

Una propuesta de autoevaluación-reflexión para potenciar la responsabilidad individual

V. Javier Traver Juan Carlos Amengual
Dep. Llenguatges i Sistemes Informàtics
Universitat Jaume I, 12071-Castellón
[vtraver, jcamen]@uji.es

Resumen

Pese a sus potenciales beneficios, la evaluación continua puede desvirtuarse en la práctica. Por ejemplo, se puede tender a enfatizar el producto (la entrega que realiza el estudiante) sobre el proceso (el aprendizaje del estudiante). En este trabajo se propone un esquema de autoevaluación y, sobre todo, de reflexión, cuyo objetivo es, no sólo mejorar el aprendizaje, sino potenciar el sentido de responsabilidad y honestidad en los estudiantes. En esencia, se trata de una autoevaluación sumativa en la que el resultado de la autoevaluación está condicionado al resultado de una prueba verificadora posterior. Este esquema se ha ensayado de modo preliminar en una actividad de una asignatura del primer curso de un grado en ingeniería informática. Dado el limitado ámbito de este estudio, no es posible asegurar los beneficios que se desprenden de la propuesta. Sin embargo, los indicios recopilados sugieren que una aplicación continuada de este esquema podría resultar fructífera a medio-largo plazo. Además, la combinación de autoevaluación y prueba verificadora permitiría la detección de casos diferentes y, en consecuencia, facilitaría la atención a la diversidad. Creemos pues que la idea es prometedora y susceptible de mejora en cursos venideros.

Abstract

Deliverable-based regular assessment of learning holds promise but hides pitfalls. As an example, the product (a deliverable) may be emphasized over the process (the learning itself). This work explores a reflection-based self-assessment scheme aimed at promoting the student's sense of responsibility where the result of a summative self-assessment is subject to a subsequent validation test. Although still limited, the evidence suggests that a consistent application of this mechanism would lead to good results. Furthermore, this combination of self-assessment plus a validation test might allow detecting students with different needs

and, consequently, better attention to diversity. We believe the idea is promising and can be improved in the coming years.

Palabras clave

Evaluación sumativa, evaluación continua, autoevaluación, prueba verificadora, responsabilidad individual

1. Introducción

En los últimos años se ha podido constatar un uso creciente de la evaluación continua en los estudios universitarios, en parte debido a la implantación de los créditos *European Credit Transfer System* (ECTS) y al fomento de las metodologías centradas en el estudiante. Sin embargo, la forma en que se plantea o desarrolla la evaluación continua puede resultar en ocasiones cuestionable o mejorable. El esquema típico, basado en entregables que los profesores corrigen y puntúan presenta, a nuestro entender, las siguientes deficiencias:

❶ **Coste excesivo.** Las entregas pueden suponer un esfuerzo desmesurado de corrección a los profesores, sobre todo en grupos numerosos [3], o de realización a los propios estudiantes, con el riesgo de la sobre-evaluación [5].

❷ **Dudosa utilidad.** Las notas en sí mismas no necesariamente ayudan al estudiante y, de hecho, los estudiantes leen el feedback con mayor atención cuando no hay calificaciones [2].

❸ **Se prima el producto sobre el proceso.** Debido a cierta inercia, se tiende a considerar la entrega en sí misma (producto) como lo que importa, en detrimento del proceso (aprendizaje). En relación con este hecho, se pueden producir comportamientos poco éticos que, por ejemplo, no garantizan la autoría de las entregas.

Respecto a este tercer aspecto, un indicio en nuestra propia docencia en los dos cursos anteriores que creemos ilustrativo es el siguiente. En una de las entregas se pide a los estudiantes un *shell script* (fichero

(A) POR PARTE DEL DOCENTE:

| Ahora | Deseable |
|---|--|
| Papel verificador, notario, controlador, policial | Facilitador, orientador, guía |
| Apenas puede corregir y dar feedback de calidad y se limita a cuantificar y justificar la nota otorgada a los estudiantes | Orienta la mejora en el proceso de aprendizaje, especialmente a quien lo pide o está en condiciones y actitud favorables para ello |
| Se invierte un tiempo y un esfuerzo cuya utilidad resulta discutible | La inversión en tiempo y esfuerzo resulta más productiva, creativa, razonable y útil |
| «Espera» que el estudiante sea honrado al entregar la actividad y aproveche el feedback recibido | No debe asumirlo (tanto). La responsabilidad debe interesar más al estudiante que al profesor |

(B) POR PARTE DEL ESTUDIANTE:

| Ahora | Deseable |
|---|--|
| Persigue entregar la actividad y obtener puntos | Valora su aprendizaje, busca comprender y «saber hacer» |
| Se rige por «el fin justifica los medios» y, por tanto, su aprendizaje resulta superficial o se hacen copias entre compañeros | Identifica los puntos fuertes y débiles de su aprendizaje. Sabe qué sabe y qué no |
| Espera que sea el profesor quien revise, califique y certifique | Aprende a valorar la calidad de su aprendizaje y de las actividades entregadas. Emite juicios de valor y/o cuantifica su entrega |
| Ignora el feedback del profesor; se queda con la nota numérica | Identifica qué debe mejorar en su aprendizaje y cómo hacerlo. Reflexiona y actúa en consecuencia |

