Usability of the mixture building section of Wolffia

Melissa López
University of Puerto Rico
Río Piedras Campus
San Juan, Puerto Rico
melissa.lopez3@upr.edu

ABSTRACT

We are presenting an early phase of a usability evaluation of *Wolffia*: an open-source application with Graphical User Interface (GUI) to prepare and monitor classical molecular dynamics (CMD) simulations. For this purpose, a modified cognitive walkthrough evaluation of the *Build Tab* section was performed. The amount of time required to perform tasks after a fifteen minutes tutorial by four high school students was measured. In order to measure user satisfaction, a Standard Usability Score questionnaire was also administered. The users were able to perform most of the functionalities without a significant amount of errors and were able to finish the set up of a mixture in a short amount of time. Finally, the user satisfaction from using the system was very well rated.

INTRODUCTION

Wolffia [8] is an open-source application (shown in figure) with Graphical User Interface (GUI) to prepare and monitor classical molecular dynamics (CMD) simulations, the simulation of the movement of atoms whose interactions are defined by classical mechanical forces. The need for this type of applications has been mentioned in CMD review papers that "*One of the more time-consuming aspects of conducting a molecular simulation is generating an initial equilibrated system.*"[7] Assembling simulations is a difficult and long process in part because of the amount of steps needed to complete each part of the process. The user needs to use a combination of different applications designed for a specific phase of CMD simulations, and even sometimes the user needs to manually modify a lot of files with hundreds or thousands of lines in which the coordinates, type and force fields parameters of thousands of atoms in the system needs to be specified.

With the use of *Wolffia*, the user can easily set up the simulation, without having to modify any file, if desired. Its user friendly design was inspired in the separation of phases required for preparing and running a simulation. Each part of the process is divided in major steps, which suggest an order to a novice user, but at the same time, an experienced user can navigate through the interface as needed, provided that the required actions does not interfere with current ones. The interface is organized based on five stages of the preparation of CMD simulations: building the system, defining force field parameters, defining the periodic boundaries conditions, minimizing energies and the simulation. It addresses the need of integrated and easy to use CMD setup tools.
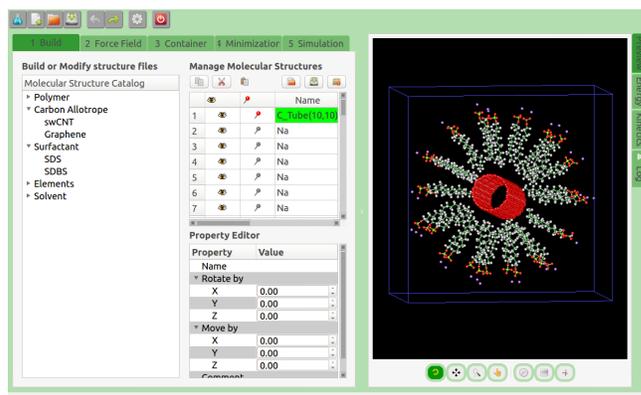

Figure 1. The environment at the system build stage.

Wolffia has been actively used by our group and collaborators at the UPRH for research projects. For example, A. de Jesús, uses Wolffia to model the movement of ions in activated carbon varying the pore size, in order to find a way to measure the capacitance of an electrical double layer capacitor [4]. Simulations assembled using Wolffia also provided insight on the dispersion of carbon nanotubes in polymers by I. Ramos' group at the Physics and Electronics Department in the University of Puerto Rico at Humacao (UPRH) [1].

Although the use of *Wolffia* for research has been successful so far, in this report we present an early phase of an usability evaluation with the aim of measuring the strengths and the weakness of *Wolffia* on novice users. The evaluation is based in well established principles in the field of Human-Computer Interaction (HCI). HCI is "*an area of research and practice that emerged in the early 1980s, initially as a speciality area in computer science embracing cognitive science and human factors engineering*" [3] . Intensive research in HCI has greatly influenced to the dramatic simplification of interfaces. For this evaluation we performed an analysis of the setting-up process of the mixture, which is the first step of a molecular dynamics simulation.

OBJECTIVES

Our goal was to identify the strengths and the weakness of *Wolffia* for novice users not necessarily familiar with CMD.

We wanted to measure some of the usability dimensions which, according to Nielsen [10], are defined as

learnability: how easy is to learn the attributes of a system in order to start using it,

efficiency: how efficient is the system to use,

errors: what is the ability of the system to minimize the number of errors a user makes while using the system and their recovery speed after making a mistake and

satisfaction: what is the subjective users' satisfaction from using the system.

