

Análisis comparativo de la Escala de Usabilidad del Sistema (EUS) en dos versiones

Comparative analysis System Usability Scale (SUS) in two versions

Análise comparativa usabilidade do sistema Scale (SUE) em duas versões

María Isolde Hedlefs Aguilar

Universidad Autónoma de Nuevo León, México

lab.cognitive@gmail.com

Andrea Abigail Garza Villegas

Universidad Autónoma de Nuevo León, México

agarza.v@hotmail.com

Resumen

Se adaptó la EUS en sus dos versiones: original y positiva, en una población mexicana donde se habla el español, con el objetivo de que después puedan ser utilizadas en otras investigaciones. Para cada versión se utilizaron dos muestras, la versión original se aplicó a 243 participantes y la positiva se aplicó a 173 participantes. El coeficiente de Alpha de Cronbach para la versión original fue de .59 y para la positiva de .92, indicando que esta última tiene una muy buena confiabilidad. El AFE en la versión positiva mostró la presencia de dos factores, como menciona la literatura, que fueron Usabilidad y Facilidad de Aprendizaje, cumpliendo así con la validez de constructo. Se concluye que la versión positiva es mucho más confiable a la hora de su aplicación y que muestra una estructura factorial más acorde con la literatura.

Palabras clave: usabilidad, EUS, estandarización, confiabilidad, validez.

Abstract

The SUS was adapted both original version and the positive version to Spanish language in a Mexican population, with the purpose to be used in further research. For each version, were used two samples, in the original version applied to 243 participants and the positive version applied to 173 participants. The Cronbach's alpha coefficient for the standard version was .59, while the positive version of the Cronbach alpha coefficient was .92, indicating a good statistical reliability. The factor analysis in the positive version showed the presence of two factors as mentioned in the literature, these were usability and Learnability, fulfilling the construct validity. It is concluded that the items in positive version is much more reliable when applied and shows a factorial structure consistent with the literature.

Key words: usability, SUS, standardization, reliability, validity.

Resumo

EUS foi adaptado em duas versões: original e positivo, em uma cidade mexicana, onde o espanhol é falado, com o objectivo de que mais tarde pode ser usado em outras investigações. Para cada versão foram utilizadas duas amostras, a versão original foi aplicado a 243 participantes e o positivo foi aplicado a 173 participantes. O coeficiente de Cronbach Alpha para a versão original foi 0,59 e 0,92 para o positivo, indicando que este último tem uma boa confiabilidade. A AFE na versão positiva mostrou a presença de dois fatores, como mencionado na literatura, que foram usabilidade e facilidade de aprendizagem, cumprindo assim a validade de construto. Conclui-se que a versão positiva é muito mais confiável quando a sua aplicação e que mostra uma estrutura fator mais consistente com a literatura.

Palavras-chave: usabilidade, EUS, padronização, confiabilidade, validade.

Fecha recepción: Enero 2016

Fecha aceptación: Julio 2016

Introdução

Uma das primeiras chamadas que vieram para avaliar a usabilidade de uma interface e onde não era necessário para os testes de conduta participante do laboratório, foi Scale usabilidade do sistema (SUE), cujo nome Inglês é Scale usabilidade do sistema (SUS) e ele foi desenvolvido por Brooke em 1986, que consistia em 10 itens (5 positiva e 5 itens itens negativos).

A construção dos itens EUS originado a partir da definição de usabilidade baseado na ISO 9241-11: "o grau ao qual um produto pode ser usado por utilizadores especificados para atingir objectivos especificados com eficácia, eficiência e satisfação um contexto particular de uso ". A eficiência é definida como o grau de precisão e fácil integralidade para atingir objectivos específicos, enquanto a eficiência é definida como o grau em que os recursos são utilizados para o usuário a atingir os seus objetivos com precisão e integridade, e satisfação é definida como a liberdade do usuário para mostrar desconfortável ou mostrar atitude positiva usando o produto (ISO 1998). Isto é, a eficiência significa que o utilizador realizar a tarefa, a eficiência do utilizador executar a tarefa tão rapidamente quanto possível, e a satisfação que os utilizadores sentem quando se utiliza o sistema ou interface.