Cuadro 1: Actuaciones (estereotipadas) actual y deseable del docente y del estudiante

| | |
|---------------------------------|-------------------------------|
| 1. Entornos Linux y Windows | 5. Google Drive |
| 2. Eclipse | 6. Tablas y gráficas |
| 3. Bases de datos (LibreOffice) | 7. Conexión y transf. remotas |
| 4. Servicios internet y HTML | 8. Intérprete de órdenes Unix |

Cuadro 2: Prácticas planteadas en la asignatura.

de texto con instrucciones a ejecutar por un intérprete de órdenes del S.O.) de baja complejidad (unas 15 líneas). Ante varias entregas dudosas o scripts que directamente no funcionaban, calificamos dichas entregas con un cero y, sorprendentemente, los estudiantes que hicieron dichas entregas no mostraron su disconformidad. A nuestro juicio, dicha actitud podría indicar fraude. Rectificar esta situación nos parece particularmente importante en estudiantes de primero para evitar que perpetúen este comportamiento en lo sucesivo. Además, el problema resulta más grave, si cabe, en actividades en equipo, donde parece que no importa que todos aprendan, sino que se entregue un producto aunque sólo sea labor de un miembro del equipo.

El Cuadro 1 recoge de forma resumida algunas reflexiones adicionales. Las conductas indicadas son generalizaciones, pero ilustran bien situaciones que pueden plantearse. Como resultado de una o varias de estas situaciones, el sentido de la evaluación continuada queda desvirtuado. Algunas soluciones propuestas pasan por organizar y automatizar el proceso [3], o acudir a la co-evaluación (evaluación entre pares) o la auto-evaluación [9, 10].

La autoevaluación se presenta como una posible solución a (una parte de) las dificultades comentadas. Aunque la utilidad de dicha autoevaluación resulta más evidente como herramienta de evaluación *formativa*, despierta naturales inquietudes si se desea utilizar, además, como mecanismo de evaluación *sumativa*, pues se considera que el estudiante puede, voluntariamente o no, cuantificar mal su nota. Por ejemplo, en algún caso se ha visto [6] que si la autoevaluación no es sumativa, ésta tiende a coincidir con la valoración del profesor, pero la discrepancia aumenta (en un sentido u otro) cuando la evaluación es sumativa. En general, la mayor responsabilidad que se otorga a los estudiantes en el aprendizaje, no va acompañada de mayor responsabilidad en la evaluación [4]. Aunque la problemática de la evaluación continua y las propuestas de soluciones son un tema recurrente en los foros educativos, es este contexto más específico, el de la «autoevaluación sumativa», el que motiva nuestra propuesta, cuya novedad creemos que estriba en potenciar y concretar procesos de reflexión encaminados a hacer al estudiante más responsable de su propio aprendizaje.

2. Metodología

Nuestra propuesta se enmarca en las prácticas de una asignatura básica de primer curso de ingeniería informática cuyo contenido se muestra en el Cuadro 2. Se describe a continuación el esquema de evaluación propuesto y su puesta en práctica este curso.

| Caso (<i>i</i>) | C_a, C_p | Probable interpretación | Peso $w_i \times 100$ (%) |
|-------------------|------------|-------------------------|---------------------------|
| 1 | Baja, Baja | Sinceridad/honestidad | 5 |
| 2 | Baja, Alta | Mejora/superación | 10 |
| 3 | Alta, Baja | Anormalidad/sospecha | -10 |
| 4 | Alta, Alta | Consistencia/excelencia | 10 |

Cuadro 3: Casuística de resultados en autoevaluación (C_a) y en prueba de verificación (C_p).

2.1. Esquema propuesto

Para articular nuestra idea de autoevaluación sumativa, planteamos, por un lado, la prueba de autoevaluación en sí misma, para la que es posible obtener una nota N_a ; por otro lado, consideramos una prueba de verificación o validación que permita contrastar la autoevaluación, y que también tendría asociada su nota N_p . Para que esta propuesta tenga sentido y no contradiga los supuestos bajo los que se propone, la prueba de verificación no debería implicar un sobre coste adicional y convendría que se tratase de algún tipo de actividad que se hubiera planteado en cualquier caso. Por ejemplo, pensamos que, relacionado con los contenidos de una entrega, un ejercicio en un examen podría servir como prueba de validación de la autoevaluación asociada a dicha entrega. Como puede apreciarse, este esquema es flexible y se presta fácilmente a muchas variantes que se ajusten a las condiciones concretas de cada contexto educativo. A nuestro juicio, un posible e interesante planteamiento es que el estudiante vea su nota de la autoevaluación como la nota provisional, «condicionada» al resultado de la prueba de validación posterior. De esta forma, el estudiante puede estar motivado por realizar una buena autoevaluación, pues su nota N_a será contrastada con su nota en la prueba N_p .

