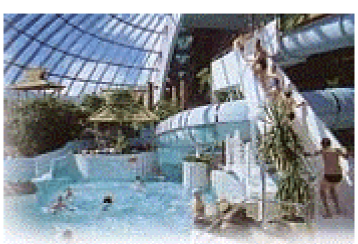
A controlled experimental study of the Personal Waterslide Process (PWP): Results and Interpretations

Philip Johnson, Audris Mockus, Larry Votta
Waterslide Engineering Institute
Technical Report WEI-99-01
Oulu, Finland
June, 1999

Introduction

Even before the introduction of the Personal Waterslide Process in 1999, watersliding was a topic of intensive research by the experimental software engineering community. Rombach suggested watersliding to be a major influence on software quality, object-oriented inspection, technology transfer, and perspective-based reading. Prior to these revelations, only elite software engineers knew that watersliding could stimulate development insights not achievable by any other means[1]. The installation of waterslides at major development centers such as Lucent Technologies, Ericsson, and the Fraunhofer IESE attests to its growing recognition as an intrinsic component of high quality software development. For example, in the 5ESS switch development group at Lucent, a retrospective analysis of 15 years of configuration management data by Votta found that watersliding had a consistent impact upon the cycle time for enhancement requests. Work by Schwinn at DaimlerChrysler showed that the introduction of watersliding effectively eliminated group process problems surrounding inspections. Basili, in work at the Software Engineering Lab, was able to show that the natural washing effect of watersliding led to increased adoption of CLEANROOM methods. However, subsequent work by Zelkowitz showed that without continuous investment in watersliding, developers would revert back to their previously unwashed state with obvious negative consequences for the CLEANROOM method.

In response to these successes, the Waterslide Engineering Institute (WEI) was founded to promote scientific watersliding, and facilities were built adjacent to VTT in Oulu, Finland. The WEI quickly established itself as a leader in the promotion of watersliding, producing the Watersliding Capability Maturity Model (wCMM). The wCMM, designed by Oivo, establishes key practice areas for watersliding. These include “take off”, “lying down”, “holding nose”, and “making a big splash”. The 5 level model allows organizations to assess and improve their maturity with respect to watersliding. The WEI-sponsored EDEN SPA[2] (European Distance Education Natatorium for the Scientific Practice of Aquasliding) facility has been acclaimed for its research in this important area.



The experimental apparatus of the EDEN SPA research facility at left. The waterslide is powered by a RAID array of high performance water jets ensuring 24/7 up time. High performance workstations were used to design the slide geometry for optimum throughput. Wireless telecommunications (cell phones) ensure that subjects are in contact with experimenters at all times.

---

[1] For best results, each waterslide run should be followed by 10 minutes in a Greco-Roman sauna.
[2] Further information available at: http://www.ouluneden.fi/

Despite the uncontested benefits, problems remain in applying watersliding at the personal level. This gap in methodology was plugged with the introduction of the Personal Watersliding Process (PWP) by Torii. Practitioners of the PWP learn advanced skills in waterslide defect management, estimation of the size and time of waterslides, and waterslide design checklists.

Despite the overwhelming presence of anecdotal data attesting to the utility of watersliding in software engineering, surprisingly, no controlled experimental studies have yet been attempted to understand the factors underlying watersliding performance. This paper is our own modest contribution toward overcoming this glaring hole in the software engineering literature. The remaining sections of the paper report on what we fully expect to be seminal and ubiquitously cited research on this important topic.

The experiment

The goal of the experiment was to determine factors that affect performance of an individual on the waterslide. It is well known that the software productivity of an individual is closely related to their performance on the waterslide. Consequently, the predictors of the waterslide performance would also predict software productivity. We asked if any of 16 independent factors measuring physical and psychological state of an individual predict the waterslide performance. The waterslide performance was operationalized as the time it takes for an individual to go down the waterslide. The actual data sheet is shown in Appendix A.

The experiment was performed on 16 randomly selected subjects. To be selected the subject had to be a participant of 1999 ISERN meeting that took place in Oulu, Finland; to show up at the waterslide at the instructed time; to be wearing a bathing suit; to provide the interviewer with the relevant information; and to be timed by another experimenter while going down the waterslide. Each subject had at least 3 years of experience in software engineering, although some subjects had an order of magnitude more experience. The subjects' prior experience in water sliding was also quite varied, but we assumed this would not affect our results.