Portanto, a escala EUS tem dois objetivos principais. A primeira é que os pesquisadores podem obter uma medida da percepção da usabilidade de um sistema, e a segunda escala EUS não requer muito tempo para implementação (Brooke, 2013). Assim, esse instrumento fornece uma avaliação da percepção da usabilidade de um sistema em um curto espaço de tempo.

As medições única relatados Brooke (1996) eram de que esta escala tinham altos níveis de correlação entre os 10 itens que a compõem. Estes foram os $r = 0,7$ a $0,9$, correlação que serviu para seleccionar os itens que compõem a escala final, mas aqui nenhuma confiabilidade ou validade da escala foi relatado.

Um dos primeiros dados psicométricas foram relatados medições escala fiabilidade. Com uma participação de 77 pessoas, os resultados mostraram um coeficiente Alpha de 0,85 (Lucey, Kirakowski 1991, 1994). Então, em 2008, Bangor, Kortum e Miller aplicaram a escala a uma amostra maior, que consistia em 2324 e aqui descobrimos que o coeficiente Alpha foi .91, outra muito boa confiabilidade. Posteriormente, Lewis e Sauro (2009) fez 324 aplicações e obtiveram coeficiente Alpha de 0,92. Como pode ser visto, o EUS demonstrado desde o início têm muito bons coeficientes alfa, o que significa que apresenta uma boa fiabilidade.

Outra medida que era necessário saber se dentro dos fatores de escala ou em falta. O primeiro que realizou estas medições foram Bangor, Kortum e Miller (2008), na mesma amostra (2324) para a confiabilidade. Eles realizaram uma análise fatorial e encontrou um fator, eles chamaram Usabilidade geral (Lewis e Sauro, 2009; Borsci, Federici e Lauriola, 2009). Lewis e Sauro disse em 2009 que era provável que o grupo de Bangor não vê a possibilidade de esta escala pode ter mais de um fator.

Mais tarde, eles próprios (2009) fez uma investigação para conhecer a estrutura fatorial do EUS, com uma amostra de 324 sujeitos, que, em seguida, em comparação com os de Bangor et al (2008, Lewis e Sauro, 2009); Dessa forma, eles tentaram obter novos dados com a confiabilidade e validade da EUS. Nesta investigação, verificou-se que a escala existem dois fatores, e não um grupo de Bangor, disse. Daí o primeiro fator que corresponde a itens 1, 2, 3, 5, 6, 7, 8, 9, chamado de usabilidade; eo segundo fator para os itens 4 e 10, que chamou Facilidade de Aprendizagem (Lewis e Sauro, 2009).

Finalmente, Borsci, Federici e Lauriola (2009), com uma amostra de 196 participantes italianos, mostrou que os dois fatores (de usabilidade e facilidade de aprendizagem) são independentes; eles apresentaram um modelo menos restritivo com os fatores correlatos cuja informação se opôs à obra de Lewis e Sauro, que não encontraram tal correlação entre os fatores. O que eu observei foi que a aplicação da escala original (com 5 positivo e 5 itens itens negativos) tinha algumas desvantagens, e alguns participantes relataram no final que pontos negativos respondeu fortemente concorda ou mal interpretado o item. Aconteceu também que os pesquisadores codificado maus pontos negativos, que, em seguida, instaurado um inquérito para comparar a escala original (versão positivo e negativo) com a versão positiva (apenas os itens positivos), analisando os seguintes pontos: 1) se o questionário original aquiescência viés (tendência a responder sim a alguma coisa, independentemente do conteúdo da pergunta), 2) sabe quão grande esse viés e 3) se o desenvolvimento de uma escala com itens positivos podem ser removidos ocorre viés aquiescência e respostas extremas. Os resultados obtidos por esta pesquisa foram que não houve diferença significativa entre a escala original e positiva, de modo que a reformulação (transformar negativo para itens positivos) não afetou as medidas de alfa de EUS e Cronbach para ambas as versões (original e positiva), o qual foi elevado ($> 0,90$); Ele não foi encontrado na aquiescência de dados ou um forte viés nas respostas extremas. No entanto, se foi

apresentado o problema recorrente de má interpretação dos artigos pelos participantes e os erros de codificação pesquisador, causando viés nas respostas da escala.

