Editors

St Anne's Academic Review

VOLUME 5, 2014

SOCIAL SCIENCES

Editor's Letter

Appearing in this fifth volume of the *St Anne's Academic Review* are ten pieces from current or recent postgraduates in the humanities, sciences, and social sciences at St Anne's College, Oxford.

There is much to admire in these pieces, above all that they uniformly strike a delicate balance between accessibility to non-specialists and technical rigour. We have gone beyond interdisciplinarity to include a blend of full-length research articles, research summaries, field reviews, and book and play reviews.

In spite of their thematic and methodological differences, these pieces nonetheless seek to more intimately *understand* their respective objects. This is as true of the pieces peering into the opaque subjectivity of human experience as it is of those wading through the intricate workings of human physiology or the universe. It is remarkable how much they share when considered in this light.

And yet, in deference to their interdisciplinarity, we have foregone journal-wide referencing consistency to maintain conventions unique to the sciences and humanities and social sciences. So you will find slight differences throughout in this respect.

It has been a pleasure working with our editorial staff and most of all with the authors to produce an absorbing academic collection. I hope you also find in it much to lose yourselves in.

Sincerely,
Kevin Busch
Editor in Chief

Play Review
Caucasian Chalk Circle

BENEDICT NICHOLSON

Benedict is an MSt student in Spanish, who moved to St Anne's college having done his undergrad at LMH. His research interests include 13th century Spanish literature and the role chroniclers had to play in the representation of early Iberian history. He is also active in the theatre scene at the university, having acted in and directed a number of shows during his time here. This review article is an examination of Bertolt Brecht's Caucasian Chalk Circle *inspired by a February 2014 student version performed at the Oxford Playhouse. The article discusses the themes of the play and how they are represented on stage, with an introduction to the principles of the style of theatre and its moral message. The main focus of the article, encompassing all the above, is a discussion of why the Brechtian mode of theatre remains relevant over fifty years after the play was first performed.*

In the programme notes of Screw the Looking Glass's production of Bertolt Brecht's *Caucasian Chalk Circle*, the playwright summons a quote to describe a significance of the circle:

'Of course the wheel just goes on turning / What's once on top cannot remain.'

This extract, from the *Ballad of the Waterwheel*, also directs our understanding of the narrative arc of *Caucasian Chalk Circle*, embodied in the rise and fall of people in society and the fluidity of the roles they play. After a brief prologue, where villagers discuss how best to use resources they have recently acquired, we are invited to watch an ancient parable, presented as drama by the singer Arkadi (Jack Sain), alongside the residents of two kolkhoz villages. The source material for this parable, as with much of Brecht's theatre and style, is adapted from Chinese theatrical traditions, more specifically a Fourteenth Century play called 'Circle of Chalk'. Whilst, through the character of Arkadi, the

playwright acknowledges that his Chalk Circle is different, its relation to its predecessor remains obvious. The story follows a very similar plot line, with the only significant change being the precise moral message in the play.

The first act tells the story of Michael, a child of aristocracy, who is abandoned by his callous mother Natella (Gráinne O'Mahoney) during a coup and rescued by a servant girl, Grusha (Constance Greenfield), on the same day that she promised herself to the Soldier (Leo Suter). Michael and Grusha flee the revolution-gripped city for the countryside and are forced to overcome a series of obstacles including disease and an unwanted marriage, only to be caught and brought back to the city. There, Grusha is made to stand trial; Natella has returned to the city demanding her child back in order to reclaim her late husband's estates, a reminder of the whimsical nature of the nobility and how vulnerable those without power can be. The second act follows the career path of Azdak (Luke Rollason), a clerk who has become an unpredictable judge in the time between Grusha's escape and return. The two narrative threads come together for the Solomon-style judgment, from which the play takes its name, during which Michael will be returned to the person deemed by the judge as his rightful mother. This judgment ultimately echoes the prologue and reinforces the parable being told; that resources should go to those who are most able to use them.

The production did justice to Brecht's script through James and Tania Stern's fluent and intelligible translation; one often forgets that *Caucasian Chalk Circle* was not originally written in English, high praise for any translator. There was an abundance of thought provoking, poetic moments that transpose Brechtian modes of expression into typically English yet innovative aphorisms: lines like 'Terrible is the temptation to do good' and

'Even in the bloodiest times there are good people' stick in the mind.

The play was presented in classic Brechtian style, with the audience always aware that they are watching a spectacle rather than a depiction of reality. The actors were permanently visible at the side of the stage along with all of the crew that were involved. This, taken, together with the sharp, high contrast lighting changes and outbursts of song interspersed throughout the piece, ensured that the audience was always aware of what they were watching through the transparency in staging that it achieved.

The set was relatively sparse for an Oxford Playhouse production, but this allowed the audience to focus on the actors' craft and ability to convey the moral message of the story. Of note, in the role of Azdak, Luke Rollason managed the difficult task of carrying the humour and controlling the pace of the second act with apparent ease. Likewise, Dominic Applewhite's versatility in the characters that he played, from grovelling government assistant to bawdy soldier, was extremely impressive. Indeed, the vocal and physical range of the ensemble at large was thoroughly deployed, though Jack Sain's physicality as the Singer was particularly striking as he stalked around the stage as the master puppeteer, controlling events and observing from afar.

Shadow puppetry, that is the projection of moving shadows onto a white background, whether they be the shadows of the actors themselves or puppets, was also used effectively. These were presumably a nod to the play's Oriental roots in the 'Circle of Chalk' and accompanied the scenes of revolution, allowing the production to imply violence more effectively than would otherwise be possible on stage. The production managed to avoid the often distracting and clumsy effect of stage gore by simply neglecting to show the most violent parts, allowing the audience to

imagine it much more explicitly than is possible to depict on stage. Two puppets were used to represent Michael in his different phases of life, ingeniously made of quartz, with the gold vein running through it representing his suppressed nobility. The puppets, moving in a way one might expect a human baby to, are a credit to their creator Suzi Battersby.

    Caucasian Chalk Circle is a notoriously difficult play to produce. Incorporating many of the traditionally ignored aspects of the play, such as the prologue and the songs, this production pulled them off with enough aplomb to justify their inclusion. Screw the Looking Glass offers a presentation of Brecht's play in a way that reveres his anti-realist traditions while still engaging with its twenty first century audience. Brecht's Twentieth Century update of an ancient parable of human nature and the rise and fall that occurs in a manufactured society serves as a timely, anti-realist reminder in an era marked by the ubiquity of realism in theatre. This production was ultimately successful in its ability to carry the moral message of a play of significant length, some three and a half hours, without the audience ever having their minds wander away from the action. It is a permanently engaging piece with the characteristic that defines all great theatre, the ability to create thought and dialogue through the simple means of entertainment.

The Ballad of Su Forty-three and the Problems of 'Historical Memory'

HANNAH THEAKER

Following an undergraduate degree in Chinese at the University of Edinburgh and a total of three years in Asia, Hannah has spent the past year honing the grasp of classical Chinese and Manchu under the watchful eyes of the China Institute here in Oxford. Her research is part of a wider academic project seeking to re-conceptualize Qing dynasty history through reintegration of perspectives from the margins of empire. Her own research interests include Sufi Muslims in China, rebellion, structures of empire and authority, and the relationship between ecology, religion and power in late imperial China. This article addresses the search for 'alternative histories' within oral and folk sources, by means of discussion of 'The Ballad of Su Forty-three', a piece recorded in Republican-era Gansu which recounts the events of a rebellion some one hundred and fifty years earlier. Through examination of the content of the ballad and its significance, it discusses the contested nature of historical memory and the mutually modifying nature of 'official' and 'unofficial' histories.

All was calm and all was quiet,
In the west rose Su Ahyin
Su Ahyin's got his Salar troops,
And they're off to seize Lanzhou.

Su Ah-yin,
Ain't so bright.
Follows a teacher named Ma Mingxin.
Ma Mingxin's a learned chap,
Raised 3600 Salar men,
Su Ah-yin's going to take the world.

压定压定且压定，
西头出了个苏阿吟；
苏阿昨引的撒 喇兵，

一心要夺兰州城。

苏阿吟，
没才情，
搬了个师傅是马明星;
马明星师傅才智大，
练起了三千六百老撒喇，
要保苏阿吟拿天下。

(Li Xinghua李兴华, Feng Jinyuan 冯今源 1985: 804)

'The Ballad of Su Forty-three' (Ch. *Qianlong sishiliu nian Hezhou shibian ge* 乾隆四十六年河州事□歌) presents a curious challenge to the scholar seeking fresh perspectives on the past. Collected from a Hezhou 河州 street performer in 1937 by ethnographer and historian Wang Shumin 王□民 as he travelled through the region, the ballad runs to some 15 pages of dense text. Clearly designed for professional performance, it is a lively, vivid piece recounting with a high level of detail the events of the 1781 uprising. (*ibid*: 802-219)

On the Ballad

The ballad is generally ascribed to the local tradition of *daojiangshui* 倒□水 singing, a folk form distinguished by its fast pace, rhythm, rhyme and humour. (*ibid*: 802) As Feng and Li explain in their introduction to the edition consulted here, the name of the tradition can be loosely translated as 'pouring sauce ballads'. They theorize that the name comes from the fluency required to perform one well: a string of words flowing like sauce from a bottle. (*ibid*) There are relatively few collections of ballads

attributed to the *daojiangshui* tradition, but descriptions suggest it was a highly localized tradition from the city of Hezhou, dedicated to recounting events and stories (Wang Shumin 王□民 1940: 271).

Fieldwork conducted by Deng Jingsheng □靖声 in tracing the footsteps of famed local performer Wan Shicang (万世□), more commonly known as Old Wan (*Lao Wan* 老万), 1880-1960, suggests that the *daojiangshui* name for the tradition, as collected by Wang Shumin, was in fact Wan's own coinage, and that it replaced the older name *kuaibanshu* 快板□, a tradition of clapper storytelling found in regional variants across northern China. (Deng Jingsheng □靖声 1989: 133-135) Wang Shumin's usage of the *daojiangshui* term in preference to an alternative name with a wider currency suggests that there is a chance it was even Old Wan he met or, perhaps, a performer associated with him.

Deng's portrait of Old Wan portrays a highly creative individual, in some respects highly eccentric and in others a typical itinerant folksinger, noted in particular for his fast tongue and disrespect for authority. To illustrate the above points, Deng recounts an occasion on which Old Wan was summoned to meet Ma Bufang □步芳, the Republican-era Qinghai warlord, to answer for his satirical lyrics. Old Wan responds fast and wittily, to the warlord's delight. Whilst the tale shows the hallmarks of being apocryphal – folk singer speaks truth to power – the anecdote indicates the satirical function this song tradition played in local life. (*ibid*: 137) Old Wan may also have played a key function as an interpreter of news for more remote communities as an itinerant performer – titles he sung included content on current affairs, notably the Japanese bombing of Xining. (*ibid*: 134)

The detailed nature of the ballad and the very existence of the *daojiangshui* tradition indicate the importance of remembrance of such issues. Following Maris Gillete's work on mourning rituals held among Xi'an Hui populations to commemorate the suffering inflicted on those communities by the 1860s rebellions, I would suggest that such ballads served a similar purpose here: a performed remembrance of violence and its consequences that structures a collective memory. (Gillete 2008: 1014) The mere fact of its performance in 1937 – over one hundred and fifty years after the events recounted – suggests the manner in which the events of 1781 have shadowed the interpretation of subsequent events. Each account of the Muslim rebellions of the 1860s, 1895 and 1928 begins with 1781, making it a touchstone that then defines and influences the present.

As an oral history of an event is normally only accessible through the records of those responsible for its repression, it is tempting to view the ballad as an 'alternative history', the truth, however, is rather more complex. The ballad weaves official terminology with local concerns, demonstrating rather that there is no straightforward opposition between official and unofficial histories. Official/unofficial realms instead are mutually modifying, interactive discourses wherein memory of the past is structured by experience of the present. The Bakhtian conception of 'social heteroglossia' is informative here: the idea that no language is a unified system of norms. It is instead a chaotic project, consisting of multiple overlapping languages, each of which represents a different way of conceptualizing the world. 'Social heteroglossia' refers to the idea that speakers creatively utilize multiple levels of discourse to their own ends. (Morson, Emerson 1990: 139-143) The ballad can be seen as existing at the intersection of numerous levels of discourse: mediating between official history, local

memory, between audience and performer, and between the interests of the ethnographer who collected it and the assumptions of the performer with regards to those interests. Before moving onto a more detailed discussion of the content of the ballad, a brief explanation of its historical context is necessary.

