preston et al (2020) en las discusiones de grupos focales con estudiantes de sexto año de medicina de una universidad regional de Australia, donde los mismos manifestaron que sería bueno mantener baja la carga de exámenes en el sexto año, ya que una menor presión les da tiempo a los estudiantes para consolidar el conocimiento y prepararse para la práctica profesional.

Por otro lado, en el estudio de Hashimat et al (2008), los estudiantes de medicina expresaron que las evaluaciones y la carga de trabajo del curso fueron muy extensas y se identificaron como fuentes comunes de estrés y ansiedad por los exámenes.

Para Gruppen et al (2018), lo más adecuado y efectivo para modelos curriculares basados en competencias es que la evaluación sea continua y flexible, pues si el tiempo no es fijo en el proceso de aprendizaje, tampoco se puede fijar en la evaluación, y debido a que los juicios de competencia se pueden hacer en cualquier momento, los datos de evaluación deberían recopilarse (más o menos) de forma continua. Aunque para implementar este tipo de evaluación son necesarios cambios organizativos muy significativos.

## **5.2.2. Efectos de aprendizaje**

Explora la naturaleza de las actividades de procesamiento cognitivo y de regulación metacognitiva en cada evaluación (cantidad y distribución del esfuerzo, elección de recursos y contenidos; estrategias de monitoreo); y mecanismos de las devoluciones.

### **5.2.2.1. ¿Qué procesos cognitivos y metacognitivos aplicaron en las distintas evaluaciones?**

Analizando las actividades de procesamiento cognitivo desarrolladas en las evaluaciones, se observaron diferencias entre los alumnos principiantes y los estudiantes avanzados, en los niveles cognitivos que aplicaron con mayor frecuencia e intensidad (según la Taxonomía de Bloom).

Los alumnos principiantes, en los exámenes de MHS, aplicaron principalmente operaciones mentales como: recordar, identificar, reconocer, ordenar, interpretar, seleccionar, explicar, comparar, esquematizar, graficar, valorar y otras; que si bien son procesos mentales que abarcan casi todos los niveles taxonómicos de Bloom, en la nube de palabras predominaron los niveles cognitivos de Orden Inferior (recordar, comprender y

aplicar). Este alcance cognitivo se corresponde con las respuestas de los mismos estudiantes cuando se les preguntó sobre el tipo de proceso mental que habían utilizado con mayor frecuencia para rendir las pruebas y la mayoría contestó: “Solo reconocer o recordar conceptos teóricos” (48%), e “Identificar o explicar mecanismos y procesos” (34%).

El alcance cognitivo de las evaluaciones de MHS, concuerda con los principales objetivos de aprendizaje de la materia que están dirigidos al conocimiento y comprensión de mecanismos, procesos y problemas relativos a la salud individual y colectiva; y aunque el predominio de niveles taxonómicos de orden inferior en las tareas de evaluación pueden parecer limitados para el aprendizaje, tal debilidad se compensa adecuadamente con la integración de contenidos básicos y clínicos que caracteriza la estructura de las pruebas aplicadas.

La integración de las ciencias básicas con contenidos clínicos o epidemiológicos, entendida como vínculos específicos y explícitos entre conceptos (observada en las diferentes evaluaciones de MHS) puede fortalecer las asociaciones entre los principios biomédicos subyacentes y las manifestaciones de enfermedad, y permitir a los estudiantes desarrollar una mayor comprensión del material cuando se integra con éxito. (Bierer et al, 2008; Kulasegaram et al, 2013)

Los estudiantes avanzados, realizaron principalmente operaciones como: identificar, interpretar, analizar, integrar, decidir, aplicar, resolver, planificar, fundamentar, juzgar, evaluar, justificar, elaborar informes, reflexionar, entre otras; y en las nubes de palabras predominaron los niveles cognitivos de Orden Superior (analizar, evaluar y crear), lo que también coincide con las respuestas de los estudiantes sobre el tipo de proceso mental que habían utilizado con mayor frecuencia en las evaluaciones, donde la mayoría contestó: “Razonar y aplicar lo aprendido a casos o problemas” (84%).

Estas actividades de procesamiento cognitivo realizadas por los alumnos al resolver los casos y problemas de las evaluaciones, son operaciones mentales muy valiosas para el aprendizaje y muy adecuadas para el nivel de formación que tienen los estudiantes avanzados, pues como lo destaca Anique de Bruin: *“Mientras el aprendizaje del novato se caracteriza principalmente por hacer malabarismos con los límites de la capacidad de su memoria de trabajo, el alumno experto puede aprovechar los beneficios de su formación en fragmentos, esquemas y guiones”*. (de Bruin et al, 2019:30)

Los fragmentos, esquemas y guiones<sup>39</sup> son construcciones cognitivas que los estudiantes desarrollan a lo largo de su formación y aunque suelen analizarse de forma independiente, todas están relacionados en la perspectiva asociacionista de la memoria. Es decir, las piezas separadas de información pueden interrelacionarse debido a la recuperación conjunta y repetida de las mismas, y de esa forma se accede a ellas como una sola unidad (o bloque) en la memoria de trabajo, lo que reduce la carga cognitiva y maximiza la capacidad de procesamiento durante una evaluación. En consecuencia, aplicando estos “bloques de construcción” los estudiantes pueden formar y reconocer patrones a través de la asociación; mejorar la velocidad y precisión del procesamiento cognitivo; y desarrollar a lo largo del tiempo habilidades más complejas de razonamiento. (de Bruin et al 2019)

La aplicación del reconocimiento de patrones conlleva también algunos riesgos para los estudiantes, pues como los patrones se reconocen por asociación rápida y automática (sin esfuerzos de razonamiento), constituyen una forma de procesamiento de “Tipo 1” (automático)<sup>40</sup> y se contraponen con el procesamiento de “Tipo 2” (analítico), por lo cual, el reconocimiento de patrones como procesamiento heurístico es propenso a sesgos como el cierre prematuro del razonamiento. (Croskerry et al, 2013)

Sin embargo, según Bowen (2006), es importante que los estudiantes comiencen a reconocer patrones y a desarrollar “script” de enfermedad, elaborados a partir del conocimiento de la presentación típica de los problemas; porque tanto el razonamiento automático como el analítico son efectivos y se usan simultáneamente en forma interactiva. El primero, es esencial para la experiencia diagnóstica y es una habilidad que se desarrolla a través de la experiencia clínica; mientras que el razonamiento analítico es la estrategia principal cuando un caso es complejo o mal definido, o el estudiante o residente tiene poca experiencia clínica con esa entidad en particular.

---

<sup>39</sup> **Fragmento** (chunk) es una fracción de información vinculada a un diagnóstico o tratamiento. **Esquema** (schema) son fragmentos que se organizan de una manera que facilita la distinción entre diagnósticos. **Guion** (script) representa un prototipo de una enfermedad que incluye sus características comunes, presentación clínica y respuesta al tratamiento. (de Bruin et al 2019).

<sup>40</sup> Según la teoría del procesamiento dual, hay dos tipos o sistemas de razonamiento: El **Tipo 1** es un sistema automático, subconsciente, más rápido y de menor esfuerzo, que se basa en las experiencias almacenadas en la memoria a largo plazo (patrones o esquemas construidos). El **Tipo 2** es un sistema analítico, más lento y de mayor esfuerzo, que combina funciones ejecutivas intensivas en recursos, incluida la memoria de trabajo y la atención. (Young et al, 2014)

### 5.2.2.2. ¿Cómo fueron las devoluciones en las diferentes evaluaciones?

Explorando los efectos catalíticos de las evaluaciones aplicadas en los casos de estudio, se observaron diferencias importantes en los tipos y mecanismos de retroalimentación que recibieron los alumnos principiantes y los estudiantes avanzados.

Según su programa, la asignatura MHS tenía un sistema de evaluación sumativo parcial (por tramos) con cinco herramientas aplicadas en un lapso de dos semanas, que eran evaluaciones de alto riesgo ya que debían ser aprobadas (tres de las cinco) por los alumnos para no perder su regularidad en el cursado. En ese contexto, el mecanismo de las devoluciones fue “diferido” hasta completar todas las instancias de evaluación, y consistió en publicar las calificaciones que obtuvieron en cada prueba y habilitar un mail mediante el cual los estudiantes pudieron plantear sus consultas y reclamos, y recibir información escrita y detallada sobre su desempeño en las pruebas. Este mecanismo de devolución “diferida” se vio reflejado en las opiniones de los estudiantes en el cuestionario, donde el 85% señaló haber recibido solo la calificación obtenida en las pruebas.

Asimismo, varios alumnos manifestaron que esperaban en los exámenes devoluciones “más específicas” y “personales”, mecanismo que no era aplicable en el contexto de la materia, por el gran número de cursantes y el propósito sumativo (de alto riesgo) que tenían las evaluaciones, ya que no había recuperatorios ni otras oportunidades para rendir.

Las evaluaciones sumativas (finales) sin retroalimentación posterior, son prácticas docentes habituales en la educación médica tradicional donde pueden verse carreras que no planifican instancias de evaluación formativa, y estudiantes que no perciben o reconocen como “feedback” los comentarios que reciben en sus evaluaciones. (Contzen et al, 2018)

Respecto a las devoluciones por escrito que se ofrecieron a los alumnos de MHS en forma diferida, son para Bartlett, Crossley & Mckinley (2017) un problema en varias escuelas de medicina, porque la mayoría de los estudiantes evaluados sintieron que los comentarios escritos eran de baja calidad, muy breves o incluían una serie de "ticks" con muy poco o ningún comentario sobre cómo mejorar.

En consecuencia, la ausencia de devoluciones oportunas y efectivas, puede interpretarse como una debilidad del sistema de evaluación de MHS, pues numerosas evidencias indican que la retroalimentación personalizada, constructiva, específica y detallada conduce a un mayor compromiso y aprendizaje de los estudiantes de grado. (Burr et al, 2013; Telio et al, 2015; Lefroy et al 2015; Preston et al, 2020).

Por su parte, MED-III utilizaba un sistema de evaluación longitudinal multimétodo con cinco instrumentos de evaluación aplicados en forma secuencial a lo largo del cursado, y cuyos resultados se usaban con un doble propósito, formativo y sumativo<sup>41</sup>.

A los efectos formativos, no hubo puntaje mínimo de aprobación para las pruebas y los estudiantes recibieron devoluciones en todas las evaluaciones, con formatos diferentes según el instrumento aplicado. En los exámenes de opción múltiple, se discutieron con los alumnos las preguntas de mayor dificultad; en los mini CEX se efectuaron devoluciones constructivas inmediatas a cada estudiante sobre su desempeño; para el Trabajo Final se realizaron tutorías para facilitar la redacción de las historias clínicas y el trabajo bibliográfico; en el Cafecito infectológico se aplicaron tutorías online; y en el Portfolio se realizaron devoluciones individuales después de su evaluación.

Estas formas de devolución se corresponden con las opiniones aportadas por los estudiantes en el cuestionario, donde el 78% expresó haber recibido explicaciones sobre sus aciertos y errores en las evaluaciones.

La combinación de diferentes formas de retroalimentación según el instrumento de evaluación aplicado, como las devoluciones verbales inmediatas al miniCEX, los ejemplos y explicaciones en las tutorías, y los comentarios escritos de los ensayos, han sido según Aguis & Wilkinson (2014), muy bien valorados por los estudiantes en otros estudios, especialmente cuando se incluyen sugerencias enfocadas en la mejora, porque no solo impacta en las calificaciones obtenidas sino también en la práctica clínica futura.

Para Bowen (2006), el profesor clínico debe proporcionar al alumno una retroalimentación cognitiva específica, señalar la información significativa en los datos del caso, identificar hallazgos redundantes o irrelevantes y resaltar las características diferenciales para sacar conclusiones sobre el diagnóstico correcto. Además, debe fomentar la lectura que promueva la conceptualización en lugar de la memorización, brindar a los alumnos la oportunidad de compartir lo que han aprendido, y reforzar la importancia del aprendizaje autodirigido.

---

<sup>41</sup> Una actividad de evaluación, per se, no es formativa ni sumativa, por lo que los mismos datos de evaluación pueden usarse para juicios formativos o sumativos. Aunque los juicios y decisiones sumativas generalmente requieren más cantidad de datos de evaluación y de mayor calidad, que para los juicios formativos. (Gruppen et al, 2018)

### 5.3. Tercer Eje de Análisis: Evaluación como aprendizaje

#### 5.3.1. Antes de la evaluación

Abarca los comportamientos de estudio (hábitos, habilidades y actitudes) que los estudiantes aplicaron en su preparación para las evaluaciones, incluyendo gestión del tiempo, establecimiento de metas, selección de qué, cómo, y dónde estudiar; y factores que motivaron y desanimaron el estudio.

##### 5.3.1.1. ¿Cómo influyeron las evaluaciones en los comportamientos de estudio de los estudiantes?; ¿Que ajustes hicieron en sus estrategias de estudio?; ¿Por qué lo hicieron?

Los cambios realizados por los estudiantes en sus comportamientos de estudio para preparar las evaluaciones se agruparon para su análisis en tres categorías: A) Cambios en las habilidades y técnicas de estudio; B) Cambios en los hábitos de estudio; y C) Elección de los materiales de estudio.

#### A) Cambios en las habilidades y técnicas de estudio

- **Releer los materiales hasta comprender**

La lectura es un proceso que requiere tiempo y dedicación, puesto que, después de leer un documento, el estudiante debería realizar un análisis para asegurar la comprensión, dando respuesta a preguntas como: ¿puedo resumir el significado del texto con mis palabras?, ¿puedo dar ejemplos?, ¿qué me queda claro y qué necesito aclarar?, ¿puedo conectar las ideas centrales del texto con otras ideas que entiendo? (Enríquez et al, 2015)

Esta técnica fue aplicada principalmente por los alumnos principiantes: *"...tenía que leer en voz alta varias veces los apuntes de la cátedra para poder comprender..."; "...analizaba y trataba de entender lo que estaba leyendo... me cuesta mucho memorizar"*.

Releer varias veces el material de libros o apuntes es una práctica de estudio muy común de los estudiantes. Sin embargo, según de Bruin et al (2019), la exposición repetida al mismo material sin intentos adicionales de construir asociaciones significativas solo puede conducir a una sensación de familiaridad. Lograr solo una sensación de familiaridad al releer conducirá a una falsa sensación de confianza con el contenido y una falsa creencia de que la relectura es efectiva, pero los estudiantes se aferran a esta creencia incluso cuando se les presenta evidencia de que la relectura por sí sola no es efectiva.

- **Elaborar resúmenes de cada material**

Según Enríquez et al (2015), un resumen es un texto que transmite la información de otro texto en forma abreviada. Hacer resúmenes es una técnica de estudio fundamental, y

exige una lectura comprensiva para identificar la información más importante incluida en el texto. Por lo tanto, hacer un resumen requiere ciertas habilidades para la comprensión y redacción de textos pues implica transformar un documento base en otro que reproduzca el cuerpo de ideas principales del primero en forma global y breve, dejando de lado las ideas accesorias.

Los estudiantes de los dos casos de estudio aplicaron esta técnica: *“...a medida que iba leyendo iba escribiendo lo más importante y así me quedaba un resumen... para repasar en los exámenes”* (principiante); *“...fuimos haciendo resúmenes de los temas más importantes y eso nos facilitó mucho la preparación para los parciales”* (avanzado).

Al ingresar a una carrera universitaria, es natural que los alumnos utilicen estrategias y técnicas de estudio con las trabajaron antes (en el nivel medio) y piensan que no hay necesidad de cambiar estos hábitos porque continúan siendo efectivos. En tal sentido, muchos estudiantes principiantes utilizan la elaboración y memorización de notas de clase o resúmenes como una herramienta principal para estudiar en la universidad, y en este contexto esa técnica no es del todo efectiva. (Zhao & Hu, 2021)

- **Hacer esquemas, gráficos y dibujos**

Un esquema contiene de forma sintética las ideas principales, las secundarias y los detalles de un texto; y permite comprender y organizar la información, facilitando de este modo la percepción y el recuerdo de las semejanzas, diferencias y las relaciones entre ideas y conceptos, así como el rápido repaso del tema tratado. (Enríquez et al, 2015)

Estas técnicas fueron utilizadas por los estudiantes en ambos casos de estudio: *“para entender... tuve que hacer muchos esquemas y dibujos...”* (principiante); *“...hacer esquemas y mapas conceptuales me ayudó mucho a comprender...”* (avanzado)

Según Ainsworth et al (2011), cuando los estudiantes dibujan para explorar, integrar y justificar sus conocimientos en ciencias, están más motivados para aprender que con la enseñanza tradicional. A través del dibujo, los alumnos razonan de varias maneras, alineando su trabajo con la observación, la medición y las ideas emergentes. Este razonamiento creativo es distinto, pero complementario, del razonamiento argumentativo, y permite a los estudiantes organizar su conocimiento de manera más efectiva e integrar la comprensión nueva con la existente, haciendo que su pensamiento sea explícito y específico, lo que conduce a oportunidades para intercambiar y aclarar significados entre sus compañeros.

Estas imágenes, según los investigadores en neurociencias, se procesan en la memoria sensorial dentro de la memoria de trabajo. La memoria de trabajo es una conceptualización de la memoria a corto plazo, que se refiere a la capacidad de mantener y procesar información de manera intencional, y puede depender de la memoria sensorial para retener información el tiempo suficiente para decidir qué hacer con ella. Esta memoria sensorial, puede conscientemente, copiar una entrada (imagen o sonido) y generar un “cuaderno de bocetos visuoespacial” o un “bucle fonológico” que se retienen como copias momentáneas durante un breve tiempo. (de Bruin et al 2019)

## **B) Cambios en los hábitos de estudio**

- **Aprendizaje colaborativo entre pares**

El aprendizaje cooperativo, según Duran (2015), es un recurso didáctico que convierte el trabajo grupal en aprendizaje, y fue una de las estrategias utilizadas tanto por los alumnos principiantes: “... con mis compañeras nos íbamos haciendo preguntas... y discutíamos las respuestas...”, “...estudie mejor con mis compañeros en la academia..., nos daban preguntas y ejercicios para responder...”; como por los estudiantes avanzados: “...con mi compañero de estudio simulábamos un ateneo, uno presentaba el caso y el otro hacía las preguntas...”; “...aprendimos mucho trabajando en grupos... y escuchando a los demás compañeros...”.

Estudiar con un/a compañero/a, puede ser equivalente a una “tutoría entre pares con rol recíproco” que para Duran (2015), tiene la ventaja de alternar los roles de tutor y tutorado entre cada alumno, permitiendo una construcción, multidireccional y simétrica del conocimiento.