Devido a isso, Lewis e Sauro (2011) apontou que você pode usar a versão positiva do EUS com confiança como as escalas positivos o usuário é menos propensos a cometer erros ao responder eo pesquisador não vai fazer um erro de codificação mas o mais importante é que dezenas de versão positiva são semelhantes aos padrões da versão original.

Por outro lado, Finstad (2006) realizou uma investigação que comparou duas amostras de 18 participantes cada; na primeira amostra que tinha falantes nativos de inglês e nativa mostra a segunda língua estrangeira (em russo, alemão, chinês, filipino, espanhol e hebraico). Verificou-se que não sejam o inglês nativos de língua esforçou para entender o item 8 ("Eu encontrei o sistema muito complicado de usar"), especialmente a palavra pesado, então pediu ajuda a compreendê-lo. Neste sentido, Finstad (2006) e Lewis e Sauro (2009) solicitou que a palavra Cumbersome por inábil, porque se os participantes não veio para entender bem o item poderia afetar os resultados devem ser substituídos. Finalmente, Finstad (2006) mencionou que a escala SUS não se deve aplicar em Inglês e eletronicamente para pessoas cuja língua nativa é outra, porque alguns termos induzir respostas erradas.

O nosso objectivo é adaptar a tanto a escala original em espanhol (5 positivo e 5 itens itens negativos) como a versão positiva; e investigar a confiabilidade das versões ambos os itens positivos e negativos.

Método

Participantes

A versão original do EUS (itens positivos e negativos) foi aplicado a 243 participantes, dos quais 129 eram do sexo feminino e 114 do sexo masculino, com uma idade média de 21 anos. Os participantes estudaram quatro raças: 121 Gestão de Sistemas de Engenharia (IAS), 4 Software Engineering Technology (ITS), 33 Administrativo Engenharia Mecânica (IMA) e 85 Licenciatura em Psicologia. A versão positiva para 173 participantes, 99 do sexo feminino e 74 do sexo masculino foi aplicado, com uma idade média de 21 anos. Eles estudaram duas corridas: 103 Sistemas de Gestão de Engenharia (IAS) e 70 graus em Psicologia.

Ferramentas

Na primeira fase do EUS original, composto por 5 elementos positivos e 5 pontos negativos (Brooke, 1996), apresentando uma confiabilidade 0,91 e validade de dois fatores (usabilidade e facilidade de aprendizagem) (Lewis e aplicado Sauro , 2009). Na segunda fase, foi aplicado a versão positiva EUS, que consiste de 10 itens positivos, com uma fiabilidade de 0,96 (Lewis e Sauro, 2012) e a mesma estrutura fatorial mencionado por eles.

o processo de tradução reversa foi usada, com a ajuda de três especialistas usabilidade bilingues para realizar a adaptação destas duas escalas; dois deles traduzido para o espanhol as duas escalas (original e versão positivo) e traduções, em seguida, cross-verificados. O terceiro perito traduzida as duas escalas traduzidas para o Inglês. Finalmente, os três especialistas verificados contra suas versões para chegar a um acordo, para obter as versões finais das duas escalas.

As duas traduções (a versão original e positivos) foram on-line em plataforma Forms Google, a fim de obter o maior número de participantes. uma escala de Likert de 5 níveis de resposta, que vão desde discordo totalmente (1) para concordar fortemente (5) foi utilizado.

algumas perguntas foram adicionados na seção de dados participante, que tinha a ver com a idade, tempo de experiência que eles estavam usando a plataforma e, finalmente, a corrida que eles estavam estudando (ver ambas as versões dos Anexos A e B).

Processo

Tanto a versão original e para estudantes positivos foram convidados a participar voluntariamente da pesquisa, informando-os de que suas informações pessoais serão tratadas de forma confidencial. O convite foi feito pessoalmente pela rede social Facebook, bem como e-mail. Através da mídia eletrônica que foi enviado um endereço web a partir do qual eles poderiam acessar a plataforma Google Forms, onde as duas escalas foram avaliados. Os participantes tomaram menos de cinco minutos para responder tanto a versão original ea versão positiva.