1781 in Historical Context

Hezhou is today known as Linxia 临夏 and lies in the remote province of Gansu 甘肃, north-western China, where the interior gives way to desert, steppe and Tibetan plateau. The area is ethnically and linguistically mixed, including populations of Sino-Muslims, Salars (Turkic-speaking Muslims), Han Chinese, Tibetan Muslims, Tibetan Buddhists and Mongolians, set amid the forbidding topography of southern Gansu. Hezhou then was a bustling entrepôt, and also an Islamic centre of learning known as 'Little Mecca'. However, for the officials charged with its governance Hezhou represented a frontier of 'China proper' – the gateway to the Tibetan, Salar and Mongol communities of Qinghai lay at Jishi Pass (积石关). In 1781, all three of them would have been governed under a separate legal code from their Han Chinese neighbours, namely the Tibetan statutes. Their cultural and linguistic differences were thus encoded into the models of governance by which the Qing approached the problems of administering their new domains. (Ma Haiyun 2007: iii)

The 1781 uprising the ballad recounts was a complex localized series of events, with origins in the local power dynamics of the area and which began with a dispute that occurred between the followers of rival Muslim *menhuan* (口宦, Chinese Sufi lineages). Sufism began to enter China in approximately the 16th century, travelling with movement of caravans along trade routes

heading west and greatly facilitated by the Qing conquest of Xinjang in 1759. Sinophone Muslims would travel to study in the centres of learning in the Islamic world, and return having been initiated into a Sufi order. Their religious charisma (Ar. *baraka*) would attract converts whose donations and support gave *menhuan* leaders (Ch. *jiaozhu* 教主, Ar. s*haykh)* significant temporal power. Religious authority - and its attendant temporal status - would be passed along to a successor on death of the shaykh. *Menhuan* networks stretched across China, but largely remained centred on Hezhou and the surrounding Muslim communities (including Sinophone, Turkish, Mongol and Tibetan-speaking communities). Although Islamic in nature, the connotations of the word *menhuan* itself, most probably derived from the Chinese *huanmen* 宦□ meaning a powerful family with close bureaucratic or official connections, emphasize the importance of the manner in which they entwined themselves with local politics. Competition for converts and inheritance disputes resulted in numerous branches and factions which divided even those groups ostensibly within the same Sufi tradition. (Guan Lianji 关□吉 1984: 56)

In 1781, the conflict involved two *menhuan,* the Jahriyya and the Khafiyya, competing for adherents in the local Salar communities. Prior to the outbreak of armed hostilities, lawsuits had been presented in local courts in which both sides had accused the other of heterodoxy. Lack of a conclusive legal judgement allowed the conflict to simmer on, culminating in the murder of a Khafiyya headman (Han Sanshiba □三十八) by Jahriyya followers. The Qing response designated the Khafiyya, who had arrived first to the area as 'the Old Teaching' (*laojiao*老教; ie old, traditional and orthodox), and the Jahriyya as 'the New Teaching' (*xinjiao* 新教; ie new, dangerous and heterodox), and on this basis,

arrived at the conclusion that it was the Jahriyya who were at fault for causing the present unrest. They arrested and eventually executed the leader of the New Teaching, Ma Mingxin □明心, prompting an armed uprising by his followers, primarily Salar Muslims from the Xunhua 循化 region of modern Qinghai province. The rebels were eventually defeated after a protracted campaign. The ferocity of the subsequent crackdown devastated the Salar communities south of the Yellow River, reducing the population so drastically that the prior twelve administrative clusters were merged into a mere eight. (Dwyer 2007: 21)

However, the complexity of these events should not be underestimated: the rebellious forces were not purely composed of Salars, nor purely of New Teaching Muslims. Similarly, the multi-ethnic force that eventually put the rebellion down also included Muslims. The uprising is often viewed by historians as the beginning of a hundred-and-fifty year period in the history of Gansu and Qinghai characterised by feuding *menhuan,* local power struggles, violence, and state repression.

The Rebellion As Told

As portrayed in the memorandum of the Qing officials, and in the numerous historical accounts, the 1781 rebellion is exclusively religious in character. The scholarship of Joseph Fletcher and Ma Tong, respectively, brings focus to bear on the mystical, Sufi nature of the *menhuan*, foregrounding the importance of loyalty to shaykh and the rebellious influence of Sufi teachings. (Fletcher 1995: 21-51; Ma Tong 马通 1983: 101-117) In such accounts, the execution of Ma Mingxin at Lanzhou is a key point in the escalation of the conflict, prompting Su Forty-three and the New Teaching Muslims into a full-scale insurrection. Crucially,

Ma's death represents the moment in which the Qing view the rebellion as a threat to their authority.

By contrast, Ma Mingxin is mentioned precisely once in the ballad, as Su's teacher. In the narrative constructed by the ballad the motivation of Su Forty-three is to 'take the world' (*na tianxia* 拿天下), with the backing of fellow Salar headman Hann Two 韩二个. To this end, he is prepared to do anything it takes – including ally with the Old Teaching headman, Hann Thirty-eight (*Han Sanshiba* 韩三十八). It is only when Hann refuses to join him that Su murders him and the Qing are thus drawn into the conflict when Hann's wife seeks justice. The focus on key individuals who chose either to join Su's campaign, or not to join reflects the course of the revolt as viewed from a local perspective: the death of headman Hann Thirty-Eight is a turning point to which the ballad devotes considerable time. His murder is presented as the key cause for the escalation of the violence rather than intractable religious feuding. (Li, Feng 1985: 805)

By the 1930s, intermittent Muslim violence was viewed as a part of life. Salar Muslims were stereotyped as violent fanatic brigands, to the extent that Salars were the bogeymen the Han women of Hezhou used to scare their children with. (Lipman 1990: 77) Successive revolts meant that for many, violence no longer needed explaining save by reference to past events such as the 1781 uprising. Sayings emphasized the recurrent, inevitable nature of revolt: 'A small rebellion every 30 years, a large one every sixty' [六十年一大乱，三十年一小]. (Mu Shouqi 慕寿祺 1972 (Vol 7): 24.42b) However, within the ballad, the conflict is not presented as being precisely Islamically-inspired. The rebels of the ballad are clearly Muslims – see, for example, the verses in which Salar rebels are tricked into eating pig fat and promptly are transformed into

pigs – but this is not the most important aspect of their identity. Hann Thirty-Eight instead complains of their status as *fan* 番 ('aborigines', a term used for Tibetans, Mongols and others which in some contexts carries connotations of barbarism) and their lack of opportunity occasioned by that status. The authority of Su and Hann was grounded in an imbrication of religious and political authority within Salar and Sinophone Muslim communities. Competition over limited local resources was thus expressed in religious terms, even when grounded in the desire of one man or faction to expand the sphere of their influence. To Qing authorities, the salient feature of the rebels was their Muslim identity and creed, a judgement predicated on the experience of 1781. As the first major uprising, 1781 then set the pattern by which all other conflicts were understood – a pattern which was then creatively adopted by the Muslim populations of Gansu to their own ends.

The adoption of the terms 'New Teaching' and 'Old Teaching' is an interesting demonstration of the intersection of official discourse with the realm of oral history. Their incorporation into the ballad suggests not so much their validity – the meanings of the terms has been shown to vary across time even in usage by Qing officials (Lipman 1997: 138) - as their utilization by individuals to present themselves. Ildiko Beller-Han's recent study emphasizes the dynamic nature of the dialogue between 'official' and 'unofficial' histories. Historical knowledge is produced in the junctures, and the resultant texts can be contradictory in themselves. (Beller-Han 2012: 311)

Whilst the terms 'New Teaching' and 'Old Teaching' feature in the ballad, Su Forty-three holds discussions at the 'Da Gongbei Mosque' (*da gongbei si* 大拱北寺), a building associated

with the Da Gongbei *menhuan,* a third Sufi lineage. (Li, Feng 1985: 810) Despite the ballad's attempts to gloss the conflict as New Teaching versus Old, as Salar troops versus innocent villagers, the self-imposed categories cannot be sustained even with the discursive field of the ballad. The complexity of the history involved defies the retrospective application of the state-defined categories as the ballad appeals to multiple levels of discourse. In a similar fashion, Su is never presented purely as a villain – although the violence of his actions is deplored, he is portrayed as initially reluctant to raid Han villages and there is a certain admiration for the manner in which he defeats each of the Qing forces sent to bring him to justice. It is only when the local Tibetan troops are called up (using *fan* to fight *fan*) that he is finally quelled – a recognition, perhaps, of the manner in which Chinese and Manchu administrators were viewed as outsiders.

Within the ballad, the villagers are presented primarily as law-abiding adherents of the 'Old Teaching'. Performed in a diverse area ruled by Muslim warlords with their own religious agendas and wherein multiple *menhuan* were active, the choice to valorize the Old Teaching is one could that speak to multiple agendas. Such a choice may represent a simple desire to portray the villagers as 'orthodox', 'conformist' and 'non-rebellious'; the manner in which local religious factions often competed to claim the label of 'Old Teaching', or it may represent a tacit opposition to the Muslim warlords then forcibly promoting a reformist Islam (a modern 'New Teaching') in preference to the what they perceived as the Sinicized practices of Chinese Sufism. The memory of 1781 as presented in the ballad was contested territory, given the role of history in interpreting the present, and thus the ballad contains multiple interpretations. It is not a unified text and its

contradictions can be resolved only by understanding of the context into which it was sung.

Women and Local Geographies

The emphasis the ballad puts on the role of women in the conflict, and local reactions to the arrival of the Salar troops foregrounds a geography at odds with the official accounts of the sieges of the major cities. In the transcription of the ballad, entire pages are devoted to the progress of the rebels through each of the nearby hamlets and the reactions of the residents:

The Salars went to Shuangcheng – o,
Women hurried to pack their children into pots and pans
And followed their husbands to the caves.

The Salars went to Jiatangli - o,
Scared a woman, she nailed fast her door and crawled in her *kang,*
Someone said: "It's no use you going inside,
Crawl in a *kang* and you'll get burnt - o."

Someone said: "Burnt is burnt,
Crawl in - oh
When you come out, the Salar Muslims'll be caught and the chaos gone."

撒喇反到双城呢，
忙了妇人把尕娃塞的着缸盆哩，
跟上丈夫钻洞去。

撒喇反到夹塘里，
吓得婆娘顶住柴门钻炕洞。
有的说:"我说不中你要钻，

钻着炕里是火烫呢。"

有的说:"烫是烫，
钻上呢，
出来时撒喇回子捉住胡藏呢。"

(Li, Feng 1985: 308)

The ballad lists the varied reactions the rebels gained in each town: from a massacre of monks at Balin Temple to the scenes when Su and his men arrive in Tangwangchuan and are fed breakfast by a rich relative who then joins their cause. Massacres of monks are recounted in graphic detail, yet so are episodes of generosity in which Su and his men distribute food to villages they pass through. By contrast, the pitched battles fought at the end of the ballad are recounted in comparatively little detail. (*ibid*: 808-809, 815-816) Whilst the Salar identity of Su and his men is emphasized throughout, the make-up of the villages they pass through – a mixed zone of Salars, Han and Sinophone Muslims - is not. Each village is simply a village with its own reaction to the arrival of the Salars, be that flight, fight or joining the rebellion. Each place, however, comes coded into a local geography of power which would have structured the understanding of the listeners: for example, Tangwangchuan was a Sinophone Muslim stronghold of the New Teaching connected to the Salar communities by marriage which suffered greatly in the post-1871 crackdown. However, reconstructing such a local geography is perhaps only possible in fragments given the incomplete nature of the historical record and required sensitivity to the various contexts of ballad and history.

Women function throughout as the moral compass of the society the ballad depicts – they are portrayed both as the principal

victims of the violence and as those who attempt to stop it. Even as the resistance Su offered the Qing troops is admired for its heroism, the wives of Hann Two, Hann Thirty-eight and Su Forty-three are united in lamenting the desire of their husbands' for temporal power. Hann Thirty-Eight's wives mourn their husband's ambition that had led him to his end thus:

> "The two wives heard and came running,
> Held their husband and cried
> "Who told you to be a headman!
> Last year you were the new headman,
> This year you're the old headman,
> What good did being headman do – o?"

> 两个妇人听着跑者来，
> 拉住丈夫动哭声:
> "谁教你给百姓当户长!
> 年年你要当户长，
> 去年你当的新户长，
> 今年你当了旧户长，
> 这户长当着有甚么利益呢?

<div align="right">(ibid: 804)</div>

In a previous episode, the wife and mother of Hann Two commit suicide in an attempt to persuade him of the futility of fighting the Qing armies. Differences of religion, language or ethnicity vanish in the presentation of women: Salar wives implore their men to give rebellion up as a lost cause, or seek justice on their behalf whilst village women co-ordinate the defense of their towns or flee to the hills.

Conclusions

The ballad itself ends with the heads of the leaders displayed in the towns of He-Huang following their execution, Buddhist temples built on the key battle sites and the use of corvee labour to rebuild Hezhou after its destruction – an utter defeat for Su and his Salar troops. Fatalism permeates the entire text: it begins with a comet, an omen of destruction, seen hanging over Hezhou. Despite Qing awareness of its presence, official attempts to avert disaster do nothing.

> An imperial city built on Hualin Mountain
> Built it up for two and a half years
> Now it's finished the empire's at peace.