EVALUATION METHODS

To perform the analysis, 4 high school students (2 from tenth grade and 2 from eleventh grade) participated in the study. We decided to use this amount of users based on results presented by J. Nielsen and T. K. Landauer[5], in which they established that five is a good number for user-based qualitative evaluations.

The participants were part of the *Experimenta con PREM* (ECP) celebrated each summer at UPRH which consists of workshops on materials science and nanotechnology as well as four laboratory experiences. Therefore, based on how the workshop was organized, we decided to use four participants for this early phase of the evaluation. CMD simulations is one of the laboratory experiences. Participants of ECP are selected by UPRH-PREM faculty among a pool of applicants from schools in the south-eastern region of Puerto Rico. Acceptance rate is around 50%. Parent consent to the student participation was obtained during the ECP admission process. The consent was not a requisite for admission nor it was considered as a factor in applicant evaluation.

Five days previous the evaluation, the participants received an introductory workshop on molecular dynamics simulations by an expert in the field from the University of Pennsylvania, our PREM partner institution. As part of the CMD laboratory experience, minutes previous to the evaluation, they also received a 15 minutes introduction of the main aspects of *Wolffia* and had the opportunity to familiarize with the GUI.

Figure 2. The laboratory where the evaluation was made. A summary of the experimental procedure is being delivered.

Experimental procedure

The experimental procedure followed by the students consisted of performing a simulation of a single wall carbon nanotube (CNT) surrounded by 35 molecules of the sodium dodecyl sulfate (SDS) surfactant. For purposes of our analysis, we only evaluated when users set up the first layer of 7 SDS molecules. The aim of the initial setup was to build a system similar to the one shown in figure 1. The evaluation was performed in our research laboratory. Figure 2 shows the moment when the general explanation of the experimental procedure was made and the laboratory setting where the usability evaluation was made. The instructor gave to the participants a quick demonstration of the experimental procedure that can be summarized as follows:

1. Select a nanotube of length 20 Å from the molecules catalog.

2. Use the *pin widget* to fix the coordinates of the atoms of the nanotube.

3. Using the load dialog, import a SDS molecule from the NCI-CADD [9] repository.

4. Select the complete molecule (because it consist of two parts: a chain and a sodium ion) and place it perpendicular to the surface of the nanotube with the polar side of the SDS facing outwards.

5. Copy and paste the selected SDS molecule until the nanotube was surrounded by 7 SDS molecules (when paste action is performed, only the pasted molecule remains selected).

Users had access to a detailed version of this experimental procedure as an online document during the evaluation process as shown in figure 3. The actual text (in spanish) can be found in appendix 1 below.
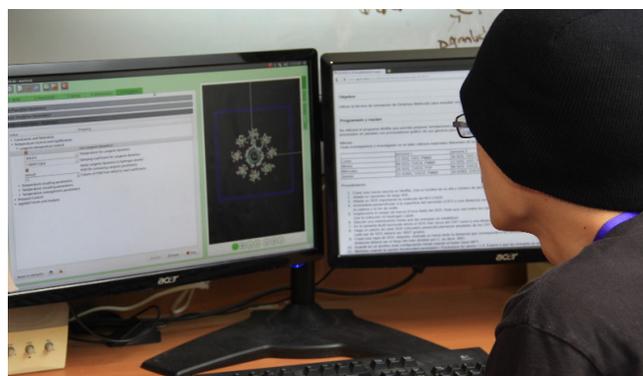
Figure 3. A user using *Wolffia* (left screen) and the experimental procedure.

A video of the computer desktop was recorded after the user's consent during the session. Users had access to the experimental procedure enumerated above as an online document during the evaluation process. The analysis was made using some observations taken at the time of the evaluation and observing the videos mentioned above. A satisfaction questionnaire was provided to the participants after the completion of the evaluation.

Usage times and errors

Tasks

In our evaluation there were a total of 5 tasks (described in the previous section). Because defaults values are used for the first task, and in combination with the second task it can be performed very fast (aprox. 2-3 seconds both), for purposes of our study we decided to evaluate the remaining tasks, which are the most complicated ones.

Time

The time for each task was measured from the desktop video recorded at the session. We measured the time it took users to complete the three tasks evaluated (all together) and also, we measured the time taken for completing tasks individually.