Resultados

Na versão original do coeficiente EUS Alpha Cronbach de 0,59 foi obtido, enquanto o teste de Bartlett de esfericidade veio significativo $p < .001$ e índice de Kaiser-Meyer-Olkin (KMO) foi 0,84, portanto, procedeu para realizar a análise fatorial exploratória (AFE). Isto mostrou que três

factores e que a percentagem de variância explicada foi de 62,80%. No fator 1 foram os itens 1, 3, 5, 7 e 9, ou seja, os itens positivos, enquanto o fator de dois itens 8, 4 e 10 foram, e fator de 3 itens 2 e 6 . Ou seja, os pontos negativos foram distribuídos entre os fatores 2 e 3.

Dessa forma, procedeu-se a fazer uma AFE, o que resultou em dois fatores, como os pontos de literatura. percentagem de variância explicada de 52,78, menor do que o primeiro AFE Os seguintes resultados foram encontrados. Neste fator de análise 1 itens todos positivos foram colocados em fator 2 e todos os itens negativos.

Na versão positivo de EUS um Coeficiente de Cronbach alfa de 0,92, enquanto que o teste de Bartlett de esfericidade para fora significativa $p < 0,001$ e índice de KMO era 0,90, o que significa que é possível realizar a análise foi obtido fatorial. Em seguida, ele passou a fazer a AFE, sendo um fator e uma percentagem de variância explicada de 58,07%. A Figura 1 mostra a sedimentação de factores e Tabela I mostra os pesos dos factores.

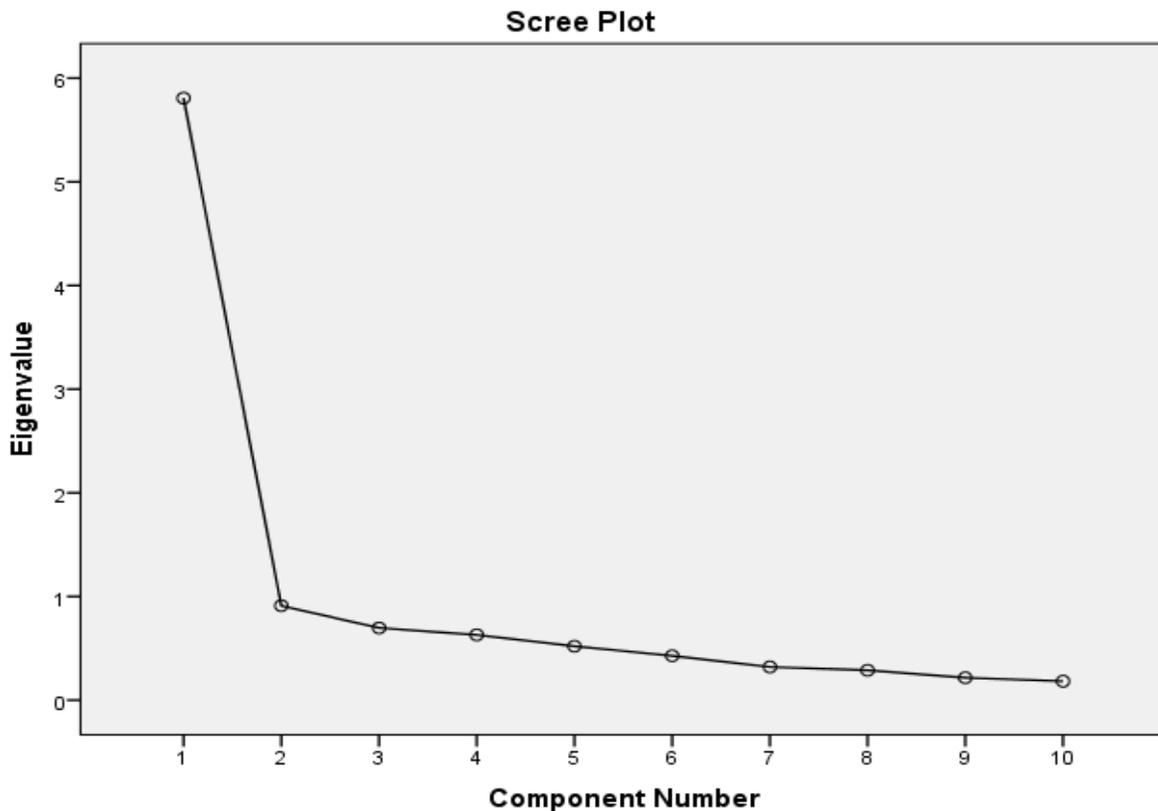


Figura 1. Graph versão sedimentação EUS positiva.