> 皇城筑在华林山。
> 皇城打了二年半，
> 打完了皇城乐平安 。

<div align="right">

(*ibid*: 319)

</div>

A bitter irony lies in these closing lines, for three years after the events of 1781, another follower of Ma Mingxin, Tian Wu 田五, took up arms against the state in another bloody affair that left the countryside still further in ruins. The fatalism of the ballad rails against the futility of such armed struggle: a reminder to listeners that 1781 brought nothing but death, same as the affairs of 1784, the 1862-1874 uprising, 1894-5 and 1928. There is no mention of the victors – the Old Teaching Muslims who gained the property of the New Teaching rebels when the violence was over. The memory of 1781 thus takes on yet another dimension as an interpretative frame for later history: a warning against ambition

and attempts to encroach on the power of the state.

'The Ballad of Su Forty-three' remains a truly fascinating survival. In the complexities and contradictions of the text and its multiple interpretative frames can be seen the difficulties of dealing with 'historical memory'. The ballad provides fragments of an account of 1781 that lies outside the official sources: of alternative geographies, of the imbrication of religious and political power, of different priorities. However, that memory is structured by experience of the present, seen in the strategic use of the terms 'Old Teaching' and 'New Teaching', now incorporated onto a past they do not entirely fit. Despite the difficulties raised by working with such material, the richness of this and other ballads from similar traditions rewards further study. At present, very little research sensitive to historical past has been conducted on Chinese traditions of clapper storytelling, perhaps due to the comparative scarcity of well-edited collections of such ballads. Space has allowed merely a short exploration of some of the issues raised. However, the importance of interpreting and working with such material, hopefully, emerges as incontrovertible.

REFERENCES

Beller-Han, Ildiko (2012) "Feudal villains or just rulers? The contestation of historical narratives in eastern Xinjiang" in *Central Asian Survey,* 31 (3), 311-325.

Bulag, Uradyn (2010) "Can the subalterns not speak? On the regime of oral history in socialist China" in *Inner Asia,* 12 (2010), 95–111.

Deng Jingsheng □靖声 (1989) "Guanyu daojiangshui yishuren – Lao Wan shengping shiji de diaocha shouji" 关于倒□水□□人 - □方生平事迹的□□手□ [On a '*daojiangshui*' artist – notes on an investigation into the life and work of Old Wan] in *Xineng chengzhong wenshi ziliao,* Vol. 2 (1989), 131-138.

Dwyer, Arienne M. (2007) *Salar : A Study in Inner Asian Language Contact Processes.* Wiesbaden: Harrassowitz.

Fletcher, Joseph (1995) "The Naqshbandiyya in North-west China" in *Studies on Chinese and Islamic Inner Asia.* Edited by Beatrice Forbes. Aldershot: Variorum, 21-51.

Gillete, Maris (2008) "Violence, the State and Chinese Muslim Ritual Remembrance" in *Journal of Asian Studies*, 67(3), 1011-1037.

Guan Lianji, 关□吉 (1984) "Lun 1781 nian Su Sishisan lingdao de zhehelinye qiyi" 论1781年□四十三□□的哲赫林耶起□ [On the Su Forty-three-led 1781 Jahriyya rebellion] in *Lilun xuexi* 理□学□ 6, 49-61.

Li Xinghua 李兴华, Feng Jinyuan 冯今源, eds. (1985) *Zhongguo yisilanjiao shi cankao ziliao xuanbian* 中国伊斯兰教史参考资料选编 1911-1949 上册 [*Selected Reference Materials on the History of Islam in China].* Yinchuan: Ningxia Renmin.

Lipman, Jonathan (1997) *Familiar Strangers: A History of Strangers in Northwest China.* Seattle, London: University of Washington Press.

Lipman, Jonathan N., and Stevan Harrell, eds. (1990) "Ethnic Violence in Modern China: Hans and Huis in Gansu, 1781-1929" in *Violence in China: Essays in Culture and Counterculture*. Albany: State University of New York Press.

Ma Haiyun (2007) "New Teachings and New Territories: Religion, Regulations, and Regions in Qing Gansu, 1700--1800." Ph.D., Georgetown University.

Ma Tong □通 (1983) *Zhongguo Yisilan jiaopai yu menhuan zhidu shilue* 中国伊斯□教派与□宦制度史略 [*A Brief History of Islamic Sects and the Menhuan System in China].* Yinchuan: Ningxia Renmin Chubanshe.

Morson, Gary Saul and Emerson, Cary (1990) *Mikhail Bakhtin: Creation of a Prosaics.* California: Stanford University Press.

Mu Shouqi 慕寿祺 (1972) *Gan Ning Qing shilue* 甘宁青史略 [*An Outline History of Gansu, Ningxia and Qinghai].* Edited by Zhao Yuanzhen 趙元贞 and Li Bing 李炳, 10 vols. Taipei: Guangwen shujuTan Qixiang 谭其骧 ed. (1987) *Zhongguo lishi ditu ji, di ba ce* 中国历史地图集,第八册 [*Collected Chinese Historical Maps, Vol. 8*]. Shanghai: Ditu.

Wang Shumin 王□民 (1940) "Gan Qing wenjian ji" 甘青闻见记 [Gansu and Qinghai Diaries], in *Gansu wenshi ziliao xuanji,* Vol. 28 (1988).

Yang Haiying (2001) "The politics of Writing History in China: A Comparison of Official and Private Histories" in *Inner Asia*, 3(2), 127-51.

The Origin of the Moon Revisited

HELEN ASHCROFT

Helen is a second year DPhil student in Earth Sciences specialising in Experimental Petrology, and her research interests involve looking into the formation, differentiation and evolution of the terrestrial planets. Helen also contributes to the online and print editions of Bang! Science Magazine, and in her free time enjoys making and listening to music.

The Moon is the closest object to us in the solar system. We observe it every night in our sky and we know it controls the rotation and orientation of the Earth's axis, causing the tides and making possible a stable climate in which life can thrive. But how the Moon formed, and what it formed from, remain among the biggest mysteries in planetary science.

Any theory of how the Moon formed from the Earth has to explain the physics and dynamics of the alleged collision with and subsequent changes to the Earth and its orbit, as well as the chemical similarities and differences between the Earth and the Moon.

According to the Giant Impact Theory, now widely accepted, the Moon formed as a result of a Mars-sized impactor colliding with the early Earth and creating a debris disk around the Earth which formed the Moon. This hypothesis was first suggested at the first 'Origin of the Moon' meeting in 1984, and gained popularity and strength over the next decade as lunar meteorites and terrestrial rocks were being analysed for their geochemistry. For it was then that the chemical similarities between the Moon and the Earth were becoming apparent. Such geochemical advances also suggested constraints on the age of the Moon. The

Giant Impact Theory is supported by the fact that intense bombardment of the accreting planets by smaller planetary building blocks was a commonly occurring process early on in Solar System history.

At first, these chemical similarities seemed to support the Giant Impact Theory by suggesting that some kind of mixing of material between the Earth and the impactor occurred during the formation of the Moon. However the discovery that the Earth and the Moon have nearly identical oxygen, iron, hydrogen, silicon, magnesium, titanium, potassium, tungsten and chromium isotope compositions threw a spanner in the works. Isotopes are atoms of a chemical element that have the same number of protons, but different numbers of neutrons, and the ratios between two isotopes of one element are governed by a variety of chemical and physical processes. In fact isotopes are so sensitive to a range of factors that they provide chemical fingerprints for different planets – Earth has a very different isotope signature from Mars, for example, and this fact is used in the classification of meteorites.

Although the Earth and the Moon have similar isotope ratios, there are clear differences in their composition, for instance the Earth is comprised of 30 % iron, whereas it is believed the Moon only contains 10 % iron. The Moon also has fewer volatile elements, suggesting that they may have been lost during formation.

Isotope similarities could suggest that the Earth and the Impactor were formed from the same material, or that the Earth and the Impactor equilibrated fully. However both of these processes require additional explanation or information. For example numerical modelling of timescales of isotope equilibration and the timescale of the impactor don't match up, and that different planetary bodies show different isotope signatures

suggests that two planetary bodies, however closely they form, should show some differences.

Therefore the similarity between the isotope signatures of the Earth and the Moon places constraints on the impact and the Impactor, and raises more questions than it answers. Initially the similarity of some of the isotope systems supported the Giant Impact theory, which suggested that there was some equilibration, or mixing, between the Earth and the Impactor, however the similarity of silicon and tungsten isotopes that are controlled by very different processes requires additional explanation.

Any theory of Moon formation will have to explain the dynamics of the Earth–Moon system, which have to be accounted for in terms of the angular moment and spin of the Earth. Many numerical models of the impact have been run, and the optimal situation appears to be one in which an impactor around $1/10^{th}$ of the size of the Earth grazes it and forms a debris disk predominantly of impactor material and to a lesser degree of terrestrial material. Hence the Moon is mainly formed from impactor material. More recent modelling suggests that a larger impactor may be responsible; however the probability of this occurring is much lower than the probability of impact by a smaller body.

In some ways the evolution of the Earth-Moon system will remain enigmatic, as a lot of the hypothesized events or processes are relatively improbable. This does not however mean that they cannot occur, but currently we do not have the information that would support the Giant Impact Theory.

The problem remains that all of the current models require secondary processes overprinting the primary formation information. Sometimes the simplest model is the best, but as a lot of the unknown factors are untestable we may never know.

However the future for lunar research is not bleak, as was evident at this year's Royal Society Meeting in London, which encompassed a lively range of debates and discussions. The general consensus from the meeting was that the Giant Impact Theory is the most probable mechanism for moon formation; however more information regarding the other terrestrial planets, for example Venus and certain asteroids is required to shed more light on early solar system processes.

Pivotal disruption? Abnormal activity in motor control regions in stuttering

EMILY L CONNALLY*, DAVID WARD+, CHRISTOS PLIATSIKAS+, KATE E WATKINS*

By day Emily is a 3rd DPhil student in Experimental Psychology. By night she is the co-founder of Imaginary Husbands Productions and acts as advocate, agent, friend, lyricist, songwriter, and general unicorn for a few local bands, but mostly for One Wing Left, which started with fellow StAnners in the Annie Barnes Room in Trinity 2013. In her free time Emily plots to become the first 3rd party President of the United States. She stands for sustainable living, education reform, global alliance, and the power of free music.

Introduction

Stuttered speech can contain repetitions of sounds, or inability to initiate sounds and often occurs simultaneously with facial-motor disruption similar to tics seen in patients with movement disorders. People who stutter do not typically have other gross movement difficulties nor do they exhibit any sort of cognitive deficits. Components of the speech and motor systems are thus considered the most likely candidates for anomalous development in people who stutter. As there are no post-mortem studies to inform the debate on causes of stuttering, currently, the bulk of our knowledge about potential brain abnormalities comes from two sources: studies comparing structure and function in people who stutter and fluent controls. These studies typically use magnetic resonance imaging, or MRI, and when the same tool is used to look at changes in blood flow related to specific tasks it is called functional MRI or fMRI.

Structural imaging gives some evidence for disrupted grey matter development in the speech and motor systems in the form

of abnormal shape of the auditory cortex[1], reduced thickness in speech-motor planning regions[2], and increased volume in left basal ganglia[3]. Some groups have shown anomalous white matter development in portions of the corpus callosum[4,5]. A consistent finding across other diffusion studies show disrupted white matter underlying speech-motor planning regions[6] that also show disrupted function[7]; disrupted white matter in major speech and motor pathways including the major language pathways connecting the brain from front to back[8–10] and all three pairs of cerebellar peduncles that facilitate communication between parts of the brain and spinal cord, and integration of behaviours with different body parts[10].

Functional imaging literature likewise supports disruption in the speech and motor systems in event-related and resting-state studies. A meta-analysis of overt speech in stuttering compared activation maps in people who stutter and in fluent controls in eight early neuroimaging studies[11]. The review concluded that all participants largely recruit the same brain regions during speech, though a diagnosis of stuttering was related to increased activation in lateral speech-motor regions and other components of the motor system including the supplementary motor area (SMA), cingulate cortex, and the vermis of cerebellum. The meta-analysis identified three "neural signatures" of stuttering 1) overactivation of the right frontal operculum, anterior insula, or both (Figure 1:A); 2) overactivation of the cerebellar vermis (Figure 1:C); and 3) an "absence" of activity in auditory cortex (Figure 1:B). Treatment studies combined with fMRI to evaluate outcome due to therapy also implicate the frontal operculum, which is resistant to interventions both during tasks and at rest[2,12], whereas abnormal cerebellar activity appears to be related to treatment both during speech tasks and at rest.