Errors

Errors were divided in three categories: slips and mistakes. *Slips* are those errors that were made because the steps outlines in the procedure were not followed (e.g. did not moved the sodium ions). *Mistakes* are those errors in which the goal is not reached (e.g. did not remembered how to rotate a molecule). The term *total errors* refers to the sum of slips and mistakes at a certain task.

We also performed a modified version of a *cognitive walk-through* evaluation. In this evaluation, "an evaluator, preferably a usability expert evaluates a user interface by analysing the cognitive processes required for accomplishing tasks that users would typically carry out supported by the computer"[6]. In other words, an evaluator simulates a novice user and explores all the possible options that a user has in order to accomplish the explicitly guided task. The modification to the traditional approach consisted in defining some tasks and instead of simulating a novice user, we evaluated if the participants accomplished the tasks. For each step, four criteria were formulated:

(a) the user achieved the correct effect,

(b) the user noticed that the correct action is available,

(c) the user associated the correct action with the desired effect, and, if the user performed the right action, and

(d) the user noticed that progress is being made toward accomplishment of his goal.

A step has no usability problems if the answer of all these four criteria were satisfied [6].

User satisfaction

The *Standard Usability Score* (SUS) questionnaire, a recognized and validated mature questionnaire developed by J. Brooke [2], was used for measuring the satisfaction of the participants. It consist of 10 items in a likert 5-point scale. The obtained values are scaled from 0 to 4 and then converted to the 0 to 100 range instead of from 0 to 40 by multiplying the scores by 2.5. According to J. Sauro, an average score above 68 would be considered above the average SUS score, calculated from 500 previous usability studies [11].

User	Time (sec)	Slips	Mistakes
1	409	1	0
2	767	1	0
3	439	0	1
4	740	2	0

**Table 1. Time to accomplish all tasks and errors committed**
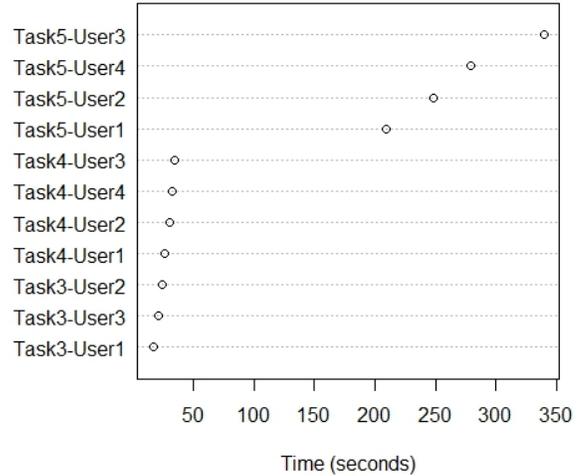
Figure 4. Amount of time it took users to complete their tasks.

RESULTS AND DISCUSSION

Usage times and errors

Table 1 shows the total time that the participants needed to complete the tasks that we were evaluating as well as the amount of slips and mistakes made by the participants. With these data we measured the efficiency of the system. It can be observed that a novice user, on average, can set up a simple mixture in 565 seconds (10 minutes) with a 95% confidence interval (332.17 ,961.04) seconds (5 to 16 minutes). The evaluation revealed that the average for both slips and mistakes is one. From table 2 we can observe the average time that users took for each task. The first two tasks were performed faster than the third one, and if we add them up, both are performed in less than a minute (see figure 4). The task 5 is performed in approximately 5 minutes, but the users were able to perform it without committing any error.

J. Sauro and J. Lewis [12] established that "*aggregating data from an usability test had a significant effect on the magnitude of the resulting correlations*" using task-level aggregation and test-level aggregation. Therefore, we wanted to analyze the

Task	TimeMean	Slips	Mistakes	Errors
3	20.0	1	1	2
4	30.2	2	1	3
5	265.0	1	0	0

Table 2. Time's geometric mean, slips, mistakes and total errors per task.

Task	User	Time (sec)	Comp	Mistakes	Slips	Errors
3	1	16.5	1	0	0	0
3	2	23.5	1	0	1	1
3	3	20.6	1	0	0	0
3	4	–	0	1	0	1
4	1	26.2	1	0	1	1
4	2	30	1	0	0	0
4	3	33.8	1	1	0	1
4	4	31.5	1	1	0	1
5	1	208.8	1	0	0	0
5	2	248.9	1	0	0	0
5	3	339.8	1	0	0	0
5	4	279.1	1	0	0	0

Table 3. Detailed description of the time, completion rates, slips, mistakes and the total number of errors committed by the participants.