Este curso hemos implementado la idea considerando las notas en dos niveles (**Alta** y **Baja**), de modo que asociadas a N_a y N_p se tiene las notas «cualitativas» C_a y C_p : $C_a = \text{Alta}$ si $N_a \geq \theta_a$ y $C_a = \text{Baja}$ si $N_a < \theta_a$, donde θ_a es un umbral para distinguir los dos niveles; C_p se obtiene por análogo procedimiento a partir de N_p y otro umbral θ_p . Como se verá después, estos umbrales no tienen por qué coincidir con la nota de corte de aprobado ni tienen por qué ser iguales para la autoevaluación y la prueba.

Con este planteamiento, son posibles cuatro casos (Cuadro 3) para los 2×2 valores de (C_a, C_p) :

1. **Baja, Baja**: No se aprecia mejora académica, pero puede valorarse la sinceridad o la honestidad de la actuación del estudiante en su autoevaluación.
2. **Baja, Alta**: En la prueba de verificación se ha obtenido un mejor resultado y parece proceder «premiar» dicha superación.
3. **Alta, Baja**: Resulta sospechoso que se obtenga una menor puntuación en la prueba de verificación.

Este caso es quizás el más conflictivo porque puede originarse por dos situaciones diferentes. Por un lado, es posible que el estudiante haya «inflado» su autoevaluación y la prueba refleje mejor su grado de aprendizaje. Por otro lado, es posible que la autoevaluación sea precisa y justa y el pobre resultado en la prueba obedezca a causas ajenas al aprendizaje en sí (ansiedad durante la prueba, por ejemplo). Si se tiene en cuenta esta doble interpretación, no conviene penalizar excesivamente esta situación.

4. **Alta, Alta**: Los resultados de la prueba tienden a confirmar que el estudiante ha aprendido realmente y merece valorarse su actuación consistente.

Quedan por definir algunos aspectos, como la obtención de la nota final que combine la nota de la autoevaluación y la de la prueba. También aquí son posibles diferentes propuestas; indicamos la nuestra en el Apartado 2.3. En cuanto a la autoevaluación, la hemos planteado mediante cuestionarios similares en algunos aspectos a las rúbricas, y fueron realizados en Moodle (moodle.org).

Para este curso nos planteamos probar este esquema para una única entrega, hacia final del semestre, a modo de primera experiencia. No obstante, consideramos oportuno ensayar la idea en una actividad preparatoria que nos permitiera descubrir dificultades y ofreciera a los estudiantes una oportunidad de conocer la dinámica antes de su aplicación en la prueba real. Se describe a continuación la experiencia tanto en la actividad preparatoria (A) como en la real (B).

2.2. Actividad preparatoria (A)

Asociado a la entrega en grupo de la práctica de bases de datos (en LibreOffice), preparamos un cuestionario (Q_A , Cuadro 4) para su cumplimentación individual. Puesto que entendemos la autoevaluación en sentido «amplio», las preguntas no son sólo de competencia técnica (Apartado C) sino también de actitudes y comportamientos (Apartados A y B), en la línea de lo que se desea inculcar y promover. También se daba la opción de comentar otros aspectos en una pregunta abierta (Apartado D).

Puesto que era una actividad preparatoria, no se utilizó la nota de la autoevaluación para obtener la nota de

A. Estudio y desarrollo de la práctica

- A₁. He leído todo o casi todo el material de apoyo antes de hacer los ejercicios del boletín
- A₂. Como tengo base suficiente, no he leído en detalle el material de apoyo, pero he recurrido a él si lo he necesitado [...]
- A₃. He realizado/seguido alguno/todos los ejercicios que se explican en el material de apoyo
- A₄. Además del ejercicio que había que entregar, he realizado otros ejercicios del boletín
- A₅. Al realizar los ejercicios o estudiar el boletín he recurrido a otras fuentes (libros, internet) para resolver dudas o ampliar información
- A₆. He consultado dudas con algún compañero de otros equipos
- A₇. He ayudado a otros miembros de mi equipo a entender la materia de esta práctica
- A₈. Hay aspectos de esta práctica que sigo sin entender bien
- A₉. Para resolver las dudas que aún tengo, sé qué necesito hacer y voy a actuar pronto para resolverlas

B. Honestidad y labor de equipo

- B₁. El trabajo entregado ha sido realizado íntegramente por los miembros del grupo
- B₂. Todos los miembros del equipo entienden el ejercicio entregado y sabrían hacer otro similar
- B₃. He comprobado «manualmente» que el resultado de las consultas es el que debe ser

C. Competencia técnica

- C₁. Sé crear registros sin dificultad
- C₂. Sé crear tablas y definir los campos con las propiedades adecuadas conforme a lo exigido
- C₃. Entiendo qué es una clave primaria y cómo definirla en las tablas
- C₄. Entiendo qué relaciones debo establecer entre las tablas y cómo hacerlo en LibreOffice
- C₅. Sé cómo definir y ejecutar consultas de la dificultad de los ejercicios del boletín

D. Comenta cualquier aspecto que consideres conveniente o pertinente en relación a cualquiera de las tres partes del cuestionario [...]Cuadro 4: Cuestionario (Q_A) de reflexión-autoevaluación en la actividad preparatoria (A).

esa entrega sino que en este caso simplemente se utilizó directamente la nota correspondiente a la corrección de la entrega. A efectos puramente de autoevaluación, se cuantificaron los Apartados B y C.

