

Using the Speed-Accuracy Operating Characteristic to Visualize Performance with Pointing Devices

Richard C. Simpson*

Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA, USA

Human Engineering Research Labs, VA Pittsburgh Healthcare System, Pittsburgh, PA, USA

Abstract: There is a nearly limitless variety of alternatives to the traditional computer mouse, including trackballs, trackpads, joysticks, head-mounted mouse emulators, and eye-gaze systems, each with its own unique strengths and limitations. Two key challenges in selecting the most appropriate pointing device are (1) collecting accurate, comparable data describing the client's performance with each device under consideration and (2) aggregating, analyzing and displaying the data in a format that allows the client, clinician and other interested parties to make an accurate decision. Tools exist for collecting data, but there has been limited research on how to visualize and compare performance data. The investigator has been exploring the use of the speed-accuracy operating characteristic (SAOC) as a means of comparing performance on computer pointing tasks between different pointing devices (e.g., mouse, trackball, touchpad). The SAOC is a graphical representation of the trade-off between speed and accuracy that has been used in a range of timed reaction tasks, including pointing tasks. The SAOC is easy for clinicians, clients and third-party payers to interpret, and provides an efficient means of summarizing and documenting performance but does not replace clinical judgement or eliminate the need for clinician-client interaction.

1. INTRODUCTION

Choosing the most appropriate assistive technology (AT) is a collaborative decision-making process involving the client, family members, clinician(s), and third-party payers. Clinical experience and AT outcomes studies [1-4] tell us that the process is, at best, imperfect and open to improvement. One avenue for improvement is through application of the lessons learned from decision-making research. In particular, there is a general need for tools to support accurate decision-making as part of the AT evaluation process [5].

consideration and (2) aggregating, analyzing and displaying the data in a format that allows the client, clinician and other interested parties to make an accurate decision [5].

As shown in Table 1, several tools are now available for collecting quantitative performance data with a pointing device. These tools are distinguished from others in that they also record the point-by-point movement of the cursor. Some of these programs were developed specifically for use in assistive technology assessments and others were developed by the human factors community for use in experiments. While it is possible to collect data in the clinic, the challenge

Table 1. Data Collection Software

Compass [6]	http://www.kpronline.com/compass-intro.html
Movement Time Evaluator [7]	http://research.cathris.com/mte.html
Cursor Motion Test	http://www.almaden.ibm.com/u/basmith/testapp.html

This paper focuses on the specific problem of choosing the most appropriate pointing device for computer access. There is a nearly limitless variety of alternatives to the traditional computer mouse, including trackballs, trackpads, joysticks, head-mounted mouse emulators, and eye-gaze systems, each with its own unique strengths and limitations. Two key challenges in selecting the most appropriate pointing device are (1) collecting accurate, comparable data describing the client's performance with each device under

of visualizing the data in a useful form remains.

The format in which information is displayed can impact which choice is made [5] because people tend to make worse decisions when they have to work harder to interpret the underlying data. A graphical representation can help just by formatting and summarizing data in a way that allows direct comparisons between multiple devices. This paper demonstrates how the Speed-Accuracy Operating Characteristic (SAOC) [8-10] can be used to provide such a graphical representation.

*Address correspondence to this author at the Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA 15260, USA; Tel: 412-383-6593; Fax: 412-383-6597; E-mail: ris20@pitt.edu

2. BACKGROUND

2.1. Measuring Performance on Pointing Tasks

Numerous investigators (e.g., [11-13]) have documented the performance of individuals with disabilities when performing point-and-click tasks on a computer. Across investigators, however, there have been large differences in task implementation and performance measures. These differences make it difficult to compare results between studies [14], reducing their clinical utility.

The human factors community moved to address this problem by incorporating a specific evaluation protocol for pointing devices into the ISO9241-9 standard [15]. The standard specifies a circular pointing task and the use of throughput (measured in bits/second) as the dependent measure compared between devices. Throughput is given by the formula [14]:

$$TP = \frac{ID}{MT} \quad (1)$$

where MT is movement time and ID is index of difficulty, given by the formula [14]:

$$ID = \log_2 \left(\frac{D}{W} + 1 \right) \quad (2)$$

where D is the distance between the cursor's starting position and the center of the target and W is the target's width. Soukoreff and MacKenzie [14] recommend averaging throughput over each index of difficulty within each subject and then averaging over each subject within each device, resulting in a single grand throughput value for each device.