	Criterion			
Task	a	b	c	d
User 1				
3	Y	Y	Y	Y
4	Y	Y	Y	Y
5	Y	Y	Y	Y
User 2				
3	Y	Y	Y	Y
4	Y	Y	Y	Y
5	Y	Y	Y	Y
User 3				
3	Y	Y	Y	Y
4	Y	N	N	Y
5	Y	Y	Y	Y
User 4				
3	N	N	N	Y
4	Y	N	N	Y
5	Y	Y	Y	Y

Table 4. Cognitive evaluation results

Users	SUS score	Converted score
1	40	100
2	36	90
3	37	92.5
4	37	92.5
Mean	37.5	93.7

Table 5. SUS questionnaire results to evaluate users' satisfaction.

results of the evaluation in general, but also try to analyze in more detail what happened at the task-level. We constructed table 3 in which it can be observed in more detail the amount of time that each user took to complete each task, if it was completed, the amount of slips, mistakes and the total of errors (the sum of slips and mistakes) that were made. Note that there is a missing value, which correspond to a user that did not followed the outlined step. The user imported the SDS molecule from the molecule catalog from *Wolffia*, instead of importing it from the repository. Therefore, because only one task could not be completed, the completion rate for all tasks is 95%, meaning that we can expect that 95% of future users can complete succesfully all tasks.

From the cognitive walkthrough analysis (table 4), we found that two users were able to complete all tasks without any usability problem. We can also see, that although task 5 is the longest, it was the only one that users performed without any error. An error found in the first task was that one user imported the molecule from the catalog instead from the repository. Another problem found was that one user did not selected the NCI-CADD repository and used the default one (Protein Data Bank). Maybe a solution to the previous two errors could be to incorporate a search function into the Build Tab section, and when a user wants to add a molecule either from the repositories or from the catalog of molecules, *Wolffia* show the available adding options to the user. For the second task, two of the users had problems when they were trying to move the SDS molecule. Neither recalled how to move the molecule, and we had to explain them so they could finish their work. One of them moved the chain but did not move the ions, and had to move it after having moved the chain. A solution to this problem could be giving the option to the user to "merge" or "group" the molecules. As seen, the problems found in this phase were minors, and users were able to recover easily from all the errors they had made.

User satisfaction

The table 5 shows our findings of user satisfaction. The average score was 93.7 with a 95% confidence interval equal to (87.1,100). We calculated the percentile rank associated with our average score, and our raw SUS score of 93.7 has a higher SUS score than 86.7% of all products from 500 studies [11].

CONCLUSIONS

With a 15 minutes tutorial the users could use *Wolffia* without major difficulties. The users were able to recall most of the functionalities and without a significant amount of errors, were able to finish the set up of a simple mixture in an average time of 10 minutes. Furthermore, the satisfaction questionnaire revealed a good users' satisfaction from using the system. Finally to conclude, we would like to refer again to E. Maggin [7] because he makes clear the current need to decrease the time required to assemblage a simulation, as he emphasizes that today " *setting up a particular calculation can often take longer than the simulation itself*" and in this evaluation we showed that with the use of *Wolffia* it can just take 5 to 16 minutes (time required to stabilization not included) to assemblage a system that to simulate a nanosecond could take all night.

FUTURE WORK

This was an early phase of an usability evaluation. We would like to repeat the evaluation, including a greater number of users, in order to obtain some quantitative measures such as correlation between time and errors. We also woul like to fix the found problems, and to extend the evaluation for the remaining phases of CMD simulations that can be performed in *Wolffia*.

REFERENCES

1. X. Barbosa, S. Rosa, V. Capel, F. Martínez, R. Quiñones, W. Cuadrado, I. Ramos, and J. Sotero-Esteva. A study of the effect of carbon nanotubes dispersion on the properties of polymer composite nanofibers. In *Proc. Int. Conf. Adv. Nano Mat.*, 2013.

2. J. Brooke. SUS: A 'quick and dirty usability scale. usability evaluation in industry. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L.McClelland, editors, *Usability evaluation in industry*, pages 189–194. 1996.

3. J. Carroll. Encyclopedia of human-computer interaction. electronic version, 2009. Retreived from: `http://www.interactiondesign.org/encyclopedia/human_computer_interaction_hci.html`.