2.3. Actividad real (B)

Asociado a la entrega de un shell script, se preparó otro cuestionario (Q_B , Cuadro 5). Las respuestas a algunas de estas preguntas se asimilaron a una escala Likert¹ de 5 puntos, una escala psicométrica habitual en el análisis de respuestas de sujetos. Por aligerar al estudiante su cumplimentación, solo las dos primeras preguntas tenían que ver con cuestiones éticas, como la autoría de la entrega.

En este caso, la prueba de validación consistió en una prueba con ordenador, de 1 hora de duración, que tuvo lugar pocos días después de la entrega del script y de la autoevaluación. En ella se planteó el desarrollo de un script para un problema muy similar al planteado para la entrega. La nota final N de la práctica, cuyo mecanismo de cálculo concreto los estudiantes desconocían, se obtuvo a partir de la nota de la prueba N_p y la de la autoevaluación N_a como

$$N = N_p + w_i \cdot N_a, \quad (1)$$

siendo w_i el peso que asociamos a cada caso i (Cuadro 3). Considerando notas sobre 100, los umbrales para los dos niveles de la autoevaluación y la prueba fueron $\theta_a = 70$ y $\theta_p = 50$, respectivamente, que reflejan distintas exigencias para las distintas condiciones de desarrollo de los scripts. Estos pesos deben adaptarse a cada situación e incluso pueden ser variables para

¹http://en.wikipedia.org/wiki/Likert_scale.

diferentes actividades, puesto que, como argumentaba un revisor, es cuestionable si resulta oportuno favorecer tanto notas N_a bajas, sean o no sinceras.

3. Resultados**3.1. Actividad preparatoria (A)**

De los 97 estudiantes que contestaron la parte A del cuestionario Q_A , sobre dos terceras partes afirman haber utilizado bastante el material de apoyo (A_1 – A_4). La mitad (entre un 50 % y un 60 %) dice haber recurrido a otras fuentes, o a otros compañeros (A_5 – A_6), o ellos mismos han ayudado a otros (A_7). Sólo una cuarta parte reconoce que hay algún aspecto de la práctica que no entiende bien (A_8), y casi el 90 % sabe qué puede hacer para resolver las dudas que le quedan (A_9).

La mayoría (entre el 70 % y el 83 %) de los 94 que respondieron la parte B de Q_A está completamente de acuerdo con las afirmaciones sobre honestidad y buen hacer (B_1 – B_3). Algunos (16–22 %) lo afirman con matices y una minoría (1–7 %) tiene mayores dudas.

De los 91 que contestaron el apartado C de Q_A , casi todos (83–93 %) afirman entender bien los aspectos de creación de registros, definición de tablas y claves primarias (C_1 – C_3). Una cuarta parte de los estudiantes afirma con menor firmeza que comprende el tema de las relaciones entre tablas y de las consultas (C_4 – C_5).

En relación a la pregunta abierta (apartado D), bastantes estudiantes plantearon preguntas sobre aspectos de la práctica de los que tenían algunas dificultades. Es posible que estudiantes que de otra forma no hubieran preguntado en clase o en tutorías aprovecharan esta oportunidad para intentar aclarar sus dudas. De di-

1. Salvo ayudas puntuales de compañeros o profesores, he realizado individualmente el script entregado
2. Entiendo bien la solución que he propuesto y sabría realizar un script para resolver un problema similar
3. He comprobado el correcto funcionamiento del script con el fichero que se proporciona de ejemplo
4. He realizado comprobaciones adicionales del buen funcionamiento del script (por ejemplo, funciona para otros ficheros de entrada)
5. Mi script consigue leer correctamente todas las líneas del fichero de entrada y obtener el valor de cada campo
6. Entiendo cómo hacer operaciones aritméticas básicas (suma, producto, etc.) con el shell y mi script las realiza correctamente
7. La solución propuesta utiliza un bucle y entiendo su propósito, o no utiliza ninguno y sé por qué no es necesario
8. He conseguido que el script determine y muestre correctamente el tiempo total de conexión
9. He conseguido que el script determine el tiempo máximo de conexión y el host y la fecha correspondientes
10. He conseguido separar correctamente el día del mes a partir de la variable que recoge la fecha
11. Entiendo cómo desglosar los minutos totales en horas y minutos, y mi script lo hace bien
12. La salida de mi script obedece al formato que se muestra en el boletín (ejercicio 59)
13. En mi solución utilizo una sentencia condicional («if») y entiendo su utilidad, o no utilizo ninguna y sé por qué no se necesita [...]
14. He conseguido que el script muestre el mes en letra (por ejemplo, «Agosto» para el mes 8)
15. Mi script acepta como entrada cualquier nombre de fichero y funciona si este cumple el formato esperado

Cuadro 5: Cuestionario (Q_B) de reflexión-autoevaluación en la actividad real (B).

chas dudas, sintetizamos y creamos una FAQ de cuatro preguntas-respuestas que intentara ayudarles. En el Apartado 3.4 se analiza si la consultaron y les ayudó.