2.2. The Speed/Accuracy Trade-Off

One problem with relying exclusively on throughput is that it provides no insight into the trade-off between speed and accuracy. Clients can change their performance under the same experimental condition (i.e., using the same device to perform tasks of equal difficulty) by adjusting their emphasis between speed and accuracy between trials [9]. The index of difficulty calculation can be modified to accommodate this to some extent by substituting the actual target width with the *effective* target width, which is based on the standard deviation of the end point of each motion [14]. This effective width is theoretically larger when motion is faster and smaller when motion is slower (and more precise).

There are two problems with using this in the clinic, however. First, by compressing the results into a single number, throughput effectively hides the effect of the speed/accuracy trade-off. However, clients and clinicians may be interested in performance under high and low speed conditions because some real-world pointing tasks require greater precision than others, and this should be considered during clinical assessments. For example, working with computer-aided design (CAD) programs or drawing programs that require precise targeting place a much greater than normal emphasis on accuracy.

Second, the nature of the speed/accuracy trade-off dictates that, at low error rates, small changes in error rates can be accompanied by extremely large variations in movement time [9]. In other words, the difference in movement time

between two devices may be much greater when one device is operated at 2% error and the other is operated at 3% error than when one device is operated at 20% error and the other is operated at 21% error. Since many clients attempt to minimize errors when trying devices in the clinic, this effect has relevance for clinical practice.

The solution is to have clients use each device under a range of speed/accuracy settings, and to present data from these trials separately for each device, rather than as a single statistic. This requires, of course, an appropriate mechanism for displaying the trade-off between speed and accuracy, and an appropriate definition of what "accuracy" means.

2.3. Alternative Performance Measures

A second drawback to relying exclusively on throughput as a basis for comparing devices is that throughput bases accuracy entirely on the end point of the cursor's motion and ignores the cursor's path of travel during each trial [16]. A device that is slightly slower but produces a much straighter path of travel may be preferable to a slightly faster device with a less optimal path of travel for tasks like menu navigation and free-hand drawing, where the cursor's actual path of travel is just as important as the final destination. Furthermore, it is reasonable to expect that analysis of the cursor's path and velocity profile could provide additional insight into the differences between devices that could be relevant for clinical decision making.

Several investigators have recognized the limitations of using throughput as the only performance measure, and have proposed additional, complementary measures. MacKenzie and colleagues [16] developed error measures that compare the cursor's actual path of travel to the "ideal" path of travel. Keates and Hwang [11, 17-20] extended this work by developing error measures that incorporated the "instantaneous" optimal cursor path. They argued that these measures were potentially more relevant for users with mobility impairments, since the "ideal" path of travel is likely to change if the cursor is thrown off course due to a spasm or tremor.

These measures can provide the client and clinician with a great deal of additional information, at the cost of greater cognitive load required to understand and compare these measures across devices. Furthermore, it is likely that a distinct subset of measures will be of greatest interest for each client. A means of comparing and visualizing these measures would therefore be useful.

2.4. The Speed-Accuracy Operating Characteristic

The Speed-Accuracy Operating Characteristic (SAOC) is a graphical representation of the trade-off between speed and accuracy [10]. Pew [8] originally used the SAOC to describe accuracy in choice reaction tasks, but it has since been used (though not widely) in a range of timed reaction tasks. Rival [21] is one of the few investigators who has used the SAOC to show performance in a pointing task under different speed-accuracy settings.

As shown in Fig. (1), there are two ways to plot the SAOC. The curved relationship between speed and accuracy (shown in the left panel of Fig. (1)) is more intuitive, but is less conducive to comparisons between devices. The linear relationship between speed and the log of the ratio of accu-

(a)

(b)

Fig. (1). Speed accuracy operating characteristic. Panel (a) shows movement time plotted against accuracy. Panel (b) shows movement time plotted against the log of the ratio of correct to incorrect responses.

racy to error (shown in the right panel of Fig. (1)) lends itself more easily to device comparisons and is the form of the SAOC discussed in the rest of this article.