4. A. de Jesús. Smulation of ion movements in activated carbon-based supercapacitors. Symposium on PREM-related Research, 2013.

5. T. K. Landauer J. Nielsen. A mathematical model of the finding of usability problems. In *Proc. INTERCHI '93 Conf. Hum. Fact. Comp. Syst.*, 1993.

6. M. Jaspers. A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. 78(5):340353, 2009.

7. E. J. Maginn. From discovery to data: What must happen for molecular simulation to become a mainstream chemical engineering tool. *AIChE J.*, 55(6):1304–1310, 2009.

8. Mirgery Medina-Cuadrado, Carlos Cortés-Martínez, Melissa López, Frances Martínez-Miranda, and Radamés-Vega Alfaro. An implementation of a graphical interface for molecular simulation software using Python and Qt4. *Proceedings of The National Conference On Undergraduate Research (NCUR) 2012*, pages 554–563, March 2012.

9. NCI/CADD group. NCI/CADD. Molecule database at `http://cactus.nci.nih.gov/`.

10. J. Nielsen. *Usability Engineering*. Morgan Kaufmann, San Francisco, California, 1993.

11. J. Sauro. Measuring usability with the system usability scale (sus).

12. J. Sauro and J.R. Lewis. Correlations among prototypical usability metrics: Evidence for the construct of usability.

APPENDIX 1: EXPERIMENTAL PROCEDURE

The following pages show the experimental procedure as provided to the students (in spanish).

Procedimiento experimental

Objetivo

Utilizar la técnica de simulación de Dinámica Molecular para estudiar cómo interactúa un polímero conductor con un nanotubo de carbón rodeado de un tensoactivo.

Programado y equipo

Se utilizará el programa Wolffia que permite preparar simulaciones para ser ejecutadas con el programa NAMD. Las computadoras utilizan el sistema operativo Ubuntu. Se utilizará un procesador en paralelo con procesadores gráfico de uso general para ejecutar las corridas de producción.

Método
Cada investigadora o investigador en el taller utilizará materiales diferentes (N-tensoctivos, solvente, polímero) en la simulación de acuerdo a la siguiente tabla.

	persona 1	persona 2	persona 3	persona 4	persona 5
Lunes	12-SDS, H2O, PMMA	28-SDS, H2O, PMMA	88-SDS, H2O, PMMA	12-SDS, CHCl3, PMMA	28-SDS, CHCl3, PMMA
Martes	88-SDS, CHCl3, PMMA	12-SDS, H2O, PVP	28-SDS, H2O, PVP	88-SDS, H2O, PVP	12-SDS, CHCl3, PVP
Miércoles	28-SDS, CHCl3, PVP	88-SDS, CHCl3, PVP	12-SDBS, H2O, PMMA	28-SDBS, H2O, PMMA	88-SDBS, H2O, PMMA
Jueves	12-SDBS, CHCl3, PMMA	28-SDBS, CHCl3, PMMA	88-SDBS, CHCl3, PMMA	12-SDBS, H2O, PVP	28-SDBS, H2O, PVP

Procedimiento:

1. Crear una nueva mezcla en Wolffia. Use el nombre de su día y número de persona como nombre de la mezcla (Ejemplo: Lunes3).
2. Añada un nanotubo de largo 40A.
3. Añada un SDS importando la molécula del NCI-CADD.
4. Acomódela perpendicular a la superficie del nanotubo (CNT) a una distancia como la mitad del largo de la molécula de SDS con la parte polar haca afuera. No olvide mover las dos partes: la cadena y el ión de sodio.
5. Inspeccione el campo de fuerza (Force field) del SDS. Note que casi todos los campos tienen cero. Encuentre valores primero *detectando los tipos de átmos*, luego *buscando* los valores. Use la colección *Al-Hydrogen Lipids*.
6. Ejecute una minimización hasta que las energías se estabilicen.
7. En la pestaña *Build* acomode ahora el SDS mas cerca del CNT como a una distancia similar a dos veces el largo de un enlace.