3.2. Actividad real (B)

De los 76 estudiantes que contestaron el cuestionario Q_B , la mayoría afirma haber realizado el script ($Q_{B.1}$: $M=4.3$, $SD=1.0$), aunque es interesante comprobar que hay 11 estudiantes que admiten que no ha sido así o lo cuestionan. Con cierta lógica, el resultado es similar respecto a si entienden el script y sabrían hacer uno parecido ($Q_{B.2}$: $M=4.4$, $SD=0.7$). Aunque muchos han comprobado el script con el fichero de texto que se les daba de prueba ($Q_{B.3}$: $M=4.8$, $SD=0.8$), menos han realizado comprobaciones adicionales ($Q_{B.4}$: $M=3.7$, $SD=1.7$). Casi todos han conseguido resolver la mayoría de las partes del script (e.g. $Q_{B.8}$: $M=4.9$, $SD=0.4$), incluso para aspectos menos triviales (e.g. $Q_{B.11}$: $M=4.8$, $SD=0.6$), y se aprecia alguna dificultad en algún otro aspecto (e.g. $Q_{B.10}$: $M=4.5$, $SD=0.9$).

La Fig. 1 muestra la relación entre N_a y N_p , así como su distribución en los cuatro casos considerados. Se aprecia cierta correlación entre ambas notas, si bien hasta una tercera parte de los estudiantes obtiene mejor nota en la autoevaluación que en la prueba de validación (Caso 3). Algunos casos (parte inferior derecha) son especialmente llamativos, con mucha discrepancia entre la nota de la autoevaluación ($N_a \geq 80$) y la de la prueba ($N_p < 40$). Analizamos en detalle 8 de estos casos y comprobamos que el script que entregaron junto con la autoevaluación es correcto. Una interpretación de esta (aparente) contradicción es que su grado de autoría en el script entregado sea reducida; de hecho, 5 de estos 8 estudiantes está *parcialmente* (no *completamente*) de acuerdo con las preguntas relacionadas con la autoría ($Q_{B.1}$) o su capacidad de resolver un problema similar ($Q_{B.2}$). Algunos indicios sospechosos en alguno de estos casos es que usa órdenes

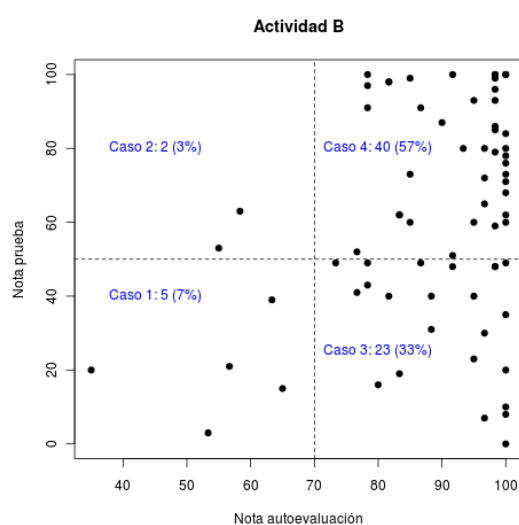


Figura 1: Distribución de notas y casos (Cuadro 3) en la Actividad B de los 70 estudiantes que la hicieron

avanzadas (no enseñadas en la asignatura), lo que contrasta con su pobre desempeño en la prueba; o el hecho de que haya alguna discrepancia entre su valoración de algún aspecto del script y lo que en realidad cumple el script de la entrega de la práctica.

Como apuntaba un revisor, sería interesante analizar si los repetidores realizan mejores autoevaluaciones. Sólo hay tres entregas correspondientes a estudiantes repetidores y los tres están en el Caso 3, pero la muestra es demasiado reducida para extraer conclusiones. Este mismo revisor sugería un seguimiento en sucesivos cuatrimestres para analizar si, con el tiempo, un estudiante se vuelve más responsable u honesto. No hemos podido realizar este tipo de estudio, pero dada su relevancia en el contexto de nuestro trabajo, podría plantearse en el futuro.

3.3. Comparación con cursos anteriores

Considerar un grupo control dentro del mismo curso es interesante para evaluar el efecto de la propuesta, pero puede resultar difícil o polémico por las diferencias de evaluación a diferentes estudiantes del mismo curso. En un futuro se podría analizar cómo establecer este grupo control en el mismo año, pero en este caso la comparación sólo se ha establecido, con obvias limitaciones, con los resultados de cursos anteriores.