When inspecting the SAOC, curves lying towards the lower right represent the least desirable performance (i.e., large movement time and low accuracy) while curves closer to the upper left corner represent the most desirable performance (i.e., short movement time and high accuracy) [10, 22]. Performance is varied *along* a single curve by changing the emphasis on speed and accuracy [10]. Performance is varied *between* two curves by changes to the task (e.g., changing the pointing device, the pointing device's operating parameters, or the difficulty of the targets) [10]. The goal of an assessment then, is to identify the solution (device, device configuration, task requirements) that produces the curve closest to the upper left.

3. APPLYING THE SAOC IN THE CLINIC

Data from a client of the University of Pittsburgh's Center for Assistive Technology (CAT) is used in this section to illustrate the use of the SAOC in the clinic. The client was a 40-year-old female with a C4 spinal cord injury. She was seen at the CAT for an evaluation of her computer access needs in preparation for employment in a call center.

3.1. Collecting Data

The first step to constructing the SAOC is collecting performance data with all of the pointing devices under consideration under multiple speed/accuracy settings. Clearly, this requires software for administering pointing tasks and collecting data. Several programs are listed in Table 1. Each implements a slightly different pointing task, but none implements a pointing task that conforms with the ISO 9241-9 standard. The clinical data described in this article were collected using Compass [6].

There are several ways to elicit different speed/accuracy behaviors [9, 23]. A simple, but effective, mechanism that was used to collect data for this article is to run trials with different deadlines (of which the client is aware) [23]. Trials with short deadlines encourage faster movement times while

trials with longer deadlines allow for a greater emphasis on accuracy.

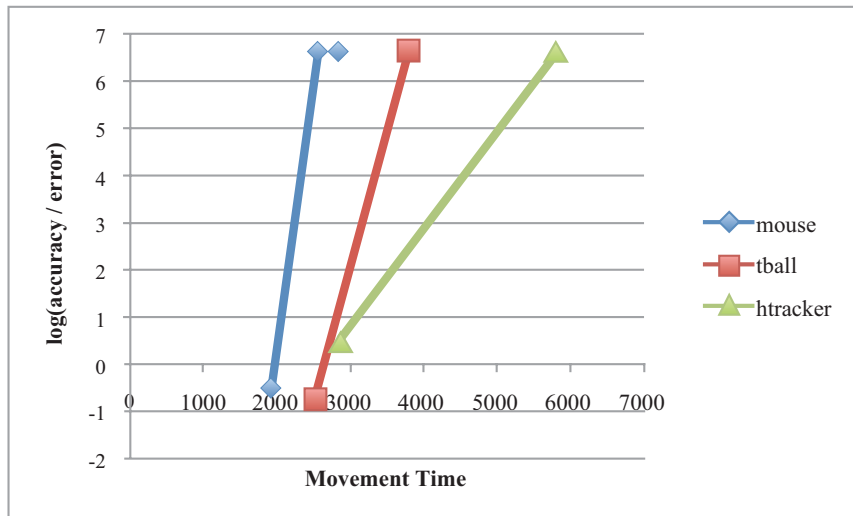
When collecting data in the clinic there are several practical considerations. Given finite limits to a client's time, endurance and attention span, it is desirable to limit the number of trials with each device. Furthermore, some of the devices being evaluated may be completely unfamiliar to the client, which can make it difficult for them to accurately produce multiple speed/accuracy behaviors. In practice, asking clients to complete two blocks of 12 to 24 trials while being either "really fast" or "really accurate" seems to be an adequate trade-off between expediency and data integrity, but future work is needed to identify the most appropriate number of trials.

The clinical data used in this article were collected over seven blocks of trials with three devices: a standard mouse, a Penny & Giles Roller II Joystick and a NaturalPoint SmartNav head-mounted mouse emulator. The client completed three blocks of 24 trials with the mouse, two blocks of 24 trials with the trackball and two blocks of 12 trials with the head tracker. For all three devices, the client was asked to complete one block of trials as fast as possible, and to complete the second block of trials as accurately as possible. For the third block of trials for the mouse, the client was asked to use the mouse as she normally would. Only 12 trials per block were used for the head tracker due to concerns about fatiguing the client.