The process of an individual trial was as follows. Subject would go up the staircase to reach the top of the waterslide, wait until their turn, nod to the experimenter, lie down on their back (at this point the experimenter would start counting the time), slide down the waterslide, drop into the pool (at this point the experimenter would stop counting the time), get out of the water (at this point the experimenter will notify the subject of their result). Finally, the second experimenter would interview the subject to obtain the remaining GQM measures.

The results of the experiment are both simple and intuitive (to an expert). Of the 16 different independent measures only two measures predict the performance – the butt area and the amount of body hair increase the time to go the waterslide, thereby decreasing the performance. However, the interaction of those predictors decrease the time, i.e., the subjects who have large values for both measures do not suffer as much degradation in performance as one would expect from a linear model.
The final model is as follows:

$$\mathrm{Time} = a * \mathrm{BUTT.AREA} + b * \mathrm{BODY.HAIR} + c * \mathrm{BA} \times \mathrm{BH} + d + \mathrm{Error}$$

Table 1. The ANOVA table.

	Value	Std. Error	t value	Pr(>|t|)
BA	1.4	0.6800	2.0733	0.0624
BH	4.0	1.1093	3.6589	0.0038
BA:BH	-1.1	0.3459	-3.165	0.0090
R square	.63			

The air and surface friction represent an obvious mechanism by which those two factors affect the waterslide performance. The mechanisms by which those factors affect software productivity remain less clear. However, the finding could significantly simplify recruiting and performance evaluation process in software organizations around the world. It would result in saving on the order of several billion US dollars per year. Furthermore, the finding could be used to reduce training costs for software engineers by replacing involved education and testing programs with two simple measures.

Despite the excellent experimental design, we have a number of threats to internal and external validity. First, only a single waterslide was used, thereby allowing for possibility that the performance predictors are unique to that particular waterslide. Some of the independent measures might be affected by the dependent measure, e.g., going down the slide too fast might reduce the amount of body hair. Finally, the software productivity of subjects was not measured before, during or after the experiment.

Acknowledgments

The authors wish to gratefully acknowledge the WEI, which paid for the results of this study. All subjects were paid professional employees of the WEI. This research was conducted during a paid visit to EDEN SPA underwritten by a generous grant from the WEI. PWP, EDEN SPA, and WEI are all trademarks of the WEI. Void where prohibited by law. This has been a test of the emergency scientific research system. If this had been real scientific research, you would have been directed to the nearest copy of the Journal of Empirical Software Engineering for further instructions.

Appendix A

This following image shows the actual data sheet used for the waterslide experiment at ISERN'99. Subjects were listed as A through O. Important measures known to impact upon software productivity were collected, including: speed, weight, butt area, height, age, gender, citizenship, shoe size, body hair, facial hair, swimming suit color, professional affiliation (academia, industry), favorite programming language, belt size, and swim suit style. To improve the experimental design, certain measurement units were randomized. For example, measurement units for weight were randomly varied between pounds and kilograms. Measurement units for height were randomly varied between centimeters and feet/inches. Shoe size and belt size measurement units were also subject to randomization. We expect "Unit Randomization" to become increasingly popular in the experimental software engineering community, once its benefits (as illustrated by this research) become more widely known.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
speed	9.5	10.5	12.5	9.5	9.0	9.0	9.5	13.5	8 1/2	13	9.5	16	9.5	10	14
weight	105 35	88	70	85	160	170	95	185	140	72	71	65	140	68	210
butt area	L	M	M	M	M	M	M	M	L	L	L	S	S	M	M
height	171	178	5'10"	183	184	5'9"	178	5'8	5'11	172	178	180	5'8"	173	6'1"
age	28	28	31	28	35	26	27	48	40	36	31	25	38	27	53
gender	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
citizenship	Ger	Ger	AUS	Ger	Lith	US	Ger	US	US	JAP	Austr	Nor	Brit	JAP	US
shoe size	42	43	10	44	41	8.5	7	10	11	26	43	42	9	26	12
body hair	L	L	L	M	L	L	M	H	M	L	L	M	L	L	H
facial hair	N	L	N	N	N	M	M	H	H	N	N	N	N	N	N
suit color	Blue	Red blk	Blue	Black	green	black red	blue	blue	blue	blue	black	black	green	blue	blue
indust/acad.	I	I	A	I	I	A	I	I	A	I	A	A	A	I	A
prog lang	guitar	pascal	C++	C	java	none	perl	Fortran	LISP	C	C	Assem	Scheme	C	pascal
belt size	42	95	33	95	N/A	34	95	35	33	82	32	92	30	75	38
bathing suit	tank	tank	b. shorts	tank	tank	tank	tank	b shorts	b shorts	tank	tank	tank	tank	tank	b. shorts