Figure 1. Regions in which we predict differences between CON and AWS based on the "neural signatures" of stuttering (A=Frontal Operculum, B=Auditory Cortex, C=Cerebellar Vermis) and components expected to be taxed by the spontaneous speech component of our tasks (D=Basal Ganglia, E=pre-Supplementary Motor Area, F=Anterior Cingulate Cortex).
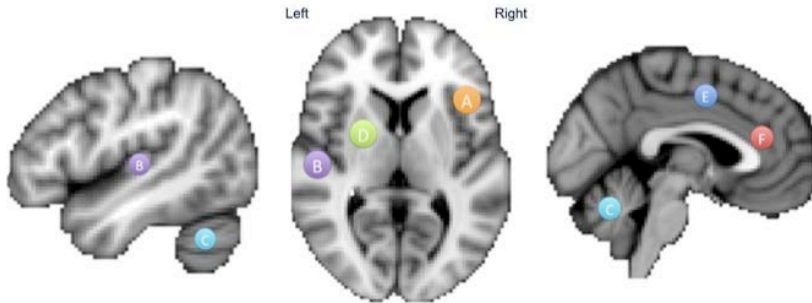
Though notably not a conclusion in the meta-analysis, abnormal basal ganglia activity is likely in stuttering, particularly as other core motor system components including cerebellum and supplementary motor area (SMA) are consistently overactive across studies. It is unlikely that overactivity in core components of the motor network would occur independent of any disruption to the hub that excites and inhibits activity of those other components. Empirical support for basal ganglia dysfunction in stuttering can be found in: treatment success of dopamine blockers (for review see[13]); early imaging studies of increased dopamine uptake in the basal ganglia components[14]; increased activity within individuals subsequent to treatment[12] or compared to fluent controls during speech-related fMRI[7]; altered connectivity at rest[2,15,16]; correlations between activity and stuttering severity or constellations of symptom[17,18]. There is strong theoretical evidence for central involvement of the motor control system in stuttering, still, there is a lack of continuity across functional imaging studies implicating specific basal ganglia structures or the direction of abnormalities relative to controls.

The primary aim of this study is to explore the influence of stuttering on overt speech-related brain activity. We use two tasks to elicit overt speech: 1) Sentence Reading and 2) Picture Description. We selected these conditions because traditionally stuttering is only elicited through more complex, longer utterances[11]. Additionally, the conditions should be well matched for speech production generally, as they place similar demands on the articulatory system. We expect recruitment of additional regions in the picture description task, which has a spontaneous speech component involving planning of both speech content and motor movement, across groups, resulting in increased activation in the components of the speech-motor planning system. This system includes regions identified as neural signatures of abnormal activity in stuttering and is likely dysfunctional in stuttering. The contrasts of interest include each condition relative to baseline and the picture description condition relative to sentence reading.

Empirical support for basal ganglia dysfunction in stuttering is ample: treatment success of dopamine blockers (for review see[13]); early imaging studies of increased dopamine uptake in the basal ganglia components (Wu et al., 1997); increased activity within individuals subsequent to treatment[12] or compared to fluent controls during speech-related fMRI[7]; altered connectivity at rest[2,15,16]; correlations between activity and stuttering severity or constellations of symptom[17,18]. There is strong theoretical evidence for central involvement of the motor control system in stuttering, still, there is a lack of continuity across functional imaging studies implicating specific basal ganglia structures or the direction of functional abnormalities (over vs. underactivity) relative to controls.

The primary aim of this study is to explore the influence of stuttering on overt speech-related brain activity. We use two tasks

to elicit overt speech: 1) sentence reading and 2) picture description. We selected these conditions because traditionally stuttering is only elicited through more complex, longer utterances[11]. Additionally, the conditions should be well matched for speech production generally, as they place similar demands on the articulatory system. Previous studies have used sentence reading (e.g.[7] and described differences in brain activity in several regions including some of those identified as neural signatures. We expected recruitment of additional regions in the picture description task, which has a spontaneous speech component involving planning of both speech content and articulatory movement, which we predicted would result in further increased activation in the components of the speech-motor planning system. Furthermore, we significantly increased the amount of speech produced in the scanner from that obtained in previous studies that was mostly fluent. By effectively doubling the amount of speech and including the picture description condition, we hoped to be able to capture utterances that were both fluent and dysfluent and compare the related patterns of activity (see next chapter). The contrasts of interest include each condition relative to baseline and the picture description condition relative to sentence reading. Within AWS we examine the effect of sex, handedness, and severity on activity in regions showing differences between groups.

We made the following predictions regarding the differences between groups:

- Across conditions we expect overactivity in right inferior frontal cortex, underactivity in auditory cortex, and overactivity in the cerebellum in Adults who stutter (AWS) relative to the control group (CON), consistent with the

reported "neural signatures" of stuttering (Figure 1; A, B, C, respectively).

- We predict that the contrast between conditions (Picture description > sentence reading) will be greater in AWS than in CON in primary speech-motor planning regions (pre-SMA and medial prefrontal cortex; Figure 1: E, F, respectively) and in regions supplying internal cuing about speech initiation to the medial frontal cortex (i.e. striato-pallidal-thalamic outputs of the basal ganglia: Figure 1: D).

Methods

Participants

Seventeen adults who stutter (AWS: 13M:4F; aged 19–54 years; 3 left-handers) and 17 age- and sex-matched fluent controls (CON: 13M:4F; aged 19–53 years; 3 left-handers) were scanned using functional MRI. No controls had a history or diagnosis of learning or speech disorders. All participants gave informed consent to their participation in the research in a protocol approved by University of Reading's ethics committee. Stuttering ranged in severity from very mild to severe, as assessed by a speech therapist using the Stuttering Severity Instrument-3[19].The data from some of the sample (CON) were used in a previous report, but further data have been added (AWS). Stuttering ranged in severity from very mild to severe, as assessed by a speech therapist (DW) using the Stuttering Severity Instrument-3[19].

Data acquisition

Functional MRI data were obtained at the University of Reading using a 3-T Siemens Trio scanner with a 12-channel head

coil. Whole-head T2*-weighted echo-planar images (TE=30ms) were acquired every 9s with a silent delay of 7s (i.e. sparse sampling) and comprised 2-s acquisition of 32 4-mm axial slices (in-plane resolution 3 mm x 3 mm). A '+' appeared in the middle of the screen during the 2-s acquisition period (Figure 2).

Figure 2. Tasks design and scanner acquisition method. Two conditions: Read (Sentence Reading) and Speech (Picture Description) were used to assess speech-related brain activity. Sparse sampling catches the final part of the stimulus-locked activity by exciting just after the task event.
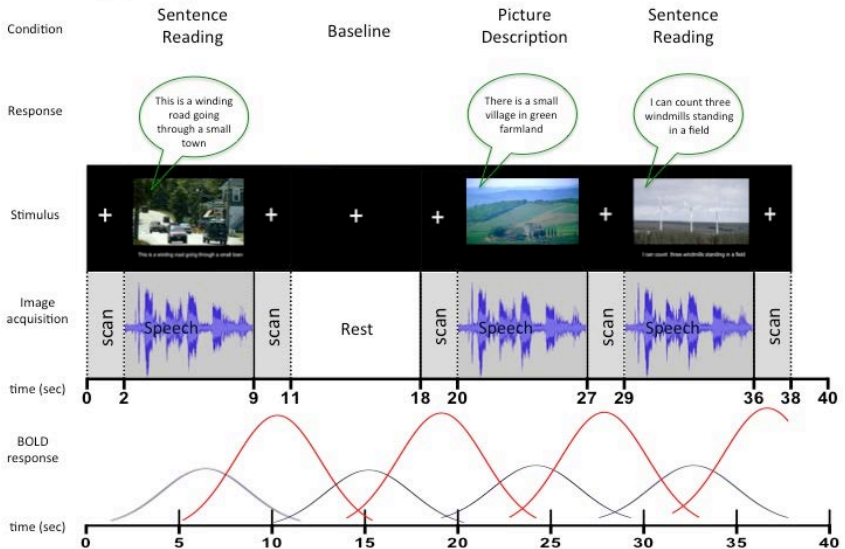
During the 7-s silent delay between measurements, subjects saw a stimulus via scanner-compatible goggles that was either a picture with a descriptive sentence below it, a picture with no text, or a '+' in the middle of the screen. Subjects were instructed to read the sentences aloud (Sentence Reading condition) or to overtly describe the pictures (Picture Description condition) and were explicitly told to stop speaking when the crosshair appeared so that there would be no speech-related movement of the head during data collection. Prior to the scan the task was explained to subjects who were allowed to practice outside the scanner. For

each of the conditions and the baseline condition, 40 volumes of data were acquired for a total of 120 volumes (18 min); the order of conditions was fixed and pseudorandom. Two runs were acquired in each subject, yielding 80 volumes of each condition (Sentence Reading, Picture Description, and Baseline).

Speech analysis

Speech was recorded using an MRI-compatible microphone. These recordings were later checked for task compliance. Most subjects produced a total of 160 task-related utterances in the scanner. In a single AWS, who was not excluded from the analysis, 37% of trials were not confirmed due to equipment malfunction, but all remaining trials were clearly compliant. Sentences were marked as normal or dysfluent.

Image analysis: Whole-brain

The functional images were analysed using the FMRIB Software Library (FSL; http://www.fmrib.ox.ac.uk/fsl[20]). In addition to the standard motion correction; which adds six head-motion parameters as covariates at the first level for each subject, volumes that were motion outliers were included as separate regressors in the model. Excessive motion (i.e. > 4mm) was observed during the end of a single scan session in one AWS and the volumes following the movement were removed from the time series (i.e. the run was truncated). The remaining data analysed normally. Each dataset was unwarped using a fieldmap and PRELUDE and FUGUE software running in FSL[20], spatially smoothed with an 8 mm full-width at half maximum smoothing kernel. A temporal high-pass filter with a cutoff of 150 seconds was used to remove low-frequency fluctuations in the signal. Two further regressors were used in the first-level analysis to remove

residual image artefacts. The mean time-courses from a 4 mm radius sphere within cerebrospinal fluid of the anterior lateral ventricle (standard space coordinates 2, 10, 8) and white matter in the dorsal posterior frontal lobe (-26, -22, 28) were extracted from preprocessed data[21]. Images were registered using boundary-based registration[22] to the individual subject's T1-weighted structural image (1 mm^3 voxels; GRAPPA sequence TR = 20.2 ms, TE = 2.9 ms, flip angle = 9^0), which in turn was registered using FNIRT (FMRIB's nonlinear registration tool) to the MNI-152 template.

In addition to the standard motion correction; which adds six head-motion parameters as covariates at the first level for each subject, volumes that were motion outliers were included as separate regressors in the model. Excessive motion (i.e. > 4mm) was observed towards the end of a single scan session in one AWS and one CON, and these volumes (< 15% of total trials) were removed from the time series (i.e. the run was truncated). The remaining data were analysed normally. Each dataset was unwarped using a fieldmap and PRELUDE and FUGUE software running in FSL and spatially smoothed with an 8-mm full-width at half maximum smoothing kernel. A temporal high-pass filter with a cutoff of 150 seconds was used to remove low-frequency fluctuations in the signal. Two further regressors were used in the first-level analysis to remove residual image artefacts. These regressors were the mean time-courses from a 4-mm radius sphere within cerebrospinal fluid of the anterior lateral ventricle (standard space coordinates 2, 10, 8) and white matter in the dorsal posterior frontal lobe (-26, -22, 28), which were extracted from preprocessed data[21]. Images were registered using boundary-based registration[22] to the individual subject's T1-weighted structural image (1 mm^3 voxels; GRAPPA sequence TR = 20.2 ms, TE = 2.9 ms, flip angle = 9^0), which in turn was registered using FNIRT (FMRIB's

nonlinear registration tool) to the MNI-152 template.

For individual subjects, statistical maps were generated to show patterns of activation during each condition relative to baseline and between the Picture Description and the Sentence Reading conditions. Contrast masking was used at the first level to examine differences between conditions only in regions where positive activation ($Z > 0$) occurs in response to the stimuli in each condition separately. The data for the two runs in each subject were combined using a fixed-effects analysis. Group averages and contrasts between groups were analysed using FMRIB's Local Analysis of Mixed Effects stage 1[23]. Masking was used for the group comparisons to show only regions in which both groups had positive activity ($Z > 0$) for each contrast of interest overall. Because AWS and CON were well-matched on age, sex, and handedness, we did not model these variables for inclusion in whole-brain group comparisons.

Results

We predicted the picture description condition would place additional demands on linguistic and motor planning and self-initiated speech processes that could be disrupted in AWS. We expected the conditions to be otherwise well matched for articulatory efforts and overt speech execution. AWS stuttered in the scanner, on average 20 times out of 160 utterances. CON showed some dysfluencies, but many fewer (2 of 160). Stuttering frequency was equivalent across conditions. Results reported in this chapter contain all speech events.

Picture description compared to baseline

Both groups activated the expected network of areas involved in overt speech production, namely bilateral posterior

superior temporal cortex, sensorimotor cortex at about the level of the face representation, SMA and preSMA, and left lateralized posterior IFG. Extensive medial and lateral occipital cortex activity was seen in occipital cortex due to the presentation of the picture stimulus for description (Figure 3).

Figure 3. Speech-related activity during picture description in adults who stutter, and controls. The expected network of auditory cortex in both hemispheres, motor cortex, pre-supplementary motor area, basal ganglia, and frontal speech-motor regions were observed.
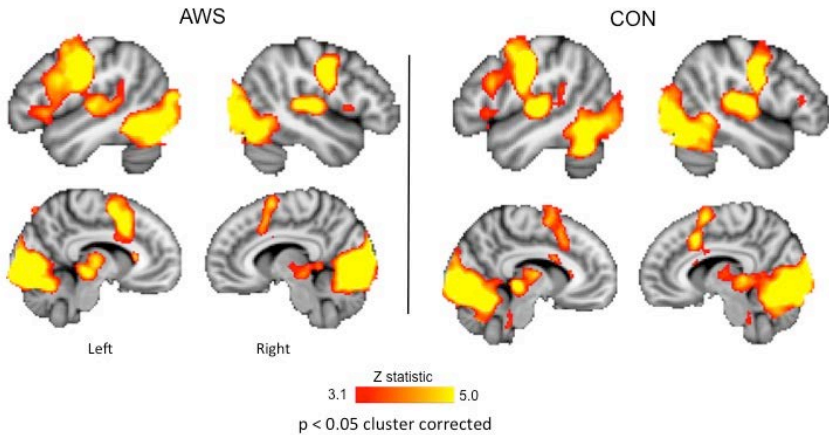
Examination of these group averages shows the AWS and CON to be strikingly similar in terms of the speech networks activated. The AWS show more extensive activation of the left IFG, lateral and medial premotor cortex relative to CON, with decreases occurring in both hemispheres while generating spontaneous speech for picture description.

Direct contrasts between the groups revealed that these cortical regions on the lateral and medial surface were significantly more active in AWS compared to CON but only at an uncorrected threshold of $p<0.01$ and extent > 30 voxels (see Table 1, Figure 4). Additionally, at the same statistical threshold, the head of the caudate nucleus and putamen in the left hemisphere were more active in AWS compared to controls (see Table 1). Furthermore,

AWS had significantly less activity than CON in the right IFG, superior temporal gyrus/sulcus, left posterior STG (including auditory cortex), and left posterior lobe of the cerebellum (Table 1, Fig 4).

Table 1: Regions where there were differences between groups in activity during Picture Description vs. baseline.

Brain Region	voxels	Z statistic	X	Y	Z
AWS > CON					
Left caudate (head)	34	2.99	-18	16	10
Left putamen	42	2.6	-20	10	-4
Left preSMA	341	3.65	-10	6	48
Left cingulate gyrus		3.33	-12	6	38
Left inferior frontal junction pars opercularis	78	2.81	-52	2	18
Left lingual gyrus	36	3.38	-28	-58	-2
Right dorso lateral occipital cortex	137	3.1	28	-68	22
Left ventrolateral occipital cortex	53	3.1	-52	-70	-20
Right occipital pole	88	3.54	42	-94	0
CON > AWS					
Right inferior frontal gyrus pars triangularis	30	2.75	52	32	8
Right superior temporal gyrus	61	3.02	64	0	-4
Left auditory cortex Heschl's Gyrus	88	2.93	-56	-12	2
Left superior temporal gyrus (ant)		2.61	-58	-10	-8
Left posterior lobe of the cerebellum	133	2.99	-46	-52	-30

Location of the highest peak in a cluster is given: **voxelwise, p < .01, uncorrected with > 30 voxel extent.** *Selected sub-peaks within the large clusters are also described. There were no areas where AWS > CON for Sentence reading. The number of voxels in a cluster is listed along with the peak height and coordinates of the peak location in MNI-152 standard space.*

Sentence Reading compared to baseline

Both groups activated the expected network of areas

involved in overt speech production, namely bilateral posterior superior temporal cortex, sensorimotor cortex at about the level of the face representation, SMA and preSMA, and left lateralized posterior IFG. Extensive medial and lateral occipital cortex activity was seen due to the presentation of the picture stimulus for description, as was the case in picture description. Examination of these group averages shows the AWS and CON to be strikingly similar in terms of the reading networks activated. The AWS show relatively less activation of the left IFG, lateral and medial premotor cortex relative to the controls.

Direct contrasts between the groups revealed that these cortical regions on the lateral and medial surface were more active in AWS compared to CON but only at an uncorrected threshold of p<0.01 and extent > 30 voxels (see Table 2, Figure 4). AWS showed relatively less activation for the sentence reading compared to CON occurring bilaterally in several speech regions including auditory and superior temporal cortex, bilaterally and in inferior frontal cortex and superior temporal cortex.

Table 2: Regions where there were differences between groups in activity during Sentence Reading vs. baseline.

Brain Region	voxels	Z statistic	X	Y	Z
CON > AWS					
Left frontal pole pars orbitalis	105	3.5	-46	26	-10
Right inferior frontal gyrus pars triangularis	33	3.01	54	34	8
Right inferior frontal gyrus pars opercularis	64	3.66	44	20	20
Left inferior frontal gyrus pars opercularis	216	3.03	-44	20	22
Left inferior frontal gyrus pars triangularis		2.74	-58	20	4
Right motor cortex	64	3.14	36	-8	40
Left superior temporal gyrus	251	3.02	-56	-10	-10
Left auditory cortex		2.8	-56	-12	2

Heschl's Gyrus					
Left temporal pole		2.73	-64	4	-6
Right insular cortex	76	3.04	40	-24	20
Right parietal operculum		2.72	46	-26	20
Right anterior vermis	69	2.6	6	-52	-18
Right posterior lobe of the cerebellum	45	3.13	16	-86	-26
Left ventrolateral occipital cortex	38	2.68	-34	-86	8
Right dorsolateral occipital cortex	56	3.41	40	-88	20

Location of the highest peak in a cluster is given: **voxelwise, p < .01, uncorrected with > 30 voxel extent.** *The number of voxels in a cluster is listed along with the peak height and coordinates of the peak location in MNI-152 standard space. Selected sub-peaks within the large clusters are also described. There were no areas where AWS > CON for Sentence reading.*

Figure 4. Group activation differences during tasks conditions in adults who stutter compared to controls. AWS show increased activity relative to CON only in the picture description condition, which has an additional spontaneous speech component. This increased activity in AWS is localized to pre-SMA, left frontal operculum, and the basal ganglia.
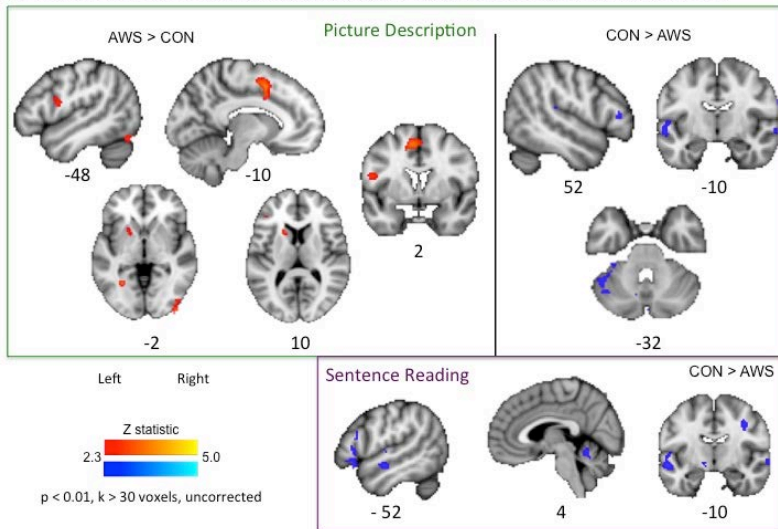
Picture Description versus Sentence Reading

Both groups showed the expected increase in motor-planning regions during the picture description condition relative to the sentence reading condition.

Major components of the classic motor-planning network showed a more pronounced change between conditions in AWS relative to CON at the uncorrected threshold: as would be expected given the patterns seen in each condition relative to silent baseline. The network included extensive portions of the inferior frontal cortex including frontal operculum, ventral premotor cortex, pre SMA, and subcortical grey matter localized in the head of the caudate nucleus and extending into putamen and thalamus on the left (Table 3).

Table 3: Regions where there were differences between groups in activity during Picture Description vs. Sentence Reading vs. baseline.

Brain Region	voxels	Z statistic	X	Y	Z
AWS > CON					
Left inferior frontal gyrus	2779	4.09	-44	36	4
pars triangularis		3.95	-42	32	6
pars opercularis		3.86	-48	10	18
ventral premotor cortex		3.83	-44	12	34
superior frontal gyrus		3.76	-24	14	56
preSMA		3.25	-4	4	52
Left caudate	297	3.35	-12	12	8
Putamen		2.5	-20	0	6
Thalamus		2.5	-12	-6	4
Left supramarginal gyrus	53	2.96	-32	-42	38
Right temporal occipital fusiform cortex	510	3.42	28	-54	-16
lingual gyrus		3.35	22	-40	-10
parahippocampal gyrus		3.03	30	-40	-12
Left temporal occipital fusiform cortex	671	4.21	-28	-58	-8
Lingual gyrus		3.09	-20	-42	-12
parahippocampal gyrus		3.21	-30	-40	-12
Precuneus	655	3.44	8	-70	54
Left occipital	230	3.46	-54	-78	12
Left occipital	37	2.71	-8	-64	72
Right occipital	292	3.64	44	-84	-4
Right occipital	33	2.74	46	-68	6
Right occipital pole	226	3.16	12	-96	8
Right occipital pole	62	2.9	34	-94	16
CON > AWS	20	3.69	-38	-20	-18
Left anterior lobe of the cerebellum	37	2.67	-32	-50	-32

Location of the highest peak in a cluster is given: **voxelwise, p < .01, uncorrected with > 30 voxel extent**. *The number of voxels in a cluster is listed along with the peak height and coordinates of the peak location in MNI-152 standard space. Selected sub-peaks within the large clusters are also described. There were no areas where AWS > CON for Sentence reading.*

Discussion

Our primary finding was of largely similar overt speech networks in adults who stutter relative to fluent controls.

We did not observe, even at lowered thresholds any clear evidence of altered cerebral dominance in stuttering. In fact, in the auditory cortex, we observed the predicted reduced activity in stuttering, bilaterally, consistent with the conclusions of the seminal meta-analysis[11]. We observed some support for theories of altered speech-motor and motor planning activity in stuttering through subtle group differences in the components of the associated networks, which were only observable at lowered thresholds. Our findings also lend support theories of basal ganglia involvement in stuttering. These structures are important for feedback processing and in particular are critical for motor learning and showed task-specific over-recruitment in AWS for internal generation of speech content relative to reading.

As the chance of our finding false positive results is quite high, we will not over speculate regarding the importance of qualitative group differences at length. The degree to which this altered activity is maladaptive, compensatory, or some combination of the two is likewise beyond the scope of this paper. However, we replicated some previous reports of functional abnormalities in the speech and motor system in stuttering. It is unclear which differences reflect general traits of the disorder and which are

related to the dysfluent state, an area begging for further exploration.

AFFILIATIONS AND FUNDING:

REFERENCES

1. Foundas, A. L., Bollich, A. M., Corey, D. M., Hurley, M. & Heilman, K. M. *Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering.* Neurology 57, 207–215 (2001).

2. Lu, C. et al. *Neural anomaly and reorganization in speakers who stutter: A short-term intervention study.* Neurology 79, 625–32 (2012).

3. Lu, C. et al. *Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers.* Cortex 46, 49–67 (2010).

4. Cykowski, M. D. et al. *Perisylvian Sulcal Morphology and Cerebral Asymmetry Patterns in Adults Who Stutter.* Cereb Cortex 18, 571–583 (2007).

5. Choo, A. L. et al. *Corpus callosum differences associated with persistent stuttering in adults.* J. Commun. Disord. 44, 470–7 (2011).

6. Sommer, M., Koch, M. A., Paulus, W., Weiller, C. & Büchel, C. *Disconnection of speech-relevant brain areas in persistent developmental stuttering.* Lance 360, 380–383 (2002).

7. Watkins, K. E., Smith, S. S. M., Davis, S. & Howell, P. *Structural and functional abnormalities of the motor system in developmental stuttering.* Brain 131, 50–9 (2008).

8. Cykowski, M. D., Fox, P. T., Ingham, R. J., Ingham, J. C. & Robin, D. A. *A study of the reproducibility and etiology of diffusion anisotropy differences in developmental stuttering: a potential role for impaired myelination.* Neuroimage 52, 1495–1504 (2010).

9. Chang, S.-E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A. & Ludlow, C. L. *Brain anatomy differences in childhood stuttering.* Neuroimage 39, 1333–1344 (2008).

10. Connally, E. L., Ward, D., Howell, P. & Watkins, K. E. *Disrupted white matter in language and motor tracts in developmental stuttering.* Brain Lang. (2013). doi:10.1016/j.bandl.2013.05.013

11. Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R. & Fox, P. T. *Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies.* Hum. Brain Mapp. 25, 105–117 (2005).

12. Neumann, K. et al. *The nature and treatment of stuttering as revealed by fMRI A within- and between-group comparison.* J Fluen. Disord 28, 381–410 (2003).

13. Maguire, G. A., Yeh, C. Y. & Ito, B. S. *Overview of the Diagnosis and Treatment of Stuttering.* J. Exp. Clin. Med. 4, 92–97 (2012).

14. Wu, J. C. et al. *A positron emission tomography [18F]deoxyglucose study of developmental stuttering.* Neuroreport 6, 501–505 (1995).

15. Lu, C. et al. *The role of large-scale neural interactions for developmental stuttering.* Neuroscience 161, 1008–26 (2009).

16. Chang, S.-E. & Zhu, D. C. *Neural network connectivity differences in children who stutter.* Brain 136, 3709–26 (2013).

17. Giraud, A.-L. L. et al. *Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering.* Brain Lang 104, 190–199 (2008).

18. Jiang, J., Lu, C., Peng, D., Zhu, C. & Howell, P. *Classification of types of stuttering symptoms based on brain activity.* PLoS One 7, e39747 (2012).

19. Riley, G. D. SSI-3 *Stuttering Severity Instrument for Children and Adults.* (Pro-Ed, 1994).

20. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. *FSL.* Neuroimage 62, 782–90 (2012).

21. Leech, R., Braga, R. & Sharp, D. J. *Echoes of the brain within the posterior cingulate cortex.* J. Neurosci. 32, 215–22 (2012).

22. Greve, D. N. & Fischl, B. *Accurate and robust brain image alignment using boundary-based registration.* Neuroimage 48, 63–72 (2009).

23. Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M. & Smith, S. M. *Multilevel linear modelling for FMRI group analysis using Bayesian inference.* Neuroimage 21, 1732–1747 (2004).

Intermediate band solar cells with AlAs-capped quantum dots

JENNA HOLDER

Jenna has recently completed her DPhil in solar energy materials, in which she worked to understand the physics of three new types if solar cells in order to aid their development. Working with solar energy was the natural result of two interests. Firstly, an interest in semiconductor technology, borne during her MPhys degree at the University of Warwick. Secondly, a passion to promote renewable energy. Jenna has actively worked with OUSU's Environment and Ethics committee to encourage more sustainable practices within the university. In her spare time, Jenna enjoys playing the piano and violin, watching films and playing squash. This article notes that intermediate band solar cells have the potential to increase solar cell efficiency. However, they have been hindered by low voltage outputs, as electrons flow out at the energy of the intermediate band, rather than the higher energy of the conduction band. Here it is shown that inserting a thin layer of aluminium arsenide (AlAs) between the intermediate band material (indium arsenide quantum dots) and the host solar cell material (gallium arsenide) can keep the voltage high and raise the solar cell efficiency. The causes for the improvement are studied, and reveal that electrical isolation of the intermediate band does not occur. Instead, the AlAs may structurally enhance the solar cell to reduce defects and electron-hole recombination.

1. Introduction

1.1 The importance of solar energy

The recent report from the intergovernmental panel on climate change states that "carbon dioxide concentrations have increased by 40% since pre-industrial times, primarily from fossil fuel emissions"[1]. The message is that the use of fossil fuels must rapidly decrease. Renewable energies, such as solar energy, are thus playing an increasingly important role in achieving a sustainable future. For solar cells to become more competitive with other electricity sources, the efficiency of converting sunlight into

electricity must increase. Intermediate band solar cells are a class of solar cell that aims to address this issue.

1.2 Increasing the current - Intermediate band solar cells

To achieve a high efficiency, a solar cell must produce both a high current output and a high voltage output.

Current is produced when electrons flow from the solar cell to the external circuit. There are two steps to current generation:

1) Light hits the solar cell. The light is absorbed if it has enough energy to raise an electron from the material's valence band to the higher-energy conduction band.

2) The high energy electron will then either:

   a) fall back to the valence band to 'recombine' with the hole it left there.

   b) contribute to electric current by being swept through the conduction band to the solar cell's electrodes and into the external circuit.

It can be seen that a limiting factor to this process is the energy gap between the conduction band and the valence band (called the 'band gap'). If this band gap is too big then a high fraction of light will have an energy that is too low to be absorbed. Intermediate band solar cells combat this problem. They are made so that there is an intermediate band of energy levels between the conduction band and the valence band. This enables lower energy light to be absorbed (Figure 1) and thus can increase the electrical current output.

The voltage output is related to the energy gained by the electrons when they absorb light. To gain as much energy as possible, electrons must be extracted from the conduction band and not the lower-energy intermediate band; in other words, current produced via the intermediate band must be obtained via a two-step process ((1) and (2) of Figure 1).
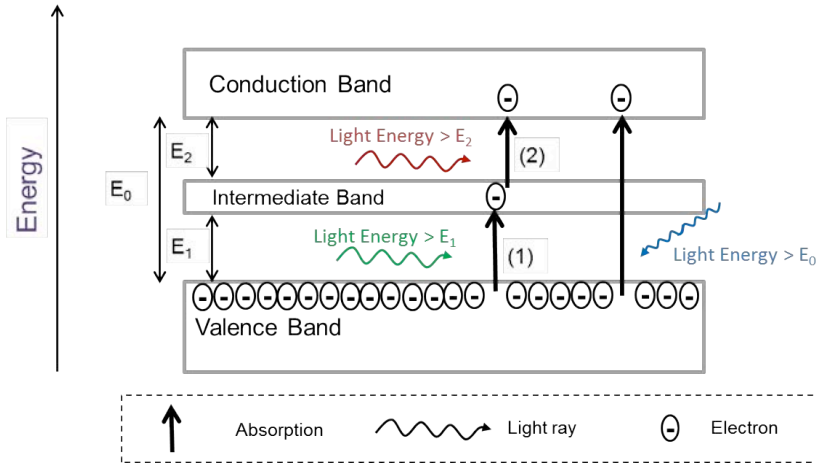

Figure 1: Energy level diagram of an intermediate band solar cell. As well as valence band - conduction band light absorption (blue light ray), light of lower energy can be used in a two-step process ((1) and (2)) to generate current.

To date, intermediate band solar cells have unfortunately had lower output voltages than standard solar cells without an intermediate band, which reduces solar cell efficiency. The low output voltages are thought to arise because energy levels form between the intermediate band and the conduction band (Figure 2A). With the assistance of heat energy[2,3] or by tunnelling[4,5] electrons can use these energy levels like a ladder. This makes it easy for electrons in the intermediate band and conduction band to mix and for the bands to become indistinct from each other. In other words, the conduction band energy is 'pinned' to the energy

of the intermediate band, which limits the output voltage of the solar cell to the difference in voltage between the valence band and intermediate.

In 2000[6], Marti et al, proposed a solution to mitigate this voltage reduction: they used layers of quantum dots to form the intermediate band. Quantum dots are materials that are smaller than a material-dependent threshold, termed the Bohr radius, which is usually on the order of nanometres. They are useful because they only have a few energy levels, unlike the continuum of levels in a conduction band or valence band. Thus, when energy levels in these quantum dots merge together to form the intermediate energy band, the band can be electronically separate from the energy bands of the host material (Figure 2B). Due to the ease of processing, most quantum dot intermediate band solar cells (QD IBSCs) are made of indium arsenide quantum dots (InAs QDs) – whose high-energy levels form the intermediate band embedded in a gallium arsenide (GaAs) solar cell – whose energy levels form the valence and conduction bands.

Despite the quantum dots, ladder levels are still present, so output voltages continue to be low. These ladder levels have been determined to be from indium gallium arsenide (InGaAs), which forms when indium from the InAs quantum dots mixes with the gallium from the GaAs. This intermixing occurs to reduce the strain that arises between the InAs and GaAs crystal lattices during solar cell fabrication[7].

1.3 Increasing the voltage – aluminium arsenide layers

The solution proposed here is to insert aluminium arsenide (AlAs) between the InAs QDs and the GaAs layers. The AlAs could restore a high output voltage via two potential means. Firstly, it has been seen to prevent the intermixing of indium and gallium

at the quantum dot edge[8,9], which would prevent the formation of InGaAs ladder levels. Secondly, it has a much wider band gap than the rest of the solar cell: AlAs has a band gap of 2.16 eV[10], whereas that of GaAs is 1.42 eV[10] and that of the InAs QDs is about 1.1 eV. The proposed band diagram is shown in Figure 2C. The wide band gap of AlAs might present an electronic barrier to electrons[4,11], isolating the intermediate band.
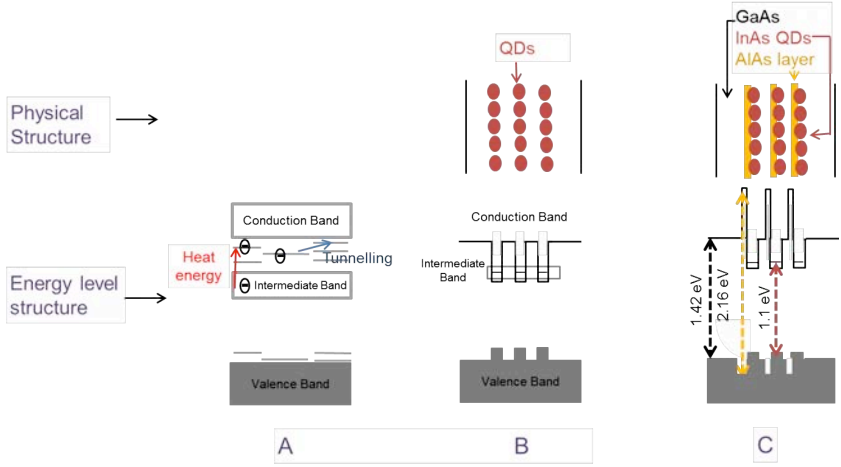

Figure 2: A) Energy level structure of intermediate band solar cell with ladder levels (grey lines), enabling electrons to escape to the conduction band with the help of heat or tunnelling. B) Physical and energy level structure of quantum dots (QDs) within a solar cell, forming an intermediate band between the conduction band and valence band. C) InAs QDs covered with wide band gap AlAs in a GaAs solar cell.

This report will study the changes induced by the AlAs layer in three steps:

1. Identify whether there are any 'ladder states' between the intermediate band and conduction band

2. Examine the mechanism of charge extraction from the intermediate band: if heat or tunnelling help the charges

from the quantum dots to escape to the GaAs bands then the energy bands are still mixed.

3. Examine the rate and means of charge recombination in the cell, as the output voltage is affected by these.

2. Devices

Two devices were constructed for this study by Mr. Frank Tutu of University College London. The first was a reference InAs QD/ GaAs device termed the 'control cell' and the second was the 'AlAs cell', which was identical to the control cell other than the inclusion of 2 monolayers (~0.57 nm) of AlAs on each quantum dot layer. The structure of the devices is illustrated in figure 3. The
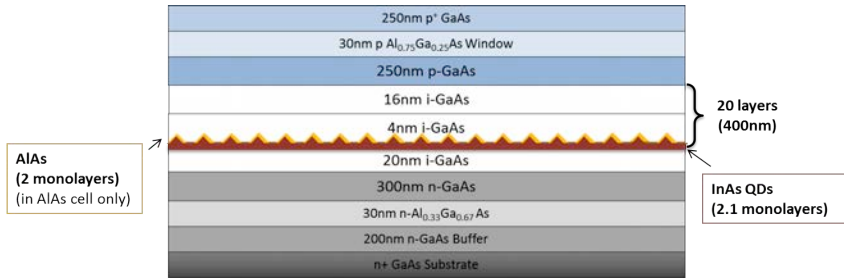
Figure 3: Device structure, giving thicknesses, doping information and the position of the AlAs layer in the case of the AlAs cell.

blue layers are p-doped (more positively charged than intrinsic gallium arsenide). The white layers are denoted with the letter 'i' for intrinsic gallium arsenide. The grey layers are n-doped (more negatively charged than intrinsic gallium arsenide). This 'p-i-n' structure is required to create an electric field across the solar cell, which drives charges towards the correct electrodes.

3. Results

3.1 Solar cell performance

Measurements of current density (J) vs voltage (V) taken under mid-latitude solar illumination provide key solar cell efficiency parameters. The results for the two solar cells are shown in Figure 4.
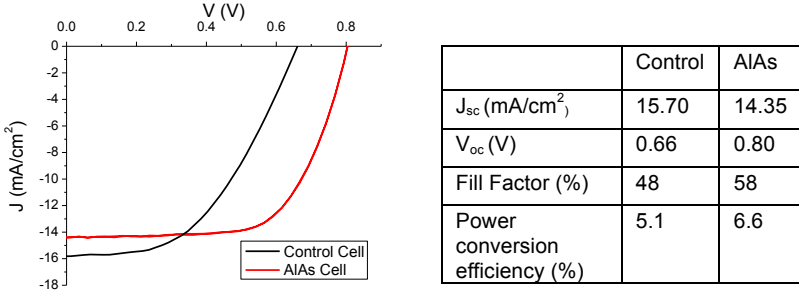

	Control	AlAs
J_{sc} (mA/cm^2)	15.70	14.35
V_{oc} (V)	0.66	0.80
Fill Factor (%)	48	58
Power conversion efficiency (%)	5.1	6.6

Figure 4: Current Density – Voltage (JV) curves at AM 1.5 illumination for the control cell and the AlAs cell. The table gives the efficiency parameters, as explained in the text.