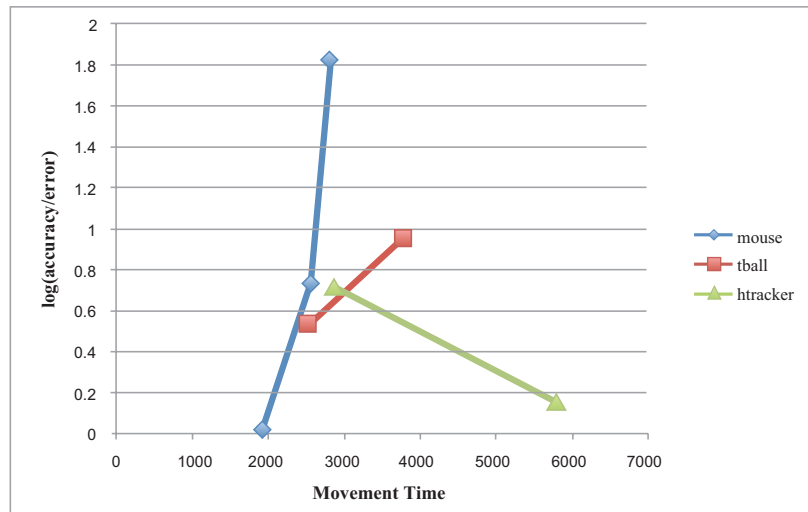
3.2. Constructing the SAOC

Since movement time is a fairly straightforward measure, the real challenge in constructing the SAOC is calculating accuracy. Fig. (2, panel a) shows an SAOC calculated using the simplest measure of accuracy, in which accuracy is 100% for trials that end inside the target and 0% for trials that end outside of the target, and error is 100% minus accuracy. To avoid division by zero, a trial with 100% accuracy was treated as 99.99% accuracy. As can be seen from the figure, by this measure two of the blocks of trials conducted with the mouse had perfect accuracy, which conceals genuine differences in performance.

(a) accuracy = 100% if cursor ends trial in target; accuracy = 0% if cursor ends trial outside of target; error = 99.99% - accuracy



(b) accuracy = ratio of length of optimal path of travel to length of actual path of travel; error = 99.99% - accuracy



(c) accuracy = product of accuracy definitions from panels (a) and (b); error = 99.99% - accuracy

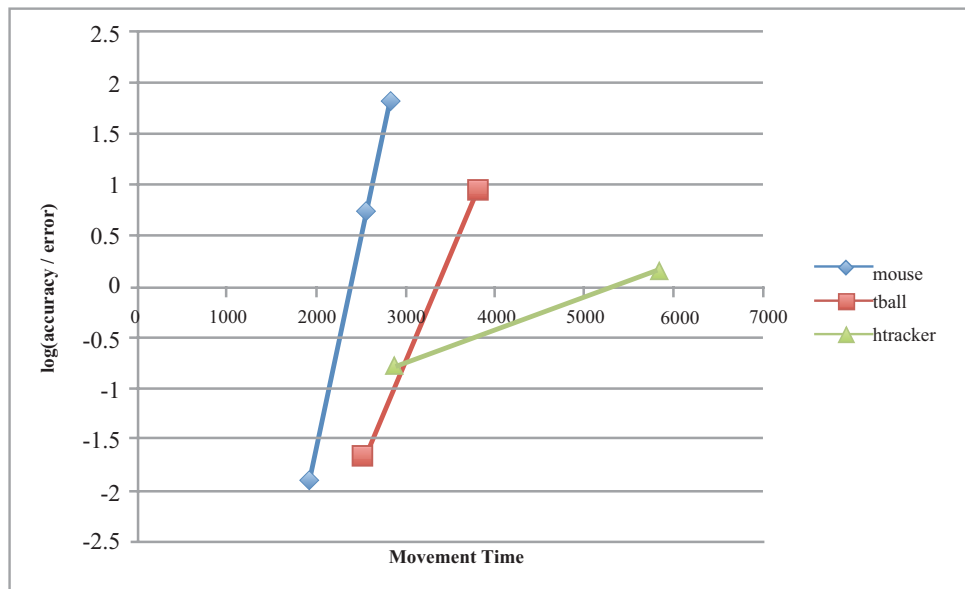


Fig. (2). Speed-accuracy operating characteristic curves.