Las discusiones conceptuales entre pares, la construcción de conocimiento compartido y la resolución conjunta de problemas, estimulan a los estudiantes a reflexionar sobre su comprensión y a coordinar el proceso de aprendizaje colaborativo, abordando directamente su regulación metacognitiva. Durante ese proceso, un alumno puede asumir un papel más instructivo para guiar la regulación metacognitiva de otro compañero, lo que resulta en la co-regulación del aprendizaje; pero también, varios compañeros pueden asumir juntos la responsabilidad del aprendizaje del grupo y regular de manera interdependiente el proceso de aprendizaje colaborativo hacia objetivos compartidos. Esta regulación metacognitiva socialmente compartida es el modo más profundo de regulación social y contribuye en gran medida al éxito del aprendizaje colaborativo. (De Backer et al, 2015)

- **Utilización del tiempo disponible para estudiar**

Las dificultades en la gestión del tiempo de estudio estuvieron presentes en ambos casos de estudio. Para los alumnos principiantes: *“...fue una locura, nunca estudié tanto en mi vida, ...no tuve ni un minuto libre”*; *“... estudiamos para todas las clases... para no atrasarnos, porque en las semanas de examen había muy poco tiempo”*; *“...con mi compañera estuvimos a full... se terminaron las salidas...”*; *“en casa no me podía concentrar, ...estudiaba mucho mejor en la academia”*. Para los estudiantes avanzados: *“...no supimos organizarnos con los tiempos y encontrar buenos momentos para estudiar”*; *“haber puesto las evaluaciones sin suspender el cursado de esa semana nos alteró muchísimo los tiempos de estudio”*; *“...el cursado me limitó mucho el tiempo de estudio, ...ya que tuve que hacer otros tipos de actividades...”*; *“Las distintas formas de evaluación y cursado, dispersaron mi objeto de estudio”*

Según Zárate-Depraect et al (2018) la mayoría de los estudiantes universitarios no poseen técnicas adecuadas de estudio, tienen dificultades para concentrarse y no organizan su tiempo adecuadamente; y observaron que el 41% de los estudiantes del área de la salud manifestó que su planificación para el estudio era insuficiente.

La “tiranía del tiempo” es un constante en casi todos los currícula basados en competencia de “tiempo fijo”, pues no todos los estudiantes necesitan el mismo tiempo para desarrollar su formación. Para revertir esta situación ha surgido el enfoque de la educación médica basada en competencias “tiempo variable” que resalta el hecho de que el aprendizaje entre diferentes personas varía en duración, base y objetivo. En consecuencia, si el tiempo no es una cantidad fija en el proceso de aprendizaje, tampoco se puede fijar en la evaluación; y dado que los juicios de competencia se pueden hacer en cualquier momento, los datos de evaluación deben estar disponibles y recopilarse de forma continua. (Gruppen et al, 2018)

Sin embargo, implementar un sistema de evaluación flexible en el tiempo conlleva un gran esfuerzo y compromiso de los docentes y demanda recursos administrativos y logísticos considerables, aunque puede ser muy beneficiosa para la formación de los profesionales de la salud. (Gruppen et al, 2018)

### **C) Elección de los materiales de estudio**

- **Guías y material didáctico elaborados por los docentes**

Estos materiales, que los estudiantes denominan “apuntes de la cátedra” fueron los más utilizados por los alumnos principiantes: *“...las guías que teníamos en el campus nos*

*orientaron bastante...”; “...usamos mucho los resúmenes de temas que hicieron los docentes”; “...no teníamos libro de referencia...tuve que buscar muchas cosas en internet” “...los ejemplos que pusieron de ejercicios resueltos fueron muy útiles...”.*

Esta elección es muy entendible pues los contenidos de la materia Medicina Hombre y Sociedad, por su diversidad epistemológica, no están todos incluidos en un libro de texto, y tratándose de alumnos sin formación previa en estos temas, requieren bases conceptuales claras y ejemplos de los ejercicios y problemas que deben resolver durante el cursado.

Algunos autores sostienen que el uso de ejemplos resueltos de problemas reduce la carga cognitiva intrínseca de aprender la mecánica de resolución de problemas. Al respecto, se ha observado un mejor rendimiento de los estudiantes de medicina cuando recibieron explicaciones utilizando problemas ya resueltos, en comparación con resolver ellos mismos los problemas por su cuenta. Sin embargo, este beneficio depende del contenido y del nivel de formación del alumno; pues el uso de ejemplos resueltos conduce a una disminución del rendimiento cuando se trata de estudiantes avanzados. (de Bruin et al 2019)

- **Búsquedas bibliográficas en internet**

Estas fuentes de información fueron más utilizadas por los estudiantes avanzados: *“hacer la H. Clínica me hizo invertir el orden del estudio, primero tomé los datos del paciente y después fui a buscar bibliografía para explicarlos...”; “...me costó mucho aprender el razonamiento clínico, tuve que buscar artículos y ver videos para poder hacerlo...”; “...engancharme en las discusiones me sirvió mucho,...te obliga a buscar evidencias y a redactar con claridad...”; “Haber cursado basada (Medicina Basada en la Evidencia) me sirvió mucho para buscar y seleccionar artículos para fundamentar mis comentarios....”.*

Este proceso de búsqueda de información científica confiable dirigida a comprender y/o resolver los problemas de los pacientes o casos estudiados, es una de las tareas que deben realizar los estudiantes para el Trabajo Final y una de las competencias enunciadas en el programa de la asignatura MED-III. Por lo tanto, constatar que algunos estudiantes han incorporado (naturalizado) la metodología de la Medicina Basada en la Evidencia a sus hábitos de estudio, es una muy buena noticia para los profesores, pues a pesar del interés establecido en la MBE como una competencia central para los médicos, la evidencia sobre la mejor manera de enseñarla y evaluarla sigue siendo débil.

Sobre este aspecto, en una revisión sistemática de 85 estudios controlados que habían investigado el efecto de las intervenciones educativas de Medicina Basada en la Evidencia, solo el 12% abordó todos los pasos de la MBE, la mayoría (74%) se enfocó

solamente en la evaluación crítica de la evidencia y el 22 % en los comportamientos. Ningún trabajo utilizó instrumentos de calidad para medir los resultados de las intervenciones. (Albarqouni et al, 2018)

- **Estudio de trabajos y exámenes de cohortes anteriores**

La revisión de trabajos/ensayos y la memorización de preguntas de exámenes anteriores son hábitos de estudio muy comunes en los estudiantes de la salud (Zárate-Depraect et al, 2018); y fueron estrategias que aplicaron tanto los alumnos *principiantes*: “...ver las respuestas de exámenes de otros años nos ayudó a comprender mejor las preguntas...”; “me gustó estudiar en la academia porque nos enseñaban a preparar los exámenes y como rendir...”; “...estudie mejor...en la academia ..., nos daban preguntas de años anteriores y ejercicios para responder...”; como los avanzados: “Con mi compañera conseguimos preguntas de exámenes anteriores y así pudimos aprender muchos temas de la materia porque todos eran casos clínicos”; “Nunca había hecho una historia clínica completa, tuve que leer varias de otros años para poder hacerla...”.

Utilizar exámenes como estrategias de estudio puede mejorar el aprendizaje de los alumnos y mejorar su rendimiento en una prueba posterior. Las pruebas pueden ayudar a fortalecer los vínculos entre señales simples, como características clínicas clave contenidas en el enunciado de la pregunta (viñeta clínica), y categorías de diagnóstico más complejas, contenidas en las opciones de respuesta a la pregunta. (de Bruin et al 2019)

El mecanismo propuesto para explicar este aprendizaje mejorado mediante pruebas (que en inglés denominan “Test-enhanced learning”), es la práctica de recuperación o la exposición repetida a claves (pistas o indicios en las preguntas) que facilitan el acceso a información relevante, y esto puede ser más efectivo cuando se hacen comentarios sobre las respuestas correctas. (de Bruin et al, 2019; Biber et al, 2020)

### **5.3.1.2. ¿Qué factores motivaron o desanimaron el estudio, en la fase de preparación para las evaluaciones?**

#### **A) Factores que motivaron el estudio:**

- **Las exigencias del sistema de evaluación**

Para los alumnos principiantes las mismas exigencias del sistema de evaluación fueron una motivación para el estudio previo: “...me parecieron bastante complicadas, porque requerían un estudio minucioso y profundo de todos los materiales... me gusta leer todo, pero no siempre hago una lectura significativa del todo, a veces para las clases dejo de

*enfocarme en algunas cosas que considero no son importantes de acuerdo a los objetivos propuestos”; “El sistema... es muy exigente y estoy de acuerdo con eso porque me ayuda a estudiar a fondo los temas para ir a las evaluaciones segura de mí misma”;*

Estas reflexiones, coinciden con los hallazgos de Preston et al (2020) donde los estudiantes a pesar de percibir una alta carga de trabajo cognitivo en las evaluaciones, consideraron que las mismas eran una gran fuerza motivadora para estudiar y preparar los exámenes y muy útiles para impulsar el aprendizaje.

- **Las condiciones de regularidad/promoción**

Por su parte, los estudiantes avanzados destacaron el efecto motivador que tuvo el mecanismo de “promoción meritoria” utilizado por la asignatura: *“Creo que lo que más te motiva a estudiar para los exámenes es la posibilidad que nos da la cátedra de aprobarla sin rendir el final cuando tenés un buen puntaje”; “... todos queríamos llegar a ser un estudiante destacado y eso nos motivó a estudiar mucho para cada examen...”*

La “promoción meritoria” es un reconocimiento especial que concede la asignatura MED-III a los estudiantes destacados, y consiste en otorgarles su promoción sin rendir el examen final de la materia (establecida por normativa).

Para ser un estudiante destacado hay que cumplir los siguientes criterios: Acreditar un 90% o más de asistencia a todas las actividades del cursado; haber rendido todas las evaluaciones; y lograr un puntaje total de 800 puntos o más (equivalente a nota 8).

## **B) Factores que desanimaron el estudio**

- **El estrés del cursado y las evaluaciones**

La mayoría de los autores coinciden en que los factores estresores que más afectan a los estudiantes universitarios son los relacionados con los procesos de evaluación y con la sobrecarga de trabajo académico. (Cabanach et al, 2016; Zárata-Depraect et al, 2018)

Algunos alumnos principiantes objetaron la planificación de evaluaciones muy cercanas en el tiempo: *“No somos genios...es imposible rendir cinco exámenes en dos semanas”; “El estrés sometido a los alumnos es exuberante...”; “...el lunes a la tarde rendimos el ALAF y el miércoles a la mañana tuvimos el SISAM, una locura”; “...al estar dos semanas presionada con 5 exámenes llegue el ultimo día agotada”; “Las evaluaciones no me dejaban dormir”; “Si no quieren que ingresemos, que nos tomen un solo examen eliminatorio así no pasamos por tanto estrés”.*

Por su parte, los estudiantes avanzados cuestionaron el estrés que generó el cursado con evaluaciones simultáneas: *“El estrés que estaba manejando fue impresionante, hasta que me superó..., fui sin leer temas por elección, no por falta de tiempo, yo decidí hacerlo así para no volver a ese grado de estrés... lo único que puedo afirmar es que no sabemos manejar las situaciones estresantes y eso resta mucho”*; *“Tantos exámenes me ponían muy nervioso... y cuando llegaba el momento, mi cabeza se ponía en blanco y no podía repetir lo que estudié”*

Estudios en neurociencias han mostrado que las emociones negativas de los estudiantes pueden dificultar el aprendizaje y el desarrollo de habilidades. Al respecto, se sabe que el estrés afecta el rendimiento de la memoria de trabajo, especialmente cuando se experimenta como una amenaza y no como un desafío, y esto sucede principalmente cuando el desempeño se evalúa en circunstancias estresantes (como las expresadas más arriba por los estudiantes). En contraposición con lo anterior, también se ha comprobado que la consolidación de la memoria mejora con niveles moderados de estrés, por lo tanto, no es el estrés en sí mismo, sino el nivel de estrés y la evaluación del estrés que hace cada estudiante, lo que determina su efecto en la memoria. (de Bruin et al, 2019)

- **Las evaluaciones grupales**

Para alumnos principiantes las evaluaciones grupales desanimaron el estudio: *“Las evaluaciones grupales no te motivan para nada, porque el resultado depende del compromiso que tenga el grupo para estudiar durante el cursado”*; *“...las personas del grupo nunca tienen la misma predisposición o compromiso y el trabajo siempre recae en unos pocos”*; *“...no me parece bien la evaluación grupal dado que es diferente el nivel de conocimiento entre los integrantes”*; *“En mi grupo, el TT presencial lo realizamos 3 integrantes y el resto no aportó nada”*; *“Las evaluaciones grupales no miden el desempeño individual”*; *“...no me parece justo que las evaluaciones sean grupales, ya que muchos aprueban a costa del esfuerzo de otros”*.

Llaman la atención estas expresiones de desánimo con las evaluaciones grupales cuando los resultados de las pruebas mostraron que los mayores puntajes obtenidos por los alumnos principiantes correspondieron a las evaluaciones de trabajos en terreno y de contenidos transversales, que son las dos evaluaciones que se realizan en forma grupal, a diferencia de las otras tres (ALAF, SISAM y EIT) que son pruebas individuales.

No obstante, es comprensible que algunos puedan percibir cierta incomodidad para trabajar en grupos, pues debido al gran número de alumnos “inscritos” en MHS, las comisiones y grupos de estudiantes se organizan en forma aleatoria (por orden alfabético) y

seguramente muchos de ellos no coincidieron en sus hábitos de estudio y enfoques de aprendizaje, con lo cual pudieron surgir grupos disfuncionales con bajo rendimiento.

- **El diseño de los instrumentos de evaluación**

Entre los factores que desanimaron el estudio, también fueron incluidos los problemas o defectos en el diseño de los instrumentos de evaluación.

Para algunos alumnos principiantes: *“Las evaluaciones no siempre se enfocan en lo que uno considera importante, es decir, en algunas ocasiones hacen hincapié en temas que algunos consideramos de poca importancia al momento de formarnos. Tal vez, eso sea algo que nos juega en contra al momento de estudiar y prepararnos para rendir los exámenes, ya que solemos pasar por alto algunas de esas cosas”; “...para que voy a romperme estudiando si en los exámenes todo es cuestión de suerte..”*

Los estudiantes avanzados también expresaron su desánimo por algunos instrumentos de evaluación utilizados: *“Desde el ingreso me sucede esto, la verdad siento que me iría mucho mejor o creo yo, que aportaría más en mí rendir exámenes escritos, creo que eso motiva mucho más a un alumno a estudiar y a entender lo que estudiar”; “...soy una persona que estudie muchísimo o poco siempre obtiene los mismos resultados en exámenes múltiple choice, por eso es que los choices no me motivan tanto”;*

Estas percepciones “negativas” que tienen los estudiantes sobre los instrumentos de evaluación son muy importantes para considerar pues constituyen factores que pueden afectar significativamente la motivación para el estudio y la preparación para los exámenes.

Diferentes autores (Epstein, 2007; Schut et al, 2018; Preston et al, 2020) sostienen que los alumnos estudian con mayor motivación y más cuidadosamente cuando perciben que los instrumentos con los que van a ser evaluados son adecuados para valorar lo que quieren demostrar a sus profesores sobre sus conocimientos o habilidades. Esto puede estar indicando cierta concordancia entre las intenciones de los estudiantes y los propósitos de los profesores, y esta coincidencia es un elemento muy valioso para mejorar los efectos de aprendizaje de las evaluaciones. (Watling & Ginsburg, 2019)

### **5.3.2. Durante la evaluación**

Abarca las experiencias de los estudiantes en las evaluaciones, los beneficios obtenidos en su formación y aprendizaje, y las oportunidades que tuvieron de mostrar su progreso y reflexionar sobre lo aprendido; así como los factores del sistema que facilitaron u obstaculizaron dichas oportunidades.

#### **5.3.2.1. ¿Qué beneficios aportaron las herramientas de evaluación a la formación y aprendizaje de los estudiantes?**

La mayoría de los estudiantes, en ambos casos de estudio, encontraron en las diferentes evaluaciones, oportunidades y beneficios para enriquecer sus aprendizajes y mejorar su formación.

Los alumnos principiantes (en mayor proporción que los avanzados) expresaron estar “de acuerdo” y “muy de acuerdo” en que preparando y rindiendo las evaluaciones de MHS habían aprendido más que cursando la materia, habían comprendido conceptos que no entendieron durante el cursado y lograron relacionar y articular mejor los diferentes contenidos de la misma.

Por su parte, los estudiantes avanzados (en una proporción más alta que los principiantes), estuvieron “de acuerdo” y “muy de acuerdo” en que las evaluaciones de MED-III los llevó a buscar y ampliar sus fuentes de información más allá de lo trabajado durante el cursado de la materia, y les dio la oportunidad de identificar sus necesidades de aprendizaje, dos habilidades que son esenciales para el desarrollo continuo y autónomo del aprendizaje.

Sin embargo, los efectos de aprendizaje que ocurren durante una evaluación no siempre son previstos por los estudiantes antes de la misma. Al respecto, Centeno et al (2007) aplicando entrevistas a 46 estudiantes de tercer año después de un examen oral, observaron que el 91% no tenía expectativas previas de aprender nada, pero después del examen, el 41% reconoció haber aprendido contenidos de la disciplina, el 40% había mejorado su integración de conocimientos con otras disciplinas y su proceso de razonamiento clínico, y el 40% había aprendido habilidades y técnicas de examen.

Los beneficios que obtuvieron los estudiantes en los casos de estudio, se discuten con mayor profundidad más adelante (punto 5.3.3.3.) donde se analizan los aportes de los sistemas de evaluación aplicados, al aprendizaje de cada materia.

### 5.3.2.2. ¿Qué factores o elementos de las evaluaciones facilitaron y cuales obstaculizaron el buen rendimiento?

#### A) Facilitaron el buen rendimiento

- **La diversidad de los instrumentos de evaluación aplicados**

Los alumnos principiantes manifestaron que la asociación de diferentes instrumentos facilitó el buen rendimiento en las evaluaciones: *“Creo que lo mejor de las evaluaciones es la relación que tenemos que realizar durante el examen, capaz cuando los estudiamos por separado no tiene todo el sentido pero cuando lo podemos unir a otra área resulta más fácil de comprender”; “...sentí que al dar exámenes tan distintos entre un horario y otro, el aprobar se vuelve una cuestión de suerte, pero puedo decir que me sirvieron de ayuda para estudiar y afianzar contenidos”; “...ya que los integradores de ejes fueron individuales y no grupales como el año pasado, mi desempeño académico y calificación final no se vieron afectadas por errores o falta de estudio ajenas”; “son mucho más prácticas que las pruebas comunes, no se pueden estudiar de memoria, te hacen pensar y relacionar conceptos para aplicarlos a la práctica, y si estudiaste así los resultados van a ser buenos”*

La diversidad de los instrumentos aplicados fue también un aspecto destacado en el trabajo de Preston et al (2020), donde los estudiantes en los grupos focales indicaron que apreciaban la variedad de los instrumentos de evaluación ofrecidos durante el curso y reconocieron que tal variedad era importante para adaptarse a los diferentes estilos de aprendizaje y a las necesidades de los diferentes alumnos.

En el estudio multicéntrico de Schut et al (2018) la mayoría de los programas educativos utilizaban múltiples evaluaciones complementarias destinadas a ofrecer más oportunidades para que los estudiantes muestren progreso y mejora. Esta diversidad de instrumentos redujo los riesgos de la evaluación “única” y dio a los alumnos una mayor sensación de tener el control, y más oportunidades para mostrar y mejorar su desempeño.

- **Las cualidades de los instrumentos utilizados**

Los estudiantes avanzados encontraron en las características de los instrumentos de evaluación un factor facilitador del buen rendimiento: *“...rendir diferentes exámenes me abrió la cabeza... ya no pienso solo en la respuesta correcta, aprendí por ejemplo a ver y encarar un caso clínico y como discernir entre lo importante y lo que no lo es, ...creo que por eso saque buenos puntajes”; “...las dos que me parecen más útiles son: los parciales que ayudan hacer un balance de lo que me cuesta y lo que no. Y la elaboración de la historia clínica*

*como ejercicio para el futuro”; “Considero que el minicex es una buena herramienta que demuestra nuestro desempeño frente al paciente y el utilizar los casos clínicos para la evaluación parcial refleja nuestra capacidad de comprender los temas. Ya que no podríamos aplicar a un contexto lo aprendido a menos que lo hayamos comprendido”*

La valoración diferenciada de los instrumentos de evaluación que hicieron los estudiantes, es un análisis muy interesante y coincide con los fundamentos de la ELM enunciados por los profesores en el programa. Esta conjunción entre los “efectos de aprendizaje” y los “propósitos de la evaluación”, es una combinación muy beneficiosa para que los estudiantes mejoren su rendimiento académico, pues los alumnos estudian más cuidadosamente cuando anticipan ciertos formatos de examen (Epstein, 2007), y los educadores tienen la oportunidad de impulsar el aprendizaje a través de la evaluación (Dannefer, 2013; Norcini et al., 2018).

## **B) Obstaculizaron el buen rendimiento**

- **Fallas en el diseño de los instrumentos de evaluación**

Para los alumnos principiantes los factores más importantes que obstaculizaron el buen rendimiento en las pruebas fueron las fallas en el diseño de algunos instrumentos: *“...algunas consignas no son completamente claras y generan confusión, sobre todo las del parcial integrador”; “...cuando no entendemos que es exactamente lo que piden y preguntamos a algún profesor este no nos resuelve las dudas tampoco. Y capaz sabíamos la respuesta, pero no entendimos la consigna”; “Podrían esclarecer un poco las preguntas, algunas son difíciles de comprender lo que facilita que se responda mal”; “Las pruebas integradoras son muy buenas para relacionar los contenidos entre sí, sin embargo, en vez de preguntar sobre temas importantes y centrales, algunos preguntan detalles que se encuentran en la última frase de un apunte, y eso puede hacerte salir mal”; “Considero que los trabajos transversales y de terreno suelen ser confusos, y que al momento de realizarlos no siempre se está en lo correcto, y por ende no se logra una buena calificación”.*

Estos comentarios coinciden con las observaciones efectuadas por Preston et al (2020) donde alumnos de primero a sexto año de la Universidad de “James Cook” (Australia) identificaron varios problemas relacionados con la calidad de las evaluaciones, que junto a los defectos en el diseño y claridad de las consignas, incluyeron también, la escasa relevancia de las tareas solicitadas en las evaluaciones, la pobre estandarización de los instrumentos y las discrepancias en las calificaciones otorgadas.

- **Fallas en la aplicación y corrección de los instrumentos**

Los estudiantes avanzados señalaron como obstáculos para el buen rendimiento en las evaluaciones a fallas en la aplicación y/o corrección de algunas pruebas: *“El tiempo que dieron para los parciales me pareció muy poco, ya que necesitaba razonar más los encabezados, además para pasar a la grilla también te lleva mucho tiempo”*; *“...el mecanismo de evaluación utilizado en el choice es muy complejo y nunca nos hemos enfrenado a ese estilo de examen por lo cual no se puede pretender aprobar con buena nota, lo cual no indica que el alumno no haya estudiado o este mal preparado”*; *“...presentaron contenidos que debían ser respondidos en base a la experiencia, y eso fue muy difícil ya que no contamos con ella...”*; *“...el sistema de corrección no es del todo objetivo... y a veces confuso, me encontré con devoluciones que no coinciden o se contraponen ... con las notas expuestas”*.

Estas observaciones de los estudiantes son “alertas” que deben ser consideradas en todo sistema de evaluación por competencias, pues autores como Lockyer et al (2017) y Gruppen et al (2018) concuerdan que la selección de los instrumentos y tareas de evaluación tienen profundas consecuencias para la validez de los juicios de evaluación; por ello, la cuidadosa planificación y aplicación de los procedimientos de evaluación mejora la calidad de los juicios de evaluación que son muy valiosos y necesarios para tomar decisiones válidas con respecto a los estudiantes en el ciclo clínico de una carrera.

### **5.3.3. Después de la evaluación**

Abarca las percepciones de los estudiantes dentro del entorno de aprendizaje y evaluación, como perciben su capacidad intelectual para afrontar las evaluaciones y los riesgos percibidos (no regularizar o promocionar, bajo promedio, etc); y cuanto aportaron las evaluaciones al aprendizaje de la materia.

#### **5.3.3.1. ¿Cómo percibieron su capacidad y sus habilidades intelectuales para afrontar las evaluaciones?**

- **Limitaciones por el cambio del nivel medio a la universidad**

Algunos alumnos principiantes percibieron que su formación en el nivel medio no fue suficiente o adecuada para afrontar las evaluaciones en la universidad: *“...fue muy fuerte el cambio, en el colegio nunca rendimos este tipo de exámenes..., y la verdad es que me sentí muy inseguro...”*; *“...las actividades y los exámenes fueron demasiados para el tiempo que tuvimos... ninguno de nosotros viene preparado para un ritmo de estudio tan intenso”*; *“...comencé en la academia... tres meses antes del cursado y por eso pude aprobar, si lo*

*hubiera intentado solo no habría llegado”; “...considerando que la gran mayoría no tenemos una formación básica en estos temas, no pueden pretender que en 4 meses pensemos y razonemos como médicos”; “No me gusta la matemática... no entiendo por qué tengo que hacer cálculos y gráficos para ingresar a la carrera de medicina... no es ingeniería”*

La “falta de preparación” de nivel medio que demuestran los estudiantes que ingresan a la universidad, es una percepción frecuente entre los profesores universitarios. Al respecto, analizando datos de México, Perú, Colombia y Chile, Zárate-Depraect et al (2018) observó que la mayoría de los alumnos ingresantes no poseen adecuados hábitos de estudio para enfrentarse a las exigencias académicas universitarias, y a lo largo de su formación muchos de ellos no los desarrollan. En general, no elaboran preguntas de estudio ni listado de palabras que desconocen, no leen comprensivamente el material que analizarán en clase, no hacen apuntes, no logran concentrarse en las tareas, no organizan su tiempo de estudio, y otras debilidades que pueden afectar su rendimiento en las evaluaciones y los aprendizajes.

Desde otra mirada, Donnon & Hecker (2010) sostienen que el rendimiento de los estudiantes en la educación postsecundaria depende en gran medida del enfoque de aprendizaje adoptado para las tareas y las evaluaciones dentro de cualquier curso o programa de estudio en particular, pues los resultados del aprendizaje están vinculados a las concepciones del aprendizaje de los estudiantes, sus percepciones del entorno de aprendizaje y sus enfoques hacia el aprendizaje. En consecuencia, si los enfoques que traen del nivel medio no se corresponden con el entorno de aprendizaje que les ofrece la universidad pueden percibir limitaciones en su capacidad para afrontar las evaluaciones.

- **Limitaciones en las capacidades personales**

Algunos estudiantes avanzados percibieron con incertidumbre ciertas limitaciones en sus habilidades personales para afrontar las evaluaciones: *“Me gustan las evaluaciones, creo que son muy útiles, pero creo que algo falla en mí que siempre me termino desilusionando de mí misma por más que me esfuerce”; “Me apasiona la carrera que elegí y trato de aprender cosas nuevas todos los días, pero me cuesta muchísimo demostrar en los exámenes todo lo que se... tal vez si fueran orales sería mejor para mí”; “Realmente no sé qué me pasó y estoy MUY decepcionada de mi desempeño. Mi sacrificio fue en vano. Mas allá de eso, el sentimiento de miedo y culpa que me generó me ayudó a cambiar la modalidad de estudio”*

Estos comentarios cargados de incertidumbre pueden explicarse a partir de las percepciones que tienen estos alumnos sobre el propósito de las evaluaciones. Según algunos autores (Harrison & Wass, 2016; Watling & Ginsburg, 2019) los estudiantes interactúan con la evaluación basándose en como comprenden su propósito, mientras los educadores tienen una noción clara de la intención que tienen las evaluaciones (formativa o sumativa). En ese contexto, las percepciones de los alumnos no siempre coinciden con las intenciones de los educadores. Los estudiantes de medicina a menudo perciben a la evaluación formativa (facilitar el aprendizaje) como si tuviera una intención sumativa (aprobar o reprobar), esa percepción puede interferir con el efecto educativo pretendido, y en consecuencia, la evaluación puede no reflejarse de la misma manera y no impulsar el aprendizaje del estudiante como se esperaba.

Las percepciones que tienen los estudiantes sobre las evaluaciones pueden ser angustiosas y hasta riesgosas para su formación. Cuando un sistema de evaluación es percibido como orientado a detectar y castigar a los alumnos incompetentes, los estudiantes no se sentirán apoyados para explorar los límites de su conocimiento, y esta exploración es exactamente el tipo de aprendizaje que se espera fomentar. (Schuwirth & Ash, 2013)

#### **5.3.3.2. ¿Qué consecuencias percibieron como riesgosas en el sistema de evaluación? (Riesgos percibidos)**

- **No poder ingresar a la carrera de Medicina**

La mayoría de los alumnos principiantes percibieron como el riesgo más importante del sistema de evaluación de MHS: “desaprobar los exámenes del segundo tramo y perder la regularidad en la asignatura” pues en la condición de “libres” no podían presentarse al Examen Final de la materia, que debían aprobar para ingresar como alumnos regulares a la Carrera de Medicina.

Este riesgo percibido está implícito en las condiciones “Inscripción e Ingreso” plasmadas en la Reglamentación del Plan 2000 de la Carrera de Medicina de la UNNE (Res. N°2400/02-CD), que en su Artículo 2° establece que: *“Serán considerados alumnos ingresantes aquellos que aprueben la primera asignatura de la Currícula de la Carrera”*.

Para aprobar la primera asignatura (MHS), los estudiantes deben aprobar el Examen Final de la materia y solo pueden rendirlo aquellos alumnos en condición de “regulares”.

Esta regularidad se logra aprobando los dos tramos de cursado que abarca la asignatura MHS, cada tramo culmina con cinco instancias de evaluación (dos grupales y tres individuales), donde los estudiantes deben aprobar al menos tres de las cinco evaluaciones.

- **No lograr promocionar la asignatura MED-III sin rendir el examen final**

Algunos estudiantes avanzados percibieron como un riesgo del sistema de evaluación de MED-III, no lograr la “promoción meritosa” que ofrecía la asignatura: *“...no había un puntaje establecido para aprobar los exámenes, ...teníamos que sacar la mayor cantidad de puntos posibles para no rendir el (examen) final”; “...el ultimo parcial me estreso muchísimo, tenía que sacar 90 puntos para ser estudiante destacado, ...estudí como loco pero lamentablemente no lo logré”; “...en las últimas semanas de cursado mi ansiedad fue enorme, ...todavía no tenía el resultado de la historia clínica y de ese puntaje dependía rendir o no el examen final”; “Estar tan pendiente de un puntaje te motiva a estudiar mucho, pero también te carcome los nervios..., no sé cuál es el beneficio”*

La pérdida de la “promoción meritosa” como riesgo percibido, concuerda con lo expresado por los estudiantes en un apartado anterior donde la “promoción meritosa” fue considerada como uno de los “factores que motivaron el estudio”.

Estas percepciones indican que la “promoción meritosa”, como reconocimiento al esfuerzo y desempeño demostrado por los estudiantes durante el cursado, ha cumplido en gran medida el objetivo esperado por los profesores de MED-III que fue: “estimular el rendimiento académico de los estudiantes”.

- **No poder continuar con la Practica Obligatoria Final**

A otros estudiantes avanzados les preocupaba no alcanzar la regularidad en MED-III: *“Yo sabía que promocionarla (por mérito) era difícil, ...solo me interesaba regularizarla, porque rendirla libre es mucho más difícil...”; “El principal riesgo es quedar libre, ...el final es muy complicado y si no lo aprobas no puedes hacer la POF al año siguiente”.*

Este riesgo percibido también se desprende de la Reglamentación del Plan 2000 de la Carrera de Medicina de la UNNE (Res. N°2400/02-CD), que en su régimen de correlatividades establece que: *“Para cursar la Práctica Obligatoria Final de 1600 horas, los alumnos deberán tener aprobadas la totalidad de las asignaturas de la Carrera (obligatorias y optativas)”.*

El temor de los estudiantes en este punto es el tiempo que podría demandarles preparar y aprobar el final de MED-III en condición de “libres”, ya que percibían que en tal

condición el examen final podía ser muy dificultoso de aprobar y esto podría retrasar su admisión al cursado de la Practica Obligatoria Final.

### 5.3.3.3. ¿Qué aportó el sistema de evaluación aplicado, al aprendizaje de la materia?

Según Schut et al (2018), el impacto de cualquier sistema de evaluación sobre el aprendizaje está mediado por las percepciones de cada estudiante, por ello es fundamental conocer estas percepciones para comprender qué significado tienen para los estudiantes, y cómo y porqué estas evaluaciones permitieron o no, que los mismos puedan optimizar y autorregular su aprendizaje a partir de las evaluaciones.

Al respecto, los alumnos principiantes reconocieron tres aportes significativos del sistema de evaluación aplicado en MHS:

- **Ampliar y profundizar los conocimientos de la asignatura:** *“Creo que las evaluaciones de MHS me permitieron comprender mucho más algunos temas que durante el cursado no los supe interpretar o poner en práctica”; “...hicieron que comenzara a prestar más atención a temas que muchas veces uno no le da tanta importancia”.*
- **Identificar las necesidades de aprendizaje:** *“...me permitieron darme cuenta los temas que me faltan ampliar y mi capacidad para estudiar, los cambios en la metodología de estudio fueron efectivas”; “si bien algunas me resultaron difíciles, fueron muy útiles para saber que conocimientos tengo que reforzar”.*
- **Desarrollar habilidades específicas:** *“La metodología de los exámenes me pareció bastante buena... ya que me permitió desarrollar el pensamiento analítico”; “Me permitieron afianzar los conocimientos y además incentivaron el desarrollo óptimo del trabajo en equipo y la importancia de éste”.*

A partir de las percepciones enunciadas por estos alumnos principiantes se puede inferir que las evaluaciones fueron muy beneficiosas para impulsar y desarrollar su aprendizaje de la asignatura. Esta valoración “positiva” que hicieron los estudiantes sobre el sistema de evaluación de la materia, según Preston et al (2020) es un factor que puede mejorar la motivación y el compromiso de los mismos para estudiar y aprender la asignatura, ya que también concuerdan con las intenciones y propósitos formativos de los docentes.

Lo interesante de esta observación, es que se trata de alumnos principiantes y de acuerdo al trabajo de Lynam & Cachia (2018), el compromiso de los estudiantes con la evaluación se alinea positivamente con el aumento de la madurez académica, como reflejo

de la creciente autenticidad de la evaluación a medida que el plan de estudios avanza y las actividades y evaluaciones se enfocan en tareas y prácticas clínicas, en contraste con las pruebas del conocimiento científico y el desarrollo de habilidades genéricas de los primeros años.

Por su parte, los estudiantes avanzados también percibieron en las evaluaciones de MED-III, efectos beneficiosos para el aprendizaje de la asignatura:

- **Ampliar y profundizar los conocimientos de la materia:** *“Las formas de evaluación son todas muy productivas. La combinación del foro, realización de portfolio, trabajos prácticos, minicex y examen parcial ayudan a profundizar más en todos los temas y no solo lo que se debería estudiar para un solo examen aislado”; “Creo que todos los métodos sirven, porque te hacen dudar de todo, y a partir de ahí nace el hecho de buscar más información y así adquirir más conocimiento”; “Me pareció una forma novedosa de evaluar..., hay que estudiar muchísimo, ...pero estoy muy contento, aprendí y conocí enfermedades que jamás había visto”.*

Es muy importante que los estudiantes hayan percibido a las evaluaciones como un “conjunto”, pues más allá de las preferencias que puedan tener por alguno de los instrumentos utilizados, los beneficios percibidos están dirigidos al “sistema de evaluación multimétodo” que aplica la asignatura, y este es un rasgo de “madurez académica” (Preston et al, 2020)

- **Desarrollar habilidades para reflexionar sobre el aprendizaje:** *“...la mayoría de nosotros necesitamos algo como este cursado para valorar nuestro esfuerzo y valorar hasta donde podemos llegar, porque es el esfuerzo lo que le da valor a las cosas y las facilidades se lo quitan”; “...sirvió para plantearme muchas cosas primero la forma de estudiar, pero también demostró mis falencias como estudiante, sin embargo, está haciendo que me esfuerce un poco más”*

La capacidad de los estudiantes de medicina para reflexionar, según Hoffman et al (2016), es fundamental para su formación profesional y su capacidad para trabajar en entornos complejos. Pero para que los mismos reconozcan la importancia de la reflexión, las actividades deben ir acompañadas de una cultura que valore la reflexión y apunte a desarrollar practicantes reflexivos.

La reflexión es un proceso metacognitivo complejo que implica el reconocimiento del propio funcionamiento cognitivo de los estudiantes permitiéndoles aprender

activamente de sus experiencias y evaluar una gama de conocimientos, habilidades y competencias tales como: profesionalismo, comunicación, empatía, razonamiento clínico y otras. (Preston et al, 2020)

Finalmente, algunos alumnos principiantes tuvieron percepciones más pesimistas: *“...las evaluaciones no te enseñan nada, las clases deberían ser donde se aprende”*; *“...lo único que me dejaron los exámenes de MHS fue un fuerte dolor de cabeza y ...”*; y una estudiante avanzada percibió en el sistema de evaluación de MED-III una intencionalidad encubierta: *“En fin creo que esta cátedra tiene como objetivo que nos vaya mal, tienen maldad, nos odian, no quieren que nos recibamos y disfrutan que nos vaya mal. Nunca me sentí tan basureada y pisoteada como en esta materia. El sistema de evaluación no es bueno. Todas las evaluaciones que se realizan obviamente son útiles y nos obligan a ir a estudiar. ¿Pero es necesaria tanta maldad? ... somos personas no somos máquinas y queremos estudiar pero no es necesario que nos hagan sentir tan mal”*

Estas apreciaciones de los estudiantes pueden interpretarse como la percepción de un “currículum oculto” que ellos mismo construyen a partir de los significados que intuyen en las evaluaciones. Este fenómeno fue estudiado por Sambell & McDowell (1998) y observaron que los estudiantes no responden de manera homogénea a la evaluación, ni la evaluación actúa como un disparador homogéneo para la acción. En consecuencia, cada estudiante puede construir su propia versión del “currículum oculto, desde sus vivencias y resultados de la evaluación, que nunca son del todo predecibles y por ende, la búsqueda de un sistema de evaluación “perfecto” puede estar condenada al fracaso desde el principio.



# CAPITULO 6

## CONCLUSIONES





## Capítulo 6: CONCLUSIONES

### 6.1. Conclusiones de la investigación

En este punto se hará una revisión de todo el proceso de diseño y elaboración de esta tesis, recuperando e integrando los interrogantes que guiaron el proceso, los aspectos metodológicos y elementos empíricos de la investigación, y los resultados, interpretaciones y reflexiones que fueron emergiendo en los diferentes momentos del recorrido investigativo.

El trabajo se inició con un interrogante que constituyó el problema de investigación:

**¿Qué impacto educacional puede producir en los estudiantes de medicina la aplicación de sistemas multimétodos para evaluar competencias?**

La investigación se realizó con un enfoque metodológico mixto (cuali-cuantitativo) aplicando la estrategia de complementación (Bericat, 1998), aunque con amplio predominio de la perspectiva cualitativa sobre la cuantitativa. De ese modo, se pudo contar simultáneamente con dos imágenes distintas de la realidad social investigada, donde cada método reveló aspectos diferentes, permitiendo enriquecer la comprensión de los hechos y ampliar el conocimiento del fenómeno estudiado.

El diseño metodológico corresponde a un estudio de casos de tipología múltiple y diseño incrustado -Tipo 4- con múltiples unidades de análisis (Yin, 2003), que si bien no constituye una muestra representativa de la que se puedan inferir generalizaciones, permitió explorar y comprender la especificidad de cada caso; hacer una lectura del conjunto, identificando características comunes y diferencias; y reconocer un arco de propiedades en torno a cada categoría que sirvió para interpretar el impacto educacional de los sistemas de evaluación multimétodo. En esa dirección, se eligieron como casos de estudio las dos asignaturas que aplicaban dicho sistema de evaluación en la Carrera de Medicina de la UNNE: Medicina Hombre y Sociedad (MHS), primera materia de la carrera; y Medicina III (MED-III) que se cursa en quinto año; y que eran las únicas que (hasta 2018) aplicaban sistemas de evaluación multimétodo.

Adoptando un alcance interpretativo se pudo indagar las diferentes dimensiones y perspectivas relacionadas al entorno aprendizaje-evaluación e identificar procesos, contextos y situaciones que ayudaron a comprender mejor la realidad investigada; y a partir de las experiencias y reflexiones de los estudiantes se pudieron construir interpretaciones sobre los efectos de aprendizaje que tuvieron las evaluaciones.

Las técnicas e instrumentos de recolección de la información aplicados: análisis de documentos y registros curriculares, inventarios estandarizados, y cuestionarios autoadministrados, permitieron recoger y triangular información sobre: el entorno de aprendizaje y evaluación en los casos de estudio; los enfoques de aprendizaje y aspectos motivacionales de los estudiantes, sus comportamientos de estudio, y sus experiencias y percepciones en las evaluaciones.

El marco analítico-metodológico resultó de mucha utilidad para ordenar y procesar los datos recogidos de las diferentes fuentes de información, y permitió realizar un proceso recursivo entre el encuadre teórico-conceptual y las indagaciones realizadas sobre los casos de estudio para profundizar en las diferentes categorías explicativas del fenómeno estudiado. De esa forma se pudieron abordar las tres dimensiones de las prácticas de evaluación de competencias: la evaluación del aprendizaje, la evaluación para el aprendizaje y la evaluación como aprendizaje, indagando en cada una, diferentes categorías de análisis a través de preguntas exploratorias.

Habiendo abordado los aspectos metodológicos, es posible responder ahora a los tres interrogantes que sintetizan los objetivos específicos enunciados en el capítulo introductorio, y que son las preguntas principales de los tres ejes de análisis que comprende la matriz analítica-metodológica construida para esta investigación:

- **¿Cómo se integran las evaluaciones al entorno de aprendizaje de los estudiantes?**

En el marco teórico construido para esta tesis, la mayoría de los investigadores expertos en evaluación de competencias coincide en que los sistemas de evaluación deben estar configurados a lo largo de la carrera para que todo el cuadro de competencias pueda ser evaluado con instrumentos adecuados al nivel de formación de los estudiantes, aplicando criterios claramente establecidos y métodos combinados cuidadosamente elegidos.

Trasladando estas pautas de la dimensión “currículum” a la “unidad curricular” se observó que había concordancia con los casos de estudio (MHS y MED-III) ya que diseñaron sus sistemas de evaluación multimétodo en forma integrada a los procesos de aprendizaje, articulando instrumentos diferentes para evaluar todo el espectro de las competencias abordadas en cada materia, que a su vez estaban definidas en el perfil de graduación. De esa manera, podían obtener información más amplia y completa sobre el desarrollo de las competencias en el mismo escenario del proceso formativo, superando las limitaciones que tienen los instrumentos de evaluación aplicados en forma aislada. Todas las herramientas de

evaluación fueron elaboradas y validadas por los docentes de cada materia, siguiendo guías establecidas por los profesores y aplicadas en forma consensuada entre los docentes evaluadores. Luego de su aplicación, los profesores responsables de las evaluaciones analizaron los resultados y el desempeño de los estudiantes en cada prueba y calcularon los indicadores psicométricos para evaluar el funcionamiento de los instrumentos utilizados.

Sin embargo, aunque las conductas y prácticas adoptadas por los docentes para el diseño y desarrollo de las evaluaciones en sus asignaturas son comparables, los escenarios de aprendizaje y evaluación fueron muy diferentes en cada materia. En MHS, el alto número de cursantes, la baja relación docente/alumno, el trabajo en grandes comisiones de alumnos y la amplia utilización de la virtualidad, limitaron el potencial formativo de las evaluaciones y el acompañamiento tutorial; mientras que MED-III con una óptima relación docente/alumno pudo trabajar en pequeños grupos de estudiantes, con mayor apoyo tutorial y supervisión docente, lo que facilitó la aplicación de estrategias de evaluación formativa.

A pesar de los contextos académicos tan diferentes, los estudiantes de ambos casos de estudio compartían algunas características comunes en sus metas de logro y enfoques de aprendizaje. Tanto los alumnos principiantes (MHS) como los estudiantes avanzados (MED-III), tenían una fuerte orientación por metas de “aproximación al aprendizaje”, que indicaba alta predisposición por aprender, aunque las razones que los motivaba a adoptar dicha orientación eran diferentes en cada caso. En los primeros, la expectativa que les producía la “nueva experiencia” de ingresar a la carrera de medicina y el interés por aprender contenidos “nuevos” fueron importantes motivaciones; y en los avanzados, porque muchos tenían autoeficacia alta por sus logros académicos a lo largo de la carrera y estaban dispuestos a esforzarse los últimos años, para completar el desarrollo de las competencias necesarias para su formación profesional. Por otra parte, la orientación por las metas de “aproximación al rendimiento” fue más intensa en los alumnos principiantes que en los avanzados, y esto puede explicarse por la fuerte intención dominante en los alumnos de MHS de regularizar y aprobar esa asignatura para poder ingresar a la carrera de medicina.

En relación a los Enfoques de Aprendizaje, se pudo comprobar en los dos grupos de estudiantes, un amplio predominio de “enfoques de aprendizaje profundo” que tenían mayor intensidad en los estudiantes avanzados, esta diferencia puede explicarse por la “madurez académica” de estos últimos y los escenarios (auténticos) de práctica clínica que les ofrecía MED-III, condiciones que pueden aumentar su motivación intrínseca y su compromiso con la asignatura, llevándolos a adoptar enfoques más profundos de

aprendizaje. Sin embargo, el “enfoque superficial” fue más frecuente en los alumnos principiantes reflejando el temor que tenían por el fracaso académico de no poder ingresar a la carrera elegida.

Con respecto a los resultados de las evaluaciones, se observó, tanto en los alumnos principiantes como en los avanzados, una importante diferencia de los puntajes según el instrumento de evaluación aplicado, lo que es esperable en cualquier sistema de evaluación de competencias multimétodo, pues cada instrumento explora atributos diferentes de la competencia y por ello, el resultado obtenido con un instrumento no es predictivo del resultado logrado en otro instrumento.

Los alumnos principiantes obtuvieron puntajes más altos en las evaluaciones grupales que en las pruebas individuales, lo que contrasta con sus propias expresiones de disconformidad y rechazo por los exámenes grupales, atribuida a la disparidad en el compromiso y predisposición de los integrantes del grupo. De estos resultados es posible inferir que, si bien la disfunción grupal pudo afectar el “proceso” de la evaluación, no siempre se reflejó en el “producto”.

Por su parte, los estudiantes avanzados lograron los puntajes más altos en la evaluación de la Historia Clínica con discusión clínica, en la búsqueda de evidencias y en el portfolio reflexivo. Esta observación coincide con la de otros autores que comprobaron, en estudiantes avanzados, que las evaluaciones en contexto clínico, los trabajos de ensayo y los ejercicios reflexivos, estimulan a los mismos a adoptar enfoques más profundos de aprendizaje, que puede explicar su mejor desempeño con estos instrumentos.

**De lo expuesto se puede concluir** que los casos estudiados habían diseñado sus propios sistemas de evaluación de competencias en forma integrada a los procesos de aprendizaje, con diversos instrumentos adecuados a los objetivos de cada evaluación, de acuerdo al nivel curricular y al contexto de cursado de cada asignatura; y con propósitos diferentes, sumativo en MHS y formativo-sumativo en MED-III. En esos escenarios, la mayoría de los estudiantes tenían orientación por metas de aproximación al aprendizaje (predisposición por aprender) y aplicaban enfoques de aprendizaje profundo, lo que se vio reflejado en los resultados que obtuvieron en las evaluaciones.

- **¿Como influyen las evaluaciones sobre el aprendizaje de los estudiantes?**

Existe consenso entre los diferentes autores que uno de los preceptos clave para mejorar los efectos de aprendizaje de las evaluaciones es clarificar y compartir con los

estudiantes los objetivos de aprendizaje y los criterios para obtener el mejor resultado en cada evaluación.

En tal sentido, los casos de estudio abordados en este trabajo, tenían explícitos en las guías didácticas e instructivos que ofrecían a los estudiantes al inicio del cursado, los objetivos de aprendizaje explorados y los criterios de evaluación y calificación que se aplicaba en los exámenes. Sin embargo, la mayoría de los alumnos principiantes no tenía conocimiento previo de dicha información, a diferencia de los estudiantes avanzados que mayoritariamente tenían amplio conocimiento de los objetivos de aprendizaje y de los criterios de evaluación, antes de afrontar las pruebas. Esta divergencia entre los alumnos principiantes y avanzados tiene su explicación en las diferentes estrategias de comunicación que utilizaron los docentes en MHS y MED-III para informar a los estudiantes al respecto. En la primera, por el alto número de cursantes, la información fue entregada “a demanda” mediante un foro habilitado para consultas, aunque también estuvo disponible en el aula virtual; en cambio, en MED-III la información fue explicada en detalle en la clase inaugural y durante todo el cuatrimestre tuvo un espacio exclusivo en el aula virtual con presentaciones, guías e instructivos, que los estudiantes debían visitar como un recurso más del cursado.

En cuanto a la interacción que tuvieron con los docentes, hubo percepciones dispares entre los estudiantes principiantes y los avanzados, porque sus experiencias en los escenarios de aprendizaje y evaluación fueron muy diferentes en ambas asignaturas en cuanto al número de cursantes y a la relación docente/alumno como ya fue explicado anteriormente.

La planificación de las instancias de evaluación fue otro aspecto cuestionado por los estudiantes en los dos casos. En MHS, por la proximidad temporal entre cada evaluación, ya que las cinco evaluaciones se aplicaron en dos semanas con uno a dos días de intervalo entre cada examen, y suspendiendo en ese período las otras actividades del cursado; mientras que en MED-III, criticaron la sobrecarga de tareas simultáneas y limitación del tiempo, pues no se interrumpieron las actividades académicas en las semanas de evaluación que fueron secuenciales a lo largo de todo el cursado.

En su exposición a los instrumentos de evaluación, hubo diferencias entre los estudiantes principiantes y los avanzados en cuanto a las actividades de procesamiento cognitivo que aplicaron para resolver o completar las pruebas. En los exámenes de MHS, los alumnos utilizaron procesos mentales que abarcan principalmente los niveles cognitivos de “orden inferior” como recordar, comprender y aplicar (Taxonomía de Bloom); en tanto que

los estudiantes avanzados, realizaron operaciones mentales en todos los niveles taxonómicos a predominio de los niveles cognitivos de “orden superior” (analizar, evaluar y crear).

También hubo disparidad en los tipos y mecanismos de retroalimentación que recibieron los alumnos principiantes y avanzados. Los primeros, recibieron devoluciones “diferidas” luego de completar todas las instancias de evaluación, que eran de alto riesgo (propósito sumativo) ya que condicionaban la regularidad en el cursado. Los estudiantes avanzados en cambio, al atravesar un sistema de evaluación de doble propósito (formativo y sumativo), recibieron devoluciones en cada una de las evaluaciones, con formatos acordes al instrumento aplicado.

**Por lo expuesto se concluye que**, tanto en los factores de la evaluación como en los efectos de aprendizaje, hubo diferencias entre los casos de estudio, debidas principalmente al elevado número de cursantes de MHS que contrastó con la adecuada relación docente/alumno de MED-III. En ese contexto “masivo” de cursado, los alumnos principiantes: no conocían (antes de las pruebas) los objetivos de aprendizaje y los criterios de evaluación; tuvieron poca interacción personal con los docentes; afrontaron instancias de evaluación muy cercanas en el tiempo; y no recibieron devoluciones oportunas y efectivas. Estas experiencias fueron inversas en los estudiantes avanzados. En cuanto a los procesos cognitivos realizados durante la exposición a los instrumentos de evaluación, las diferencias entre principiantes y avanzados estuvieron relacionadas a su nivel de formación en la carrera, los primeros aplicaron principalmente procesos mentales de “orden inferior” (Bloom) como recordar, comprender y aplicar; y los avanzados, operaciones mentales de los niveles cognitivos de “orden superior” como analizar, evaluar y crear.

- **¿Qué puede aportar cada momento de la evaluación al aprendizaje de los estudiantes?**

En la fase de preparación para las evaluaciones, tanto los estudiantes principiantes como los avanzados, hicieron algunos ajustes o cambios en sus comportamientos de estudio (habilidades, hábitos y actitudes) para afrontar los desafíos que les planteaban los sistemas de evaluación multimétodo. Entre las habilidades de estudio, las técnicas más utilizadas fueron: la “relectura de los materiales” de estudio, con la intención de lograr una mayor comprensión a través de la exposición repetida al mismo texto; la “elaboración de resúmenes” de cada material, para identificar la información más relevante y contar con un documento de repaso; y la “realización de esquemas, gráficos y dibujos”, para mejorar la comprensión, representando y relacionando ideas y conceptos relevantes.

Como hábitos de estudio, algunos estudiantes de ambos casos, destacaron la importancia que tuvo para su aprendizaje, estudiar en grupos o con algún/a compañero/a, ya que las discusiones entre pares, la construcción de conocimiento compartido y la resolución conjunta de problemas, estimulan la reflexión y contribuyen al aprendizaje colaborativo. Aunque para otros, la gestión del tiempo de estudio fue una gran dificultad por las exigencias y la intensidad del cursado en ambos casos de estudio.

En cuanto a la elección de los materiales de estudio, las guías y los materiales didácticos elaborados por los docentes de MHS fueron los preferidos por los alumnos principiantes; mientras que las fuentes de información más utilizadas por los estudiantes avanzados fueron las búsquedas bibliográficas en internet. Por otro lado, los dos grupos de estudiantes valoraron muy bien la estrategia de revisar trabajos y exámenes de cohortes anteriores, pues analizando la redacción de los ensayos y las respuestas a las preguntas de examen, pudieron lograr un mejor desempeño en las evaluaciones.

Entre los factores que motivaron el estudio, las “exigencias del sistema de evaluación” fueron para algunos alumnos principiantes una gran fuerza motivadora para preparar los exámenes estudiando en profundidad. En tanto que los estudiantes avanzados destacaron el efecto motivador que tuvo el mecanismo de “promoción meritoria” utilizado por MED-III que los estimuló a estudiar en forma continua y minuciosa para alcanzar el estatus de “estudiante destacado” que es uno de los criterios para lograr la promoción.

Como factores que desanimaron el estudio, la proximidad temporal entre las evaluaciones constituyó para muchos alumnos principiantes una gran sobrecarga de trabajo que desalentó el ritmo de estudio; un impacto similar, sufrieron los estudiantes avanzados relacionado a la intensidad del cursado y a las evaluaciones simultáneas que fueron factores estresores muy difíciles de controlar. Otros alumnos principiantes se sintieron desanimados con las evaluaciones grupales porque fueron integrantes de grupos disfuncionales. Además, ambos grupos de estudiantes consideraron como factor de desánimo para el estudio a la utilización de instrumentos de evaluación con defectos en su diseño.

Entre los factores que facilitaron el buen rendimiento durante los exámenes, la diversidad de los instrumentos de evaluación aplicados fue para los alumnos principiantes un factor importante; mientras que, para los avanzados, las cualidades particulares de algunos instrumentos facilitaron el buen rendimiento. En contraposición, como factores que obstaculizaron el buen rendimiento, los principiantes señalaron las fallas en el diseño de

algunos instrumentos utilizados; y los avanzados, las fallas en la aplicación y corrección de ciertas evaluaciones.

En relación a las percepciones que tuvieron los estudiantes sobre su capacidad y sus habilidades intelectuales para afrontar las evaluaciones, los principiantes expresaron que su formación en el nivel medio no fue suficiente o adecuada para las exigencias de la universidad, en tanto que los avanzados percibieron con incertidumbre ciertas limitaciones en sus habilidades intelectuales personales para afrontar las evaluaciones. En cuanto a los riesgos percibidos en el sistema de evaluación, la mayoría de los alumnos principiantes expresaron que el riesgo más importante era perder la regularidad en MHS y no poder ingresar a la carrera de medicina; mientras que a los avanzados les preocupaba no lograr la “promoción meritoria” en MED-III y tener dificultades para continuar con la POF.

Por último, entre los beneficios que aportó el sistema de evaluación aplicado al aprendizaje de la materia, los alumnos principiantes reconocieron como más significativos: ampliar y profundizar los conocimientos de MHS, identificar sus necesidades de aprendizaje, y desarrollar habilidades para el trabajo equipo; en tanto que los estudiantes avanzados, además de ampliar y profundizar los conocimientos de la materia, lograron desarrollar habilidades metacognitivas para reflexionar sobre su aprendizaje.

**Como respuesta a la pregunta se puede concluir que**, los tres momentos de la evaluación tuvieron efectos de aprendizaje. En la fase de pre-evaluación, los estudiantes adecuaron sus comportamientos de estudio para mejorar su preparación, valorando el aprendizaje grupal colaborativo y los beneficios de revisar exámenes de otras cohortes; los principales motivadores para el estudio fueron: “las exigencias del sistema de evaluación de MHS” y la posibilidad de “promoción meritoria” en MED-III; y como factores de desánimo, el escaso tiempo entre las pruebas (MHS) y el cursado intenso con evaluaciones simultáneas (MED-III). Durante las evaluaciones, facilitaron el buen rendimiento: la diversidad de instrumentos aplicados y las cualidades de algunas herramientas; y obstaculizaron el buen rendimiento, fallas en el diseño de los instrumentos y, en la aplicación y corrección de ciertas evaluaciones; algunos alumnos tenían autoconcepto académico “negativo” autoatribuido a la formación inadecuada en el nivel medio (principiantes) y a ciertas limitaciones intelectuales personales (avanzados); y como principales riesgos percibidos, no ingresar a la carrera de medicina y no poder cursar la POF. Entre los beneficios que aportó el sistema de evaluación al aprendizaje de la materia, reconocieron como más significativos: ampliar y

profundizar los conocimientos, identificar las necesidades de aprendizaje, y desarrollar habilidades para el trabajo equipo y para reflexionar sobre el aprendizaje.

## 6.2. Implicaciones de la investigación para la práctica docente

En este punto se abordarán las posibles implicaciones que pueden tener los resultados de esta investigación en los procesos de formación de grado de los profesionales de la salud, retomando desde la perspectiva de la “evaluación como aprendizaje” los principales factores identificados en este trabajo de tesis que pueden impactar en forma positiva o negativa sobre los efectos de aprendizaje de las evaluaciones.

Dichos factores fueron agrupados en tres categorías según los efectos que tuvieron sobre el aprendizaje en los casos estudiados: A. Factores que promovieron el aprendizaje; B. Factores que dificultaron el aprendizaje; y C. Factores con efectos ambivalentes sobre el aprendizaje. La Figura 6.1. ofrece una representación gráfica de las interacciones observadas entre los diferentes factores.

### A. Factores que promovieron el aprendizaje

- **Estudio en grupos o en parejas pedagógicas** (aprendizaje colaborativo)

Esta estrategia de estudio que aplicaron los dos grupos de estudiantes (principiantes y avanzados), puede ser muy valiosa para el aprendizaje en cualquier nivel de la carrera, porque a través de las discusiones conceptuales, la resolución conjunta de problemas y la construcción de conocimiento compartido, se convierte el trabajo grupal en aprendizaje (Duran, 2015), y se estimula la reflexión y la regulación metacognitiva. (De Backer et al, 2015)

- **Revisión de Trabajos y Exámenes de otras cohortes**

La utilización de trabajos, ensayos y/o exámenes anteriores (observada en los dos casos estudiados) es un hábito de estudio muy común en los estudiantes de la salud (Zárate-Depraect et al, 2018), y constituye una estrategia que puede mejorar el aprendizaje de los alumnos y su rendimiento en pruebas posteriores. Este efecto de aprendizaje mejorado mediante pruebas o “Test-enhanced learning”, tiene su explicación en las ciencias cognitivas a través de un mecanismo de exposición repetida a claves que facilitan el acceso y la recuperación de información relevante. (de Bruin et al, 2019; Biwer et al, 2020)

- **La diversidad de instrumentos aplicados en la evaluación**

Esta característica (común en los sistemas de evaluación multimétodo) ha sido reconocida también por otros autores (Schut et al, 2018; Preston et al, 2020), como un factor

que promueve el aprendizaje, ya que la variedad de instrumentos, permite a los estudiantes aplicar en las evaluaciones sus diferentes estilos de aprendizaje y sus enfoques personales. Por otro lado, reduce los riesgos de la evaluación “única” y da a los alumnos una mayor sensación de tener el control, y más oportunidades para mostrar y mejorar su desempeño.

- **La utilización de instrumentos de evaluación de buena calidad**

Esta cualidad, particularmente destacada por los estudiantes avanzados, es coincidente con las opiniones de algunos autores (Schut et al, 2018; Preston et al, 2020) que sostienen que los alumnos estudian con mayor motivación e intensidad cuando perciben que los instrumentos con los que van a ser evaluados son adecuados para valorar lo que quieren demostrar a sus profesores sobre sus conocimientos o habilidades. Igualmente, cuando los alumnos anticipan ciertos formatos de examen, estudian más cuidadosamente (Epstein, 2007). Por otro lado, cuando las intenciones de los estudiantes (que quieren demostrar) coinciden con los propósitos de los docentes (que quieren evaluar), los efectos de aprendizaje de las evaluaciones son muy significativos. (Watling & Ginsburg, 2019)

- **Las tareas de aplicación o práctica “clínica” comprendidas en la evaluación**

La inclusión en las evaluaciones de viñetas clínicas y problemas para resolver, y la observación del desempeño clínico, fueron muy bien valoradas por los estudiantes para su aprendizaje. Este efecto de aprendizaje de las evaluaciones con enfoque “clínico” ha sido observado por otros autores tanto en el ciclo básico como en el ciclo clínico. En el primer nivel, la integración de las ciencias básicas con contenidos clínicos puede fortalecer las asociaciones entre los principios biomédicos subyacentes y las manifestaciones de enfermedad, y permitir a los estudiantes una mayor comprensión de los contenidos básicos (Kulasegaram et al, 2013); mientras los estudiantes avanzados pueden reconocer patrones y desarrollar “script” de enfermedad a través de la asociación, mejorar la velocidad y precisión del procesamiento cognitivo, y desarrollar progresivamente habilidades más complejas de razonamiento. (Bowen, 2006; de Bruin et al, 2019)

- **El reconocimiento al buen desempeño académico**

Los estudiantes avanzados destacaron el efecto motivador para el estudio que tuvo el mecanismo de “promoción meritosa” en MED-III. Es un reconocimiento que reciben los estudiantes “destacados”, que consiste en otorgarles su promoción sin rendir el examen final de la materia. Los criterios para ser estudiante “destacado” son: acreditar un 90% o más de asistencia a todas las actividades del cursado; haber rendido todas las evaluaciones; y lograr

un puntaje total de 800 puntos o más (equivalente a nota 8). Por consiguiente, reconocimientos académicos de este tipo pueden ser potentes factores motivadores que promueven el aprendizaje.

- **Los riesgos percibidos en el sistema de evaluación**

En ambos casos de estudio hubo riesgos percibidos. En los alumnos principiantes el temor a perder la regularidad en MHS y no poder ingresar a la carrera de Medicina, y en los avanzados la posibilidad de no lograr la promoción meritatoria en MED-III. Estos riesgos y posibles consecuencias, constituyen “lo que está en juego” en la evaluación, que los estudiantes pueden percibir como un “obstáculo” o como una “oportunidad” para el aprendizaje (Schut et al, 2018). Por lo tanto, cuando los riesgos son percibidos como oportunidad, constituyen una gran fuerza motivadora para estudiar y preparar los exámenes y muy útiles para impulsar el aprendizaje. (Preston et al, 2020)

## **B. Factores que dificultaron el aprendizaje**

- **Contexto de cursado masificado**

Esta dificultad fue identificada por los alumnos principiantes que habían atravesado una experiencia muy compleja de cursado en la asignatura MHS debido al elevado número de inscriptos, con baja relación docente/alumno, grandes comisiones de estudiantes y amplia utilización de la virtualidad. En este contexto, disminuyen las posibilidades de interactuar en forma personal con los docentes en los escenarios de aprendizaje y evaluación, las actividades de supervisión docente se ven limitadas al igual que las instancias de retroalimentación, haciendo muy dificultosa y poco efectiva la evaluación para el aprendizaje. (Schut et al, 2018)

- **Desconocimiento de los objetivos y criterios de evaluación**

Este es otro aspecto que señalaron los alumnos principiantes y en gran medida, relacionado también a la masificación. Desconocer los objetivos de aprendizaje y los criterios de evaluación es “navegar a ciegas”, por ello, uno de los criterios clave para mejorar los efectos de aprendizaje de las evaluaciones es clarificar y compartir con los estudiantes las intenciones de aprendizaje y los criterios para el éxito (Black & Wiliam, 2009); y es un elemento esencial en todo programa formativo cuando el propósito de las evaluaciones está enfocado en “evaluar para el aprendizaje”. (Yin et al, 2008; Wiliam & Thompson, 2007)

- **La proximidad temporal entre las evaluaciones**

Esta situación fue reconocida por los alumnos principiantes como una dificultad para el aprendizaje, ya que tuvieron que afrontar diferentes evaluaciones de alto riesgo con intervalos de tiempo muy cortos entre las mismas. Cuando una evaluación es inminente la percepción de “falta de tiempo para estudiar” es, según Cilliers et al (2012), un efecto inevitable, los estudiantes perciben que no saben lo suficiente e intentan memorizar tantos datos como pueden para aprobar el examen con muy poca comprensión y análisis, pudiendo afectar el desempeño en la prueba y disminuir los efectos de aprendizaje. Por ello, algunos autores sugieren mantener baja la carga de exámenes en el ciclo clínico, y dar tiempo a los estudiantes para consolidar el conocimiento y prepararse para la práctica profesional. (Preston et al, 2020)

- **La sobrecarga de tareas académicas y el estrés del cursado**

Estos aspectos fueron identificados como factores que desanimaron el estudio en los dos grupos de estudiantes y pueden ser vinculados porque las evaluaciones y la carga de trabajo académico muy intensas son causas frecuentes de estrés y ansiedad en los estudiantes (Hashimat et al, 2008; Cabanach et al, 2016; Zárate-Depraect et al, 2018). Dichas emociones son importantes en la evaluación ya que pueden afectar el rendimiento de la memoria de trabajo, dificultando el aprendizaje, y eso ocurre principalmente cuando los alumnos perciben el estrés como una amenaza y no como un desafío. (de Bruin et al, 2019)

- **Fallas en el diseño o construcción de los instrumentos**

Estos defectos fueron reconocidos por los alumnos principiantes como obstáculos para el buen rendimiento en las pruebas y por ende pueden dificultar el aprendizaje en las evaluaciones. Estas observaciones coinciden con las efectuadas por Preston et al (2020) donde los alumnos percibieron dificultades de aprendizaje relacionadas con la calidad de las evaluaciones, como defectos en su diseño, poca claridad de las consignas, escasa relevancia de las tareas solicitadas y pobre estandarización de los instrumentos. Por lo tanto, cuando el propósito es impulsar el aprendizaje a través de la evaluación, es esencial utilizar herramientas de buena calidad. (Norcini et al., 2018).

- **Desaciertos en la aplicación y/o corrección de las evaluaciones**

Estos desaciertos que percibieron algunos estudiantes avanzados fueron reconocidos como obstáculos para el buen rendimiento en los exámenes, pero al mismo tiempo, al afectar la validez de los juicios de evaluación pueden dificultar los efectos de

aprendizaje de las evaluaciones. Por estas razones, es muy importante la cuidadosa planificación, aplicación y corrección de las evaluaciones, pues mejora la calidad de los juicios de evaluación que son muy valiosos y necesarios para tomar decisiones válidas con respecto a los estudiantes en el ciclo clínico de una carrera. (Lockyer et al, 2017; Gruppen et al, 2018)

- **Autoconcepto académico “negativo”**

Esta autopercepción estuvo presente en algunos alumnos principiantes y avanzados, en los primeros fue atribuida a su formación inadecuada en el nivel medio y en los segundos a ciertas limitaciones intelectuales personales. Pero más allá de estas diferencias, un autoconcepto académico “negativo” puede reducir las expectativas y motivaciones de los estudiantes, disminuyendo su capacidad objetiva para las tareas y en consecuencia, afectar el rendimiento académico y dificultar el aprendizaje. (Yeung et al, 2014)

### **C. Factores con efectos ambivalentes sobre el aprendizaje**

- **Las evaluaciones grupales**

Para algunos alumnos principiantes las evaluaciones grupales desanimaron el estudio; para otros, trabajar en grupo con diferentes perspectivas en un objetivo común, fue una experiencia muy valiosa. Este contraste de opiniones fue atribuido a la conformación aleatoria de los grupos que permitió la presencia de grupos disfuncionales muy heterogéneos, que tuvieron muchos conflictos y bajo rendimiento. Sin embargo, más allá de estas diferencias, los resultados de las evaluaciones grupales fueron los puntajes más altos de los alumnos principiantes, lo que sugiere que las dificultades no están en el “instrumento” sino en el “grupo”.

- **El mecanismo de “retroalimentación”**

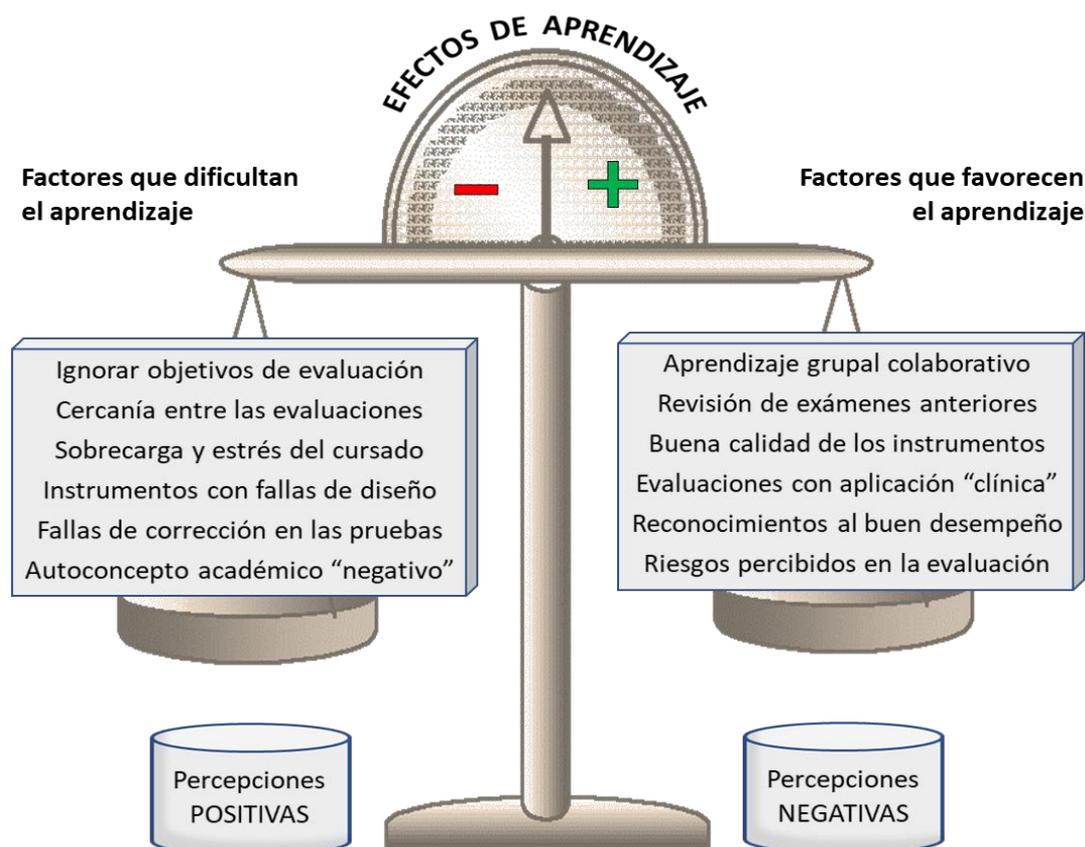
Las diferencias observadas entre los casos de estudio en sus mecanismos de retroalimentación guardan relación con los propósitos de evaluación de cada uno (sumativo en MHS y formativo-sumativo en MED-III). Sin embargo, también hubo diferencias en las percepciones de algunos estudiantes sobre las devoluciones que recibieron en la misma asignatura, lo que refleja la disparidad que puede existir entre las “imágenes” que construye cada estudiante del mismo fenómeno (Perera et al, 2008; Bartlett et al, 2017; Contzen et al, 2018), ya que para algunos estudiantes las devoluciones fueron muy valiosas para su aprendizaje y otros las percibieron como críticas punitivas poco constructivas. En consecuencia, los efectos de aprendizaje vinculados a la retroalimentación pueden ser más

o menos significativos según la percepción que tenga cada estudiante de dicho proceso. (Wood, 2019; Watling & Ginsburg, 2019; Tuma & Nassar, 2022)

- **Las percepciones sobre el sistema de evaluación**

Como se puede ver en los resultados de esta investigación, las percepciones de los estudiantes sobre las evaluaciones fueron muy diferentes y hasta contrapuestas, porque cada uno percibió (valoró) el mismo proceso desde sus propias expectativas e intenciones (Harrison & Wass, 2016; Watling & Ginsburg, 2019). Por lo tanto, teniendo en cuenta que el impacto de cualquier sistema de evaluación sobre el aprendizaje está mediado por las percepciones de cada estudiante (Schut et al, 2018), estas percepciones (positivas y negativas) funcionan más como “limitadoras” que como “facilitadoras o inhibidoras” de los efectos de aprendizaje que operan dentro del sistema de evaluación (Figura 6.1.)

**Figura 6.1.:** Interacción de los factores que tienen efectos sobre el aprendizaje



Nótese que las percepciones funcionan como “limitadoras” de los efectos de aprendizaje

Elaboración propia

### 6.3. Limitaciones de la investigación

Reconocer las limitaciones de una investigación, es una forma de delinear con mayor precisión los alcances que tienen sus resultados, y ello requiere una profunda reflexión sobre las acciones y decisiones tomadas antes y durante el proceso investigativo para identificar sus debilidades.

