Scientific
Research
Publishing

Reflections on Long-Term Development and Use of Automated Scoring Technology in a Sport (Modified Boxing) Context

Paul Perkins[1,2]*, Richard J. N. Helmer[3], Colin Mackintosh[4], Allan G. Hahn[1,5,6,7,8]

[1]University of Canberra Research Institute for Sport and Exercise, Canberra, Australia
[2]Boxing Australia Limited, Canberra, Australia
[3]Superinteractive Pty Ltd, Geelong, Australia
[4]AppSen, Canberra, Australia
[5]Queensland Academy of Sport Centre of Excellence for Applied Sport Science Research, Brisbane, Australia
[6]Griffith University School of Engineering, Brisbane, Australia
[7]Australian Institute of Sport, Canberra, Australia
[8]Victoria University College of Engineering and Science, Melbourne, Australia
Email: *boxeo68@live.com

Abstract

Technology is being increasingly used to aid judging in sport, but its employment as the primary means of scoring is rare. We have developed and implemented a fully automated scoring system in the context of a modified, low-risk form of boxing. The system, which requires contestants to wear vests and gloves incorporating sensor fabrics, has been used in multiple settings over the past five years. During that period, it has undergone progressive iteration guided by action research methodology. Here, we summarise that iteration, reflect on present status and identify possible future directions. We have found that concept of automated scoring has wide appeal, and the wearable technology is almost universally considered comfortable. Nevertheless, some issues remain to be addressed. Use of the technology requires considerable prior and subsequent commitment of time. Apparently valid contacts occasionally fail to score. Causative factors include the configuration of electrical circuitry in the vests and deterioration of that circuitry with repeated vest use and washing. Also, false positive scores are sometimes generated by vest self-shorting and effects of sweat. Many contestants adopt unorthodox styles aimed at exploiting the characteristics of the automated scoring methodology, affecting the aesthetics of the modified sport. There is an expectation that technologically-based scoring should have much greater accuracy than human judging, and should be essentially fail-proof. Disillusionment can occur in situations where this expectation is not met. We have identified potential solutions to all the existing issues, with some now being actively explored.

Continuation of the quest seems justified by popular dissatisfaction with subjective human judging of boxing and other sports, but we have come to realise that purely technological judging can introduce unforeseen complexities. Our observations could be relevant to various sports interested in the notion of technological judging.

Keywords

Automated Scoring Technology, Box'Tag, Smart Fabric, Sport Technology, Wearable Sensors

1. Introduction

Over the past decade, a modified, low-risk form of boxing known as Box'Tag has emerged in Australia [1]. A key feature of the modified sport is that impacts to the head and neck are prohibited, as recommended by various experts seeking to enhance boxing safety [2] [3] [4] [5]. Additionally, automated scoring technology is employed. This technology is comprehensively described elsewhere [1] [6]. In summary, contestants wear specialised vests incorporating sensor fabrics and boxing gloves that have patches of electrically conductive material affixed to their surface. The vest sensor fabrics include stripes of silver nylon yarn forming a circuit that can be connected to a transceiver located in a pocket on the upper back. The transceiver directs a low-level electrical current through the circuit. When a conductive patch on a glove bridges two vest stripes, a change in the electrical resistance of the vest occurs and is detected by the transceiver, which sends the data by wireless mechanisms to a ringside computer. A customised algorithm is then applied to determine whether a score should be registered. Scores can be displayed in real time. The major wearable components of the technology are shown in Figure 1.

The accuracy of scores produced through the above technology was evaluated by Bruch *et al* [7], who used frame-by-frame video analysis of 32 rounds of Box'Tag competition as the criterion for determining valid impacts. It was reported that the automated scoring system correctly identified ~90% of all legitimate impacts, and that of the ~10% of legitimate impacts not successfully detected,

Figure 1. Box'Tag contest showing instrumented vests and conductive patches on gloves. Wires connecting vest to transceiver can be just seen behind the neck of the red boxer.

more than a third were directed to target zones located on the shoulders. False positive scores were rare. Based on these findings, the system was considered adequate for use in the Box'Tag context.

We have employed the automated scoring technology extensively in multiple different situations over a 5-year period and have implemented various actions aimed at continuous improvement. This has led to a range of insights beyond those provided by the initial, much more isolated analysis.

Our observations are the subject of this paper and, since technology is being increasingly adopted to assist with sports judging [8], may be pertinent to the future plans of various sports.

2. Methods

In an effort to gain both broad and deep understanding of the evolution of the automated scoring technology, reactions to it, issues emerging from its use and its potential for further development, we integrated and examined qualitative and quantitative data from a range of sources. The data were collected over five years as part of an action research process [9] that guided iterations of the technology. In keeping with the principles of action research [9] [10] [11] [12], that process entailed repeated cycles of observation, reflection, planning, action and evaluation conducted in collaboration with end-users of the technology and designed to enable progressive, holistic identification and resolution of real-world problems.

Materials gathered through the action research methodology included a comprehensive journal maintained by the first author of this paper, who was not only a member of the research team but also the coach of a Box'Tag program operating at a Police Community Youth Club (PCYC) in Canberra, where he was a primary user of the automated scoring technology. He was therefore a 'practitioner-researcher'. Campbell [13] has noted that practitioner-researchers can facilitate high-quality research outcomes because they are uniquely positioned to provide an insider's perspective on practical problems and the success of attempts to address them.

In our case, the practitioner-researcher 'lived and breathed' the Box'Tag program for the whole 5-year period encompassed by the present analysis. For all but the last year of this period, he conducted almost weekly sessions in which participants engaged with the automated scoring technology. In addition to these regular sessions, he oversaw use of the technology at numerous special Box'Tag events held at the PCYC and elsewhere, demonstrated it in the context of workshops and technology showcases held at multiple Australian locations, exhibited it at amateur boxing tournaments and draft camps, participated in its operation for filming of television programs, and assisted with its use for the purposes of Box'Tag competitions at the Victorian Police & Emergency Services Games in 2012 and 2013 and the Australasian Police & Emergency Games in 2014.

Importantly, the practitioner-researcher adopted a highly systematic and

deliberate approach to the processes of observation and reflection, as recommended by action research experts [9] [14]. His journal served as a vehicle for documenting his experiences, impressions and ideas, recording comments made either formally (in interview settings or through surveys) or informally by Box'Tag participants, summarising salient discussions with other parties, logging the results of specific experiments conducted as part of the technology iteration, and making methodological notes. The journal, which was updated several times each week, eventually consisted of ~100,000 words. It related to all aspects of the involvement of the practitioner-researcher with Box'Tag, but a large volume of information concerning the automated scoring technology was included. While the continuous compilation of the journal was critical to the rigour of the action research process, the process itself was highly collaborative and produced numerous other significant historical records.

Apart from the practitioner-researcher, our research team includes the original developers of the automated scoring technology [15] and the software package that supports it. We interacted regularly throughout the whole five years through a combination of face-to-face meetings, video conferences, telephone discussions and email correspondence.

We all spent some time directly supporting Box'Tag events and sessions at which the automated scoring technology was used, and we all participated in the design and conduct of experiments aimed at obtaining information to guide its improvement.

As a consequence of this collaborative commitment to the action research methodology, it was possible for us to jointly supplement the journal with a substantial collection of video footage of Box'Tag contests and automated scoring technology trials, an archive of email correspondence in which aspects of the technology were discussed, research grant applications, 12 written project reports, software and firmware code iterations, PowerPoint presentations and a series of published papers [1] [7] [16] [17] [18].