An alternative measure is the ratio of the distance of the optimal path to the actual distance traveled. This measure has the advantage of considering the entire path of travel, but the disadvantage of not distinguishing between paths that end inside the target and those that do not. Hence, a trial that ends inside the target can be considered less accurate than a trial that ends outside of the target. The resulting SAOC is shown in Fig. (2, panel b).

The final measure considered thus far combines the prior two measures through a simple decision rule. If the cursor path ends inside the target then the accuracy is the ratio of the distance of the optimal path to the actual distance traveled. If the cursor path does not end inside the target then the accuracy is 0. The principal shortcoming of this approach is that it fails to distinguish between cursor motions that almost reached the target from cursor motions that ended far from the target. The resulting SAOC is shown in Fig. (2, panel c).

The accuracy calculations described above are but a fraction of the possible formulas that could be considered. Many measures of accuracy are possible [9], and it is also conceivable that different measures will be needed for different tasks, devices or users. The question of how to calculate accuracy when constructing the SAOC is still very much open to investigation.

3.3. Interpreting the SAOC

For this client, the mouse clearly provides greater performance than the trackball or head tracker, based on its location closest to the upper left corner of the SAOC. The trackball would appear to be distinctly better than the head tracker in most situations, although there is overlap between the two curves at low levels of accuracy. Another interesting aspect of the SAOC is the linear shape of the mouse curve, which indicates that the speed/accuracy trade-off holds strongly for this client when using the mouse.

It is interesting to consider how to interpret an SAOC when two curves overlap. The overlap indicates that accuracy at slower speeds is higher with one device and accuracy at faster speeds is higher with the other device. The question then becomes which type of behavior is more common for the user? Additional trials to generate points along the SAOC curves that correspond to the client's "usual" or "comfortable" speed/accuracy trade-off may determine which device is most appropriate.

To the extent an SAOC curve remains linear (or nearly linear), it is worth considering how to interpret the y-intercept and slope. A greater y-intercept implies better performance, because the curve will be closer to the upper left corner of the graph. A steeper slope means that small changes in movement time produce large changes in accuracy, whereas a shallow slope means that it is hard to change accuracy by changing moving time.

4. DISCUSSION

The SAOC presents several advantages to clinicians. It visualizes the trade-off between speed and accuracy [10], is easy for clinicians, clients and third-party payers to interpret, and provides an efficient means of summarizing and documenting performance. The SAOC is not, however, the sole factor by which pointing devices should be judged. The cli-

nician must still work with the client to evaluate devices on a range of factors, including: posture, pain, fatigue, the likelihood of incurring repetitive stress injuries, aesthetics, compatibility with existing equipment and AT, consumer preference, cost, training requirements and the availability of technical support.

Furthermore, several questions remain regarding the SAOC. Procedural questions include how to best induce performance at different speed/accuracy sets [9] and how many data points at each speed/accuracy setting are needed? It has also yet to be determined what task (or tasks) are most appropriate. Compass, for example, provides a point-and-click task, a click-and-drag task and a menu navigation task. It would be nice to integrate results from all three tests into the SAOC, if possible, since all three tests reflect different aspects of using a pointing device.

Finally, the question of quantitative analysis must be addressed. The SAOC lends itself well to visual analysis, but how does one determine whether a clinically or statistically significant difference exists? The issue of statistical significance can be addressed through linear regression, but what constitutes a clinically significant difference in performance remains an open question.

5. CONCLUSIONS

The ultimate goal of this research is to develop a clinical tool that clinicians can apply to their practice with little effort. Several steps remain before that goal will be realized. This paper has demonstrated the potential for the SAOC and has identified some of those open research questions.