Tabla I. Componentes da matriz da versão positiva do ZEUS com as respectivas cargas fatoriais.

Matriz de Componentes	
	<u>Componentes</u>
	<u>1</u>
1	.668
2	.832
3	.818
4	.660
5	.838
6	.694
7	.820
8	.805
9	.796
10	.653

Finalmente, a AFE foi feito, forçando a análise a dois fatores como pontos de literatura. Temos uma tentativa para verificar a sua existência. A AFE descobriram que, se os fatores acima (Figura 1), com uma variância explicada de 67,18%, é ainda mais justificado forçou a AFE a dois fatores têm. os dois factores com os seus respectivos elementos (Tabela II) foram também encontradas.

Tabla II. Versão positiva matriz componente rodada da EUS.

Matriz de Componentes Rotados		
	Componentes	
	1	2
1	.821	-.007
2	.736	.398
3	.675	.462
4	.234	.817
5	.730	.419
6	.660	.268
7	.723	.297
8	.760	.317
9	.594	.539
10	.216	.829

No fator 1 (usabilidade), o coeficiente Alfa de Cronbach foi 0,92, enquanto que o fator 2 (capacidade de aprendizado) o coeficiente Alfa de Cronbach foi 0,70.

Assim, parece que a versão positiva de EUS mostra melhores resultados na variância explicada e colocando os itens com os respectivos factores, como mencionado na literatura.

Discussão e conclusões

No que diz respeito à versão positiva do EUS, este foi muito boa confiabilidade interna (0,92) e também mostrou que a AFE tem boa percentagem de variância explicada e os dois fatores mencionados literatura, que são de usabilidade (itens 1, 2, 3, 5, 6, 7, 8 e 9) e a facilidade de aprendizagem (4 e 10). Como se observou nos resultados, na versão original dos EUS não era bom coeficiente de Alpha, mas boa pontuação no teste de Bartlett de esfericidade e o índice de KMO para a AFE, jogando três fatores. Ainda que não tenha feito qualquer trabalho de investigação sobre este relatório escala que três fatores (Lewis e Sauro, 2009; Borsci, Federici e Lauriola, 2009) são apresentados. Uma das razões pelas quais o coeficiente de Cronbach Alpha não era o mesmo nesse tipo de trabalho, pode ser a inclusão de itens negativos ou que os participantes acharam difícil de entender. Por exemplo, Lewis e Sauro (2011) mencionou em sua pesquisa que esses itens negativos pode haver erros na resposta e erros de codificação.

Acreditamos que especialmente gera erros na resposta, como havia participantes mencionaram ter cometido tal erro.

Outro ponto a considerar é designado por Lewis e Sauro (2011), que argumentam que as duas versões do EUS pode ser usado porque ambos têm altos níveis de consistência interna, mas também que é melhor usar a versão positiva do EUS para evitar problemas mau erros de codificação em responder ou tem a versão positiva semelhante às partituras originais. No entanto, em nosso estudo não conseguimos obter esse efeito sobre a versão original; em vez suporta dados e Frye Stewart (2004), que descobriu que a inclusão de itens negativos e positivos tem baixa consistência interna. Enquanto isso, Pilotte e Gable (1990) e Schmitt e Stuitts (1985) mostrou que o uso de itens mistos distorce a estrutura fatorial, algo que aconteceu na versão original aplicada.

Um fato importante é que Lewis e Sauro (2009) relataram coeficientes alfa para cada um dos fatores, o fator de 0,91 usabilidade e facilidade de aprendizagem 0,70. Nosso estudo produziu os mesmos resultados (usabilidade e facilidade de aprendizagem 0,92 0,70), de acordo com a literatura, com a única diferença de que Lewis e Sauro (2009) utilizaram o questionário original e EUS-nos a versão positiva.