A efectos de comparar resultados de este curso con los de cursos anteriores, consideramos la relación entre la nota en la entrega de una práctica y la nota en el examen del ejercicio correspondiente al mismo tema de la práctica. En analogía a las notas de la autoevaluación y la prueba verificadora, consideramos los mismo umbrales θ_a y θ_p para las entregas y el ejercicio del examen que los considerados para el esquema en la prueba real de este curso (i.e. $\theta_a = 70$, $\theta_p = 50$). Con ello se obtiene la distribución de casos que se muestra en el Cuadro 6. En analogía al análisis del apartado anterior, se constata un elevado porcentaje de estudiantes en el Caso 3 para la prueba del script, lo que sugiere una posible falta de honestidad en la entrega de la práctica. Sin embargo, respecto al análisis anterior (Apartado 3.2), la distribución de los Casos 1 y 4 está invertida, lo cual apunta a un dominio de este tema insuficiente para afrontar el script del examen (de dificultad similar al de la práctica). Otros factores que pueden explicar esta discrepancia son (i) que la prueba verificadora se celebró muy pocos días después de la fecha límite de la autoevaluación; (ii) el script de la prueba era de una naturaleza similar a la de la práctica; y (iii) que, a diferencia de un examen final en el que los estudiantes no saben qué ejercicios van a plantearse, en el caso de la prueba verificadora era una prueba específica sobre ese tema, y podían prepararse mejor.

Respecto al tema de bases de datos (BBDD), se observa que los estudiantes lo dominan más (casi o más de la mitad están en el Caso 4). Una forma de analizar el efecto de la reflexión-autoevaluación sería comprobar si, en relación a otros años, se da una mayor correlación entre la nota de la práctica y la del examen. Se puede comprobar que no es así y que, de hecho, hay más porcentaje de estudiantes en el Caso 3 que en otros años. No tenemos una explicación clara a esta discrepancia respecto al patrón de otros cursos, aunque parece que los estudiantes encontraron cierta dificultad en el ejercicio del examen de este año.

3.4. Opinión de estudiantes y profesores

Días después del examen se planteó una consulta anónima a los estudiantes acerca de su experiencia en las entregas de prácticas y las autoevaluaciones (de ambas actividades, la preparatoria y la real). Las respues-

| Caso | 10-11 | | 11-12 | 12-13 |
|---------------|--------|------|-------|-------|
| | Script | BBDD | BBDD | BBDD |
| 1 | 48 | 10 | 6 | 3 |
| 2 | 0 | 7 | 6 | 3 |
| 3 | 47 | 28 | 29 | 48 |
| 4 | 5 | 54 | 59 | 46 |
| # estudiantes | 73 | 81 | 105 | 92 |

Cuadro 6: Distribución de casos (%) en los tres cursos académicos y los dos temas (script y BBDD).

tas para casi todas las preguntas expresaban un grado de acuerdo en una escala Likert (desde «Completamente de acuerdo» hasta «Completamente en desacuerdo») y se analizaron asociando valores 1-5 a estas respuestas (mayor número, mayor acuerdo). Respondieron la encuesta 33 estudiantes.

Correcciones de las entregas. Hubo consenso en que no necesitan que el profesor corrija una entrega para saber cómo está su entrega o qué saben ($M=4.0$, $SD=1.0$), y en que la nota del profesor suele coincidir con la que ellos esperan ($M=3.8$, $SD=0.7$). Suelen encontrar útiles y a tener en cuenta en el futuro los comentarios o justificación de nota del profesor ($M=4.0$, $SD=1.1$). Las respuestas a estas tres preguntas señalan la escasa utilidad de una corrección cuantitativa y la preferencia por comentarios cualitativos.

Autoevaluaciones sumativas. Existe cierta discrepancia en si las autoevaluaciones les ayudan a reflexionar sobre el aprendizaje o la calidad de las entregas ($M=3.4$, $SD=1.2$) o a cuestionarse su forma de estudiar o trabajar, ya sea individualmente ($M=3.0$, $SD=1.3$) o en equipo ($M=2.8$, $SD=1.1$). También hay opiniones contrapuestas sobre la conveniencia de que la autoevaluación contribuya a definir parte de la nota ($M=3.3$, $SD=1.5$) y por eso no hay una evidencia clara de que nos lo recomienden hacer en otros cursos o extenderlo a otras actividades ($M=3.4$, $SD=1.1$). Por tanto, no está clara la utilidad de las autoevaluaciones sumativas, pero creemos que el proceso de reflexión es beneficioso.

Otras formas de realimentación. No todos los que responden consultaron la FAQ sobre la práctica de bases de datos ($M=3.3$, $SD=1.6$), pero los que lo hicieron (25 de los 33, $\approx 75\%$) sí la encontraron clarificadora ($M=3.9$, $SD=0.9$). Muchos estudiantes echan de menos un seguimiento más personalizado que les orientara en aprender más o mejor ($M=3.9$, $SD=1.0$). Estas evidencias sustentan la conveniencia de formas de realimentación alternativas a las correcciones sumativas.

Quiénes contestaron. La consulta incluía una pregunta sobre los rangos de sus notas. La mayoría de los que contestaron ($\approx 67\%$) son los del Caso 4 (notas altas tanto en la autoevaluación como en la prueba verificadora). A los del Caso 3 (nota alta en la autoevaluación pero baja en la prueba) se les preguntó a qué creen que puede deberse esta discrepancia. De las cuatro res-

*E*₁: «Los cuestionarios de autoevaluación sólo sirven para hinchar la nota y no para evaluarnos a nosotros mismos. Mucha gente entre la que me incluyo los ha rellenado muy positivamente para alcanzar la máxima nota en ese apartado, a pesar de mentir en ciertas respuestas (aunque por suerte, no todas).