REFERENCES

- [1] Hocking C. Function or feelings: factors in abandonment of assistive devices. *Technol Disabil* 1999; 11: 3-11.
- [2] Johnston SS, Evans J. Considering response efficiency as a strategy to prevent assistive technology abandonment. *J Spec Educ Technol* 2005; 20: 45-50.
- [3] Phillips B, Zhao H. Predictors of assistive technology abandonment. *Ass Technol* 1993; 5: 36-46.
- [4] Verza R, Carvalho MLL, Battaglia MA, Uccelli MM. An interdisciplinary approach to evaluating the need for assistive technology reduces equipment abandonment. *Mult Scler* 2006; 12: 88-93.
- [5] Simpson RC. Making better decisions. *IEEE Eng Med Biol Mag* 2008; 27: 23-8.
- [6] Koester HH, Simpson RC, Spaeth D, LoPresti EF. Reliability and validity of compass software for access assessment. *RESNA 2007 Annual Conference*, Phoenix, AZ: RESNA Press 2007.
- [7] Schedlbauer MJ. An Extensible Platform for the Interactive Exploration of Fitts' Law and Related Movement Time Models. In: Rosson MB, Gilmore D, Eds. *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, San Jose, CA: ACM Press 2007; pp. 2633-8.
- [8] Pew RW. The speed-accuracy operating characteristic. *Acta Psychol* 1969; 30: 16-26
- [9] Wickelgren WA. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol* 1976; 41: 67-85.
- [10] Wickens CD, Hollands JG. *Engineering Psychology and Human Performance* Upper Saddle River, NJ: Prentice Hall 2000.
- [11] Hwang F. A study of cursor trajectories of motion-impaired users. In: Terveen L, Wixon D, Minneapolis MN, Eds. *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. ACM Press 2002; pp. 842-3.
- [12] Keates S, Trewin S. Effect of Age and Parkinson's Disease on Cursor Positioning Using a Mouse. In: Scars A, Pontelli E, Baltimore MD, Eds. *7th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM Press 2005; pp. 68-75.

- [13] Trewin S, Pain H. Keyboard and mouse errors due to motor disabilities. *Int J Hum Comput Stud* 1999; 50: 109-44.
- [14] Soukoreff RW, MacKenzie IS. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int J Hum Comput Stud* 2004; 61: 751-89.
- [15] International Standards Organization. Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices. ISO Standard 9241-9: 2000(E). Geneva, Switzerland: ISO 2002.
- [16] MacKenzie IS, Kauppinen T, Silfverberg M. Accuracy measures for evaluating computer pointing devices. SIGCHI Conference on Human Factors in Computing Systems, Seattle, WA: ACM Press 2001.
- [17] Hwang F. Partitioning Cursor Movements in "Point and Click" tasks. In: Cocton G, Korhonen P, Lauderdale FL, Eds. CHI '03 Extended Abstracts on Human Factors in Computing Systems, Ft.: ACM Press 2003; pp. 682-3.
- [18] Hwang F, Keates S, Langdon P, Clarkson J. Mouse movements of motion-impaired users: A submovement analysis. 6th International ACM SIGACCESS Conference on Computers and Accessibility, Atlanta, GA: ACM Press 2004.
- [19] Keates S, Hwang F, Langdon P, Clarkson PJ, Robinson P. Cursor Measures for Motion-Impaired Computer Users. In: Hanson VL, Jacko J, Eds. A Proceedings of the Fifth International ACM Conference on Assistive Technologies, Edinburgh, Scotland: ACM Press 2002; pp. 135-42.
- [20] Keates S, Hwang F, Langdon P, Clarkson PJ, Robinson P. The user of cursor measures for motion-impaired computer users. *Univ Acc Info Soc* 2002; 2: 18-29.
- [21] Rival C, Olivier I, Ceyte H. Effects of temporal and/or spatial instructions on the speed-accuracy trade-off of pointing movements in children. *Neurosci Lett* 2003; 336: 65-9.
- [22] Pew RA, Mavor AS, Eds. Modeling Human and Organizational Behavior: Application to Military Simulations, Washington, DC: National Academy Press 1998.
- [23] McCormick PA, Francis L. Speed-accuracy tradeoff operator characteristics of endogenous and exogenous covert orienting of attention. *Scientific WorldJournal* 2005; 5: 128-46.

Received: September 18, 2008

Revised: January 16, 2009

Accepted: January 19, 2009

© Richard C. Simpson; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.