Pertaining to the current output and voltage output are, respectively, the J_{sc} (short circuit current density) and V_{OC} (open circuit voltage). J_{sc} is the output current when there is no load connected to the solar cell and the voltage is zero. V_{OC} is the voltage across a solar cell when no current is flowing. Strikingly, V_{OC} is much higher for the AlAs cell than that of the control cell. The fill factor, which is a measure of the curve's 'squareness', is also increased, indicating more favourable resistance through the AlAs cell. These two factors result in a higher power conversion efficiency for the AlAs cell than the control cell. Unfortunately, the AlAs layer is detrimental to the J_{sc}.

3.2 Ladder Levels

Electroreflectance is used to identify the energies of electronic transitions (electrons moving from lower to higher energy levels). Electronic transitions appear as oscillations in the spectrum. Therefore, the presence of ladder levels would be identified by an oscillation at energies corresponding to the difference in valence band and ladder level energies.

Comparing the spectra of the control cell with the AlAs cell (figure 5A), the main electronic transition in both devices is at
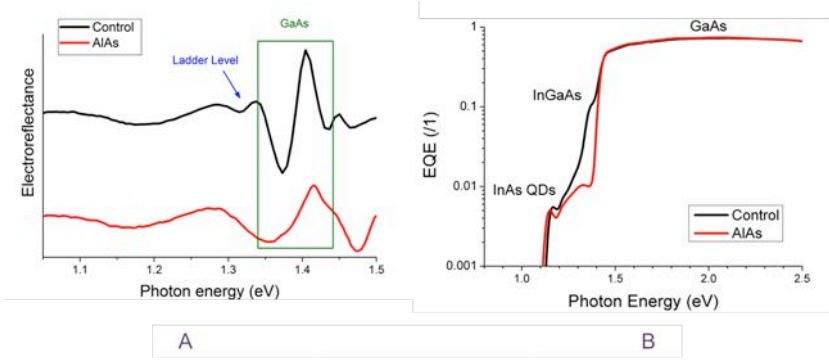
Figure 5: Electroreflectance (A) and external quantum efficiency (B) for the control cell and the AlAs cell

around 1.4 eV, which is the GaAs valence band to conduction band transition (shown in the green rectangle). The other main feature to note is at around 1.33 eV in the control cell spectrum, likely to originate from valence band to InGaAs ladder level transitions. The lack of this InGaAs feature in the AlAs cell spectrum supports the hypothesis that AlAs has reduced indium/gallium intermixing.

A negative effect of these reduced InGaAs levels is shown in the external quantum efficiency (EQE) spectra (Figure 5B),

which identify which electronic transitions contribute to the short-circuit current. In both devices, current from GaAs transitions is seen at energies greater than 1.4 eV and the valence band to InAs QD transition is seen below 1.2 eV. However, the AlAs cell outputs significantly lower current in the InGaAs ladder level region (1.2 - 1.4 eV), which explains the lower J_{sc}.

3.3 Charge extraction from the intermediate band

With fewer InGaAs ladder levels, it should be harder for electrons to get from the intermediate band to the conduction band via heat or tunnelling. Instead, secondary light would be required to excite an electron to the conduction band[12] (step (2) of Figure 1). This secondary photon must raise the electron from the intermediate band to the conduction band before it can fall back to the valence band and recombines with a hole (timescale of the order of 1ns[13]). The likelihood of this 'two photon to one electron' process occurring in this timeframe increases with greater intensity of light, so the current output from the QD intermediate band should increase superlinearly with increasing light intensity. The current output as function of incident light intensity was recorded to test this. Two light energies were used, representing different electronic transitions in the solar cell: 1.38 eV - GaAs valence band to conduction band - and 1.16 eV - valence band to InAs QD intermediate band.

Figure 6 reveals that the current increases linearly with light intensity for both electronic transitions, which is characteristic of one photon producing one electron. In other words, despite the reduction in ladder levels, the intermediate band is not isolated from the conduction band, as secondary light raise electrons from the QD intermediate band to the conduction band.
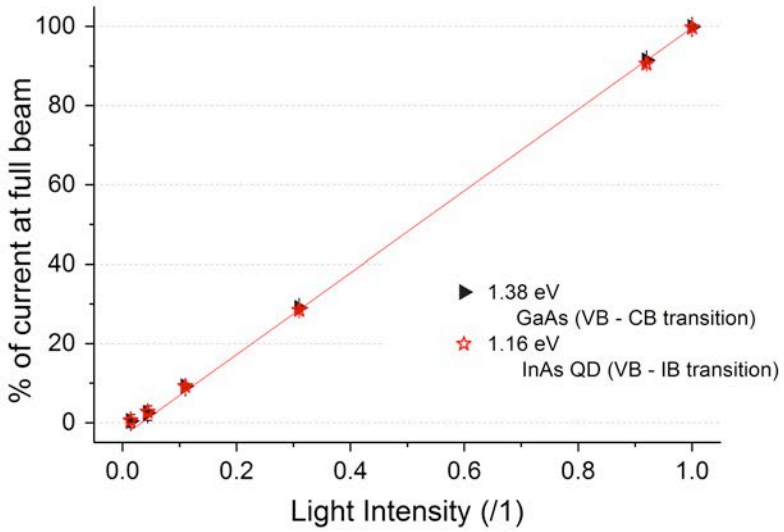
Figure 6: Current output on exposure of the cell to light of energies: 1.38 eV (exciting the valence band to conduction band transition) and 1.16 eV (exciting the valence band to intermediate band transition). Values are shown as a percentage of current output at full light intensity.

To determine whether it is heat that promotes the electrons, external quantum efficiency of the QD transitions from the AlAs cell was measured as a function of temperature (Figure 7). Current is seen from the QDs even at the lowest temperature of 102K (-171°C). There is a small rise in QD current up to 202K (-71°C) and then a drop in current at higher temperatures. Other measurements attributed this drop to an increase in charge recombination as temperatures rise past 202K (data not shown). If the charge extraction process was purely thermal, a more dramatic rise in EQE would be expected from 102K – 202K as well as a low quantum dot current at 102K[5].
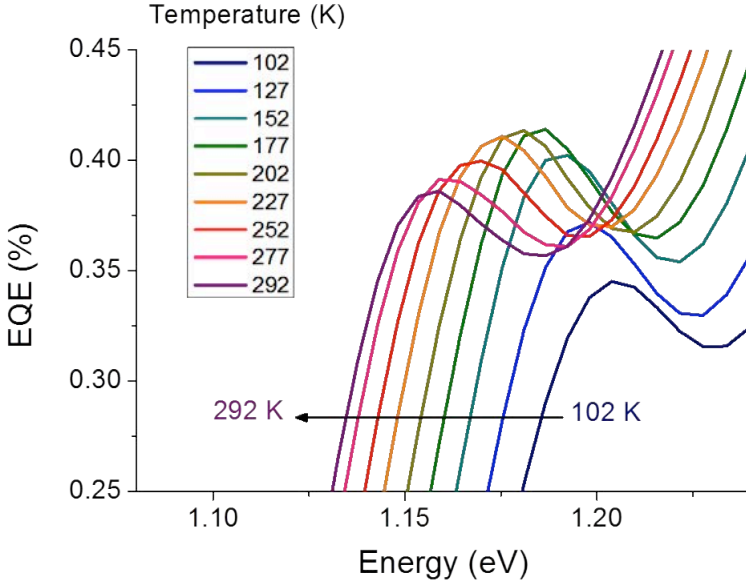
Figure 7: External quantum efficiency (EQE) around the QD excitation.

As both light and heat do not have a key role in exciting electrons from the intermediate band in the AlAs device, tunnelling must prevail. This is feasible according to other studies in the literature with thin layers (< 15 nm) between QDs[5,14]. Analysis of photoluminescence data (not shown) showed moderate tunnelling in both the AlAs and the control devices, but the AlAs cell showed a slightly higher hindrance to tunnelling movement of charges.

As the AlAs does not isolate the intermediate band from the conduction band, the investigation into the higher open circuit voltage now turns to the examination of charge recombination in the solar cell.

3.4 Charge Recombination

V_{OC} is voltage that causes charge generation to balance charge recombination[15,16] and thus cause a net current of zero.

- Charge generation: absorption of light generates free electrons in the conduction band and leaves their opposite charge ("holes") in the valence band.

- Charge recombination: at forward voltage, electrons and holes are injected into the solar cell from the electrodes. These move across the solar cell in the opposite direction to the photogenerated electrons and holes and recombine with the photogenerated holes and electrons (respectively). Therefore, these injected charges create a recombination current, acting in the opposite direction to the generation current.

A higher rate of recombination will result in a lower forward voltage being required to balance generation current, i.e. the V_{OC} will be lower. To study *just* the recombination current, current-voltage measurements were taken in the dark (Figure 8).

The AlAs cell starts injecting at a higher forward voltage than the control cell, i.e. there is initially a lower injection rate. As electrons can tunnel through the AlAs, it seems unlikely that the AlAs barrier can reduce charge injection. Alternatively, something about the AlAs cell could limit recombination. By fitting the dark JV curve to the diode equation, parameters can be extracted that quantify any recombination differences.

These parameters are the ideality factor (n) and reverse saturation current density (J_0). The higher the ideality factor, the more ways there are for electrons and holes to recombine. J_0 gives the magnitude of this recombination. These parameters are related to the V_{OC} by equation 1, where J_{ph} is the photogenerated current density, q is the charge on an electron and k_B is Boltzmann's constant.
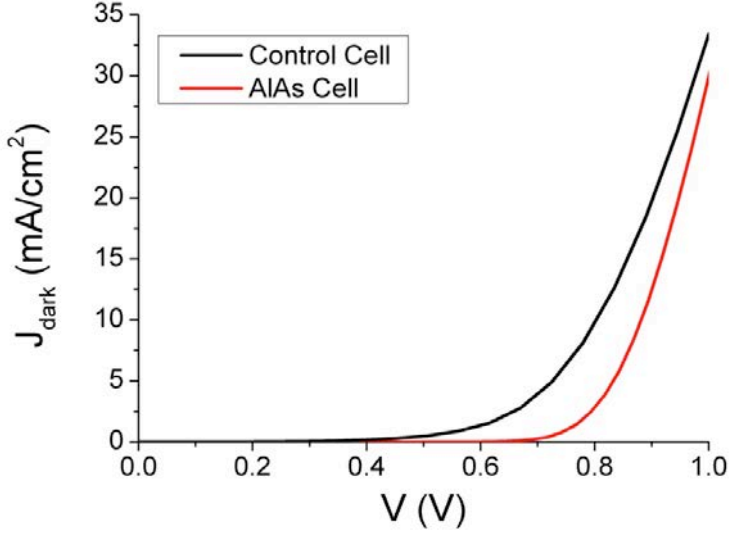
Figure 8: Dark JV curves of the control cell (black) and AlAs cell (red).

$$V_{oc} = \frac{n k_B T}{q} \ln \left(\frac{J_{ph}}{J_0} + 1 \right)$$

Equation 1

The parameters extracted from the dark JV curves are shown in Table 1. The ideality factor of 1.46 for the AlAs cell is

Cell	N	J_0 (mA/cm^2)
Control	3.74 ± 0.10	(2.4 ± 0.1) E-3
AlAs	1.46 ± 0.06	(5.3 ± 1.2) E-9

Table 1: Dark JV parameters for the AlAs and control solar cells.

comparable to other GaAs solar cells[17]. However, the much higher value for the control cell may be a sign of recombination via dislocations[17], increasing with the density of defects in the solar cell[18]. Defects are known to form when strain builds up in the device[19–21]. Supporting this theory is the six-orders of magnitude higher J_0 for the control cell compared to the AlAs cell, which is again is a possible indicator of defects[17,22,23].

The lower J_0 and n for the AlAs cell indicate decreased recombination in this device, which would lower the injection rate as seen in the dark JV curves and reduce V_{OC}. The reason for the improved recombination may be that the AlAs layer results in fewer defects within the device, or that the lack of InGaAs states help to reduce recombination.

4. Conclusions

AlAs layers were incorporated into InAs quantum dot/GaAs solar cells to try to isolate the QD intermediate energy band from the GaAs conduction and valence bands. It was hoped isolation would be achieved for two reasons: firstly, AlAs is a wide band gap material, and secondly, placing AlAs between InAs and GaAs will reduce indium-gallium intermixing. The aim of this was to increase the open-circuit voltage, and indeed it increased from 0.66 V to 0.80 V, but this was at the expense of the short-circuit current. Electroreflectance and external quantum efficiency measurements showed that the reduction in short-circuit current arose from fewer InGaAs ladder levels in the AlAs cell, so the AlAs successfully prevented the intermixing of gallium and indium.