En tal sentido, la primera debilidad identificada fue la imposibilidad de efectuar los dos “grupos focales” con los estudiantes de ambos casos de estudio, que en el proyecto de investigación estaban previstos para la última etapa del trabajo empírico de acuerdo a las recomendaciones de Stalmeijer et al (2014). Estos grupos focales estaban dirigidos a indagar los cambios en los comportamientos de estudio y las percepciones sobre los entornos de aprendizaje y evaluación con el propósito de explorar con mayor profundidad y enriquecer la información obtenida anteriormente con los otros instrumentos. Los grupos de estudiantes se constituyeron a fines de 2019 y por cuestiones inherentes al calendario académico de exámenes, las reuniones fueron diferidas para marzo 2020, el aislamiento social obligatorio y la suspensión de las actividades presenciales en el contexto de la pandemia COVID-19, invalidaron las posibilidades de realizar las entrevistas grupales, y en su defecto, se optó por efectuar “entrevistas electrónicas” (Fontana & Frey, 2015:186) a través de un cuestionario online con preguntas abiertas.

Otro aspecto que puede considerarse como debilidad, es la ausencia de las voces de los docentes en esta investigación, pues, aunque sus intenciones y acciones pedagógicas pudieron ser desprendidas de los documentos y materiales curriculares analizados, sus percepciones y reflexiones sobre el cursado de las asignaturas seguramente hubieran brindado una idea más amplia y compleja de los entornos de aprendizaje-evaluación y el desempeño de los estudiantes en los casos de estudio.

Finalmente, si bien el enfoque epistemológico y metodológico adoptado para este trabajo de tesis permitió comprender los factores y efectos de aprendizaje de los sistemas de evaluación multimétodo, quedan muchos interrogantes sobre los mecanismos que intervienen en el impacto educacional de las evaluaciones, que abren nuevas líneas de investigación para proyectos futuros.





## BIBLIOGRAFÍA





## BIBLIOGRAFIA

### A

- Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association*. 103(3), 152-153. <https://doi.org/10.3163/1536-5050.103.3.010>
- Addinsoft (2021). XLSTAT statistical and data analysis solution. New York, USA. <https://www.xlstat.com>
- Aguis, N. M., & Wilkinson, A. (2014). Students' and Teachers' Views of Written Feedback at Undergraduate Level: A Literature Review. *Nurse Education Today* 34: 552–559 [doi:10.1016/j.nedt.2013.07.005](https://doi.org/10.1016/j.nedt.2013.07.005)
- Ainsworth, S., Prain, V., & Tyler, R. (2011). Drawing to Learn in Science. *Science*. 333(6046):1096-1097. <https://doi.org/10.1126/science.1204153>.
- Al-Kadri, H. M., Al-Kadi, M. T., & Van Der Vleuten, C. P. (2013). Workplace-based assessment and students' approaches to learning: a qualitative inquiry. *Medical Teacher*, 35 Suppl 1, S31–S38. <https://doi.org/10.3109/0142159X.2013.765547>
- Al-Kadri, H.M., Al-Moamary, M.S., Roberts, C., & van der Vleuten, C.P.M. (2012). Exploring assessment factors contributing to students' study strategies: Literature Review. *Medical Teacher*, 34:S42–S50.
- Alves de Lima, A.E. (2008). Devolución Constructiva. Una estrategia para mejorar el aprendizaje. *MEDICINA (Buenos Aires)*; 68: 88-92
- Archer, J.C. (2010). State of the science in health professional education: Effective feedback. *Medical Education*; 44:101-108.
- Artino, A.R., Dong, T., Dezee, K.J., Gilliland, W.R., Waechter, D.M., Cruess, D., & Durning, S.J. (2012). Achievement goal structures and self-regulated learning: Relationships and changes in medical school. *Academic Medicine*; 87:1375–81.

### B

- Barnett, R. (2001). *Los límites de la competencia. El conocimiento, la educación superior y la sociedad*. Barcelona: Gedisa.
- Bartlett, M., Crossley, J., & McKinley, R. (2017). Improving the Quality of Written Feedback using Written Feedback. *Education for Primary Care* 28: (1): 16–22. [doi:10.1080/14739879.2016.1217171](https://doi.org/10.1080/14739879.2016.1217171).

- Bericat Alastuey E. (1998). *La integración de los métodos cuantitativo y cualitativo en la investigación social. Significado y medida.* Barcelona. Editorial Ariel SA. ISBN: 84-344-1693-X
- Bierer, S. B., Dannefer, E. F., Taylor, C., Hall, P., & Hull, A. L. (2008). Methods to assess students' acquisition, application and integration of basic science knowledge in an innovative competency-based curriculum. *Medical Teacher*, 30(7), e171–e177. <https://doi.org/10.1080/01421590802139740>
- Biggs, J. B. (1988). Assessing study approaches to learning. *Australian Psychologist*, 23, 197-206. doi: 10.1080/00050068808255604
- Biwer, F., de Bruin, A. B., Schreurs, S., & oude Egbrink, M. G. (2020). Future steps in teaching desirably difficult learning strategies: Reflections from the Study Smart program. *Journal of Applied Research in Memory and Cognition*, 9(4), 439-446.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21 (1), 5–31.
- Bonillo, A. (2013). Análisis de los Items. Universitat Oberta de Catalunya. PID: 00198631. *Psicometría* [http://openaccess.uoc.edu/webapps/o2/bitstream/10609/69325/7/Psicometr%C3%ADa\\_M%C3%B3dulo%205\\_An%C3%A1lisis%20de%20los%20C3%ADtems.pdf](http://openaccess.uoc.edu/webapps/o2/bitstream/10609/69325/7/Psicometr%C3%ADa_M%C3%B3dulo%205_An%C3%A1lisis%20de%20los%20C3%ADtems.pdf).
- Bowen, J.L. (2006). Educational Strategies to Promote Clinical Diagnostic Reasoning. *New England Journal of Medicine*; 355:2217-2225. DOI: 10.1056/NEJMra054782
- Bowen, L., Marshall, M., & Murdoch-Eaton, D. (2017). Medical Student Perceptions of Feedback and Feedback Behaviors Within the Context of the "Educational Alliance". *Academic Medicine*, 92(9), 1303–1312. //doi.org/10.1097/ACM.0000000000001632
- Brailovsky, C.A. (2001). Educación Médica, Evaluación de las competencias. En *Aportes para un Cambio Curricular en Argentina 2001.* OPS y Facultad de Medicina, UBA; 103-122.
- Bunniss, S., & Kelly, D. R., (2010). Research paradigms in medical education research. *Medical Education* 44: 358–366. doi:10.1111/j.1365-2923.2009.03611.x
- Burr, S.A., Brodier, E., & Wilkinson, S. (2013). Delivery and use of individualised feedback in large class medical teaching. *BMC Medical Education* 2013; 13: 63.
- Bzowycykj, A. S., Dow, A., & Knab, M. S. (2017). Evaluating the Impact of Educational Interventions on Patients and Communities: A Conceptual Framework. *Academic Medicine*, 92(11), 1531–1535. <https://doi.org/10.1097/ACM.0000000000001718>

## C

- Cabanach, R. G., Souto-Gestal, A., & Franco, V. (2016). Escala de Estresores Académicos para la evaluación de los estresores académicos en estudiantes universitarios. *Revista Iberoamericana de Psicología y Salud* 7, 41-50. <http://dx.doi.org/10.1016/j.rips.2016.05.001>
- Carreras Barnés, J., Branda, L.A., Castro Salomó, A., Fenoll Brunet, M.R., Gual Sala, A., Mahy Géhenne, J.N., Martínez Carretero, J.M., Nolla Domenjó, M., Palés Argullós, J., Pérez Sánchez, J., Portero, M., & Roma Millan, J. (2009). *Guía para la evaluación de competencias en medicina*. Barcelona. Agència per a la Qualitat del Sistema Universitari de Catalunya.
- Centeno, A., Primogero, C., & Llull, L. (2007). The process of learning during an examination. *Medical Education*, 41(6), 619. <https://doi.org/10.1111/j.1365-2923.2007.02773.x>
- Charmaz, K. (2008) *Reconstructing Grounded Theory*. En Alasuutari, P., Bickman, L. & Brannen, J. (edit) *The SAGE handbook of social research methods* (pp. 461-478). London: SAGE Publications Ltd.
- Charmaz, K. (2014). *Constructing Grounded Theory* (2a. ed.). California: SAGE Publications Inc.
- Cilliers, F.J., Schuwirth, L.W.T, Adendorff, H.J., Herman, N., & van der Vleuten, C.P.M. (2010). The mechanism of impact of summative assessment on medical students' learning. *Advances in Health Sciences Education. Theory and Practice*; 15(5); 695-715.
- Cilliers, F. J., Schuwirth, L. W., Herman, N., Adendorff, H. J., & van der Vleuten, C. P. (2012a). A model of the pre-assessment learning effects of summative assessment in medical education. *Advances in Health Sciences Education: Theory and Practice*, 17(1), 39–53. <https://doi.org/10.1007/s10459-011-9292-5>
- Cilliers, F. J., Schuwirth, L. W., & van der Vleuten, C. P. (2012b). A model of the pre-assessment learning effects of assessment is operational in an undergraduate clinical context. *BMC medical education*, 12, 9. <https://doi.org/10.1186/1472-6920-12-9>
- Cilliers, F. J., Schuwirth, L. W. T., & van der Vleuten, C. P. M. (2012) Modelling the pre-assessment learning effects of assessment: evidence in the validity chain. *Medical Education*; 46: 1087–1098. doi:10.1111/j.1365-2923.2012.04334.x
- Cleland, J., Leggett, H., Sandars, J., Costa, M.J., Patel, R., & Moffat, M. (2013). The remediation challenge: theoretical and methodological insights from a systematic review. *Medical Education*. 47:242–251.
- Cobb, K. A., Brown, G., Jaarsma, D. A., & Hammond, R. A. (2013). The educational impact of assessment: A comparison of DOPS and MCQs. *Medical Teacher*, 35, 1598–1607.

- Contzen, M. P., Parra, P., & Ortega, J. (2018). Rol docente-estudiante e Implementación del Feedback en Educación Médica. *Rev. Educación en Ciencias de la Salud*; 15(1): 35-38
- Cook, D. A., Castillo, R. M., Gas, B. & Artino, A. R. (2017). Measuring achievement goal motivation, mindsets and cognitive load: validation of three instruments' scores. *Medical Education*, 51: 1061-1074. doi: 10.1111/medu.13405
- Credé, M., & Kuncel, N. R. (2008). Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance. *Perspectives on Psychological Science*, 3(6), 425-453. DOI: 10.1111/j.1745-6924.2008.00089.x
- Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ quality & safety*, 22 Suppl 2(Suppl 2), ii58–ii64. <https://doi.org/10.1136/bmjqs-2012-001712>

## D

- Dannefer E. F. (2013). Beyond assessment of learning toward assessment for learning: educating tomorrow's physicians. *Medical Teacher*, 35(7), 560–563. <https://doi.org/10.3109/0142159X.2013.787141>
- De Backer, L., Van Keer, H., & Valcke, M. (2015). Exploring evolutions in reciprocal peer tutoring groups' socially shared metacognitive regulation and identifying its metacognitive correlates. *Learning and Instruction*, 38, 63-78. doi: <https://doi.org/10.1016/j.learninstruc.2015.04.001>
- de Bruin, A.B.H., Sibbald, M., Monteiro, S. (2019) The Science of Learning. En Swanwick T, Forrest K, O'Brien BC. *Understanding Medical Education. Evidence, Theory and Practice*. Third Edition. Oxford, UK; John Wiley & Sons, Inc. p 23-36
- Di Bernardo, J. J., Navarro, V., Fernández, M. G., Demuth Mercado, P. B., & Larroza, G. O. (2017). Medicina, Hombre y Sociedad: adecuando el ingreso a medicina al modelo de competencias y al perfil de graduación. *Revista de la Facultad de Medicina de la UNNE*, XXXVII (3): 5-14
- Di Bernardo, J. J., Palma, D., & Popescu, B. (2009). Sistema de Créditos y Evaluación Longitudinal Multimétodo. Su impacto sobre el desempeño académico de los estudiantes en una asignatura medicina. XVIII Conferencia Panamericana de Educación Médica y X CAEM. Universidad Austral - Pilar - Buenos Aires, octubre 2009
- Di Bernardo, J. J., Popescu, B., Navarro, V., Larroza, G. O., & Bluvstein, S. (2014). Curriculum Complementario: una estrategia para articular e integrar un curriculum basado en asignaturas. *Revista de la Facultad de Medicina de la UNNE*, Vol. XXXIV (2):15-26.

- Díaz-Plasencia, J. A., Gamarra-Sánchez, J. E., & Valencia-Mariñas, H. (2016). Impacto educativo de una evaluación basada en el desempeño en residentes de medicina del Programa de Segunda Especialización de la Universidad Privada Antenor Orrego, Trujillo, Perú. *FEM: Revista de la Fundación Educación Médica*, 19(3), 135-145.
- Donnon, T., & Hecker, K. (2010). A Model of Approaches to Learning and Academic Achievement of Students from an Inquiry Based Bachelor of Health Sciences Program. *Canadian Journal of Higher Education*, 38(1), 1-19.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14, 4–58. <https://doi.org/10.1177/1529100612453266>
- Duran, D. (2015). Tutoría entre iguales: compartir la capacidad de enseñar con los alumnos. En Mayordomo, R. y Onrubia, J. (coords.). *El aprendizaje cooperativo*. Barcelona: Editorial UOC.
- Durante, E. (2005). La evaluación de los conocimientos: Lo que parece ser, ¿es realmente lo que es?. *Revista del Hospital Italiano de Buenos Aires* 25 (1):18-23 [https://www.hospitalitaliano.org.ar/multimedia/archivos/noticias\\_attachs/47/documentos/10469\\_18-23.Durante.pdf](https://www.hospitalitaliano.org.ar/multimedia/archivos/noticias_attachs/47/documentos/10469_18-23.Durante.pdf)

## E

- Ebel, R. L. (1982). Proposed solutions to two problems of test construction. *Journal of Educational Measurement*, 19, 267–278.
- Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, 100, 613–628.
- Embo, M. P., Driessen, E. W., Valcke, M., & van der Vleuten, C. P. (2010). Assessment and feedback to facilitate self-directed learning in clinical practice of Midwifery students. *Medical Teacher*; 32 (7): e263-9. doi: 10.3109/0142159X.2010.490281.
- Enríquez, M., Fajardo, M., & Garzón, F. (2015). Una revisión general a los hábitos y técnicas de estudio en el ámbito universitario. *Psicogente*, 18(33), 166-187. <http://doi.org/10.17081/psico.18.33.64>
- Epstein, R. M., & Hundert, E. M. (2002). Defining and Assessing Professional Competence. *JAMA*; 287(2):226-35.
- Epstein, R. M. (2007). Assessment in Medical Education. *New England Journal of Medicine*, 356, 387-396. <https://doi.org/10.1056/NEJMra054784>

- Eva, K. W., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K., & Sargeant, J. (2012). Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. *Advances in Health Sciences Education: Theory and Practice*, 17(1), 15–26. <https://doi.org/10.1007/s10459-011-9290-7>
- Evans, J. S. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspectives on Psychological Science*. 8 (3): 223–241.

## F

- Farmer, E. A. & Page, G. (2005). A practical guide to assessing clinical decision-making skills using the key features approach. *Medical Education* 39: 1188–1194
- Ferguson, P. C., Caverzagie, K. J., Nousiainen, M. T., Snell, L., & ICBME Col (2017). Changing the culture of medical training: An important step toward the implementation of competency-based medical education. *Medical Teacher*, 39(6), 599–602.
- Ferrando-Castagnetto, F., Macri, E., Silva, A., Padula, D., & Garcés, G. (2019). Percepciones estudiantiles sobre las propiedades del examen clínico objetivo estructurado en Uruguay. Disponiendo un cuestionario válido para analizar la factibilidad del nuevo formato durante la transición curricular. *Educación Médica*, 20, 87-94.
- Ferreira, M. D. L. D. S. (2021). Avaliação no processo ensino-aprendizagem: uma experiência vivenciada. *Revista Brasileira de Educação Médica*, 27, 12-19.
- Fontana, A. & Frey, J.H. (2015). La entrevista. En Denzin, N.K. & Lincoln, Y.S. (comps) Vol IV Métodos de recolección y análisis de datos (pp. 140-202). Manual SAGE de investigación cualitativa. Barcelona: Gedisa.
- Fournier, J.P., Demeester, A., & Charlin, B. (2008). Script concordance tests: guidelines for construction. *BMC Medical Informatics and Decision Making*, 8:18
- Frank, J. R., Mungroo, R., Ahmad, Y., Wang, M., De Rossi, S., & Horsley, T. (2010b). Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Medical Teacher* 32: 631–637
- Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., Harris, P., Glasgow, N. J., Campbell, C., Dath, D., Harden, R. M., Iobst, W., Long, D. M., Mungroo, R., Richardson, D. L., Sherbino, J., Silver, I., Taber, S., Talbot, M., & Harris, K. A. (2010a). Competency-based medical education: theory to practice. *Medical Teacher*, 32(8):638-645.
- Freiberg Hoffmann, A. & Fernández Liporace, M. M. (2016). Enfoques de aprendizaje en universitarios argentinos según R-SPQ-2F: Análisis de sus propiedades psicométricas. *Revista Colombiana de Psicología*, 25(2), 307-329. doi: 10.15446/rcp.v25n2.51874

## G

- Garrote, D., Garrote, C., & Jiménez, S. (2016). Factores influyentes en motivación y estrategias de aprendizaje en los alumnos de grado. REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 14(2) 31-44
- Gibbs, G. & Simpson, C. (2005). Conditions Under Which Assessment Supports Students' Learning. Learning and Teaching in Higher Education (1) 3-31.
- Gruppen, L. D., ten Cate, O., Lingard, L. A., Teunissen, P. W. & Kogan, J. R. (2018). Enhanced Requirements for Assessment in a Competency-Based, Time-Variable Medical Education System. Academic Medicine; 93(3): S17-S21. doi: 10.1097/ACM.0000000000002066
- Guba, E. G. & Lincoln, Y. S. (2015). Controversias paradigmáticas, contradicciones y confluencias emergentes. En Denzin, N. K. & Lincoln, Y. S. (comps) Vol II Paradigmas y perspectivas en disputa (pp. 38-78). Manual SAGE de investigación cualitativa. Barcelona: Gedisa.