Em suma, o estudo mostrou que a versão positiva do EUS é melhor do que a versão original, para ser mais confiável em termos estatísticos e têm validade de construção adequada. Por isso, fazemos um convite para a versão positivo em vez de negativo ao avaliar as diferentes plataformas web usados.

Bibliografía

- Bangor, A., Kortum, P., & Miller, J. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24, 574-594. doi: 10.1080/10447310802205776.
- Borsci, S., Federici, S., & Lauriola, M. (2009). On the dimensionality of the System Usability Scale: a test of alternative measurement models. *Cognitive processing*, 10(3), 193-197. doi: 10.1007/s10339-009-0268-9.
- Brooke, J. (1996). SUS: A Quick and Dirty Usability Scale. In: P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability Evaluation in Industry*. London: Taylor & Francis.
- Brooke, J. (2013). SUS: A Retrospective. *Journal of Usability Studies*, 8(2), 29-40.
- Finstad, K. (2006). The System Usability Scale and Non-Native English Speakers. *Journal of Usability Studies*, 1(8), 185-188.
- ISO 9241-11 (1998). Ergonomic requirements for office work with visual display terminals (VDTs). Recuperado el 12 de abril de 2016, de <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en>
- Kirakowski, J. (1994). The use of questionnaire methods for usability assessment. Unpublished manuscript. Recuperado el 12 de abril de 2016, de <http://sumi.ucc.ie/sumipapp.html>
- Lewis, J., & Sauro, J. R. (2009). The Factor Structure of the System Usability Scale. *In Human Centered Design*. doi: 10.1007/978-3-642-02806-9_12.
- Lewis, J., & Sauro, J. R. (2012). Quantifying the user experience: Practical statistics for user research. Elsevier, USA. doi:10.1016/B978-0-12-384968-7.00001-1.
- Lewis, J., & Sauro, J. R. (2011). When Designing Usability Questionnaires, Does It Hurt to Be Positive? *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. doi: 10.1145/1978942.1979266.
- Pilotte, W. J., & Gable, R. K. (1990). The impact of positive and negative item stems on the validity of a computer anxiety scale. *Educational and Psychological Measurement*, 50(3), 603-610. doi: 10.1177/0013164490503016.
- Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, 9(4), 367-373. doi: 10.1177/014662168500900405.

Stewart, T. J., & Frye, A. W. (2004). Investigating the use of negatively phrased survey items in medical education settings: common wisdom or common mistake? *Academic Medicine*, 79(10), 18-20. doi: 10.1097/00001888-200410001-00006.

Apéndice A

Escala de Usabilidad del Sistema (versión original)

	Totalmente en desacuerdo			Totalmente de acuerdo		
1. Creo que me gustaría utilizar frecuentemente este sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Encontré el sitio web innecesariamente complejo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Pienso que el sitio web es fácil de usar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Creo que necesitaré el apoyo de personal técnico para poder utilizar este sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Encontré que varias de las funciones en el sitio web estaban bien integradas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Pensé que había demasiada inconsistencia en este sitio.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Me imagino que la mayoría de las personas podrían aprender a usar este sitio web muy rápido.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Encontré el sitio web muy difícil de usar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Me sentí muy confiado (seguro) al utilizar el sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Necesité aprender muchas cosas antes de poder usar este sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Apéndice B

Escala de Usabilidad del Sistema (versión positiva)

	Totalmente en desacuerdo			Totalmente de acuerdo		
1. Creo que me gustaría utilizar frecuentemente este sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Encontré el sitio web sencillo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Pienso que el sitio web es fácil de usar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Pienso que podré utilizar este sitio web sin el apoyo de personal técnico.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Encontré que varias de las funciones en el sitio web estaban bien integradas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Pensé que había demasiada consistencia en el sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Me imagino que la mayoría de las personas podrían aprender a usar este sitio web muy rápido.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Encontré el sitio web muy intuitivo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Me sentí muy confiado (seguro) al utilizar el sitio web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Pude utilizar el sitio web sin tener que aprender nada nuevo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>