»Si realmente quieren reflexión, deberían realizar las FAQ, que son MUY útiles, al final de cada entrega; a ser posible, un poco más extensas y detalladas. Este documento es mucho más interesante que un comentario de una línea como: “Faltan por definir las relaciones pertinentes”, que no te aclara nada e incluso puede confundirte más.

»El cuestionario de autoevaluación debería constar sólo de un cuadro de texto donde podamos enviar nuestras dudas para que sean compiladas en las FAQ. Y por supuesto, no dar ninguna puntuación al cuestionario, si lo que buscan es sinceridad.»

*E*₂: « Las autoevaluaciones son buenas si la gente es sincera, creo que hay personas que saben que cuenta para la nota y solo buscan poner lo que sea para sacar la máxima puntuación.

»Aun así, creo que son una buena herramienta, pero deberían ir acompañadas no solo de preguntas de si crees haber aprendido lo explicado en las prácticas, sino también preguntas relacionadas con lo aprendido, a modo de test sencillo sobre la propia práctica. De ese modo, si [sic] que ves si realmente sabes lo necesario sobre ese tema o práctica.»

*E*₃: «Las autoevaluaciones están muy bien, creo que deberían incluirse en todas las prácticas y ejercicios adicionales»

*E*₄: «No deberían contar para nota. Los que realizan las autoevaluaciones somos los interesados en tener una nota alta [...]»

*E*₅: «Aunque la comunicación de la situación de aprendizaje del alumno al profesor me parece buena y efectiva, he echado de menos algunos cuestionarios sobre la materia, para poder valorar, según las exigencias del profesorado si he comprendido correctamente la materia o no.»

Cuadro 7: Comentarios de diferentes estudiantes *E*_{*i*}.

| Síntesis de las opiniones de los estudiantes | Reflexión de los profesores e ideas de mejora |
|--|--|
| Los estudiantes no (siempre) son sinceros | Es lamentable que sea así y sería oportuno plantear cómo fomentar la honestidad, la responsabilidad y otros valores [7] útiles en los ámbitos personal y profesional. Por cierto, ¿son sinceros cuando dicen que no lo han sido? |
| La autoevaluación sólo sirve para hinchar notas | Es cierto que con el esquema concreto planteado este curso, todos los estudiantes, excepto los del Caso 3, ven incrementada su nota. Sin embargo, con la prueba de verificación, el efecto de la nota de la autoevaluación no es significativa para llegar a distorsionar el resultado. Hay que analizar mejor si convendría otro esquema para cuantificar las autoevaluaciones o incluso plantear que éstas sean solo formativas, no sumativas. |
| Sería mejor preguntas de autoevaluación que realmente ayudaran a comprobar si conocen los contenidos | Es curioso que los estudiantes no vean ya las prácticas como una forma de autoevaluación. Esto parece confirmar la idea de que el producto prima sobre el proceso. Una forma sencilla de reconducir esta tendencia podría ser renombrar las prácticas o las entregas como «autoevaluaciones». |

Cuadro 8: Síntesis de opiniones de los estudiantes y reflexión de los profesores.

puestas, dos coinciden en que contaron con la ayuda de los compañeros, les faltó tiempo en la prueba, o piensan que ésta fue más difícil que la práctica. Pensamos que la dificultad de ambos scripts era muy similar, pero que faltase garantizar tiempo suficiente; también convendría reforzar la autonomía del estudiante.

Comentarios adicionales. En una pregunta abierta los estudiantes podían comentar otros aspectos de interés. El Cuadro 7 recoge una muestra casi completa de estos comentarios y el Cuadro 8 resume sus opiniones acompañadas de nuestra reflexión.

Entrevistas previstas. De los 33 estudiantes, 14 mostraron su disposición a seguir comentando con los profesores, y la mitad de ellos facilitaron su dirección de e-mail. Pensamos que una entrevista con ellos podría resultar fructífera, en la línea del diseño centrado en el usuario aplicado a la educación [8].

4. Discusión

La propuesta realizada permite abordar los tres problemas indicados en la introducción, puesto que: **1 El coste es igual o menor.** La tarea adicional del estudiante es sólo rellenar cuestionarios de reflexión y autoevaluación. El profesor prepara los cuestionarios, pero no corrige entregas y la prueba de validación se hace coincidir con actividades que se harían igualmente. Más importante que dedicar menos tiempo es aumentar el sentido y la utilidad de nuestra dedicación. **2 La realimentación puede ser más útil.** El profesor no corrige las entregas; el estudiante recibe realimentación a través de su autoevaluación (tipo rúbrica + reflexión) y, más adelante, con la nota del profesor, pudiendo contrastar ambos resultados. **3 El proceso cobra mayor importancia.** Durante la autoevaluación-reflexión, el estudiante toma conciencia de lo que realmente impor-

ta, a nivel técnico (e.g. ¿sabe lo que sabe y lo que no?), profesional (e.g. ¿ha comprobado el correcto funcionamiento?) y ético (e.g. ¿es el autor de la entrega?).

En el futuro, los diferentes casos considerados podrían servir para detectar, diagnosticar y atender convenientemente a estudiantes con baja motivación o dificultades de aprendizaje (e.g. los del Caso 1), baja autoestima (quizás los del Caso 2), dificultades en autoevaluarse adecuadamente u otros problemas subyacentes (puede que los del Caso 3). A los estudiantes en el Caso 4 se les podría felicitar y animar a continuar o a mejorar y, si procede, a proponerle otros retos.

A diferencia de esta primera experiencia puntual, pensamos que la aplicación de estas autoevaluaciones a lo largo del semestre, con el seguimiento oportuno, permitiría mejorar el aprendizaje y potenciar buenas conductas. Con tal fin, se podrían añadir mecanismos que fueren al estudiante a reflexionar y tratar de explicar las discrepancias entre sus notas en la autoevaluación y en la prueba de validación, así como incentivar la concordancia entre ambas notas. Otra idea atractiva consiste en combinar la propuesta con mecanismos de co-evaluación. Y, a más largo plazo, pensamos que podría ser oportuno avanzar hacia algo similar a un «sistema de honor» [1], lo que supone afrontar otros retos.

Aunque nos hemos centrado en el aspecto reflexivo de la autoevaluación, es posible integrar pautas para una autoevaluación más efectiva [2]. A pesar de sus limitaciones, creemos que la idea del esquema ensayado es positiva, y requiere mejora y constancia. Hacemos nuestras estas palabras del filósofo Francesc Torralba: «*La educación es una inversión a largo plazo, un ejercicio difícil y constante que no da resultados inmediatos, sino frutos después de haber trabajado mucho*» (traducción al castellano de la frase en [7, p. 335]).

5. Conclusiones

Se ha explorado un mecanismo de autoevaluación sumativa para afrontar algunas de las deficiencias que tal evaluación puede presentar en la práctica. Se persigue fomentar procesos de reflexión para mejorar la honradez y responsabilidad del estudiante así como el proceso y la calidad de su aprendizaje. Esta primera experiencia, más que conclusiones definitivas, aporta indicios y recomendaciones como: ❶ Evitar evaluaciones sumativas basadas en correcciones exhaustivas; los estudiantes no suelen valorar ni beneficiarse de esta realimentación. ❷ Reorientar el esfuerzo del profesorado en la ejecución de otras formas de realimentación de mayor calidad (e.g. preparar FAQs a partir de las dudas de los estudiantes, o realizar seguimientos más individualizados). ❸ Potenciar la idea y la utilidad de las «autoevaluaciones», sean sumativas o no, y mejorar la autonomía del estudiante que depende en exceso

de compañeros. ❹ Educar a los estudiantes en valores como la honestidad, la humildad, o la responsabilidad. Es primordial que los profesionales y los ciudadanos de nuestra sociedad tengan estas cualidades humanas, además de competencias técnicas.

El esquema de autoevaluaciones seguidas de pruebas verificadoras hemos visto que parece útil para detectar anomalías de diferente etiología y atender de forma adecuada la diversidad de necesidades. Su aplicación regular a lo largo de toda la asignatura (y mejor aún, en más asignaturas y más cursos), sólo puede tener, a nuestro entender, efectos beneficiosos.

Agradecimientos: Proyecto de Innovación Educativa (curso 2012-13), «Unitat de Suport Educatiu», Universitat Jaume I.

Referencias

- [1] E. Camahort y F. Abad. Métodos alternativos de evaluación basados en el sistema de honor. En *JENUI*, 2002.
- [2] G. Gibbs y C. Simpson. *Condiciones para una evaluación continuada favorecedora del aprendizaje*, vol. 13 de *Cuadernos de Docencia universitaria*. Octaedro, 2009.
- [3] J. Ibáñez, I. Usandizaga, y A. Sánchez. Herramientas de instrucción masiva: pistas para implantar evaluación continua en grandes grupos. En *JENUI*, 2012.
- [4] D. J. Nicol y D. Macfarlane-Dick. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2):199–218, Apr. 2006.
- [5] J. Serrano-Guerrero, F. P. Romero, E. Fernández-Viñas, y J. A. Olivas. La sobre-evaluación. En *JENUI*, 2011.
- [6] R. A. Tejeiro y cols. La autoevaluación sumativa en la enseñanza superior: implicaciones de su inclusión en la nota final. *Electronic J. of Research in Ed. Psychology*, 10(27), Sept. 2012.
- [7] F. Torralba Roselló. *Cent valors per viure: la persona i la seva acció en el món*. Pagès editors, 2001.
- [8] V. J. Traver. Can user-centered interface design be applied to education? *SIGCSE Bulletin*, 39(2):57–61, 2007.
- [9] M. Valero-García y L. M. Díaz de Cerio. Evaluación continuada a un coste razonable. En *JENUI*, 2003.
- [10] M. Valero-García y L. M. Díaz de Cerio. Autoevaluación y co-evaluación: estrategias para facilitar la evaluación continuada. En *Simposio Nacional de Docencia en la Informática*, pág. 25–32, 2005.