## H

- Haladyna, T. M., Downing, S. M. & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. Applied Measurement in Education; 15: 309-334
- Harden, R. M., & Laidlaw, J. M. (2013). Be FAIR to students: four principles that lead to more effective learning. Medical Teacher; 35 (1): 27-31.
- Harden, R. M., Crosby, J. R., & Davis, M. H. (1999). An introduction to outcome-based education, Medical Teacher 21(1): 7-14
- Harrison, C. J., Könings, K. D., Dannefer, E. F., Schuwirth, L. W., Wass, V., & van der Vleuten, C. P. (2016). Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. Perspectives on Medical Education, 5(5), 276–284. <https://doi.org/10.1007/s40037-016-0297-x>
- Harrison, C. J., Könings, K. D., Schuwirth, L., Wass, V., & van der Vleuten, C. (2017). Changing the culture of assessment: the dominance of the summative assessment paradigm. BMC Medical Education, 17(1), 73. <https://doi.org/10.1186/s12909-017-0912-5>
- Harrison, C., Wass, V. (2016). The challenge of changing to an assessment for learning culture. Medical Education; 50 (7): 704–706.

- Harrison, C. J., Könings, K. D., Molyneux, A., Schuwirth, L. W., Wass, V., & van der Vleuten, C. P. (2013). Web-based feedback after summative assessment: how do students engage? *Medical Education*; 47 (7): 734-44. doi:10.1111/medu.12209.
- Hashmat, S., Hashmat, M., Amanullah, F., & Aziz, S. (2008). Factors causing exam anxiety in medical students. *Journal-Pakistan Medical Association*, 58(4), 167-170.
- Heeneman, S., Oudkerk Pool, A., Schuwirth, L. W., van der Vleuten, C. P., & Driessen, E. W. (2015). The impact of programmatic assessment on student learning: theory versus practice. *Medical Education*, 49(5), 487–498. <https://doi.org/10.1111/medu.12645>
- Heijne-Penninga, M., Kuks, J. B. M., Hofman, W. H. A., & Cohen-Schotanus, J. (2010). Influences of deep learning, need for cognition and preparation time on open- and closed-book test performance. *Medical Education*, 44(9), 884–891. <https://doi.org/10.1111/j.1365-2923.2010.03732.x>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación* (6ta edición). México D.F. McGraw Hill Education. ISBN: 978-1-4562-2396-0
- Hodges B. (2013). Assessment in the post-psychometric era: learning to love the subjective and collective. *Medical Teacher*, 35(7), 564–568. <https://doi.org/10.3109/0142159X.2013.789134>
- Hoffman, L. A., Shew, R. L., Vu, T. R., Brokaw, J. J., & Frankel. R.M. (2016). Is Reflective Ability Associated With Professionalism Lapses During Medical School? *Academic Medicine* 91 (6): 853–857. doi:10.1097/ACM.0000000000001094.
- Holmboe, E. S., Sherbino, J., Englander, R., Snell, L., Frank, J. R., & ICBME Collaborators (2017). A call to action: The controversy of and rationale for competency-based medical education. *Medical Teacher*, 39(6), 574–581. <https://doi.org/10.1080/0142159X.2017.1315067>

## K

- Kool, A., Mainhard, T., Brekelmans, M., van Beukelen, P., & Jaarsma, D. (2016). Goal orientations of health profession students throughout the undergraduate program: a multilevel study. *BMC Medical Education*, 16, 100. <https://doi.org/10.1186/s12909-016-0621-5>
- Kordestani Moghaddam, A., Khankeh, H. R., Shariati, M., Norcini, J., & Jalili, M. (2019). Educational impact of assessment on medical students' learning at Tehran University of Medical Sciences: a qualitative study. *BMJ open*, 9(7), e031014. <https://doi.org/10.1136/bmjopen-2019-031014>

Kulasegaram, K. M., Martimianakis, M. A., Mylopoulos, M., Whitehead, C. R., & Woods, N. N. (2013). Cognition before curriculum: rethinking the integration of basic science and clinical learning. *Academic Medicine* 88(10), 1578–1585. <https://doi.org/10.1097/ACM.0b013e3182a45def>

## L

Larsen, D.P., Butler, A.C., Lawson, A.L., & Roediger, H.L., 3rd (2013). The importance of seeing the patient: test-enhanced learning with standardized patients and written tests improves clinical application of knowledge. *Advances in Health Sciences Education: Theory and Practice*, 18(3):409-25. <https://doi.org/10.1007/s10459-012-9379-7>

Lefroy, J., Watling, C., Teunissen, P. W., & Brand, P. (2015). Guidelines: the do's, don'ts and don't knows of feedback for clinical education. *Perspectives on medical education*, 4(6), 284–299. <https://doi.org/10.1007/s40037-015-0231-7>

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior research methods*, 45(4), 1058–1072. <https://doi.org/10.3758/s13428-013-0334-1>

Leppink, J. & Duvivier, R. (2016). Twelve tips for medical curriculum design from a cognitive load theory perspective, *Medical Teacher*, 38:7, 669-674, DOI:10.3109/0142159X.2015.1132829

Loai Albarqouni, L., Hoffmann, T., & Glasziou, P. (2018). Evidence-based practice educational intervention studies: a systematic review of what is taught and how it is measured. *BMC Medical Education*; 18:177 <https://doi.org/10.1186/s12909-018-1284-1>

Lockyer, J., Carraccio, C., Chan, M.K., Hart D, Smee, S., Touchie, C., Holmboe, E.S., Frank, J.R., & ICBME Collaborators. (2017) Core principles of assessment in competency-based medical education. *Med Teach*; 39:609–616. doi: 10.1080/0142159X.2017.1315082.

López Cabrera, M. V., & Silva López, Y. E. (2017). Evaluación y supervisión en Educación Médica. En S.L. Olivares y J.E. Valdez García (Ed.) *Aprendizaje Centrado en el Paciente: cuatro perspectivas para un abordaje integral*. (pp. 242-261). Ciudad de México. México: Editorial Medica Panamericana.

Lörwald, A. C., Lahner, F. M., Greif, R., Berendonk, C., Norcini, J. & Huwendiek, S. (2017): Factors influencing the educational impact of Mini-CEX and DOPS: A qualitative synthesis, *Medical Teacher*, DOI: 10.1080/0142159X.2017.1408901

Lynam, S., & Cachia, M. (2018). Students' perceptions of the role of assessments at higher education. *Assessment & Evaluation in Higher Education*, 43(2), 223-234.

## M

- Maia, M. V., & Struchiner, M. (2016). Aprendizagem significativa e o portfólio reflexivo eletrônico na educação médica. *Revista Brasileira de Educação Médica*, 40, 720-730.
- Mann, K., van der Vleuten, C., Eva, K., Armson, H., Chesluk, B., Dornan, T., Holmboe, E., Lockyer, J., Loney, E., & Sargeant, J. (2011). Tensions in informed self-assessment: how the desire for feedback and reticence to collect and use it can conflict. *Academic Medicine*, 86(9), 1120–1127. <https://doi.org/10.1097/ACM.0b013e318226abdd>
- McLachlan, J. C. (2006). The relationship between assessment and learning. *Medical Education*, 40, 716–717.
- McManus, I. C., Richards, P., Winder, B. C., & Sproston, K. A. (1998). Clinical experience, performance in final examinations, and learning style in medical students: Prospective study. *British Medical Journal*, 316, 345–350.
- Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine (Supplement)* 65: S63-S7
- Miller, A. & Archer, J. (2010). Impact of workplace based assessment on doctors' education and performance: a systematic review. *BMJ Online*; 341:c5064 doi:10.1136/bmj.c5064
- Moller, A., & Elliot, A. J. (2006). The 2 x 2 achievement goal framework. An overview of empirical research. *Educational Psychological*, 307- 326
- Monroe, K. S. (2016). The relationship between assessment methods and self-directed learning readiness in medical education. *International Journal of Medical Education*; 7:75-80. doi: 10.5116/ijme.56bd.b282.
- Moore, D. E., Green, J. S., & Gallis, H. A. (2009). Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. *The Journal of Continuing Education in the Health Professions*, 29(1), 1–15. <https://doi.org/10.1002/chp.20001>
- Moore, P., Leighton, M. I.; Alvarado, C.; & Brali, C. (2016). Pacientes simulados: el lado humano de la simulación. *Revista Médica de Chile* 2016; 144: 617-625
- Morales Castillo, J. D., & Varela Ruiz, M. (2015). El debate en torno al concepto de competencias. *Investigación en Educación Médica*, 4(13), 36-41.
- Morales, P. (2012). Análisis de ítems en las pruebas objetivas. Madrid: Universidad Pontificia Comillas En <http://educrea.cl/wp-content/uploads/2014/11/19-nov-analisis-de-items-en-las-pruebas-objetivas.pdf>. Acceso el 02/10/2014.

- Morchio, I. & Difabio, H. (2019). Autopercepción de competencia para aprender en estudiantes universitarios. Precisiones conceptuales y resultados empíricos. *Psicopedagógica* 11(14) 32-76
- Morcke, A. M., Dornan, T., & Eika, B. (2013). Outcome (competency) based education: an exploration of its origins, theoretical basis, and empirical evidence. *Advances in Health Sciences Education. Theory and Practice*, 18(4), 851-863.
- Motola, I., Devine, L.A., Chung, H.S., Sullivan, J.E., & Issenberg, S.B. (2013). Simulation in healthcare education: a best evidence practical guide. *AMEE Guide N° 82. Medical Teacher*; 35 (10): e1511-30.
- Mulder, M., Weigel, T., & Collins, K. (2007) The concept of competence in the development of vocational education and training in selected EU member status: a critical análisis. *Journal of Vocational Education & Training*, 59 (1), 67-88.

## N

- Newble, D. (2004). Techniques for measuring clinical competence: objective structured clinical examinations. *Medical Education*, 38: 199-203.
- Newble, D. I. & Jaeger, K. (1983). The effect of assessments and examinations on the learning of medical students. *Medical Education*; 17:165-71.
- Nolla-Domenjó, M. (2009) La evaluación en educación médica: Principios básicos. *Educación Médica*, 12(4), 223-229.
- Norcini, J., Anderson, B., Bollela, V., Burch, V., Costa, M. J., Duvivier, R., Galbraith, R., Hays, R., Kent, A., Perrott, V., & Roberts, T. (2011). Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher*, 33(3), 206–214. <https://doi.org/10.3109/0142159X.2011.551559>
- Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher*, 29(9), 855–871. <https://doi.org/10.1080/01421590701775453>
- Norcini, J., Anderson, M. B., Bollela, V., Burch, V., Costa, M. J., Duvivier, R., Hays, R., Palacios Mackay, M. F., Roberts, T., & Swanson, D. (2018). 2018 Consensus framework for good assessment. *Medical Teacher*, 40(11), 1102–1109.
- Norcini, J.J. (2005). The mini-CEX. *The Clinical Teacher* 2(1): 25-30.
- Norman, G., Neville, A., Blake, J.M., & Mueller, B. (2010). Assessment steers learning down the right road: Impact of progress testing on licensing examination performance. *Medical Teacher*, 32:496–499.

## P

- Palés-Argullós J. (2016). Bolonia en las Facultades de Medicina españolas, seis años después: algunas preguntas para la reflexión. *FEM-Fundación Educación Médica*, 19(3), 109-111
- Palés-Argullós, J. (2010). ¿Cómo elaborar correctamente preguntas de elección múltiple?. *Educación Médica*, 13 (3): 149-155
- Paniagua, M.A., & Swygert, K. A. (Editores). (2016). *Cómo elaborar preguntas para evaluaciones escritas en las áreas de ciencias básicas y clínicas*. Philadelphia. National Board of Medical Examiners (NBME). Cuarta Edición. En [https://www.nbme.org/sites/default/files/2020-01/DownloadingtheGoldBook\\_ES.pdf](https://www.nbme.org/sites/default/files/2020-01/DownloadingtheGoldBook_ES.pdf).
- Pelgrim, E. A., Kramer, A. W., Mokkink, H. G., & van der Vleuten, C. P. (2013). Reflection as a component of formative assessment appears to be instrumental in promoting the use of feedback; an observational study. *Medical Teacher*. 35(9):772-778. DOI: 10.3109/0142159X.2013.801939
- Peräkylä, A. (2015). Análisis de la conversación y el texto. En Denzin, N.K. & Lincoln, Y.S. (comps) Vol IV Métodos de recolección y análisis de datos (pp. 462-483). Manual SAGE de investigación cualitativa. Barcelona: Gedisa.
- Perera, J., Lee, N., Win, K., Perera, J., & Wijesuriya, L. (2008). Formative feedback to students: the mismatch between faculty perceptions and student expectations. *Medical Teacher*. 30(4): 395-399.
- Pérez, G., Kattan, E., Collins, L., Wright, A. C., Rybertt, T., González, A., Sirhan, M., Solís, N., Pizarro, M., Arrese, M., Sarfatis, A., Lustig, N., Arab, J.P., Labarca, J., & Riquelme, A. (2015). Evaluación para el aprendizaje: experiencia en un curso teórico de pregrado en medicina. *Revista médica de Chile*, 143(3), 329-336. <https://dx.doi.org/10.4067/S0034-98872015000300007>
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review*, 16(4), 385-407.
- Pippitt, K. A., Moore, K. B., Lindsley, J. E., Cariello, P. F., Smith, A. G., Formosa, T., Moser, K., Morton, D. A., Colbert-Getz, J. M., & Chow, C. J. (2022). Assessment for Learning with Ungraded and Graded Assessments. *Medical science educator*, 32(5), 1045–1054. <https://doi.org/10.1007/s40670-022-01612-y>
- Planas-Coll, J. (2013). El contrasentido de la enseñanza basada en competencias. *Revista Iberoamericana de Educación Superior*, 4(10) 75-92. <https://ries.universia.net/article/view/117>

Preston, R., Gratani, M., Owens, K., Roche, P., Zimanyi, M., & Malau-Aduli, B. (2020) Exploring the Impact of Assessment on Medical Students' Learning. *Assessment & Evaluation in Higher Education*; 45(1):109–124 <https://doi.org/10.1080/02602938.2019.1614145>

## R

Raupach, T., Andresen, J. C., Meyer, K., Strobel, L., Koziol, M., Jung, W., Brown, J., & Anders, S. (2016). Test-enhanced learning of clinical reasoning: a crossover randomised trial. *Medical Education*, 50(7), 711–720. <https://doi.org/10.1111/medu.13069>

Roberts, C., Khanna, P., Bleasel, J., Lane, S., Burgess, A., Charles, K., Howard, R., O'Mara, D., Haq, I., & Rutzou, T. (2022). Student perspectives on programmatic assessment in a large medical programme: A critical realist analysis. *Medical Education*, 56(9), 901–914. <https://doi.org/10.1111/medu.14807>

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*; 24(2): 3-13.

Ruiz de Gauna, P.; González Moroa, V.; & Morán-Barrios, J. (2015). Diez claves pedagógicas para promover buenas prácticas en la formación médica basada en competencias en el grado y en la especialización. *Educación Médica*; 16(1):34-42

## S

Sambell, K., & McDowell, L. (1998). The construction of the hidden curriculum: messages and meanings in the assessment of student learning. *Assessment & Evaluation in Higher Education*, 23(4), 391-402.

Sánchez Rosas, J. (2015). Validation of the Achievement Goal Questionnaire-Revised in Argentinean university students (A-AGQ-R). *International Journal of Psychological Research*, 8 (1): 10-23

Sánchez-Mendiola, M. (2015). La educación médica basada en competencias: ¿santo remedio o vino viejo en nueva botella? *Investigación en Educación Médica* 4(13):1-2

Sanders, J. (2009). The use of reflection in medical education: AMEE Guide No. 44. *Medical Teacher*; 31: 685-95.

Schuelper, N., Ludwig, S., Anders, S., & Raupach, T. (2019). The Impact of Medical Students' Individual Teaching Format Choice on the Learning Outcome Related to Clinical Reasoning. *JMIR Medical Education*, 5(2), e13386. <https://doi.org/10.2196/13386>

- Schut, S., Driessen, E., van Tartwijk, J., van der Vleuten, C. & Heeneman, S. (2018). Stakes in the eye of the beholder: an international study of learners' perceptions within programmatic assessment. *Medical Education*; 52: 654-663. doi: 10.1111/medu.13532
- Schuwirth, L. W. T., & van der Vleuten, C. P. M. (2004). Different written assessment methods: what can be said about their strengths and weaknesses? *Medical Education* 38: 974-979
- Schuwirth, L.W.T., & van der Vleuten C.P.M. (2011a). General overview of the theories used in assessment: AMEE Guide No. 57. *Medical Teacher*, 33:783-797
- Schuwirth, L.W.T., & Van der Vleuten, C.P.M. (2011b). Programmatic assessment: from assessment of learning to assessment for learning. *Medical Teacher*, 33(6), 478–485. doi: <http://dx.doi.org/10.3109/0142159X.2011.565828>.
- Schuwirth, L., & Ash, J. (2013). Assessing tomorrow's learners: in competency-based education only a radically different holistic method of assessment will work. Six things we could not forget. *Medical Teacher*; 35 (7):555–9.
- Schuwirth, L. W. T., Van der Vleuten, C. P. M., & Durning, S. J. (2017a). What programmatic assessment for learning in medical education can learn from healthcare. *Perspectives Medical Education*, 6(4): 211-215. DOI: 10.1007/s40037-017-0345-1
- Schuwirth, L., Valentine, N., & Dilena, P. (2017b). An application of programmatic assessment for learning (PAL) system for general practice training. *GMS Journal for Medical Education*, 34(5), Doc56. <http://doi.org/10.3205/zma001133>
- Segers, M., & Dochy, F. J. R. C. (2006). Introduction Enhancing student learning through assessment: Alignment between levels of assessment and different effects on learning. *Studies in Educational Evaluation*, 32, 171–179.
- Segers, M., Nijhuis, J., & Gijsselaers, W. (2006) Redesigning a learning and assessment environment: the influence on students' perceptions of assessment demands and their learning strategies. *Studies in Educational Evaluation*; 32 (3):223–42.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Shumway, J. M., Harden, R. M., & Association for Medical Education in Europe (2003). AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Medical teacher*, 25(6), 569–584. <https://doi.org/10.1080/0142159032000151907>

Stalmeijer, R. E., McNaughton, N., & Van Mook, W. N. (2014). Using focus groups in medical education research: AMEE Guide N°91. *Medical Teacher*, 36:1-17  
DOI:10.3109/0142159X.2014.917165

Stegers-Jager, K. M., Cohen-Schotanus, J., & Themmen, A. P. (2012). Motivation, learning strategies, participation and medical school performance. *Medical education*, 46(7), 678–688. <https://doi.org/10.1111/j.1365-2923.2012.04284.x>

## T

Tapia Castillo, F. (2017). Metas de logro, estrategias de aprendizaje y rendimiento académico en alumnos universitarios. [Tesis de Licenciatura]. Pontificia Universidad Católica del Perú: Lima. <http://hdl.handle.net/20.500.12404/9961>

Telio, S., Ajjawi, R., & Regehr, G. (2015). The "educational alliance" as a framework for reconceptualizing feedback in medical education. *Academic Medicine*, 90(5), 609–614. <https://doi.org/10.1097/ACM.0000000000000560>

Ten Cate, O. (2013). Nuts and bolts of entrustable professional activities. *Journal of Graduate Medical Education*, 5(1), 157–158. <https://doi.org/10.4300/JGME-D-12-00380.1>

Thrien, C., Fabry, G., Härtl, A., Kiessling, C., Graupe, T., Preusche, I., Pruskil, S., Schnabel, K. P., Sennekamp, M., Rüttermann, S., & Wünsch, A. (2020). Feedback in medical education - a workshop report with practical examples and recommendations. *GMS Journal for Medical Education*, 37(5), Doc46. <https://doi.org/10.3205/zma001339>

Tuma, F., & Nassar, A. K. (2022). Feedback in Medical Education. In StatPearls. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK544311/>

## U

Uygun, J., Stuart, E., De Paor, M., Wallace, E., Duffy, S., O'Shea, M., Smith, S. & Pawlikowska, T. (2019) A Best Evidence in Medical Education systematic review to determine the most effective teaching methods that develop reflection in medical students: BEME Guide No. 51, *Medical Teacher*, 41:1, 3-16, DOI: 10.1080/0142159X.2018.1505037

## V

Valle Arias, A., Gonzalez Cabanach, R., Nuñez Perez, J.C., & Gonzalez-Pienda, J.C. (1998). Variables cognitivo-motivacionales, enfoques de aprendizaje y rendimiento académico. *Psicothema*, 10 (2): 393-412

- van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J., & Van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher* 34: 205-214
- van der Vleuten, C. P. M. (2016) Revisiting “Assessing professional competence: from methods to programmes”. *Medical Education* 50: 885–888
- van der Vleuten, C. P. M. & Heeneman, S. (2016). A new holistic way of assessment: programmatic assessment. *FEM Revista de la Fundación Educación Médica*,19 (6): 275-279
- van der Vleuten, C., Sluijsmans, D., & Joosten-ten Brinke, D. (2017). Competence Assessment as Learner Support in Education. In: Mulder M. (eds) *Competence-based Vocational and Professional Education. Technical and Vocational Education and Training: Issues, Concerns and Prospects*, vol 23. Springer, Cham. [https://doi.org/10.1007/978-3-319-41713-4\\_28](https://doi.org/10.1007/978-3-319-41713-4_28)
- Van Merriënboer, J. J. G. & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education*, 44, 85–93.
- Van Tartwijk, J., & Driessen, E.W. (2009). Portfolios for assessment and learning. *AMEE Guide No.45. Medical Teacher*, 31(9): 790-801.
- Van Wynsberghe, R., & Khan, S. (2007). Redefining Case Study. *International Journal of Qualitative Methods*, 6(2): 80-94

## W

- Wass, V., van der Vleuten, C. P. M., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *The Lancet*, 357, 945-949
- Watling, C. J. & Lingard, L. (2012). Grounded theory in medical education research: *AMEE Guide No. 70. Medical Teacher*, 34(10): 850-861.
- Watling, C.J., & Ginsburg, S. (2019). Assessment, feedback and the alchemy of learning. *Medical Education*; 53: 76–85. doi: 10.1111/medu.13645
- Wen, C.C., Lin, M.J., Lin, C.W., & Chu, S.Y. (2015). Exploratory study of the characteristics of feedback in the reflective dialogue group given to medical students in a clinical clerkship. *Medical Education Online*. 20:25965. DOI: 10.3402/meo.v20.25965. PMID: 25661500; PMCID: PMC4320997.
- William, D., & Thompson, M. (2007). Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.) *The future of assessment: shaping teaching and learning* (pp. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Wilson, K. & Fowler, J. (2005). Assessing the impact of learning environments on students' approaches to learning: comparing conventional and action learning designs. *Assessment & Evaluation in Higher Education* 30 (1):87-101

Wood, D. F. (2019). *Formative Assessment: Assessment for Learning*. En Swanwick T, Forrest K, O'Brien BC. *Understanding Medical Education. Evidence, Theory and Practice*. Third Edition. Oxford, UK; John Wiley & Sons, Inc. p 361-373

## Y

Yan, V. X., Thai, K. P., & Bjork, R. A. (2014). Habits and beliefs that guide self-regulated learning: Do they vary with mindset?. *Journal of Applied Research in Memory and Cognition*, 3(3), 140-152.

Yazan, B. (2015). Three Approaches to Case Study Methods in Education: Yin, Merriam, and Stake. *The Qualitative Report*, 20(2), 134-152. <http://nsuworks.nova.edu/tqr/vol20/iss2/12>

Yeung, A. S., Li, B., Wilson, I., & Craven, R. G. (2014). The role of self-concept in medical education. *Journal of Further and Higher Education*, 38(6), 794-812. DOI: 10.1080/0309877X.2013.765944

Yin, R.K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, California: Sage Publications. ISBN 7619-2553-8

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Tomita, M. K., & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, 21(4), 335-359.

Young, J. Q., Van Merriënboer, J., Durning, S., & Ten Cate, O. (2014). Cognitive Load Theory: implications for medical education: AMEE Guide No. 86. *Medical teacher*, 36(5), 371–384. <https://doi.org/10.3109/0142159X.2014.889290>

## Z

Zárate-Depraect, N.E., Soto-Decuir, M.G., Martínez-Aguirre, E.G., Castro-Castro, M.L., García-Jau, R.A., & López-Leyva, N.M. (2018). Hábitos de estudio y estrés en estudiantes del área de la salud. *FEM Revista de la Fundación Educación Médica*; 21 (3): 153-157

Zhao, C., & Hu, Y. (2021). Reflections on Study Strategy Modifications Using Cognitive Load Theory and Dual Processing Theory in the First Year of Medical School. *Medical Science Educator* 31:813–818



**ANEXOS**





## ANEXO 1.1: Instrumento A (1)

### Cuestionario de Metas de Logro de Elliot y Murayama (2008)

Versión **A-AGQ-R**<sup>42</sup>, cuestionario revisado, traducido y validado para estudiantes universitarios argentinos por Sánchez Rosas (2015).

El cuestionario comprende doce (12) proposiciones sobre las cuales los estudiantes deben expresar su grado de acuerdo con cada ítem en una escala de 1 (muy en desacuerdo) a 5 (totalmente de acuerdo):

1. Mi objetivo es llegar a dominar completamente el material presentado en esta materia.
2. Mi propósito es que me vaya bien comparado con otros alumnos.
3. Mi meta es aprender tanto como sea posible.
4. Mi objetivo es desempeñarme bien con respecto a otros alumnos.
5. Mi objetivo es evitar aprender menos de lo que podría.
6. Mi meta es evitar desempeñarme mal comparado con otros.
7. Mi propósito es comprender el contenido de esta asignatura tan completamente como sea posible.
8. Mi meta es desempeñarme mejor que los otros alumnos.
9. Mi meta es evitar aprender menos de lo que es posible aprender.
10. Mi propósito es evitar desempeñarme peor que otros.
11. Mi propósito es evitar una comprensión incompleta del material de la asignatura.
12. Mi objetivo es evitar que me vaya peor que a otros alumnos.

#### Análisis de los resultados:

Los ítems que representan a cada categoría de metas son los siguientes:

Metas de aproximación-aprendizaje	(ítems 1, 3, 7)
Metas de evitación-aprendizaje	(ítems 5, 9, 11)
Metas de aproximación-rendimiento	(ítems 2, 4, 8)
Metas de evitación-rendimiento	(ítems 6, 10, 12)

Para determinar el grado de inclinación de los estudiantes por cada categoría, se calculan en cada una, las medidas de tendencia central (media y mediana), DE y rango.

---

<sup>42</sup> The Argentinean Achievement Goal Questionnaire – Revised. (Sánchez Rosas, 2015)



## ANEXO 1.2: Instrumento A (2)

### Enfoques de Aprendizaje: modelo de Biggs et al (2001)<sup>43</sup>

#### Cuestionario Revisado de Procesos de Estudio, Dos Factores (R-SPQ-2F)

Validado y adaptado para estudiantes universitarios argentinos por  
Freiberg Hoffmann & Fernández Liporace (2016).

El cuestionario comprende 20 proposiciones sobre las cuales los estudiantes deben expresar su grado de acuerdo con cada ítem en una escala de 1 (muy en desacuerdo) a 5 (totalmente de acuerdo):

1. Tengo momentos en los que estudiar me produce gran satisfacción.
2. Debo estudiar bastante un tema para poder formar mis propias conclusiones y así quedar conforme.
3. Mi objetivo es aprobar la materia haciendo el menor trabajo posible.
4. Solo estudio lo que se ve en clase.
5. Me parecen muy interesantes todos los temas cuando los empiezo a estudiar.
6. La mayoría de los temas nuevos me parecen interesantes y con frecuencia dedico tiempo extra a conseguir más información sobre ellos.
7. Dedico un mínimo esfuerzo al estudio de las materias que no me interesan.
8. Aprendo algunas cosas mecánicamente, repitiéndolas una y otra vez hasta que ya las sepa de memoria aunque no las comprenda.
9. Me parece que estudiar algunos temas académicos puede ser tan interesante como leer una buena novela o ver una buena película.
10. Cuando estudio temas importantes me autoevalúo tanto como sea necesario hasta lograr comprenderlos por completo.
11. Puedo aprobar la mayoría de los exámenes memorizando partes claves de una materia sin intentar comprenderlas.
12. Generalmente me limito a estudiar solo lo que se me pide, porque creo innecesario hacer cosas extra.
13. Trabajo duro en mis estudios cuando creo que el material o el contenido son interesantes.
14. Dedico gran parte de mi tiempo libre a buscar más información sobre temas interesantes que fueron expuestos en clase.

---

<sup>43</sup> Biggs, J.B., Kember, D., & Leung, D.Y. (2001) The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*; 71:133-149.

15. Creo que no es conveniente estudiar los temas a fondo. Eso produce confusión y pérdida de tiempo, cuando lo único que hace falta es conocer los temas por encima para poder aprobarlos.
16. Creo que los profesores no deberían esperar que los estudiantes dedicaran mucho tiempo a estudiar los contenidos que se sabe que no van a evaluarse en el examen.
17. Voy a la mayoría de las clases con preguntas en mente, que quiero responderme.
18. Me esfuerzo por leer la mayor parte de los materiales que me recomiendan en clase.
19. No entiendo por qué tengo que aprender contenidos que no se exigen en el examen.
20. La mejor forma de aprobar los exámenes es memorizar las respuestas de las preguntas más probables.

#### **Análisis de los resultados:**

El R-SPQ-2F fue diseñado para evaluar principalmente dos dimensiones de Enfoques de Aprendizaje: “Profundo” (ítems 1, 2, 5, 6, 9, 10, 13, 14, 17, 18) y “Superficial” (ítems 3, 4, 7, 8, 11, 12, 15, 16, 19, 20). De los 10 ítems que componen cada dimensión, cinco representan a las “estrategias” y cinco a los “motivos”, así surgen las subescalas: “estrategia profunda” (ítems 2, 6, 10, 14, 18), estrategia superficial (ítems 4, 8, 12, 16, 20), motivo profundo (ítems 1, 5, 9, 13, 17) y motivo superficial (ítems 3, 7, 11, 15, 19). Pero dado que los análisis psicométricos realizados por Biggs et al (2001) no permitieron confirmar a estas subescalas como dimensiones, solo se utilizan las dos dimensiones principales (profunda y superficial)

Para establecer en los estudiantes el tipo enfoque de aprendizaje aplicado y su intensidad, se utiliza la sumatoria de los puntajes consignados en las dos dimensiones principales y se calcula la diferencia de puntos entre ambas.

La dimensión con mayor puntaje indica el enfoque de aprendizaje dominante, que puede ser profundo, superficial o neutro; y la diferencia de puntajes, marca la intensidad, cuando es entre 1 y 13 es “Baja”, entre 14 y 26 es “Media” y entre 27 y 40 es “Alta”.

### **ANEXO 1.3: Instrumento A (3)**

#### **Exploración de los efectos de las evaluaciones**

##### **Cuestionario sobre el Sistema de Evaluación de MHS**

El cuestionario comprende ocho proposiciones sobre las cuales los estudiantes deben expresar su grado de acuerdo en una escala de 1 (nada de acuerdo) a 5 (muy de acuerdo); tres preguntas de varias opciones; y una pregunta abierta:

1. Para los exámenes tuve que estudiar varios materiales que no había leído para las clases.
2. Preparando y rindiendo los exámenes siento que aprendí más que cursando la materia.
3. En las evaluaciones pude comprender conceptos que no entendí durante el cursado.
4. Las evaluaciones NO me motivaron a estudiar más de lo que estudio habitualmente.
5. Gracias a las evaluaciones pude relacionar y articular mejor los contenidos de MHS.
6. En los exámenes pude aplicar mejor los conocimientos de MHS a situaciones prácticas.
7. A través de las evaluaciones NO he aprendido más de lo que aprendo en las clases.
8. A través de las evaluaciones pude identificar mejor lo que todavía me falta aprender.
9. ¿Qué tipo de proceso mental utilizaste con mayor frecuencia para rendir las pruebas?
  - a) Solo reconocer o recordar conceptos teóricos
  - b) Identificar o explicar mecanismos y procesos
  - c) Razonar y aplicar lo aprendido a casos o problemas
10. ¿Qué tipo de devolución recibiste de los docentes después de rendir las pruebas?
  - a) Solo la calificación obtenida
  - b) Pude ver las respuestas correctas
  - c) Me explicaron los aciertos y errores
11. Indica cuál de estos instrumentos de evaluación aportó MAS a tu aprendizaje de MHS
  - a) Las evaluaciones grupales (TT y Transversales)
  - b) Las pruebas integradoras de ejes (AL-AF y SIS-AM)
  - c) El examen integrador de Tramo (multiple choice)
  - d) Los cinco instrumentos combinados
  - e) NINGUNA de estas pruebas aportó a mi aprendizaje
12. ¿Cuál es tu percepción o valoración sobre el sistema de evaluación aplicado en MHS?

### Cuestionario sobre el Sistema de Evaluación de MED-III

El cuestionario comprende nueve proposiciones sobre las cuales los estudiantes deben expresar su grado de acuerdo en una escala de 1 (nada de acuerdo) a 5 (muy de acuerdo); tres preguntas de varias opciones; y una pregunta abierta:

1. Para los exámenes tuve que estudiar varios materiales que no había leído para las clases.
2. Preparando y rindiendo los exámenes siento que aprendí más que cursando la materia.
3. En las evaluaciones pude comprender conceptos que no entendí durante el cursado.
4. Las evaluaciones NO me motivaron a estudiar más de lo que estudio habitualmente.
5. En las evaluaciones pude relacionar mejor mis conocimientos de Cs. básicas con la clínica.
6. En los exámenes pude aplicar mejor los conocimientos a situaciones de la práctica.
7. A través de las evaluaciones NO he aprendido más de lo que aprendo en las clases.
8. A través de las evaluaciones pude identificar mejor lo que todavía me falta aprender.
9. Las múltiples evaluaciones orientan más el aprendizaje que los exámenes aislados.
10. ¿Qué tipo de proceso mental utilizaste con mayor frecuencia para rendir las pruebas?
  - a) Solo reconocer o recordar conceptos teóricos
  - b) Identificar o explicar mecanismos y procesos
  - c) Razonar y aplicar lo aprendido a casos o problemas
11. ¿Qué tipo de devolución recibiste de los docentes después de rendir las pruebas?
  - a) Solo la calificación obtenida
  - b) Pude ver las respuestas correctas
  - c) Me explicaron los aciertos y errores
- 12.Cuál de estos instrumentos de evaluación aportó MAS a tu aprendizaje de Medicina III
  - a) La elaboración de la Historia Clínica y Trabajo bibliográfico
  - b) Los mini-CEX aplicados en los Trabajos Prácticos
  - c) Los exámenes parciales escritos (multiple choice)
  - d) El portfolio y los cafecitos en el aula virtual
  - e) TODOS los instrumentos combinados
  - f) NINGUNA de estas pruebas aportó a mi aprendizaje
13. ¿Cuál es tu percepción o valoración sobre el sistema de evaluación de Medicina III?

## ANEXO 2: Instrumento B

### Cuestionario abierto para las entrevistas electrónicas

Este cuestionario consta de ocho (8) preguntas abiertas dirigidas a indagar los efectos de las evaluaciones sobre el aprendizaje de los estudiantes, sus hábitos, habilidades y actitudes de estudio, los cambios que adoptaron para afrontar los sistemas de evaluación, los factores que motivaron y desanimaron el estudio, y sus experiencias y percepciones en las evaluaciones aplicadas en los casos analizados.

1. ¿Qué rol tuvo la evaluación en tu proceso de aprendizaje? ¿Fue un componente esencial para aprender o solo un instrumento para comprobar lo que aprendiste?; puedes dejarnos tu impresión:
2. ¿Qué aspectos o situaciones (de los exámenes o del cursado) te motivaron a estudiar para las evaluaciones? y ¿Cuáles te desanimaron para el estudio?; explícanos lo que ocurrió:
3. ¿Qué efectos tuvieron las evaluaciones sobre tus comportamientos de estudio?; ¿Estudiaste de igual forma e intensidad durante el cursado que para las evaluaciones?; ¿Hiciste algunos cambios en tus técnicas de estudio?; puedes comentar tus experiencias:
4. ¿Qué aspectos o elementos de las evaluaciones te facilitaron lograr buenos resultados en las pruebas? y ¿Cuáles obstaculizaron o afectaron tu rendimiento?; explícanos lo que ocurrió:
5. ¿Percibiste en algún momento del cursado que tus habilidades de comprensión y/o razonamiento no eran suficientes o adecuadas para afrontar las evaluaciones?; Si esto te pasó, puedes comentar tu experiencia:
6. ¿Qué consecuencias de riesgo (académicas o personales) te preocupaban más si no lograbas obtener buenos resultados en las evaluaciones?; puedes mencionarlas:
7. En la encuesta aplicada al finalizar el cursado, algunos de tus compañeros estuvieron de acuerdo en que las múltiples evaluaciones orientaron más su aprendizaje que los exámenes aislados. ¿Qué opinas al respecto, se te ocurre alguna explicación?
8. Puedes comentarnos otras experiencias o dejarnos algún comentario sobre el sistema de evaluación que aplicó la asignatura.