Five years of action research has therefore enabled us to reflect on a very rich, diverse and extensive body of recorded information in order to facilitate the insights outlined below. Elliott [9] has noted the importance of such sources in permitting analytical reflection, or genuine "reconnaissance" as opposed to just casual observation. In advocating the value of reflective practice, Leitch and Day [19] describe reflection as the "engine" of action research and see it as critical to the instigation of change, a concept inherent in action research philosophy. They note that the process of collating and "writing up" data gives substance to reflection that otherwise could remain tacit and amorphous and produce little practical benefit. Accordingly, our latest phase of thorough reflection—undertaken for the purpose of preparing this paper—is seen as integral to the efficiency and effectiveness of ongoing decisions and actions regarding the automated scoring technology.

Access to the materials collected in association with the use of the automated scoring technology was provided by the organisations through which that use

occurred, under arrangements that were approved by the Human Research Ethics Committee of the University of Canberra, Australia.

3. Results

3.1. Appeal of Automated Scoring Technology

In our experience, the automated scoring technology makes a highly favourable first impression on almost everyone who sees it in action. The objectivity of the scoring and the excitement generated by dynamic, real-time display of scores (Figure 2) are the most commonly identified positive features. Also well-regarded is a feature that provides for occurrence of sounds to indicate the registration of scores, with the sounds differing for the two contestants. There has been some debate, though, as to whether the current sounds for the two contestants are sufficiently distinct from each other.

3.2. Comfort and "Look" of Specialised Apparel

The instrumented vests worn by contestants closely resemble conventional T-shirts and the contestants almost universally consider them to be comfortable. They have been produced in a wide variety of sizes [6] and contestants therefore feel that good "fit" is generally achievable.

The original T-style vests had lycra as the base fabric and were somewhat stretchable, a characteristic that caused them to conform quite closely to body shape. Some athletes indicated that they would prefer a looser-fitting garment, and in 2015 a new batch of vests was produced to cater for this preference. The new vests, which incorporated a micro-mesh base fabric, were well-received, although approximately half of all athletes continued to find the original vests more to their liking.

To date, the vests have been made available in just two colours—red and blue. This is in keeping with colours historically employed in conventional amateur
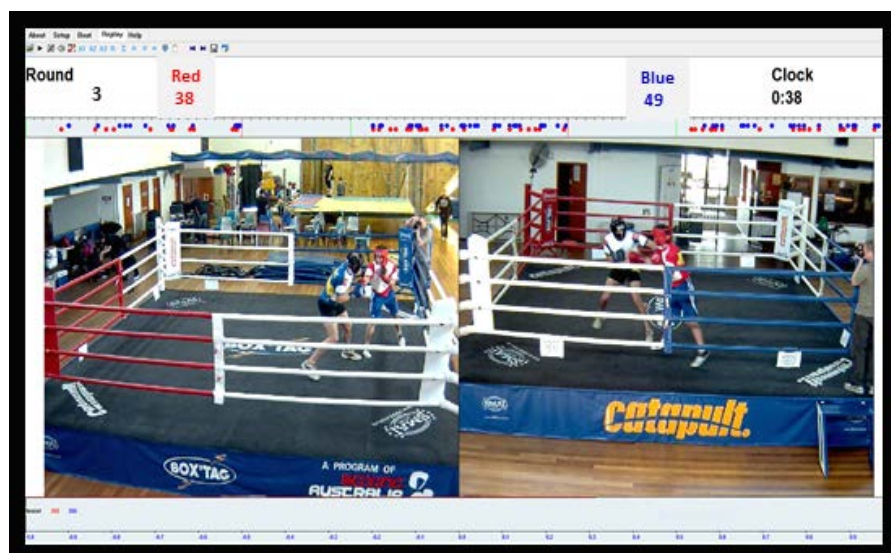

Figure 2. Real-time display of scores in a Box'Tag contest.

boxing [20], although both the red and blue vests used in the Box'Tag setting also include white sensor fabric on the front of the torso and small areas of the upper arms, defining the target regions. While the colours are widely regarded as having aesthetic appeal, there have been many suggestions that the choice should be expanded.

The transceivers that form part of the automated scoring technology are fitted into a specially designed pocket on the upper back of the vests just below the neck. Those used for most of the past five years have dimensions of 4 cm × 4 cm × 1 cm (with the last of these being the depth) and a mass of 19 grams. A new version produced earlier this year has dimensions of 5 cm × 2.5 cm × 1.3 cm and a mass of 25 g (including a cable for connection to the vest). The transceivers are shown in Figure 3. Their small size has meant that athletes have found them to be almost completely unobtrusive.

3.3. Software

The software package—named Spartan [1] [7]—developed as part of the automated scoring technology has undergone progressive iteration over the past five years. It currently runs only on Windows operating systems.

The primary purposes of the package are wireless receipt (via Bluetooth) of vest resistance data from contestants, analysis of those data to enable detection of vest contacts, and real-time display of scores. From the outset, however, it has also incorporated ability to capture video data from two simultaneously operating, orthogonally positioned ethernet cameras and to display the video footage in concert with the sensor data both in real time and subsequently in replay mode. The replay facility has provided a means for checking on the accuracy of scores recorded in the real-time situation [7].

Over time, the user-friendliness of the Spartan software has continually increased through such actions as provision of augmented operator control icons, inclusion of an on-line help manual, implementation of capacity for slow-motion and even frame-by-frame replay of the video footage, and creation of an ability

Figure 3. An original (left) and a newer transceiver employed with the automated scoring technology.

to achieve more precise temporal alignment between images from different cameras.

Additionally, it has become possible for users of the software to score three contests simultaneously, select a "countdown" scoring method in which each contestant is required to defend an allocated number of points, download all impact data recorded during a bout to an Excel file for further analysis, and produce automated summaries of impact data for each contestant (including mean contact duration, the number of impacts resulting in scores, the number of impacts not satisfying the scoring criteria and a percentage calculated from the ratio of scoring to non-scoring impacts). Similarly, monitoring of the equipment performance has been enabled, with transceiver operation evaluated in terms of number of Bluetooth data packets received and average sampling frequency, and baseline resistance of vests measured in real time and made available for post-contest display. Overall, the Spartan software package has become a powerful tool for deployment in competition, training, performance analysis and research environments.

Researchers have seen great value in the Spartan software due to its extensive analytical capabilities. However, some people with less technological background apparently have been overwhelmed by the range of possibilities that it provides, and therefore have been reluctant to use it in the absence of technical support. For several months, there were regular reports of the Spartan software sometimes "crashing", and indeed we encountered this problem ourselves. We initially suspected that it might be due to momentary losses of transceiver function caused by mechanical or other stresses, but when we subsequently ran multiple transceivers for hours at a time and exposed them to various mechanical stresses including dropping on to floor from a height of ~30 cm, mildly forceful impact with one another and application of tension and minor jerking forces to cables connecting the transceivers to vests, we were unable to induce a single electrical interruption in any of the units. The crashing of Spartan was eventually rectified following identification of an intermittent software "bug" that had an effect only when the capacity of the package for video collection was being deployed, but it likely contributed to a perception of complexity in regard to the successful operation the automated scoring technology. Another perceived disadvantage of the Spartan software is that it requires the use of a dedicated laptop computer, which entails significant expense.

3.4. Apps for iPhone and iPad

To facilitate use of the automated scoring technology, an app that can be downloaded on to an Apple iPhone or iPad has recently been developed. Employment of the app, called ModBox, currently requires use of an additional bridging app to select and route transceiver signals and thereby enable pre-configuring of bouts.

Within the ModBox app, the bout can then be selected from a list. Once the selection has occurred, scoring via ModBox can be initiated, paused and stopped

via the pressing of a single button. There is an option for selecting either "count-up" or "countdown" scoring methodology but all other parameters are fixed so as to provide maximum possible simplicity of operation. Scores are displayed in real time. The app does not provide for collection of video data. Field testing of the app has only just commenced and has been confined to a small group, but initial reactions have been positive.

3.5. Time Requirements Associated with Use

We have found that optimal deployment of the automated scoring technology requires significant commitment of time, both in the lead-up and afterwards. For events involving multiple contests, prior checking of vests and transceivers to ensure their full functionality can easily take half a day. Even before this checking can be conducted, electrical charging of transceivers is necessary.

Time is needed also for computer set-up and (in some situations) positioning and adjustment of Ethernet cameras. Post-event washing, drying and re-examination of vests are time-consuming tasks. For people focused primarily on the coaching of athletes and/or the general running of community exercise programs incorporating modified boxing, the high time cost inherent in using the automated scoring technology has emerged as a definite disincentive to engagement.

3.6. Accuracy of Scoring and Robustness of Equipment

Over the years, we have encountered a number of technical difficulties in the use of the automated scoring technology.

3.6.1. Vest Shorting

Contact between stripes on a vest can cause an electrical short circuit and consequent acute reduction of vest electrical resistance, leading to registration of false positive scores. The problem typically occurs when the fit of the vest is not ideal, but occasionally the movements of a contestant can bring a stripe on a shoulder sensor fabric of an apparently well-fitting vest into contact with one on the torso region of the vest. The loose-fitting micromesh vests produced in response to participant feedback regarding vest comfort and appearance are more prone than the originals to this issue. "Flapping" of the sections of these vests that cover the shoulders is a primary causative factor, but contact between torso stripes also can sometimes occur.

3.6.2. Changes in Electrical Resistance of Vests

Pre-contest checking of vests sometimes revealed losses of impact detection sensitivity. These were most common in the shoulder target regions. In 2014, we employed a multi-meter to measure the base electrical resistance of 78 vests that had been in operation for periods ranging from 3 months to more than 3 years. One multi-meter electrode was placed on a press-stud that was the connection point of the vest to the transceiver, while the other was placed on the most lateral stripe of silver nylon yarn on the vest shoulder target that formed the other end of the electrical circuit. The measured resistance consequently related to the entire

circuit length. All of the vests were also subject to impacts with a glove incorporating a conductive fabric. The impacts were to the target areas on the torso and both shoulders. The results showed that:

- 44 vests had an overall electrical resistance of less than 2000 Ohms (Ω) and all were fully functional.
- 18 vests had resistance readings of more than 3600 Ω and none were fully functional, with the great majority failing to register impacts to the left shoulder.
- 16 vests had resistance readings between 2000 and 3600 Ω. Of these, seven were fully functional and nine were not.

The two categories were not distinguishable on the basis of resistance readings, with the average values for the functional and non-functional vests being almost the same. The lowest resistance reading for a vest with left shoulder insensitivity was 2260 Ω and the highest reading for a fully functional vest was 3600 Ω. It was therefore evident that resistance readings could be used to define a zone of definite vest functionality and a zone of definite non-functionality, with these zones separated by one in which functionality is unpredictable.

There was a tendency for vest resistance to become higher as vest size increased (see Figure 4), reflecting longer path lengths of the electrical circuitry in larger vests. Important here is the fact that, because of the methods used to determined to determine valid impacts, even fully functional vests with different levels of baseline resistance could differ in terms of impact sensitivity, giving one contestant an advantage over the other. Impacts are registered when vest resistance falls (for a specified period that can be set in the Spartan software) to less than 80% of a running 10-second average of vest resistance determined during non-contact periods. The upper limit of the specified range within which the transceivers can reliably measure electrical resistance is 2500 Ω. This will be the baseline value recorded when true vest resistance is actually higher, meaning that for an impact to be recorded, resistance must decrease to below 2000 Ω. If
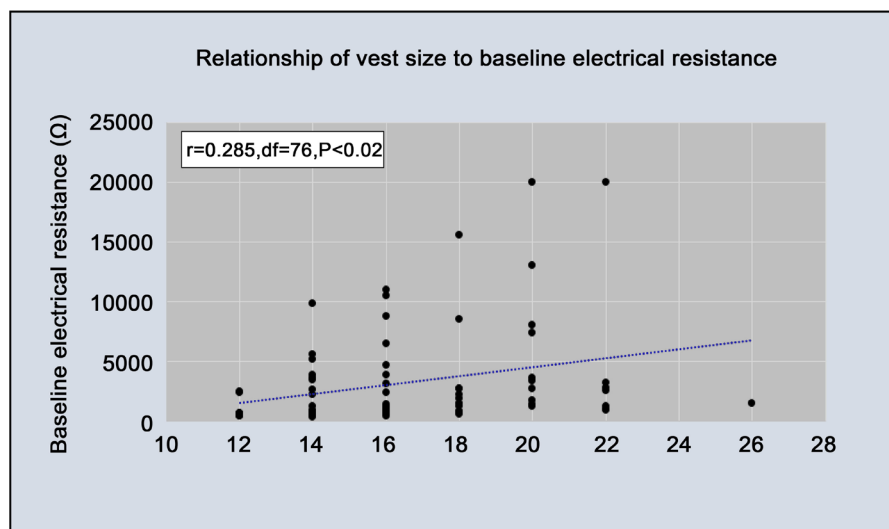

Figure 4. Relationship between vest size and baseline electrical resistance.

the true baseline resistance is 3000 Ω, a decrease to below 2000 Ω would amount to ~33%, compared with the threshold of just 20% required when the baseline resistance is within range.

Increases in vest resistance across time probably have a range of causes. Mechanical stresses produced by impacts, washing and/or storage conditions may cause breakages in individual silver nylon fibres as well as reductions in the silver content of intact fibres. Regardless of the source, resistance differences between vests means that, to ensure the fairness of a contest, the two vests involved should have similar resistance readings at the outset, but this situation can be logistically difficult to achieve.

It is noteworthy that as the overall baseline resistance of a vest rises, the areas of the vest furthest from the source of the low-level electrical current directed through the vest are typically the first to lose impact sensitivity. Variation in sensitivity from one region of the vest to another can therefore occur and is obviously problematic. Also, failure of a region of the vest makes the whole vest effectively non-functional, and repair is complex and expensive.

3.6.3. Effect of Sweat

Because the electrolyte content of sweat makes it electrically conductive, sweat can cause "shorting" of vest electrical circuitry, thereby reducing the electrical resistance of the vests employed in Box'Tag and producing false positive scores. The use of a running 10-second average to calculate baseline electrical resistance of the vest, and the identification of valid impacts through evaluation of short-term reductions in resistance relative to this baseline, was originally implemented with a view to overcoming this issue [1] [7].

That solution is predicated, however, on a premise that the influence of sweat in lowering vest resistance generally will be gradual. We have quite frequently experienced situations in which the premise has proven unjustified. For a vest in the process of becoming sweat-soaked, there can be a point at which athlete and/or vest movement can cause intermittent bridging of two sensor stripes by sweat. This can result in 'rapid-fire' recording of false positive scores for the opponent.

The outer surface of the sensor fabric incorporated into vests has a hydrophobic coating [6] designed to prevent dripping sweat (and also water that may be splashed on to contestants by trainers) from affecting vest performance. We have found, though, that this does not prevent the diffusion of sweat from the body surface of contestants through the base fabric of vests and into the sensor fabric. The effect is likely exacerbated by adverse vest washing.

We have sought to prevent the problem in several different ways, including restricting the duration of bouts (usually to a maximum of three 2-minute rounds separated by 1-minute rest intervals), encouraging contestants to wear their own T-shirt underneath the instrumented vest, and counselling against wearing of the instrumented vest during warm-up. In the most recently produced vests, the inner surface of the sensor fabric has been laminated to provide a barrier against sweat ingress. Although combination of these approaches has

limited the extent of the occurrence of sweat-induced failures of the automated scoring technology, it has not yet entirely eliminated them.

3.6.4. Spacing of Vest Stripes

Slow-motion analysis of video footage that we have collected over the past five years has confirmed the published finding of Bruch *et al.* [7] in that visible contacts to vests do sometimes fail to register scores.

Apart from diminished vest sensitivity, this apparently can be due to a glove happening to contact the vest at a point that prevents its conductive patch from bridging two vest stripes. Because the resultant occurrence of false negative outcomes occurs with sufficient frequency to be perceptible to Box'Tag contestants, there have been suggestions that vest stripes, which are currently ~4 cm apart, should be considerably closer together. Such action would of course have a downside through increasing the total length of the electrical circuit formed by the vest stripes and consequently increasing the vest electrical resistance.

3.6.5. Conductivity of Glove Patches

In the first iteration of the automated scoring technology, the conductive patches affixed to the surfaces of the gloves consisted of a conductive rubber material [1]. Since it was difficult to bond this material to the glove surface through use of readily available glues (and since eventual deterioration of the conductive rubber material started to cause discolouration of vests contacted by it), a move to silver-containing conductive fabrics soon followed.

Initially, a prototype conductive fabric manufactured by the Commonwealth Scientific and Industrial Research Organisation (Australia's leading Government-funded research agency) was employed [1]. This fabric had a higher silver content on one side than the other. We tried some gloves on which the 'high silver' side was placed downward (*i.e.* on to the glove surface) and some on which it was placed upward.

Both versions were found to work almost perfectly in initial trials, but after only a few rounds of use the patches with high-silver side down showed substantial losses of conductivity, such that contacting vests with them no longer led to reliable registration of scores. The reason for the loss of conductivity was unclear, but we suspect that impacts may have caused silver particles to be driven into the glue matrix and consequently inactivated. The patches with high silver side placed upward continued to work effectively for many years. Eventually, we moved to making patches from either of two commercially available silver-containing fabrics, ArgenMesh and Silverell (Less EMF Inc, Latham, NY, USA), and to sewing the patches in place rather than gluing them.

This has provided excellent results but our experience has demonstrated that the selection of the conductive material, its orientation and perhaps also the method of its attachment to glove surfaces can be critical to the successful operation of the automated scoring technology Progression in the development of conductive glove patches is illustrated in Figure 5.
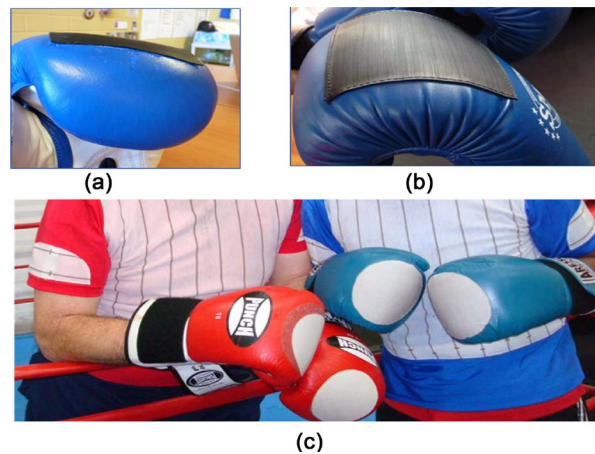
Figure 5. Evolution in the development of conductive patches for gloves, showing move from gluing to sewing as the method of fixing rubber patches to glove surfaces, and then progression to use of patches made from silver-containing fabric.

3.6.6. Possible Influence of Impact Mechanics

In 2014, two of the authors of this paper completed 23 simulated Box'Tag contests in which they alternated impacts, using specialised gloves incorporating air-containing bladders designed to minimise peak impact forces in keeping with the requirements of modified boxing (16). Each person used the same fully functional vest throughout, but in combination with 9 different transceivers.

The outcomes are presented in Table 1. Although the number of impacts delivered by each participant in each bout was identical, the "red" participant recorded a higher score in 14 of the 23 simulations. Five of the simulations produced tied scores, and the "blue" participant recorded the higher score on just four occasions. In total, 1015 valid impacts were recorded for the red participant and only 983 for the blue. The mean scores per simulation were therefore 44.1 and 42.7 points respectively.

An independent samples t-test showed that the difference in scores, despite being small, was consistent enough to be statistically significant ($P = 0.026$). It is possible that the blue vest was slightly more sensitive than the red, but since both vests had baseline electrical resistance levels below 2000 Ω this explanation is considered unlikely.

The mean contact time for impacts delivered by the red contestant was 102.1 (SD = 8.6) msec, while that for impacts delivered by the blue contestant was 81.7 (SD = 15.1 msec). This difference too was statistically significant ($P < 0.000001$). The lower mean contact time and greater variability for the blue contestant meant that a few contact times were less than the minimum duration required to register a score. The most likely reason for this outcome is that the two contestants differed slightly in terms of their impact delivery mechanics. Accordingly, there is a logical basis for believing that such differences between contestants may have at least small effects on scores produced by the automated scoring technology.

3.6.7. Effect of Automated Scoring Technology on Athlete Technique

Our long-term experience suggests that the automated scoring technology has

Table 1. Results of 23 simulated bouts in which red and blue "contestants" alternated impacts and so delivered equal numbers of impacts to designated scoring regions. Columns 2 and 3 identify the specific transceivers worn by the two contestants. Columns 4 and 5 show the final scores, and Columns 6 and 7 show the average glove contact times in msec.

Bout Number	Red Transceiver #	Blue Transceiver #	Red Score	Blue Score	Red Contact Time (Ave)	Blue Contact Time (Ave)	Bout winner
1	9	5	40	37	107.9	63.4	Red
2	10	5	40	39	97.5	69.9	Red
3	10	11	38	35	98.4	79.2	Red
4	29	11	43	38	92.7	79.8	Red
5	29	9	40	36	93.4	83.8	Red
6	1	2	40	38	102	82.8	Red
7	23	2	39	34	109.7	89.4	Red
8	23	24	35	39	122.5	127.3	Blue
9	25	24	41	43	92.3	98.2	Blue
10	25	26	44	44	106.8	74.4	Draw
11	1	26	48	51	101.5	75.4	Blue
12	18	19	46	46	101.6	57	Draw
13	21	19	43	47	90.5	75.3	Blue
14	21	22	51	50	93.8	86	Red
15	18	22	48	46	100.2	68.8	Red
16	1	2	47	47	97.9	81.9	Draw
17	23	2	49	46	100.3	75.6	Red
18	23	24	45	42	95.8	70.5	Red
19	25	24	45	42	94.8	68.6	Red
20	25	26	44	39	111.9	99.7	Red
21	1	26	46	41	114.2	103.6	Red
22	30	31	52	52	117.3	82	Draw
23	30	31	51	51	105.8	87.2	Draw
Mean			**44.1**	**42.7**	**102.1**	**81.7**	
SD					**8.6**	**15.1**	

caused contestants to adopt techniques aimed simply at rapidly registering numerous points and depending more on capacity to sustain all-out attack rather than on the execution of highly-refined skills.

3.7. Other Observations

For much of the history of our involvement with the automated scoring technology, it has been necessary to have multiple sets of gloves set aside purely for use in Box'Tag contests, as there was a risk that the conductive patches glued or sewn on to the gloves could be damaged during training activities so rendering them non-functional when automated scoring was required. This situation has

Figure 6. Gloves with pull-on Lycra covers incorporating sewn-on silver-containing conductive patches.

now been addressed through the development of "pull-on" glove covers incorporating conductive patches. The covers have been specifically designed to fit over the specialised impact-damping gloves that we have designed and produced with a view to maximising the safety of modified boxing, but are usable also with standard glove types (**Figure 6**).

At present, the cost of the automated scoring technology is an apparent impediment to its uptake. Our investigations have shown that, with low-volume manufacture, vests would have to retail for ~AUD 250 (USD 190) in order to make their supply commercially viable. Several parties that have shown initial interest in obtaining the automated scoring technology have become deterred when advised of this price.

4. Discussion

Our finding that first exposure to the automated scoring technology elicits highly positive reactions is perhaps unsurprising, since the technology appears to offer a tangible solution to a problem that has long plagued many subjectively judged sports—the potential for outcomes to be influenced by conscious or unconscious human bias. In boxing, disputed judging decisions have been common over many years [22] [23].

Boxing competitions at Olympic Games have been affected by highly controversial results over more than a century [22], and the 2016 Olympics in Rio de Janeiro were no exception [24] [25]. There have been frequent accusations of official corruption [22] [26] [27] and, where judging is entirely dependent on expressions of a small number of human opinions, the risk of such corruption is difficult to exclude. Additionally, evidence exists that human judging may be prejudiced by subliminal factors including knowledge of the prior competitive records of athletes and their nations of origin [28], crowd noise [29] [30], and even the colours of athlete uniforms [31]. Against this background, a demonstrable possibility of technology-enabled, completely objective scoring has obvious allure.

A further enticement of the automated scoring technology is the availability of a scoreboard that changes in real-time so that progress toward the final result can be continually monitored by athletes, their immediate supporters and spectators.

This contrasts with the current situation in both amateur and professional boxing where official scores remain unknown until post-contest announcement. In amateur boxing, real-time scoring was implemented during the period between 1992 and 2012 but was contingent upon human ability to perceive legitimate contacts [26]. Each of five judges had a keypad incorporating a red and a blue button, and was required to press the appropriate button to indicate a perception that a contestant wearing that colour had registered a valid impact to the designated target region of the opponent. If three of the five judges pressed a button of the same colour within a 1-second window, a point was awarded to the corresponding boxer. Over time, shortcomings of this method became apparent, including the possibility that all five judges could press the button more times for one competitor than the other only to have that competitor lose because of the vagaries of the timing window. The use of the timing window meant that judges who pushed buttons frequently could have a disproportionate influence on results, but also led on several occasions to contests finishing in nil-all draws with outcomes then having to be determined through the casting votes of officials. The fundamental weakness was the inability of even the trained human eye to cope with the rapidity of action, and the problems eventually forced abandonment of the method [32] in favour of the present system in which each of five judges simply forms an overall impression as to which contestant is the winner of a round and awards that contestant 10 points, with the other contestant awarded a lesser number based on the perceived extent of the win [26]. At the end of the contest, the round-by-round scores are summed to produce final scores. The scorecards of three of the five judges are then randomly selected and the athlete with the higher score on at least two of these is declared the victor. The collation process typically takes several minutes.

It is arguable that this delay and the real-time invisibility of the scores is detrimental to the sport, particularly as it prevents build-up of excitement associated with empirically close contests. With the fully automated scoring technology described in this paper such issues are overcome, and our experience shows that this is commonly regarded as a major advantage. It may be particularly important in the context of modified boxing, where entertainment value deriving from the prospect of highly forceful impacts and knockouts is (by design) absent.

Over the past five years, though, it has become clear to us that in its current form the automated scoring technology is not without problems of its own. Cost is a significant barrier to its uptake but could be addressed if demand for the technology was eventually sufficient to enable manufacture of its components in higher volumes. Of more immediate concern are issues relating to the performance of the technology.

For the judging of sport through entirely technological means to be viable, it needs to be almost fail-proof, and this is a very difficult criterion to meet. We initially thought that, with Box'Tag having a community sport orientation rather than a high-performance focus, a scoring system capable of detecting and registering ~90% of all valid impacts—as determined by Bruch *et al* [7]—would be

acceptable, particularly when compared with the very limited precision of alternative scoring methods utilised in conventional boxing settings.

It soon became evident, however, that even people competing in friendly community environments often have a fierce desire to win, and that the very notion of using technology for judging engenders an expectation that results should be unquestionable. We have regularly received complaints from contestants who have felt that some valid impacts to the target zones of their opponents have failed to score (possibly due to failure of the conductive patch on the glove to bridge two vest stripes).

On one occasion following a close contest, we recorded a comment to the effect that "if you're going to be cheated by technology, you might just as well be cheated by human judges". It is thus evident that requisite standards for technological judging are higher than those for human judging.

While the above applies to situations in which the technology appears to be working perfectly, there are far greater ramifications when the system overtly fails, such as when impacts to a shoulder target region clearly stop scoring or sweat causes sudden generation of a series of false positive scores. Failures of this nature have sometimes caused participants to express profound disappointment and have exerted long-term effects on their enthusiasm. When considered in conjunction with the large time commitment that we have found necessary for the organisation of Box'Tag events, the risk of technological failures and consequent negative feedback creates a sense of caution with regard to the holding of such events and, in our opinion, has therefore restricted their frequency. In the context of the events, a single technological failure within a program of numerous bouts can leave an adverse impression. For this reason, it can easily happen that the focus of the events becomes directed more to the operation of the technology than to other vital aspects of participant experience.

A question obviously arises as to how the practical issues that we have encountered might be addressed. We have conceived and explored a wide range of possibilities. One of these has entailed experimentation with an alternative to the automated scoring technology. A method has been developed that allows judges and/or audiences to use a simple interface (see Figure 7), accessible via either

Figure 7. On-line interface for subjective judging of modified boxing. contests.

mobile phone or tablet, to communicate their subjective impressions of a contest to a dedicated website. The communication occurs at the end of each round and requires use of a maximum of three screen touches.

The first touch identifies both the perceived winner of the round and (on a 5-point scale) the margin, while the second denotes the perceived technical quality of the contest (on a 7-point scale). The third touch leads to submission of the assessment. Once the information is received via the website, an algorithm is applied to convert the inputs of each judge to a score.

The number of points assigned to the perceived winner of the round is determined as the quality rating (which progresses in increments of 0.1429 corresponding to the 7 equally-spaced points on the rating scale) multiplied by 20. The score assigned to the opponent is one point below that of the perceived round winner if the margin is considered slight and five points less if it is considered extreme. Other ratings of margin lead to differences between these values. Scores calculated from the inputs of each individual judge are averaged to determine the official score for the round, and round scores are summed to provide the overall bout score.

The method is designed to reward technical proficiency, in that the loser of a high-quality contest can receive more points than the winner of a bout of lesser standard. It also permits selection of a 'quality only' mode in which judges give no opinion as to the winner of a round or the margin but instead enter just the quality rating. This feature provides for the possibility of competition between bouts as opposed to competition between individuals within a bout, a model comparable that which exists in, for example, dance sports, figure skating, synchronised swimming and synchronised diving.

This alternative to the automated scoring technology obviously does not require the use of instrumented vests or gloves and capitalises on the ubiquity of mobile technology. Although its employment of subjective human judgements exposes it to the same criticisms of potential bias as those directed to the scoring methodology presently employed in conventional boxing, it does have the advantage of being able to incorporate inputs from very large numbers of judges, thereby exploiting the notion of the "wisdom of crowds" [33]. In circumstances where live streaming of video footage is available, the judges need not even be physically present at the contest, with capacity for remote voting having at least some potential to minimise occurrence of "home-town" decisions. We have tried this crowd-based scoring approach in a variety of settings.

At a special Box'Tag session involving participants aged less than 12 years, the "quality only" option was utilised. The participants showed great enjoyment of the session and quickly realised that there was reward for skill and for working with rather than against their partner. Feedback from parents, some of whom took part in the judging, was highly positive.

Later we employed the method for the purposes of a Box'Tag competition associated with a Draft Camp held at the Australian Institute of Sport for athletes interested in moving into boxing from other sports. Again, the response was

favourable, but when a similar camp was held a year later, the organisers specifically requested the use of the automated scoring technology, largely because they felt it would better contribute to the competitive atmosphere that they wanted to create.

Our most recent major trial of the alternative scoring method was at a 2015 Box'Tag event held at the Sydney club where the Box'Tag concept was first implemented. After three contests, the principals of that club chose to return to the automated scoring technology for the remainder of the program, based on an emergent realisation that the real-time scoring and the sounds associated with it were critical to maximising participant and audience enjoyment of the event.

Also, when another club expressed interest in establishing a Box'Tag program and the potential for it to save on costs by initially adopting the alternative scoring method was raised, its owner indicated that in his view, "Box'Tag without the automated scoring technology would no longer be Box'Tag". These experiences strongly suggest that, despite its current limitations, the automated scoring technology remains valued by users, and that effort to surmount the limitations is therefore merited. Objective and subjective approaches to scoring need not be mutually exclusive, with the latter possibly becoming a social media metric review or "check" with quite wide applications.

What, then, can be done? We reason that if increases in the electrical resistance of vests over time are largely or even partly attributable to mechanical stresses occurring during washing, decreasing the requirement for washing should prolong effective vest life. Accordingly, we have recently produced two prototype vests in which the base fabric is a thin mesh so that the capacity of the vest to hold sweat is greatly diminished. This could allow post-contest recovery of the vests to often occur by simple drying, rather than being always dependent on washing. The availability of the new prototype vests will permit testing of the idea in the near future.

As another possible way of restricting increases in vest resistance to levels consistent with continued vest functionality, the silver nylon vest stripes could be arranged into several electrical circuits of lesser length, rather than comprising a single circuit. This would necessitate quite substantial vest redesign, and transceivers would have to modified to enable integration of multiple input/output channels rather than just one, but if this could be accomplished it would seem highly likely to yield a positive outcome.

It would be particularly good to have the shoulder target regions on easily replaceable separate circuits, so that any loss of sensitivity of those targets could be immediately redressed, without need to take the whole vest out of commission and undertake complete refitting of the athlete.

If vests were adapted to encompass multiple electrical circuits, stripes of silver nylon yarn within each circuit could perhaps be closer together, making it easier for glove conductive patches to bridge two stripes and therefore reducing the incidence of false negative occurrences. Also, if each circuit was constrained to a length which meant baseline electrical resistance very rarely increased to levels

preventing impact detection, the time requirement for checking of vests prior to their deployment could be considerably diminished, possibly leading to an increase in deployment frequency and substantial reduction in costs and inconvenience attendant upon needs for vest repair.

False positive scores resulting from vest shorting could be largely prevented by making all vests tight-fitting and giving further attention to relatively simple aspects of their design. Currently, vest shoulder stripes are at right angles to the long axis of the arm [6]. Logic suggests that if they were instead positioned parallel to the long axis of the arm, the probability of their contact with torso stripes would be substantially decreased. The elimination of false positive scores due to vests being affected by sweat has proved to a challenging task.

To date, lamination of the body-facing surfaces of vest sensor fabrics has been successful only in delaying the onset of the problem rather than in removing it altogether. Furthermore, we have observed cracking of the laminate after only a few uses and washes of the vests. This might be due to imperfections in the lamination process, particularly as the process has been carried out by people who had little prior experience with it. Accordingly, there could be benefit in at least short-term persistence with the idea. It is of course evident that the sweat problem is resolvable through introduction of a sufficient barrier to movement of sweat from the body surface into the sensor fabric.

We are currently investigating the feasibility of employing protective vests in Box'Tag as a means of lowering impact forces, thereby further augmenting participant safety. A first prototype of such a vest has been produced by one of our associates and, in addition, we have become aware of the recent release of a potentially suitable torso protector by the New Zealand company QP Sport. Participant receptiveness to the wearing of a protective vest will undoubtedly depend on whether the vest is comfortable and allows freedom of movement.

If uptake can be achieved, fixing of a sensor fabric to the out-facing surface of the protective vest could well be the answer to the sweat predicament, since the thickness of the vest would provide a quite substantial separation distance between the body surface and the sensor fabric, and in any case the materials comprising the vest (including EVA foam and plastics) would have low sweat permeability. Notably, attachment of the sensor fabric to a relatively firm platform such as the surface of a protective vest would essentially negate any possibility of false positive scores caused by "flapping" of the vest and consequent electrical shorting though contact between vest stripes.

With regard to our observations concerning the unorthodox techniques and strategies that have been adopted by participants in contests involving use of the automated scoring technology, the seemingly common predilection to continuous all-out attack is probably explained by perceptions that the technology rewards such an approach and that the comparative safety of Box'Tag makes caution largely unnecessary. The prevalence of the attacking strategy has important consequences. One is that physical fitness attributes and physical characteristics of contestants often become primary determinants of contest outcomes.

While the development and expression of physical fitness deserves encouragement, beauty of movement and the demonstration of exquisite skill can be considered vital to the fascination of sport [34] [35] [36] and among the factors that differentiate sport and exercise.

The popular appeal of elite-level conventional boxing has been attributed partly to its inclusion of highly-evolved dance-like qualities [37] [38], and the boxing style of a current world professional champion has been described as a "symphony of movement" [39]. We consider it likely that future uptake of Box'Tag and other technology-supported forms of modified boxing will depend on their ability to capture something of this aesthetic element.

Another effect of the "constant attack" strategy employed by many Box'Tag contestants is that the occurrence of large numbers of impacts during bouts. We have observed contests in which scores at the end of three 1½-minute rounds indicate that each contestant has received ~200 impacts. Although none have been head impacts, we believe that the quest for participant safety—a quest that has underpinned the entire development of Box'Tag [1]—would be better served by reduction of impact frequency.

In general, then, a case exists for attempting to foster more skill-orientated Box'Tag styles. Several possible approaches already have been explored. An initial idea was to simply ask referees to ensure that periods of engagement between contestants were restricted to a few seconds, with instructions to separate issued at the end of each such period. Consistency and precision in referee compliance with this request proved difficult to achieve. We therefore looked into the feasibility of calculating a 'bout quality index' that could be used to adjust raw scores. The transceivers forming part of the automated scoring technology incorporate tri-axial accelerometers. We reasoned that the sum of accelerometer readings across the three axes should provide estimates of contestant work outputs, and that if the sum of the readings for the two contestants was divided by the sum of their combined scores, the result might be an indicator of bout quality. This was predicated on the assumption that the combination of high workloads and relatively infrequent scores would typically be due to the contestants displaying good defensive techniques, but the method had the disadvantage of adding an extra layer of complexity in that it demanded accuracy (and therefore pre-contest calibration) of the accelerometers, and in any case it produced a number of results that were markedly inconsistent with the subjective assessments of experts. There was also the problem that, once fully understood by contestants, the method could lead to deliberately inefficient actions aimed purely at increasing the accelerometer-derived workload estimate, and to excessively low levels of effort to score points.

In 2014, we implemented a different approach through modification of the Spartan software so that no contestant could register more than three points in any running 4-second period. The idea was that a contestant who registered three points in very rapid succession would then find it best to disengage so as to avoid the possibility of the opponent scoring during the remainder of the four

seconds. This failed to have the desired effect, perhaps because contestants found it hard to perceive exactly when they had scored three points and so tended to simply ignore the constraint.

Our next experiment was the introduction of the "countdown" scoring method under which each contestant starts each round with an allocated number of points and is required to defend them, with a point being lost every time a target area on their vest is contacted by the opponent. At the end of the round, the contestant with the most points remaining is the round winner. A round can be concluded inside its scheduled duration if the score of one contestant diminishes to zero. At the end of the bout, the contestant who has won most rounds is the bout winner (just as in tennis each set starts afresh and the player winning the most sets is the eventual match winner). To date, use of the countdown method has been the most successful of all our efforts to encourage more sets of skilled athlete technique in the modified boxing context.

It also has the advantage of enabling implementation of a handicap system since contestants in the same bout may start with different numbers of points. "A count-up" method in which each contestant was restricted to obtaining a certain number of points per round might have a similar effect.

We recognise that our perspectives on the present status and potential of the automated scoring technology have been influenced by the historical context in which the development of the technology has occurred, and by our own backgrounds. The project commenced at the Australian Institute of Sport, an organisation focused on the high-performance component of the sport spectrum. Three of us have been employed there, and the other has collaborated with the organisation on a range of research activities. One of us has an extensive background in amateur boxing, and another two have a history of interest in the sport.

Originally, the development of the automated scoring technology was aimed primarily at providing a more objective scoring system for conventional amateur boxing [40], and the project was therefore somewhat constrained by practices associated with that sport. In particular, the technology was designed for use in situations involving two highly-trained contestants competing against each other in a ring, watched by an audience, and with the result often having substantial consequences. These constraints largely persisted when the technology was later deployed to support modified boxing, since the new application was instigated by a former amateur boxer whose own vision had its roots in traditional high-performance boxing protocols. As modified boxing has evolved, it has become evident that conventional competition formats represent only one of many possible embodiments, and that there may be scope for the automated scoring technology to engage new communities of users willing to understand and accept the limitations to its accuracy and to tolerate occasional system failures while progressively developing their own games and rules in concert with the technology iteration. The perceptions of these communities regarding the aesthetics of their activities could well differ from ours since they would be relatively

uncontaminated by knowledge of techniques and movement styles typically seen as constituting the "art" of boxing. The value proposition associated with the current cost of the automated scoring technology also might be seen in a different light. Potential exists for the technology to be gainfully employed in specific non-combative but still competitive situations, such as counting the number of impacts to punching bags or training pads, and even counting the number of repetitions of standard physical fitness exercises entailing a series of contacts between two surfaces. The practical utility of the technology for applications of this nature remains to be investigated. If it was confirmed, the modified boxing concept would be augmented. At present, though, our observations are necessarily confined to the circumstances in which we have operated—circumstances dictated by habituated views concerning the characteristics of boxing.

Our experience suggests that the automated scoring technology is valuable in giving modified boxing a visible "point of difference" capable of capturing public attention. The issues that we have documented in relation to the technology could very likely be satisfactorily addressed through concerted input from product development specialists if this input could be afforded. Up to now, development of the automated scoring technology has been funded almost entirely through small research grants and the community of 'niche' users has not been required to pay for it. This obviously is an unsustainable situation. Over time, existing components of the technology will progressively deteriorate, making failures more common. Pursuit of further development will be contingent on ability to attract new investment from research agencies, the community of users and/or other sources.

Since the current users are largely community-based boxing clubs with tight budgets any inputs from them are likely to be small. Even if investment can be secured, success in addressing the issues that have emerged with respect to the automated scoring technology cannot be guaranteed. Exploration of low-cost alternatives to the technology therefore remains essential to the goal of building a long-term future for modified boxing.

Continuing refinement of the automated scoring technology will be worthwhile if it can contribute to accomplishment of that goal. Assessment of the prospects in this regard should take account of expanding notions as to the range of activities that eventually might be encompassed under the banner of modified boxing.

5. Conclusions

This paper highlights challenges associated with an attempt to implement an objective, automated scoring system in a sport context. While we have focused on experience in the specific situation of modified boxing, the insights that we have obtained may have relevance also to other sports. Although there is an increasing trend for sports to use technology as an aid to judging, and particularly as a means for "on-the-spot" resolution of uncertainties concerning human judging decisions [8], complete replacement of human judges through deployment

of technology is unusual. Despite recent rapid increase in the availability of wearable technology, we are among the first groups to attempt to use it to achieve completely objective judging of sport contests in which results would otherwise depend on subjective human determination. Indeed, we are perhaps the very first to make such an attempt in a sporting environment where scoring frequency can be high.

Our work has yielded lessons that could be relevant to other sports investigating the use of technology to address problems arising from human judging. For example, we have found that appetite does exist for removal of judging subjectivity, but that even a technological solution that initially seems quite simple can turn out to have unforeseen complexities. Also, athletes expect very high precision from a technologically-based scoring system—precision far greater than that currently demanded of human judges. To be truly viable, automated scoring technology needs to be virtually fail-proof, since any breakdown can lead to considerable disruption of an event, and to athlete and spectator disappointment. During the developmental phase, when some failures are almost inevitable, a subjective scoring method is likely to be needed as a back-up. The developmental phase can prove to be quite long. Contestants may tend to alter their techniques in response to introduction of new scoring methodology, and new interventions might then be required to protect the aesthetics of the sport. The new scoring method may favour athletes with particular physical or physiological characteristics, thereby diminishing the success rates of others and influencing their levels of enjoyment.

We have come to realise that while technology can presently provide an effective means for such basic tasks as detecting and counting impacts, it is not yet able to discern more subtle aspects of performance, such as beauty of movement and quality of skill execution. Because these subtleties are central to the appeal of sport, a strong argument exists for retaining some contribution of human perception to sports judging, perhaps in concert with a technological approach.

Finally, our efforts over the past five years have emphasised the importance of designing and progressively refining technologies to meet the needs of sports, without expectation that sports should adapt to fit with the capabilities of the technologies. At same time, however, it needs to be recognised that technological developments can yield opportunities for sports to take on new dimensions.

Acknowledgements

References

[1]   Hahn, A.G., Helmer, R.J., Mackintosh, C., Staynes, L.M. and Blanchonette, I. (2011)

Technological Foundations and Current Status of a Modified, Low-Risk Form of Competitive Boxing (Box' Tag®). *Sports Technology*, **4**, 78-184. https://doi.org/10.1080/19346182.2012.725413

[2] American Medical Association (1999) Meeting of the American Medical Association: Reports of the Council on Scientific Affairs, in 1999 Meeting of the American Medical Association: Reports of the Council on Scientific Affairs; Boxing Injuries. pp. 6-7.

[3] Beran, R. and Beran, J. (2009) The Law(s) of the Rings: Boxing and the Law. *Journal of Law and Medicine*, **16**, 684-695.

[4] Pearn, J.H. (2000) Boxing, Brains and Balls. *Medicine Today*, **1**, 146-148.

[5] Boyle, R. (1983) Too Many Punches, Too Little Concern. In: *Sports Illustrated*, Time Inc, New York, 44-67.

[6] Helmer, R.J.N., Hahn, A.G., Staynes, L.M., Denning, R.J., Krajewski, A. and Blanchonette, I. (2010) Design and Development of Interactive Textiles for Impact Detection and Use with an Automated Boxing Scoring System. *Procedia Engineering*, **2**, 3065-3070. https://doi.org/10.1016/j.proeng.2010.04.112

[7] Bruch, H., Hahn, A.G., Helmer, R.J., Mackintosh, C., Blanchonette, I. and McKenna, M.J. (2011) Evaluation of an Automated Scoring System in a Modified Form of Competitive Boxing. *Procedia Engineering*, **13**, 445-450. https://doi.org/10.1016/j.proeng.2011.05.112

[8] Officiating (2013) The Impact of Technology in Sport. https://technologyinsport.wordpress.com/tag/officiating/

[9] Elliott, J. (1991) Action Research for Educational Change. Open University Press, Buckingham.

[10] Stringer, E. (1999) Action Research. 2nd Edition, SAGE Publications, Thousand Oaks.

[11] Kemmis, S. and McTaggart, R. (2000) Participatory Action Research. In: Denzin, N.K. and Lincoln, Y.S., Eds., *Handbook of Qualitative Research*, SAGE Publications, Thousand Oaks, 567-595.

[12] Berg, B. (2001) Qualitative Research Methods for the Social Sciences. Allyn & Bacon, Boston.

[13] Campbell, K. (2013) Why We Need More Practitioner Research: A Response to "A Teacher Educator Uses Action Research to Develop Culturally Conscious Curriculum Planners." *Democratic Party on Education*, **21**, 7.

[14] McTaggart, R. (1991) Principles for Participatory Action Research. *Adult Education*, **41**, 168-187. https://doi.org/10.1177/0001848191041003003

[15] Helmer, R.J.N. and Hahn, A.G. (2009) Impact Detection Method and Apparatus. https://www.google.ch/patents/EP2493584A1?hl=de&cl=en

[16] Perkins, P., Hahn, A., Lucas, R. and Keegan, R. (2014) The Boxing Conundrum: Is There a Place for a New Variant of the Sport? *Journal of the Humanities and Social Sciences*, **2**, 9-25. http://www.questjournals.org/jrhss/papers/vol2-issue9/B290925.pdf

[17] Perkins, P., Hahn, A., Keegan, R. and Collis, I. (2014) Why Do Community-Based Athletes Choose to Participate in a Modified, Low-Risk Form of Boxing? An Interpretative Phenomenological Analysis. *Journal of the Humanities and Social Sciences*, **2**, 60-69. https://doi.org/10.9790/9467-2116069

[18] Perkins, P., Hahn, A.G. and Keegan, R. (2015) What Are the Perceived Coaching Practices that Have Contributed to Participation in a Modified Low-Risk Form of

Boxing? An Interpretive Phenomenological Analysis. *Journal of Sports and Physical Education* (*IOSR-JSPE*), **2**, 46-55.
http://www.iosrjournals.org/iosr-jspe/papers/vol2-issue5/K0254655.pdf

[19] Leitch, R. and Day, C. (2000) Action Research and Reflective Practice: Towards a Holistic View. *Educational Action Research*, **8**, 179-183.
https://doi.org/10.1080/09650790000200108

[20] HearnsMyFave (2015) "Fighting out of the Red/Blue Corner..." Anyone Truly Know the Origins?
https://www.reddit.com/r/Boxing/comments/38iljh/fighting_out_of_the_redblue_corner_anyone_truly/

[21] Lazarus, B. (2011) Quantification of Work Intensity in a Modified Low-Risk Form of Competitive Boxing (Box' Tag). Victoria University, Melbourne.

[22] Eisele, A. (2017) Biggest Controversies in Olympic Boxing History.
http://boxing.about.com/od/history/a/controversy.htm

[23] Ammer, C. (2002) We Wuz Robbed.
http://www.dictionary.com/browse/we-wuz-robbed

[24] Staniforth, M. (2016) Olympic Boxing Body AIBA Expels Referees and Judges Involved in "Handful" of Controversial Decisions. Daily Mirror, London, 17 August 2016.

[25] Agence France Press (2016) Boxing on Brink after Rio Judging Controversies—The Financial Express.
http://www.financialexpress.com/sports/rio-2016-olympics/boxing-on-brink-after-rio-judging-controversies/353764/

[26] McClintock, A. (2012) Olympic Boxing: Controversies Highlight the Failures of the Scoring System.
https://www.theguardian.com/sport/blog/2012/aug/02/olympic-boxing-controversies-failures-scoring-system

[27] Belson, K. and Blumenthal, S. (2016) Irish Fighter Accuses Amateur Boxing Body of Being Corrupt. The New York Times.
https://www.nytimes.com/2016/08/17/sports/olympics/irish-fighter-accuses-amateur-boxing-body-of-being-corrupt.html?_r=0

[28] Ste-Marie, D.M. and Valiquette, S.M. (1996) Enduring Memory-Influenced Biases in Gymnastic Judging. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, **22**, 1498-1502. https://doi.org/10.1037/0278-7393.22.6.1498

[29] Balmer, N.J., Nevill, A.M. and Williams, A.M. (2001) Home Advantage in the Winter Olympics (1908-1998). *Journal of Sports Sciences*, **19**, 129-139.
https://doi.org/10.1080/026404101300036334

[30] Balmer, N.J., Nevill, A.M. and Williams, A.M. (2003) Modelling Home Advantage in the Summer Olympic Games. *Journal of Sports Sciences*, **21**, 469-478.
https://doi.org/10.1080/0264041031000101890

[31] Hagemann, N., Strauss, B. and Leibing, J. (2008) When the Referee Sees Red. *Psychological Science*, **19**, 769-771. https://doi.org/10.1111/j.1467-9280.2008.02155.x

[32] Herman, M. (2012) Boxing to Scrap Computerised Scoring after London. Reuters.
http://uk.reuters.com/article/uk-olympics-boxing-scoring-idUKTRE80B17420120112

[33] Surowiecki, J. (2005) The Wisdom of Crowds. Reprint Edition, Anchor, New York, 336.

[34] Boxill, J. (1984) Beauty, Sport and Gender. *Journal of the Philosophy of Sport*, **11**, 36-47. https://doi.org/10.1080/00948705.1984.9714413

[35] Lacerda, T. and Mumford, S. (2010) The Genius in Art and in Sport: A Contribution to the Investigation of Aesthetics in Sport. *Journal of the Philosophy of Sport*, **37**, 182-193. https://doi.org/10.1080/00948705.2010.9714775

[36] Ryall, E. (2016) Sport Is Not Just about the Athleticism: Beauty Matters Too. https://aeon.co/ideas/sport-is-not-just-about-athleticism-beauty-matters-too

[37] Lewis, C. (1986) Thrilling Sport Boxing Is Brutal but Also Beautiful. The Inquirer, London.

[38] Chandler, D. (1996) Introduction: The Pictures of Boxing. An Anthology of Writings on Boxing and Visua Culture. MIT Press, Cambridge, 13-26. https://doi.org/10.1080/j006v16n01_01

[39] Makinde, A. (2013) Floyd Mayweather: The Aesthetics of African-American Boxing Style. http://adeyinkamakinde.blogspot.com.au/2013/10/floyd-mayweather-aesthetics-of-african_2

[40] Hahn, A., Helmer, R., Kelly, T., Partridge, K., Krajewski, A., Blanchonette, I., *et al.* (2010) Development of an Automated Scoring System for Amateur Boxing. *Procedia Engineering*, **2**, 3095-3101. https://doi.org/10.1016/j.proeng.2010.04.117