8. Haga un patrón de siete SDS colocados perpendicularmente alrededor de los CNT. Considere que si una vuelta completa es un ángolo de 360 grados, el ángulo que tendrán que hacer cada par de SDS deberá ser 360/7 grados.
9. Copie esa capa de SDS, péguela, muévala un hacia atrás la distancia que corresponda a su cantidad de SDS. Habrá C = N/7, donde N es su cantidad final de polímeros. Por tanto, la distancia deberá ser el largo del tubo dividido por C, es decir, 40/C.
10. Guarde en un archivo esta configuración inicial usando el botón *Save WFY*.
11. Minimice usando la opción *Nonbonded parameters / Exclusions for atoms / 1-4*. Espere a que las energías se estabilicen. Las configuraciones con más moléculas tardarán mas.
12. Ejecute una simulación con 100,000 pasos.
13. (receso de merienda).
14. En la pestaña *Set-up*, cree una caja de dimensiones del triple de la caja mínima en las direcciones perpendiculares al CNT y súmele 4 a la paralela al CNT de modo que el CNT queda casi justo en la caja.
15. Añada un polímero con 10 monómeros en el espacio disponible paralelo al CNT.
16. Haga copias de ese polímero y acomode alrededor hasta llegar a 8.
17. Minimice 10,000 pasos.
18. Añada el solvente que le corresponda.
19. Minimice hasta que estabilice.
20. Corra una simulación lo mas larga posible mientras quede tiempo de taller (ponga 10,000,000 de pasos como límite).

APPENDIX 2: SATISFACTION QUESTIONNAIRE

The following pages show the satisfacion questionaire as administered to the students (in spanish).

Evaluación de Wolffia

Usuario:_____   Fecha:_____

Computadora:_____

Instrucciones: Este cuestionario esta diseñado para medir la satisfacción del usuario con respecto al uso de Wolffia. El mismo es voluntario, pero sus respuestas nos ayudarán a medir la usabilidad de Wolffia. Sus respuestas son totalmente confidenciales. Seleccione la contestación mas adecuada para cada pregunta. Por favor responda todas las preguntas.

1. Pienso que podría utilizar Wolffia frecuentemente.

TOTALMENTE   **TOTALMENTE**
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **DE ACUERDO**

Comentarios:

2. Encontré Wolffia innecesariamente complejo.

TOTALMENTE   **TOTALMENTE**
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **DE ACUERDO**

Comentarios:

3. Pienso que Wolffia fue fácil de usar.

TOTALMENTE   **TOTALMENTE**
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **DE ACUERDO**

Comentarios:

4. Pienso que necesitaría soporte técnico para poder utilizar Wolffia.

TOTALMENTE   **TOTALMENTE**
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **DE ACUERDO**
Comentarios:

5. Encontré que la mayoría de las funciones en Wolffia estaban bien integradas.

TOTALMENTE
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **TOTALMENTE**
                                                                                          DE ACUERDO
Comentarios:

6. Pienso que había mucha inconsistencia en Wolffia.

TOTALMENTE
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **TOTALMENTE**
                                                                                          DE ACUERDO
Comentarios:

7. Pienso que la mayoría de las personas podrían aprender a utilizar Wolffia rápidamente.

TOTALMENTE
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **TOTALMENTE**
                                                                                          DE ACUERDO
Comentarios:

8. Encontré que Wolffia es muy incómodo de usar.

TOTALMENTE
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **TOTALMENTE**
                                                                                          DE ACUERDO
Comentarios:

9. Me sentí muy confiado utilizando Wolffia.

TOTALMENTE
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **TOTALMENTE**
                                                                                          DE ACUERDO
Comentarios:

10. Necesité aprender muchas cosas antes de poder utilizar Wolffia.

TOTALMENTE
EN DESACUERDO   **1**   **2**   **3**   **4**   **5**   **TOTALMENTE**
                                                                                                DE ACUERDO
Comentarios:

11. Wolffia provee la información necesaria para poder completar mi trabajo. (Ej. Mensajes en la
    pantalla, documentacion)
    1. Si
    2. No
Comentarios:

12. Pude recordar fácilmente como utilizar Wolffia.
    1. Si
    2. No
Comentarios:

13. En caso de haber cometido un error, pude recuperar mi trabajo fácilmente.
    1. Si
    2. No
Comentarios:

14. En caso de haber cometido un error, los mensajes de Wolffia eran claros.
    1. Si
    2. No
Comentarios:

15. En general, estoy satisfecha/o con el uso de Wolffia.
    1. Si
    2. No
Comentarios:

16. Antes de utilizar Wolffia por primera vez, tenía experiencia en simulaciones de dinámica molecular.
    1. Si
    2. No
Comentarios: