STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

QWERTY AND THE SEARCH FOR OPTIMALITY

BY

NEIL M KAY

NO 13-24

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

QWERTY and the search for optimality

Neil M. Kay

Department of Economics, University of Strathclyde
e-mail for correspondence: n.m.kay@strath.ac.uk

Abstract

This paper shows how one of the developers of QWERTY continued to use the trade secret that underlay its development to seek further efficiency improvements after its introduction. It provides further evidence that this was the principle used to design QWERTY in the first place and adds further weight to arguments that QWERTY itself was a consequence of creative design and an integral part of a highly efficient system rather than an accident of history. This further serves to raise questions over QWERTY's forced servitude as "paradigm case" of inferior standard in the path dependence literature. The paper also shows how complementarities in forms of intellectual property rights protection played integral roles in the development of QWERTY and the search for improvements on it, and also helped effectively conceal the source of the efficiency advantages that QWERTY helped deliver.

(1) Introduction [1]

This paper builds on Kay (2013a and b) and associated commentaries by Arthur (2013), Margolis (2013) and Vergne (2013). Kay 2013(a) concluded that Christopher Latham Sholes' QWERTY typewriter format "could be regarded as near-optimal for its time in terms of crucial features relating to format/device compatibility". This raises the issue of whether Sholes could have improved the efficiency of his device even further in those terms. Here we look at how Sholes did indeed pursue that objective. The evidence is contained in two patents filed separately by Sholes some years apart but only issued posthumously. The solution that Sholes found is not evident if the patents are considered independently of each other. However, they can be combined like two pieces of a jigsaw puzzle to help reveal Sholes intentions. Even then, Sholes solution only becomes clear when the result is interpreted in the light of his infrequency meta-rule or principle as explained in Kay (2013a). Had a typewriter ever been built integrating specifications set out in these two patents, it would have had the potential to outperform QWERTY machines on the crucial performance criterion that had helped generate QWERTY's own efficiency advantages and ensure its eventual adoption as dominant standard.. .

This also provides further evidence to support the arguments in Kay (2013a) that Sholes infrequency principle lay behind his original design of QWERTY because it shows he continued to apply the principle for new models of typewriters even after he ceased to have any direct association with Remington and QWERTY. At the same time, we will show how Sholes employed complementarities and timing in the use of different forms of IPR (intellectual property rights) protection in a systematic effort to pursue and protect his interests in these areas. In turn, by providing further support for the theoretical arguments in Kay (2013a), it may be regarded as reinforcing the conclusions drawn from that paper to the effect that QWERTY was deliberately designed and was as near-optimal in terms of crucial systems compatibility features as could be reasonably expected with the state of technical knowledge of his day. Given QWERTY's role as the "paradigm case" (Lewin, 2001) of a supposed inferior standard and accident of history in the literature on path dependence, this paper is also intended to make a contribution in that context.

In Section 2 we look at the development of the QWERTY keyboard and the role of the infrequency principle in its design, and then consider the genesis of the principle in Section 3. Section 4 looks at evidence from history in terms of how firms have exploited complementarities in different forms of IPR protection and Section 5 takes up that issue in the context of QWERTY and Sholes subsequent efforts to improve on it. We discuss some implications in terms of IPR and path dependence in Section 6, and then finish with a short concluding section.

(2) The evolution of QWERTY and the infrequency principle

Arthur (1983 and 1989) and David (1985) outline the basic mechanisms by which technical standards can evolve and become locked in to one dominant form, with David looking particularly at the case of QWERTY in this context. David's (1985) paper sets out the basic conditions of technical interrelatedness, economies of scale and quasi-irreversibility of investment which could lead to lock into a particular technical standard. Technical interrelatedness in the case of QWERTY arose from the need for system compatibility between the typist's skills and the keyboard format. Network externalities arose from typists trained in one keyboard format (here QWERTY) and in the training market for typing skills. Quasi-irreversibility reflected the costs skilled typists would encounter from unlearning QWERTY if they switched to another format. David argued these processes led to QWERTY standard becoming "locked-in" as dominant standard even when the 1936 Dvorak design displayed what were claimed to be more efficient properties in terms of ergonomic design and typing speed.

David (1985) also described QWERTY as reflecting the outcome of a path dependent sequence of economic changes in which "important influences upon the eventual outcome can be exerted by temporally remote events, including happenings dominated by chance elements rather than systematic forces" (p. 332). David concluded that such early random events can lead to the wrong system being adopted and his analysis of QWERTY in this connection led to its being widely cited as the paradigm case of path dependence where an inferior standard is adopted as a consequence of accidents of history.

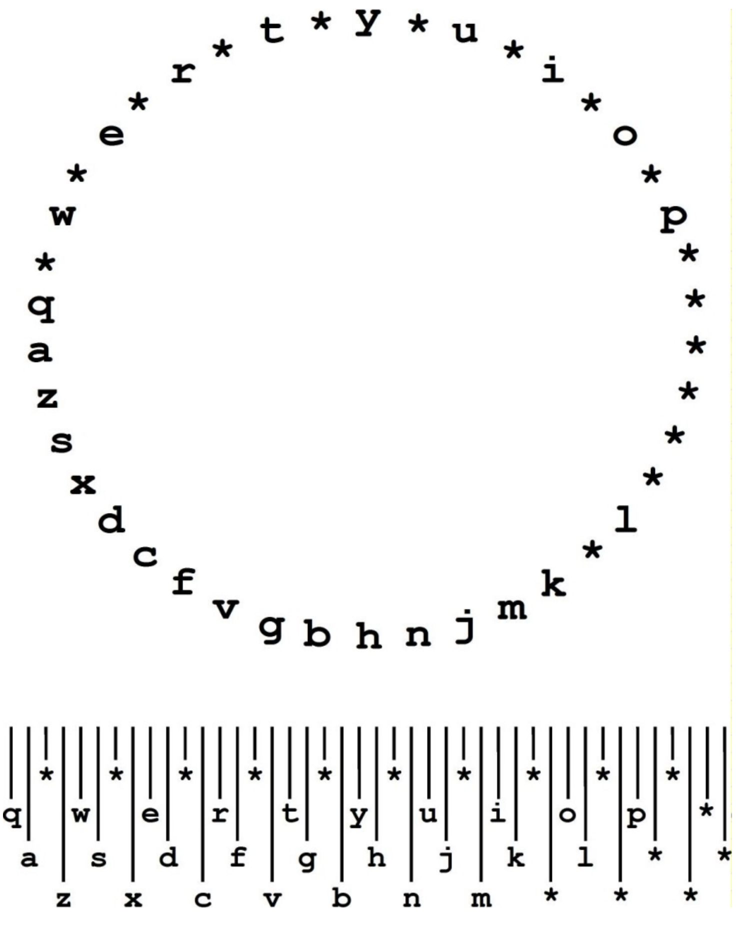

Figure1: Remington no.7 keyboard and typebasket

Before we consider these issues in more detail we consider the circumstances surrounding the early development and diffusion of QWERTY. The format was developed in 1872 by Chrtopher Latham Sholes working with his partner James Densmore and the basic relationship between format and hardware remained essentially unchanged with minor modifications over several subsequent models by Remington. Figure 1 shows a stylized top elevation representation for the keyboard and typebasket of the later (1896) Remington No. 7. The letters on the keyboard were printed with typebars, metal strips with the character to be printed on their ends. The typebars hung down in a typebasket forming a shape rather like an ice cream cone with the bottom half cut off. When a keyboard key was struck, the corresponding typebar swung upwards inside the cone to hit the appropriate point on the printed page. To simply Figure 1 we have represented all non-letter characters such as numbers and punctuation with an asterisk in the respective cases.

As noted in Kay (2013a), a problem that Sholes faced in early typewriter design was that adjacent typebars in the typebasket tended to jam (Joyce and Moxley, 1988). Rehr (1997) commented that to deal with this problem, "all Sholes needed to do was separate the letter pairs by at least one type bar." (p.4).

The absence of evidence as to how and why Sholes developed QWERTY has led to much speculation down the years, including claims that he used a list of frequent letter pairs to separate them on the typebasket and so reduce the chances of jamming. However, Kay (2013a) used a combination of basic probability theory and lists of frequent letter pairs in the English language to show this strategy would not have been much more effective than scattering the letters randomly around the keyboard, or indeed staying with the alphabetic ABCDE format that Sholes had started with.

What Sholes did to solve Rehr's problem was much simpler and more elegant and effective than trying to separate frequent letter pairs. With 26 letters and only 44 typebars (42 in the later Remington No. 7), some letters had to be placed next to each other on the typebasket. What Sholes did was to turn the problem on its head and did his best to ensure that these were <u>infrequent</u> letter pairs in the English language.

The cases where letters were adjacent to each other on the Remington No. 7 typebasket can be seen in Figure 1 as being contained within the string of letters QAZSXDCFVGBHNJMK. It is extremely difficult at first sight to think of many words that contain letter pairs associated with that sequence read in either direction (e.g. QA or AQ; FV or VF). Kay (2013a) showed with some electronic search experiments on texts that would have been popular in Sholes time that the 1873 version of QWERTY was near-optimal in separating letters on the typebasket that might be encountered together on the printed page. Mark Twain's 145,000 word "Life of the Mississippi" (or "LotM") was cited as indicative text (Manis, 1999) and in that case there were only 146 events where words contained letter pairs that were also adjacent on the typebasket, or about one such occurrence every 1,000 words. This contrasted with a similar experiment using LotM with the Dvorak format which resulted in 2358 such pairing events, about 16 times more frequent (Kay 2013a, p.1182). QWERTY was clearly far superior to Dvorak in terms of format/device compatibility expressed in terms of its ability to minimise or avoid jamming issues

Much of the basis for the supposed superiority of Dvorak over QWERTY cited by David (1985) and others was based in terms of format/user compatibility issues, notably advantages claimed in speed of touch typing for Dvorak over QWERTY. However, there are at least three difficulties with David's subsequent argument (1985, p.336) that QWERTY was the wrong system for the industry to standardise around . First, Liebowitz and Margolis (1990) cast doubt on the empirical evidence used to argue that Dvorak really was superior to QWERTY on format/user compatibility grounds. Second, modern ten-finger touch typing that Dvorak was designed for in 1936 was very much a later development, and indeed as late as 1887 (fourteen years after Remington sold Sholes' first typewriters) an editorial in a trade journal was still arguing that optimal typing speed was achieved by using only the first two fingers of each hand (Wershler-Henry, 2005, p.232). Third, even when touch typing began to be widely adopted towards the end of the 19^{th} century, there would have been little point in pursuing format/user compatibility (and typing speed) if the user repeatedly encountered format/device compatibility issues (and jamming problems).

Mahoney (2000) cites the evidence on format/user compatibility analysed by Liebowitz and Margolis (1990) as calling QWERTY's role as paradigm case for path dependence into question precisely because QWERTY may have been more efficient format compared to Dvorak all along (pp. 514-15). In fact, QWERTY would have been unequivocally demonstrably superior to Dvorak on the only efficiency grounds that mattered in the early years of the typewriter, format/device compatibility.

This tends to conflict with David's (1985) argument that QWERTY's genesis and eventual success was characterised by historical accidents, or random transient factors in the form of choices made close to the beginning of the process (p.335). But it is difficult to characterise QWERTY in such a fashion when it was clearly deliberately designed to be as near-optimal in terms of the efficiency criteria that really mattered at the time. The evidence tends to less consistent with the notion that small chance events can play major roles in the evolution and diffusion of technologies, and more consistent with those such as Garud, Kumaraswamy, and Karnøe (2010) and Vergne and Durand (2010) who place more emphasis on path creation and the role of deliberate and creative design. It is also tends to be consistent with Hossain and Morgan (2009) whose laboratory trials with QWERTY-type outcomes found that in repeated experiments the subjects never got stuck on the inefficient platform; and with Liebowitz and Margolis' (2013) broader assertion that it is difficult to find evidence that markets lock standards into inferior and inefficient solutions.

The effectiveness of the infrequency principle in dealing with jamming problems can be clearly demonstrated, though the circumstances that led to its development are less clear. In the next section we try to establish how much can be said with reasonable certainty about this critical period in the development of the technology.

(3) The genesis of the infrequency principle

The closest to an authoritative account of the development of the typewriter and the roles played by Sholes and those associated with him was set out by the academic historian Richard N. Current (1954). Current[2] faced problems in that his documented sources, including letters, memoirs and periodicals were often partial or incomplete, while the claims made by the main protagonists often contradicted each other, and sometimes even themselves. Current resolved these difficulties as far as he could with original sources and validating claims and assertions by cross-checking from alternative sources.

Current's account indicates that by early 1871 Sholes' successive models were still using an alphabetical layout for the keyboard (1954, p. 44), but by late 1872 this had been radically transformed by Sholes and Densmore into the first version of QWERTY:

> "… some of the letters were so close together that the type bars, hung as they were in a circular 'basket' collided more frequently than need be. As experienced printers, Sholes and Densmore were familiar with the type case, in which the pieces were assorted according to convenience and not according to the alphabet. The two men, working together, finally arranged the typewriter keyboard in the spirit of the printer's case, though they did not duplicate its particular arrangement. They agreed upon a pattern – q w e r t y u i o p, etc. – which was to remain almost exactly the same from that day to this" (Current, 1954, p. 55)

Perhaps the first recorded examples of potential network externalities and quasi-reversibility in the context of QWERTY was given by Densmore, who noted in November 1872 that "it is better to have (the keyboards) all alike"; that the new arrangement meant that he had to "unlearn as well as learn"; but also that "the change was better to be made than not" (Current, 1954, p.59, italics in original).

So while Sholes was primary responsible for many of the other developments that were to be fed into the typewriter, the answer to the question as to who developed QWERTY is that it was really jointly developed and agreed by Sholes and Densmore. As to the question who would have known about the infrequency principle at this point, irrespective of who actually thought of it, it is difficult to see how and why either of them would have kept the justification for such a radical redesign secret from the other. Even though QWERTY is about the only major element of these endeavours to this survive to this day, at the time it would have been just one more technical fix out of numerous others that Sholes and Densmore were grappling with to make their machine work. Also, Sholes the main inventor and Densmore the promoter and financier both needed each other and would have had incentives to share the secret and motivate the other to keep committed to an enterprise which at more than one point looked fairly hopeless. We shall see later that Sholes was aware of the principle and indeed continued to apply it, but for the above reasons it is probable that Densmore was also aware of it.

When Densmore brokered the deal to produce Sholes' typewriter with Remington in 1873, Remington made some fine tuning to QWERTY which David cites as including placing the R in the top letter line, thus “were assembled into one row all the letters which a salesman would need to impress customers by rapidly spelling out the brand name TYPE WRITER” (1985, p.333). It is quite possible that this was on advice from Densmore since Current (1954) provides strong evidence that it had been Densmore and not Sholes who coined the name “typewriter”, as a single word and unhyphenated.[3]

The application of the rule to QWERTY from the 1872 version onwards meant that QWERTY proved highly efficient in absolute terms in that it all but eliminated the typebar jamming problem arising from neigbouring typebars on the typebasket being typed in succession (Kay, 2013a, p.1180). It also led to QWERTY being highly efficient in relative terms in that it demonstrated significantly superior performance compared to alternative formats such as the alphabetic (ABCDE) format that Sholes started with, or even the Dvorak format had that been available in 1872 (Kay 2013a, pp. 1181-82). The rule also meant QWERTY achieved levels of performance in those terms that simply would not have been feasible had the popularly-believed method of using a list of frequent letter pairs been used instead (Kay, 2013a pp.1180-81). It also explains the placing of the vowels, including why “E” and “I” are the only two letters missing from the alphabetic string DFGHJKL[4] in the third row on the keyboard and why “Q” and “A” are next to each other on the typebasket (Kay, 2013a p. 1181). All this is consistent with deliberate development and application of the infrequency principle, with Sholes (and possibly also Densmore) being aware of the principle and applying it.

Sholes at least was aware of the infrequency principle in 1872 and had applied it, but was it subsequently known to those in Remington who took the 1872 version and made modifications to it? The evidence is that it was. First, all the changes made by Remington from 1872 QWERTY to “modern” QWERTY in 1878 are explicable in terms of the consistent and systematic application of the rule (Kay, 2013a, pp. 1181-82). In particular, the continuous letter sequences on the typebasket were highly sensitive to perturbations, and experiments confirmed that rearranging letters associated with these sequences typically ran the danger of creating high-frequency couplings unless the infrequency principle was followed (Kay, 2013a, p.1181). Second, experiments showed that not just QWERTY but also later modifications (specifically: AZERTY, French; QWERTZ, German; and AZERTY, Italian) were all individually crafted with format/device compatibility in mind, and with each exhibiting a high degree of sensitivity to linguistic idiosyncrasies in these regards in their respective domains (Kay, 2013d).

So the infrequency principle was known to Sholes and possibly also Densmore, while there was also awareness and use of the principle in Remington itself, as evidenced by later modifications. This was not surprising; when the initial deal was made with Remington the company’s representatives expressed only mild interest, Sholes and Densmore were running out of funds and Densmore in particular was eager to get their typewriter manufactured “at almost any cost” (Current, 1954, p. 65). On Densmore’s side, this incentive was heightened by his retaining a financial interest in QWERTY’s success in a contract with Remington, Sholes was to sell his rights for a fixed sum in 1873 (Wershler-Henry, 2005, p. 70). Concealing what would have been one of the strongest selling points for the new technology could have imperilled the whole deal. In turn, Remington would have an incentive to keep the principle secret once it was shared with them.

But just keeping the infrequency principle a trade secret would not have been sufficient protection for the widening number of interested parties who wished to exploit it commercially. In the next section we look at how different forms of IPR protection can act in complementary fashion before turning to the issue of how this is of relevance in the case of QWERTY and Sholes.

(4) Complementarities in IPR protection

IPR can involve a variety of tools including patents, trademarks, copyright and trade secrets. In the present context it is the distinction between patents and trade secrets that matters. The most obvious difference between patents and trade secrets lies in disclosure; "a trade secret is some sort of information that has value because it is not generally known" (Risch, 2007, p.6). While a patent may be disclosed publicly and granted on a new and useful design or process for a limited period of time, a trade secret is information which may give a business an advantage over competitors who do not know or use it (Besen, and Raskind, 1991).

The potential importance of complementarity in terms of IPR protection has not been really fully recognised until recently, indeed it has been more usual to portray forms of IPR protection (including patents and trade secrets) as simply substitutes for each other (e.g. Friedman, et.al. 1991; David 1993, p.31; Fisher and Oberholzer-Gee, 2013, p.160). Where complementarities at the level of IPR protection have been recognised it has tended to be along the lines of several patented inventions combining as economic complements to produce one product (e.g. Cohen et al, 2000, p.22).

However, it is now clear that potential complementarities in IPR protection can extend across forms, can include trade secrets, and indeed have been actively sought by firms in different industries from the early years of industrialisation to the present day. Arora (1997) cites several examples from the late 19th and early 20th century of chemical firms combining patents with secrecy, for example by patenting several compounds but retaining secrecy over the composition of specific dyestuffs. In other cases, decoy patents could act as chaff and help camouflage which patents were really intended to add commercial value. More recent times have seen examples of very different products and processes which have patented some elements of the underlying technology while keeping other aspects secret, such as Pilkington's float glass technology (Al-Aali and Teece, 2013, p.26), GEs process for making industrial diamonds, a hormone therapy drug produced by Wyeth, and C&F's technology for freezing ingredients for pizza toppings (Jorda, 2013, pp.28-30). Mazzone and Moore (2008, pp.57-9) note several ways in which in which patent holders may use undisclosed information strategically and to their advantage, while Sherwood (2008) notes that the contribution of many research joint ventures can be in the form of a combination of patents and trade secrets

The exploitation of complementarities can extend to other forms of IPR protection, for example Henkel et al (2013, p.67) note the case of a flash drive where the binary code was copyrighted while the source code was kept secret. Even more complex combinations are possible, for example different aspects of a single data processing system or a biotechnological diagnostic kit can be protected by a phalanx of complementary patents, copyright registration, trademarks and trade secrets (Jorda, 2008, p.13). At strategic level, Ottoz and Cugno (2008) explore the optimal patent/secret mix for complex products, while at policy level Ottoz and Cugno (2011) develop a model of the ideal scope of trade secret law where a technology comprises two complementary components, one of which is patented and one is secret.

The traditional notion that patents and trade secrets may be regarded as substitute forms of IPR protection has also been expressed in the argument that the choice between trade secret and patent is a matter of comparing relative costs and benefits; in this perspective, trade secrets are said to be preferred if the patent option would be more costly or less rewarding than the trade secret (Friedman et al,1991, p.64). However, in the case of the development of QWERTY there really was no choice over the forms and combinations of IPR protection that had to be adopted. The intellectual property of concern to Sholes and his collaborators related to the machinery; the QWERTY format; and the infrequency principle. The first two forms of intellectual property would be made public as soon as the devices that embodied them came on the market so trade secret was simply not an option here; Sholes patent for his "improvement in type-writing machines" that included the QWERTY format was eventually published as US patent 207559 in 1878. Similarly, patenting was not an option for the infrequency principle, the US Supreme Court had already ruled that an abstract idea such as the discovery of a new principle was not patentable but had to be embodied and brought into operation by machinery to produce a new and useful result (US Supreme Court. 1852)[5].

The complementarity of patent and trade secret protection was not a choice for Sholes and his partners, it was a necessity if they were to defend their intellectual property successfully. Take away the patent aspect and imitators could have simply cloned his typewriter once it began to show commercial promise. Take away the trade secret aspect and rivals could have developed alternative formats using the infrequency principle.

In the next section we explore how Sholes' search for improvements to QWERTY led him into even more sophisticated combinations of IPR protection.

(5) Sholes and the search for the "perfect" typewriting machine

Kay (2013a) notes that given the tools and the information available to Sholes in his day, any further improvements in respect of format/device compatibility would probably have been marginal and difficult to achieve" (p,1181). But as we shall see, that did not prevent Sholes from trying. A clue as to whether Sholes could have improved the performance of QWERTY in format/device compatibility terms can be found in the six asterisks clustered together on the right of the Remington No. 7 typebasket in Figure 1. In reality, the asterisks in question were punctuation marks or fractions. However, in addition to their assigned role in representing a specific character, asterisks on the typebasket could also play a secondary role as buffers separating letters from each other. This buffering role is shown to great effect along the top of the typebasket where all the vowels were buffered from immediate proximity with any other letters (except for A which was largely neutralized by being buffered between Q and Z). Given the active role that vowels can play in forming frequent letter pairs, buffering in this fashion ensured that these vowels were not adjacent to any other letter on the typebasket. But if we were to assess QWERTY in the Remington No. 7 as to how effectively that buffering function was pursued, it is clear that the cluster of six asterisks represent potentially wasted opportunities seen in those terms, with some asterisks just buffering other asterisks instead of letters.

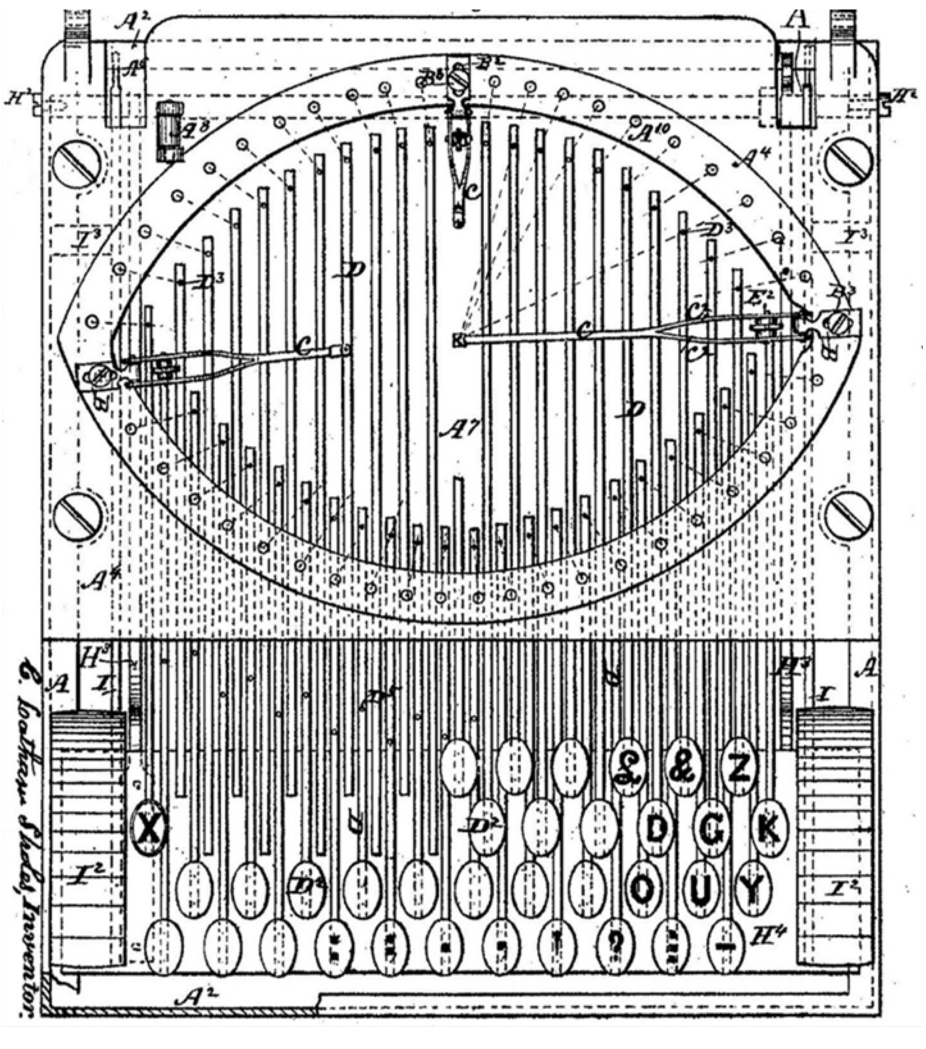

Figure 2: Top elevation from Sholes patent US 558428

The first Sholes patent we shall look at in this connection is US patent 558428, filed December 31, 1881, but only issued (posthumously) to Sholes April 14, 1896. Figure 2 shows an image from this patent with a top elevation view for a 45 character keyboard and elliptical typebasket. The keyboard is arranged in four rows of 11, 12 11 and 11 characters respectively (where the keys are not indicated their existence and position can be inferred from the levers connecting to the typebasket). The levers follow the Sholes protocol for QWERTY by connecting the characters from the top two rows of the keyboard to the top half of the typebasket in alternating fashion; so for example £&Z from the end of the first row and DGK from the end of the second row alternate with each other on the typebasket to form £D&GZK there. Similarly, the levers connecting the lower two rows of the keyboard to the typebasket also do so in alternating fashion on the lower half of the typebasket in Figure 2.

Although we can see from the diagram how the keyboard was intended to connect to the typebasket, the patent only gives a very partial indication of what the character arrangement would be. It specified only 18 out of 45 characters, such as X and K anchoring the ends of the second row.

But Sholes later filed another patent (US 568630) on September 11, 1889, this was later issued (again posthumously) September 29, 1896. This patent also describes a 45 character keyboard and elliptical typebasket with the keyboard again arranged in four rows of 11, 12 11 and 11 characters respectively. The top view elevation for this patent is shown in Figure 3.

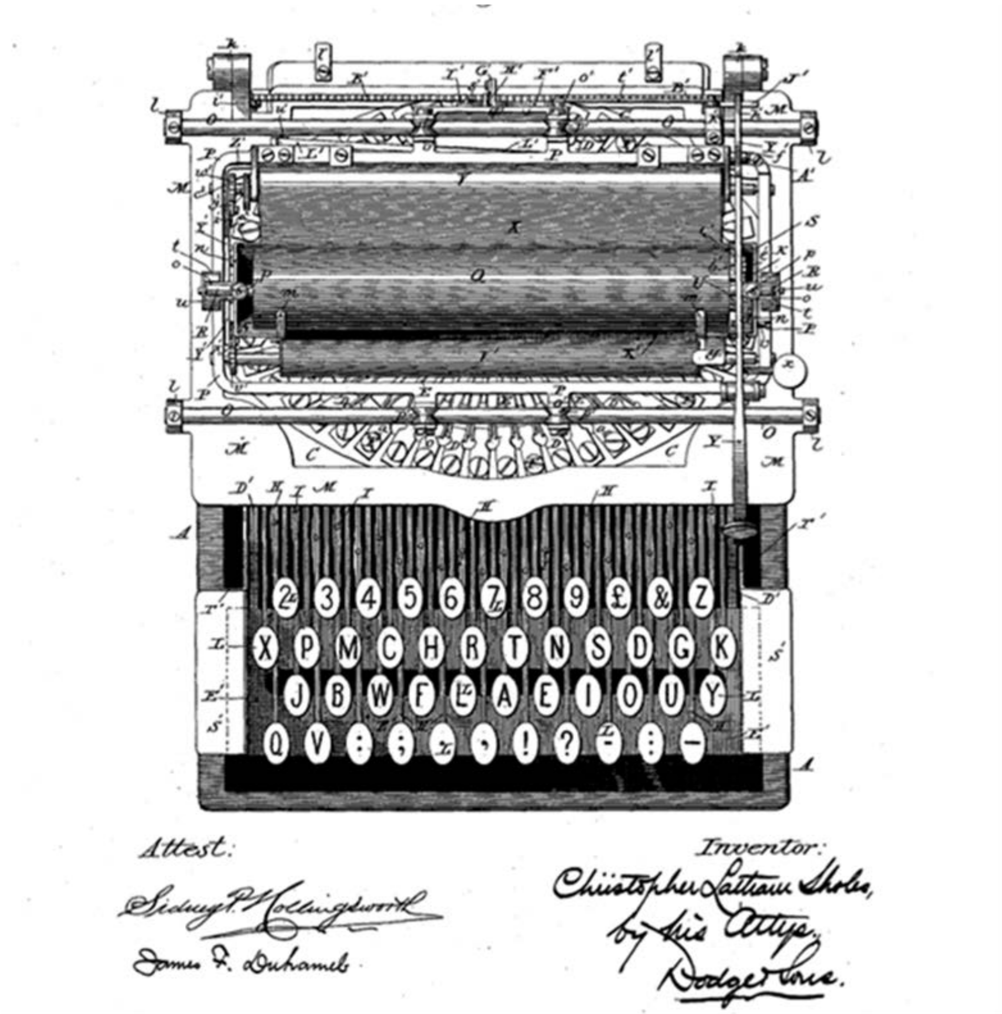

Figure 3: Diagram from Sholes patent US 568630

Not only do the technical specifications for the two patents complement each other, so do the keyboards in the respective cases with 11 keys in each row except for the second row which has 12. The full keyboard here exactly matches the ten characters in the top three rows that were actually specified in the earlier patent (£, &, Z, X, D, G, K, O, U and Y), and also the two end punctuation marks on the bottom row. There is some minor churning of other (non-letter) characters in the bottom row compared to those that were specified in the earlier patent, but in both cases they are mutually consistent insofar far as they both have non-letters on these parts of the bottom row where QWERTY had letters. The swapping of places by some punctuation marks between the two patents would have had absolutely no impact as far as the application of the infrequency principle was concerned.

(6) Ingredients and recipe for a near-perfect type-writing machine?

We can explore what Sholes typewriter would have looked like if he had built one with the hardware platform in Figure 2 hosting the keyboard arrangement in Figure 3. We show this in Figure 4. Again for simplicity we have replaced all non-letters in the figure with asterisks.

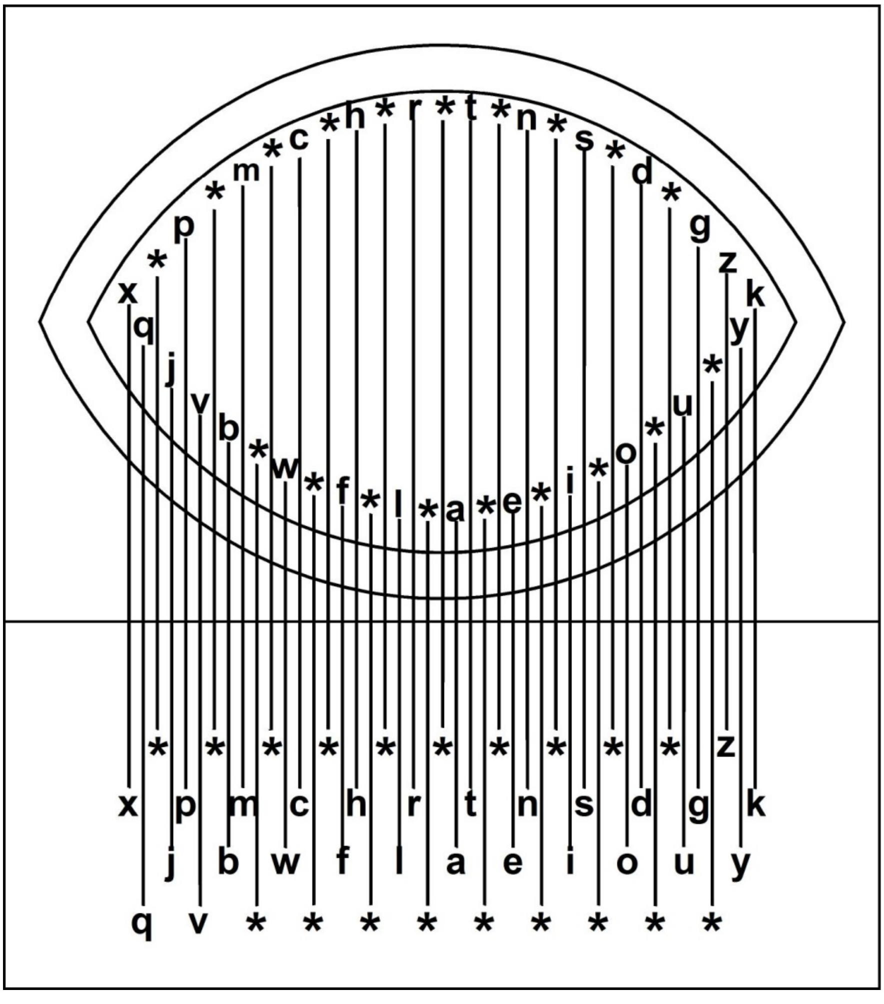

Figure 4: Patent 558428 hardware combined with patent 568630 format

The first point to note is that by populating most of the bottom row of the keyboard with non-letters (mostly punctuation), one result is that every non-letter is fully employed as a buffer preventing letters from being adjacent to other letters on the typebasket, including the five vowels A, E, I, O and U. That still means that some letters have to be next to each other on the typebasket in two sequences XQJVB and GZKY; however, it does mean that the actual numbers of letter pairs that can be formed out of these sequences on the typebasket is less than in any of the QWERTY cases developed up to that point by first Sholes and later Remington.

The next point to consider is whether what we shall describe as the QV format[6] (from the first two letters on the bottom row) would have been potentially more efficient than QWERTY in terms of format/device compatibility and reduced chances of jamming from typing letter pairs that were also adjacent on the typebasket. We compared how QWERTY on the Remington No. 7 would have performed against QV on Sholes elliptical typebasket in these terms using a number of test texts. We show the results for LotM as indicative text in Table 1. Table 1 also shows the results for the alphabetic or ABCDE format that Sholes originally started with (the alphabetic or ABCDE format would have replaced QWERTY's typebasket continuous letter string QAZSXDCFVGBHNJMK with AKTLUMVNWOXPYQZR)

ABCDE format		**QWERTY format**		**QV format**	
Pairing	**Frequency**	**Pairing**	**Frequency**	**Pairing**	**Frequency**
AK/KA	740	QA/AQ	2	XQ/QX	3
KT/TK	4	AZ/ZA	73	QJ/JQ	0
TL/LT	1166	ZS/SZ	0	JV/VJ	0
LU/UL	2214	SX/XS	1	VB/BV	2
UM/MU	953	XD/DX	0	GZ/ZG	0
MV/VM	0	DC/CD	9	ZK/KZ	0
VN/NV	141	CF/FC	0	KY/YK	54
NW/WN	788	FV/VF	0	**Total**	59
WO/OW	3919	VG/GV	0		
OX/XO	34	GB/BG	2	**Modified QV format**	
XP/PX	196	BH/HB	22	**Pairing**	**Frequency**
PY/YP	59	HN/NH	66	YQ/QY	0
YQ/QY	0	NJ/JN	45	QJ/JQ	0
QZ/ZQ	0	JM/MJ	2	JV/VJ	0
ZR/RZ	0	MK/KM	14	VF/FV	0
Total	10,214	**Total**	238	MZ/ZM	0
				ZK/KZ	0
				KX/XK	0
				Total	0

Table 1: Pairing frequencies for ABCDE, QWERTY, QV and modified QV on LotM

We can use the ABCDE format as benchmark for the performance of the QWERTY and QV formats in terms of format/device compatibility. In the case of LotM, the ABCDE format had 10,214 occurrences of letter pairs on the printed page that were also next to each other on the typebasket, or about once every 14 words. By contrast, QWERTY would have encountered such events about once every 609 words while QV would have met the same fate about once every 2458 words. These results with LotM, were found to be broadly in line with results from other 19th century American English test cases.

While both QWERTY and QV were clearly superior in terms of format/device compatibility with the original ABCDE format, the QV format was clearly superior to QWERTY with implied efficiency benefits. As such, it is also consistent with Sholes applying his infrequency principle in developing his QV format.

But was this the best Sholes could have achieved in those terms or could he have done better still? We can take his solution a stage further towards his goal with the help of modern tools. Suppose, for example, Sholes had swapped Y and X, F and B, M and G, that would have made the two continuous letter sequences YQJVF and MZKX on the typebasket. Running the pairs associated with those sequences through LotM resulted in zero events ("ex-keelboatmen" which came up as an XK pairing would in practice have had a hyphen interposing between X and K). The results are shown in Table 1 for this "modified QV" format where it achieved a perfect score with no incidences of letter pairs in LotM that were also adjacent to each on the typebasket.

The online PDF experiments were run with a series of other popular American novels roughly contemporaneous with Sholes.[7] Again, no incidences of letter pairs in text that were adjacent to each other on the typebasket were found in a total (including LotM) of about 1,110,000 words of text. While these results might suggest that modified QV might be considered at least a tentative candidate as optimal configuration in those terms, the same experiments suggested that there might not be a uniquely optimal solution. For example, replacing G with S rather than M led to the same perfect results with no letter pair events in any of the texts tested. In other cases with alternative substitutions, it was only obscure or foreign words which prevented perfect scores. However, this might be thought sufficient to demonstrate that Sholes' infrequency principle could have been used to design an optimal format in terms of this crucial design criterion..

But it is important to note that we were to achieve our "modified QV" results with the help of tools which were not available to Sholes in his day. First, we were able to access online PDFs for whole texts like LotM and use the Reader search tool to check the incidence of specific letter pairs in the text (making sure that the "whole words only" box was not ticked).

Second, we were aided by the popularity of the 20th century word game Scrabble. In a sense, what Sholes was doing could be described as playing anti-Scrabble. Instead of searching for words that contained a specific letter pair as in Scrabble, he was searching for letter pairs that were not contained in any words. We were able to access online search tools designed to help Scrabble players, with one website[8] being particularly useful in helping play "anti-Scrabble". We were able to use this tool to confirm the general absence of recognised words containing new letter pairs that would result from these swaps before we even trialled the effect on actual texts.[9]

Finally, different PDF searches with successive iterations could quickly experiment with alternative possibilities, simply eliminating those that resulted in any letter pair events. Sholes had none of these advantages in the 19th century. Sholes had done the hard creative work of developing the infrequency principle and the QV format, and it would have been difficult for him to achieve superior solutions with the limited tools available to him in his day. However, in a counterfactual world where he did have access to these tools he could easily have developed the "perfect" type-writing machine in terms of format/device compatibility and Rehr's problems of separating letter pairs on the printed page by at least one typebar on the typebasket .

But would it have been worth the effort? The QV format only had about a quarter of the incidences of pairing frequencies on LotM as did QWERTY, while the modified QV format had none. But another way of looking at the same results is that while QWERTY reduced the pairing frequencies by 97.7% on LotM compared to ABCDE, the QV format could have reduced the frequency of these pairings by a further 1.8% and modified QV by 2.3%. So while QV would have reflected quite an advance in these terms if just compared to QWERTY, the incremental gains in efficiency compared to what had already been achieved by QWERTY are actually quite modest.[10]

(6) Discussion

Sholes subsequent attempts to patent the "perfect" typewriting machine (at least in terms of format/device compatibility) helps explain more fully Sholes behavior with respect to QWERTY. Kay 2013 (a) had asked:

> " …if Sholes did deliberately apply a rule based on the conjunction of infrequency and contiguity, why did he apparently take the secret to the grave with him (he died in 1890)? It is obvious that the rule would have been concealed in 1873 and soon after to protect intellectual property, but why not eventually confide in friends or family and for the sake of posterity and his reputation?
>
> In fact, Sholes had sold his rights to QWERTY for a fixed sum in 1873, but his two partners, Densmore and Yost, retained a financial interest in QWERTY's success in a contract with Remington (Wershler-Henry, 2005, p. 70). Wershler-Henry describes Sholes as worn down by the various trials and tribulations associated with the development of QWERTY. By the late 1880s Sholes had totally disowned his own invention, refusing to own one, use one, or even recommend it (Wershler-Henry, 2005, p. 67). If you do not value your own creation, or what others had done with it, then you would not expect other people to, either" (Kay, 2013a, p.1181).

All that was strictly true, but these later patents make it clear that there was a deeper reason why Sholes had stayed silent apart from disparaging the outcome of his earlier efforts. It was not QWERTY the patent which Sholes valued as his own creation, but the trade secret that was the infrequency principle and of which QWERTY was only one of many possible embodiments. Sholes was still aiming to create a superior alternative to QWERTY and he now had every reason to regard QWERTY as a rival solution. And as long as he harboured the ambition to develop a better alternative to QWERTY either for himself or to help secure his family's financial future, it also helps explain why he took the secret of the infrequency principle to the grave with him.

Delays in filing patents here could be explained in part in that it could delay their subsequent expiration (Current, 1954, p. 134). But that still leaves the question of why the two crucial constituent elements were contained in two separate patents filed years apart, the first largely hardware-based with the Sholes protocol and an elliptical typebasket, and the second largely based around the QV format. If the infrequency principle was still not publicly obvious even after there had been opportunities to reverse engineer the Remington platforms that carried QWERTY, where were the risks in a single patent for the QV format and the hardware that would carry it, just as Sholes had done with his 1878 QWERTY patent?

The obvious risks here lay in that Sholes knew that the trade secret was now in the hands of what was now his rival, Remington, and its successors.[11] Remington had begun selling Sholes' typewriter in 1873 and by November 1875 had acquired the exclusive rights to make and sell typewriters using the Sholes patents (Current, 1954, p. 82). A single patent combining hardware and format would have made his intentions clear to those who were in possession of the trade secret. Further, even though Densmore had a financial interest in the success of Remington typewriters, he continued to cooperate with Sholes and urge him to invent an improved typewriter. Remington had been threatening Densmore with law suits for alleged breaches as early as 1880 because of Densmore efforts to develop alternative typewriters (Current, 1954, p.100), So it would have been a rational strategy for Sholes to fragment the QV solution into different patents filed well apart and with a curious elliptical typebasket (which, irrespective of any contribution it might make in terms of functionality, at least differentiated it from the original QWERTY typebasket).

This strategy would have been aided by the fact that the two patents in question here were only two of several Sholes patents filed after 1873, potentially making it more difficult to see the connections between the two that really matter here. Two other patents[12] were filed by Sholes in between the two patents we are discussing here, while five other were filed posthumously (four of them the day after his death in February 17, 1890).[13] None of these other patterns contained the crucial keyboard references we are concerned with here.

The combination of patent with secrecy in this context was first reflected in the 1878 QWERTY patent which is itself an excellent example of how patents and trade secrets can be complementary as well as alternative methods for the protection of intellectual property. It appears that Sholes took this process even further with fragmented patents and a trade secret in order to conceal his intentions from his erstwhile partner Remington. But as noted earlier, it has become clear in recent years that Sholes was not alone in framing complex IPR protection strategies involving complementary forms of IPR protection. Arora (1997) gives a number of examples of such complementarities, but there is one practice roughly contemporaneous with Sholes from the 19th century chemical industry that is particularly relevant in the present context.

> "German companies skilfully combined patents and secrecy to keep potential imitators at bay. The dyestuffs were typically composed of a number of different

> compounds. In some cases, the precise composition of the dyestuff was kept secret, but the individual compounds protected by patents" (Arora, 1997, p.393)

Sholes similarly combined trade secret with the rest of the QV solution split into two out of his several post-1873 patents, and the two patents in question filed years apart. The widely reported speculation that he used a list of frequent letter pairs to separate these pairs on the typebasket may have been deliberately started by Sholes or his partners to put potential rivals off the scent. Even if they did not start the speculation, it would have been rational not to deny it.

But what would have been Remington's reaction had Sholes attempts to use the infrequency principle to improve on QWERTY become apparent to them? Remington could not challenge for patent infringement because the principle was not patented. A reasonable fear that Sholes might have had is that by confirming that QWERTY could indeed be improved on, Remington would direct their attention towards the same end. Even more likely was that Sholes simply wanted to escape the attention of Remington and not have such a well-resourced rival focus its ire on him.

What is ironic is that Sholes efforts were doomed to failure. By the late 1880s QWERTY was rapidly moving towards lock-in as industry standard, effectively pre-empting attempts from even its own creator to dislodge it from that position, especially since any efficiency gains in terms of format/device compatibility from the QV format would have been marginal at best. In turn, soon after Sholes' later patents were issued in 1896, Underwood was to launch a revolution in typewriter design with their front-strike or front-stroke designs replacing the Remington up-strike or up-stroke design as dominant hardware standard within a few years (while still retaining QWERTY as universal keyboard standard). QV's format/device compatibility advantages were specific to up-strike design; since the front-strike typebaskets were radically different from those in up-strike even QWERTY encountered severe format/device compatibility problems in being in adopted on a front-strike platforms (Kay, 2013c), and much the same could have been expected from the QV format.

(7) Conclusions

Sholes almost achieved his aim of creating what would have been at the time the "perfect" type-writing machine in terms of the performance parameter that mattered at the time, and we have shown with the help of modern tools how he could have achieved that aim (though it is likely there would have been no single unique optimal solution). But, ironically, QWERTY had been so well designed in the first place that any efficiency gains from QV or "modified QV" would have been marginal at best and swept aside by QWERTY which was rapidly becoming locked in as industry standard. What this helps emphasize is that QWERTY was only one of many possible embodiments of the infrequency principle, and it was that underlying principle embodied in QWERTY and ingeniously integrated with the hardware in Sholes' devices that really helped win the battle of the standards.

However none of this would have been possible had Sholes not used a judicious blend of complementary trade secret and patent protection from the beginning. Most trade secrets tend to be “wasting assets” dissipating after an average of 3-5 years because of employee mobility and reverse engineering, though with some exceptions such as Coca Cola’s recipe which has been secret for over 100 years (Jorda, 2008). Sholes infrequency principle had a useful economic life of about a quarter of a century before the development of upstrike typewriter technology made it obsolete, but it still stayed secret for 140 years despite the considerable interest in the provenance and genesis of QWERTY. In view of the central role that these issues played in the evolution of this standard, it also raises interesting questions for future research as to whether such complementarities may have played key but unappreciated roles in the development of other standards.

By showing how Sholes used his infrequency principle in further efforts to improve on the QWERTY format in efficiency terms, this paper adds further weight of evidence that this was indeed the principle that he and his partner Densmore used to design QWERTY in the first place. As such, it helps to strengthen the conclusions of Kay (2013a) that QWERTY was a consequence of creative design rather than an accident of history, indeed QWERTY was as near-optimal in terms of the crucial performance criterion of format/device compatibility as could be reasonably expected with the state of technical knowledge that existed in Sholes’ time. This in turn helps reinforce arguments in Kay (2013a) that QWERTY’s role as “paradigm case” of inferior standard in the path dependence literature is not consistent with the evidence, and that instead the case is more consistent with path creation than with path dependence.

References

Al-Aali, A. Y. and D.J. Teece, 2013. Towards the (strategic) management of intellectual property; retrospective and prospective. California Management Review 55 (summer), 15-30.

Arthur, W.B., 1983. On Competing Technologies and Historical Small Events: The Dynamics of Choice Under Increasing Returns. Technological Innovation Program Workshop Paper, Department of Economics, Stanford University.

Arthur, W. B., 1989. Competing technologies, increasing returns, and lock-in by historical events. Economic Journal 99, 116-31.

Arthur, W. B., 2013. Comment on Neil Kay's paper: Rerun the Tape of History and QWERTY Always Wins. Research Policy 42, 1186-87.

Arora, A., 1997. Patents, licensing, and market structure in the chemical industry. Research Policy 26, 391-403.

Besen, S. M. and L.J. Raskind, 1991. An introduction to the Law and Economics of Intellectual Property. Journal of Economic Perspectives 5, 3-27.

Cohen, W. M., R. R. Nelson and J. P. Walsh, 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not), NBER Working Paper No. 7552, Cambridge, Mass.

Current, R. N., 1954. The Typewriter and the Men who made it. Urbana, University of Illinois Press

David, P.A., 1985. Clio and the economics of QWERTY. American Economic Review 75, 332–337.

David P. A., 1993. Intellectual Property Institutions and the Panda's Thumb: Patents, Copyrights, and Trade Secrets in Economic Theory and History, in Wallerstein, M. B. Mogee, M. E. and R. A. Schoen (eds) Global Dimensions of Intellectual Property Rights in Science and Technology National Academy Press, Washigton, D. C. 19-62.

Durham, A. L., 2011. The paradox of "abstract ideas". Utah Law Review no 3, 787-853.

Fisher III, W. W., and F. Oberholzer-Gee, 2013. Strategic management of intellectual property an integrated approach California Management Review. 55 (summer) pp. 157-183.

Friedman, D. D., W. M. Landes and R. A. Posner, 1991. Some Economics of Trade Secret Law. Journal of Economic Perspectives 5 (Winter), 61-72.

Garud, R., Kumaraswamy A. and P. Karnøe, 2010. Path Dependence or Path Creation? Journal of Management Studies 47, 760-774.

Henkel, J., Baldwin, C. and W. Shih, 2013. IP modularity: profiting from innovation by aligning product architecture with intellectual property. California Management Review 55 (summer), 65-82.

Hossain, T., and J. Morgan, 2009. The Quest for QWERTY. American Economic Review Papers and Proceedings 99, 435-40.

Jorda, K. F., 2008. Patent and Trade Secret Complementariness: An Unsuspected Synergy. Washburn Law Journal 48, 1-32.

Joyce, B. and R. A. Moxley, 1988. August Dvorak (1894-1975): early expressions of applied behavior analysis and precision teaching. The Behavior Analyst 11, 33-40.

Kay, N. M., 2013a. Rerun the tape of history and QWERTY always wins. Research Policy 42.1175-85.

Kay, N. M., 2013b. Rerun the tape of history and QWERTY always wins: response to Arthur, Margolis, and Vergne. Research Policy 42, 1195-96.

Kay, N. M., 2013c. The QWERTY problem, Paper presented at the 35th DRUID Celebration Conference 2013, Barcelona, Spain, June 17-19: http://druid8.sit.aau.dk/acc_papers/rluar3kk4i5t4gelr7j36gx9hma5.pdf (accessed 7th May 2013)

Kay, N. M., 2013d. Lock in, path dependence, and the internationalization of QWERTY, Strathclyde Discussion Papers in Economics, no 13-10, Department of Economics, University of Strathclyde, Glasgow. http://www.strath.ac.uk/media/departments/economics/researchdiscussionpapers/2013/13-10FINAL.pdf (accessed 30th August 2013).

Liebowitz, S. J. and S. E. Margolis, 1990. The fable of the keys. Journal of Law and Economics 33, 1-25.

Liebowitz S. J. and S. E. Margolis, 2013. The troubled path of the lock-in movement. Journal of Competition Law and Economics 9, 125-52.

Lewin, P., 2001. The market process and the economics of QWERTY: two views.The Review of Austrian Economics 14, 65–96.

Mahoney, J., 2000. Path Dependence in Historical Sociology. Theory and Society 29, 507-548.

Manis, J. (Ed), 1999. Life on the Mississippi by Mark Twain (Samuel L. Clemens) Pennsylvania State University, Electronic Portable Document File format http://www2.hn.psu.edu/faculty/jmanis/twain/lifeonmr.pdf (accessed June 20th, 2011).

Mazzone, J. and M. Moore, 2008. The Secret Life of Patents. Washburn Law Journal 48, 33-66.

Margolis, S.E., 2013. A tip of the hat to Kay and QWERTY. Research Policy 42, 1188-90.

Ottoz, E. and F. Cugno, 2008. Patent–Secret Mix in Complex Product Firms, American Law and Economics Review 10, 142-158.

Ottoz, E. and F. Cugno, 2011. Choosing the scope of trade secret law when secrets complement patents. International Review of Law and Economics 31, 219– 227.

Rehr, D., 1997. QWERTY revisited. ETCetera 38, pp.4-5.

Risch, M., 2007. Why do we have trade secrets? Marquette intellectual Property Law Review 11, 1-76.

Sherwood, R. M., 2008. Trade Secret Protection: Help for a Treacherous Journey. Washburn Law Journal 48, 67-106.

US Supreme Court, 1852. Le Roy v Tatham *55*, 156.

Vergne, J-P., 2013. QWERTY is dead, long live path dependence. Research Policy 42, 1191-94.

Vergne, J.-F. and R. Durand, 2010. The missing link between the theory and empirics of path dependence: conceptual clarification, testability issue, and methodological implications. Journal of Management Studies 47, 736-59.

Weber, B., 2012. Richard N. Current, Civil War Historian, Dies at 100, New York Times, November 4th http://www.nytimes.com/2012/11/04/us/richard-n-current-civil-war-historian-dies-at-100.html?_r=0 (accessed 9th September, 2012)

Wershler-Henry, D., 2005. The Iron Whim: a Fragmented History of Typewriting. Cornell University Press, Ithaca.

Endnotes

[1] I am grateful to Martin Kenney for drawing my attention to the diagram for Sholes U.S. Patent No. 568,630 which was the first clue that Sholes had indeed tried to improve on QWERTY.

[2] Richard N. Current's appointments included professorships of history at the University of Illinois and North Carolina (Greensboro), he was best known for his biographical works on Lincoln. On his death at the age of 100, the New York Times carried an obituary (Weber, 2012) in which Mark E. Neely, a Pulitzer Prize-winning Penn State professor described Current as a "demythologizer". The description is ironic given the various myths surrounding the development of the typewriter that have gained currency since Current (1954), many of which would not have gained traction had Current's work been given the attention it deserved.

[3] Densmore names it as such in 1871 in correspondence (Current, 1954, p. 48) and indeed he disliked the subsequent labelling of it in 1874 as the "Type-Writer" (Current, 1954, p. 69 and footnote), while by the late-1880s, Sholes was still describing it in correspondence as the "type-writer" (Current, 1954, p.123),

[4] In the first versions of QWERTY, M was also included at the end of this alphabetic string

[5] The notion that abstract ideas or principles are not patentable is well established and generally accepted, though what constitutes a non-patentable principle can be less clear (Durham, 2011). In the case of QWERTY it is not difficult to see problems that could have arisen from patenting such an idea, had that been permitted. If it had been patented, this could have had a stultifying effect on technical progress. For example, would such a patent force all third parties to deliberately place more frequent letter pairs next to each other on continuous sequences on their typebaskets to avoid breaches of such a patent? And how would infrequent letter pairs be identified and categorised? For such reasons patentability has generally been restricted to specific means for exploiting a principle rather than being seen as applicable to the principle itself.

[6] If we were being consistent with the naming of QWERTY we would have labelled this the XPMCHR format for the first six letters in the second row but QV seems both more appropriate and easier to say.

[7] The PDFs in each case were quickly accessible by Googling "PDF" after the tile. The texts were: *Little Women* (Louisa May Alcott, 1868); *The Last of the Mohicans* (James Fenimore Cooper, 1826); *The American* (Henry James, 1877); *Moby Dick* (Herman Melville, 1861); *Uncle Tom's Cabin* (Harriet Beecher Stowe, 1852) and *Walden* (Henry David Thoreau, 1864). The results were consistent with those for LotM with modified QV beating QV which in turn beat QWERTY in terms of letter pair events in each case. Modified QV produced no letter pair events in any of these cases.

[8] The website was http://www.scrabblefinder.com (accessed 18th May 2013)

[9] To play "anti-Scrabble" with http://www.scrabblefinder.com , first select the letters pair to be searched for, say KX. Then search viz: http://www.scrabblefinder.com/contains/kx/. The page should indicate that there are no words found that contain that letter pair. The page also gives options to search for words that begin with KX and also words that end with KX. No known words with that letter pair were identified in any of these options (searches conducted 18th May 2013)

[10] In principle, Sholes could have been able to at least match these results with a 52-character keyboard, middle two rows letters, the other two rows non-letters, and all letters flanked by an asterisk. But this would have risked making his solution obvious and one motive behind publishing the hardware and format patents separately would have been commercial secrecy. Also more characters would have meant bigger keyboards, typebaskets and casings, and early typewriters were already expensive and bulky products.

[11] E. Remington and Sons sold its typewriter business in 1886 to the Standard Typewriter Manufacturing Company, which changed its name to Remington Typewriter Company.

[12] Patent numbers 418239 and 464902

[13] Patent numbers 464903, 559621, 559755, 559756, 583156