The increase in open-circuit voltage was not due to the isolation of the intermediate band. This would have been demonstrated by light causing the transition from the QD intermediate band to the GaAs conduction band. Instead,

tunnelling is the dominant extraction mechanism, showing that the AlAs barrier does not prevent charge movement from the QDs. Dark current-voltage measurements showed that the open-circuit voltage improves in the AlAs cell because of reduced recombination, possibly due to fewer defects in the AlAs cell.

REFERENCES

1. IPCC. *IPCC Fifth Report - Mitigation of Climate Change.* (2014). at <http://www.ipcc.ch/report/ar5/wg3/>
2. Linares, P. G. *et al.* Voltage recovery in intermediate band solar cells. *Sol. Energy Mater. Sol. Cells* **98,** 240–244 (2012).
3. Luque, A. *et al.* Radiative thermal escape in intermediate band solar cells. *AIP Adv.* **1,** 022125 (2011).
4. Antolín, E. *et al.* Advances in Quantum Dot Intermediate Band Solar Cells. *IEEE* 65–70 (2010).
5. Antolín, E. *et al.* Reducing carrier escape in the InAs/GaAs quantum dot intermediate band solar cell. *J. Appl. Phys.* **108,** 064513 (2010).
6. Marti, A., Cuadra, L. & Luque, A. Quantum dot intermediate band solar cell. *Conf. Rec. Twenty-Eighth IEEE Photovolt. Spec. Conf. - 2000 (Cat. No.00CH37036)* 940–943 (2000). doi:10.1109/PVSC.2000.916039
7. Eisele, H. *et al.* Change of InAs/GaAs quantum dot shape and composition during capping. *J. Appl. Phys.* **104,** 124301 (2008).
8. Arzberger, M., Käsberger, U., Böhm, G. & Abstreiter, G. Influence of a thin AlAs cap layer on optical properties of self-assembled InAs/GaAs quantum dots. *Appl. Phys. Lett.* **75,** 3968 (1999).
9. Hu, D., McPheeters, C. C., Yu, E. T. & Schaadt, D. M. Improvement of performance of InAs quantum dot solar cell by inserting thin AlAs layers. *Nanoscale Res. Lett.* **6,** 83 (2011).
10. Adachi, S. GaAs, AlAs, and AlxGa1−xAs Material parameters for use in research and device applications. *J. Appl. Phys.* **58,** R1 (1985).
11. Song, H. Z. *et al.* AlGaAs capping effect on InAs quantum dots self-assembled on GaAs. *Phys. Lett. A* **375,** 3517–3520 (2011).
12. Cuadra, L., Marti, A. & Luque, A. Present status of intermediate band solar cell research. *Thin Solid Films* **451-452,** 593–599 (2004).
13. Dawson, P. *et al.* Relaxation and recombination in InAs quantum dots. *Phys. Status Solidi* **244,** 2803–2815 (2007).

14. Mazur, Y. I. *et al.* Interdot carrier transfer in asymmetric bilayer InAs⁄GaAs quantum dot structures. *Appl. Phys. Lett.* **86,** 063102 (2005).

15. Ruppel, W. & Wurfel, P. Upper limit for the conversion of solar energy. *IEEE Trans. Electron Devices* **27,** 877–882 (1980).

16. Luque, A., Martí, A. & Stanley, C. Understanding intermediate-band solar cells. *Nat. Photonics* **6,** 146–152 (2012).

17. Hubbard, S. M. *et al.* Effect of strain compensation on quantum dot enhanced GaAs solar cells. *Appl. Phys. Lett.* **92,** 123512 (2008).

18. Breitenstein, O., Altermatt, P. P., Ramspeck, K. & Schenk, A. The Origin of Ideality Factors n > 2 of Shunts and Surfaces in the Dark I-V Curves of Si Solar Cells. *Proc. 21st Eur. PVSEC, Dresden* 625–628 (2006).

19. Alonso-Álvarez, D. *et al.* Carrier recombination effects in strain compensated quantum dot stacks embedded in solar cells. *Appl. Phys. Lett.* **93,** 123114 (2008).

20. Tutu, F. K. *et al.* Improved performance of multilayer InAs/GaAs quantum-dot solar cells using a high-growth-temperature GaAs spacer layer. *J. Appl. Phys.* **111,** 046101 (2012).

21. Ng, J. T., Bangert, U. & Missous, M. Formation and role of defects in stacked large binary InAs/GaAs quantum dot structures. *Semicond. Sci. Technol.* **22,** 80–85 (2007).

22. Wei, G. & Forrest, S. R. Intermediate-Band Solar Cells Employing Quantum Dots Embedded in an Energy Fence Barrier. *Nano Lett.* **7,** 218–222 (2007).

23. Lu, H. F. *et al.* Temperature dependence of dark current properties of InGaAs/GaAs quantum dot solar cells. *Appl. Phys. Lett.* **98,** 183509 (2011).

On the Mathematical Representation of Stem Cells

HIMANSHU KAUL

Himanshu obtained his DPhil in Engineering Science at St Anne's College. His research focused on the development of a modeling platform that can quantify the dynamic relationship between cells and their microenvironment. As part of his research, Himanshu extended the principle of Dynamic Reciprocity as well as introduced the computational principle of Dynamic Assimilation. This article discusses the various methods that could be employed to represent stem cells mathematically. The computational ontologies thereby generated can be employed to develop mathematical models that can provide fundamental insights into the mechanics governing cellular behaviour, which for practical or conceptual reasons have, thus far, eluded observation.

Introduction

Stem cells are the less specialised, precursor cells that are capable of forming any type of more specialised cells in the body of multicellular organisms. *Stemness* can, therefore, be described as the cell's capacity to participate in this specialisation process. In other words, the richer the variety of resulting specialised cells, the more "*potent*" the stemness. Stem Cells are classified based on their source: *embryonic* stem cells are derived from the epiblast tissue of the ICM of a blastocyst[1]; *adult* stem cells are found in adult tissue and include hematopoietic (bone marrow), mesenchymal (connective as well as non-marrow tissue), and neural stem cells[2]; and *amniotic* stem cells come from amniotic fluid[3]. In terms of functionality, stem cells are either unspecialised (Embryonic Stem Cells) or slightly specialised (Mesenchymal Stem Cells, Osteoblasts, etc.). This is to say that while embryonic stem cells can form any

kind of cells (retinal, cardiac, or neural cells, for example) and thus relevant structures in the body, osteoblasts can only lead to bone formation. Beyond their ability to acquire any/a specialised phenotype, stem cells possess limitless proliferative potential[4] (i.e. they can produce a large number of similar cell types) as well as the ability to differentiate into different (specialised) cell types beyond the tissues in which they normally reside[5] ("stem cell plasticity"). This means that a fat cell (adipocyte) can be transformed into a nerve cell. The two features collectively make stem cells appealing from a therapeutic perspective, and form the cornerstone of the technique of cell therapy, which involves reinforcing the compromised tissue or cell population with relevant stem cells. This is usually achieved by delivering the cells via a catheter or syringe, though other alternatives also exist.

Stemness and the Stem Cell Niche

While the exact causes of *stemness* are not yet quite clear, two schools of thought dominate the scene. The first attributes stemness to the cell itself whereas the other claims it to be a part of the stem cell *niche* (Figure 1). Although in French, the term *niche* refers to a dog-house[6], the term *stem cell niche* refers to a dynamic environment replete with anatomical, functional and physiological cues that can promote self-renewal, reproduction, or differentiation[6,7] (the latter being the process by which a less specialised cell turns into a more specialised one). The limited functionality of adult stem cells and hematopoietic stem cells without their niche seems to support the niche concept[6], which Schofield is credited to have first[6,8] proposed in 1978[9].

This instructional relationship, whereby information acting as "cues" flow dynamically, between stem cells and their microenvironment is best captured by the principle of *dynamic*

reciprocity[13], which further extends the instructive behaviour to cells in addition to their extra-cellular matrix (ECM): "*the ECM affects the cell which in turn responds by synthetic and degradative processes causing the composition and the structure of ECM to change which in turn influences the cell and so forth.*"[13] The general inability of cells to form functional structures when cultured as monolayers or on two-dimensional substrates (with certain exceptions) also testifies to the importance of the cells' microenvironment to tissue development. The dependence of tissue micro-architecture on the ECM forming ability of cells further validates this principle[14,15]. In 2013 Kaul et al[16] extended dynamic reciprocity to include the impact of local transport processes on tissue development.
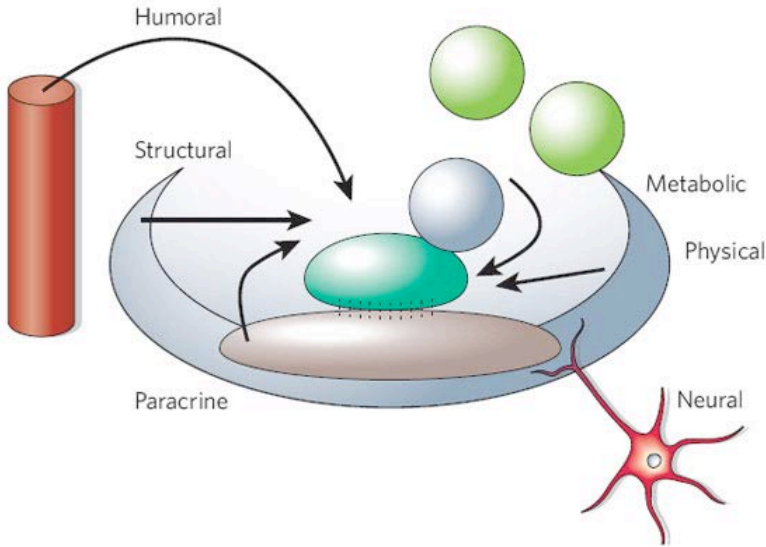

Figure 1: The stem cell niche. The figure is a schematic representation of the stem cell niche and shows the various elements in a cell's microenvironment that can influence the expression of its genotype. The figure also shows the variety of

regulatory signals a cell has to consider in the course of its lifetime. In order to capture the dynamism exhibited by the cellular system, a computational framework must be able to offer suitable ontologies to not only the dynamic structure that is the niche but also capture the broad array of regulatory cues (such as electrical, chemical, mechanical, and architectural) that a cell relies on for growth as well as instructions. *Reproduced with kind permission from Ref. 6 © (2006) Nature Publishing Group.*

Computational Modelling of Stem Cell Behaviour

The dynamic reciprocity principle provides a special perspective into stem cell dynamics. Stemness as a property is often considered an *endogenous* attribute: Either a cell has stemness, and will fulfil its role as one, or not. If that is indeed the case, then stem cells divide either:

- *asymmetrically*, producing two daughter cells of different phenotype, one stem and the other more specialised, or

- *symmetrically*, producing two daughter stem cells, thereby preserving the stemness.

Bone marrow transplantation, first performed in 1968, relies on the ability of hematopoietic stem cells to divide symmetrically and in large numbers[17] so as to reconstitute the entire hematopoietic system of the recipient. However, the stem cell *niche* theory challenges that notion and proposes that stemness is an exogenous property that is dependent on the entire microenvironment, not just the cell[6]. The failure of cells to differentiate into desired lineages *ex vivo*, due to lack of relevant chemical, mechanical, or even microenvironmental spatial cues adds credence to this theory. Accordingly, advances in material science and fabrication techniques[18-20] are being used to recreate

the architecture as well as spatiotemporal cues native to the developing tissue.

Stem cell behaviour can be simulated by *population-based*, *cell-based*, *attractor-state*, or *statistical* mathematical approaches. The cell-based approach is a form of the discrete approach, the statistical-based approach a form of the empirical approach, and the rest form the continuum method of modelling stem cell behaviour where a population of stem cells is assumed to behave in an identical manner. In assuming absence of heterogeneous behaviour within cell population these models treat cell population as a continuum. Statistical methods, such as *Bayesian networks*, are employed to mine data sets to identify principal variables impacting the overall behaviour of the population. Such methods can potentially highlight unintuitive network behaviours. However, this approach entails fitting models to experimental data that do not track individual cell responses and, therefore, do not account very well for the heterogeneity of a stem cell niche.

Population-based models rely on ordinary or partial differential equations to capture the global behaviour of stem cell populations. They involve representing stem cell behaviour with mathematical equations that may either be algebraic, or capture variability in cell behaviour temporally or spatiotemporally[21]. Both *statistical* and *population* based models usually ignore the microscopic details of the cellular microenvironment, even cellular heterogeneity, and capture average responses of a cell population[22]. As such, they exhibit large cellular variability[22] and rarely offer much beyond qualitative similarity.

The attractors approach[23], relying on differential equations, better represents the stem cell niche. In the niche context, a set of ordinary or partial differential equations describing a dynamic system will converge towards attractors as stable phenotypes.

Attractors are the equilibrium states to which dynamical systems phase space converges on[8,23]. Equations representing the system can have multiple attractors[23].

To visualise an attractor, imagine a ball rolling around in a landscape with several depressions (Figure 2): the ball represents the cell, and the depressions attractors. The ball settled in the depression indicates that based on the initial conditions, represented by the basin of the attractor[23], the equations have converged to a particular solution.
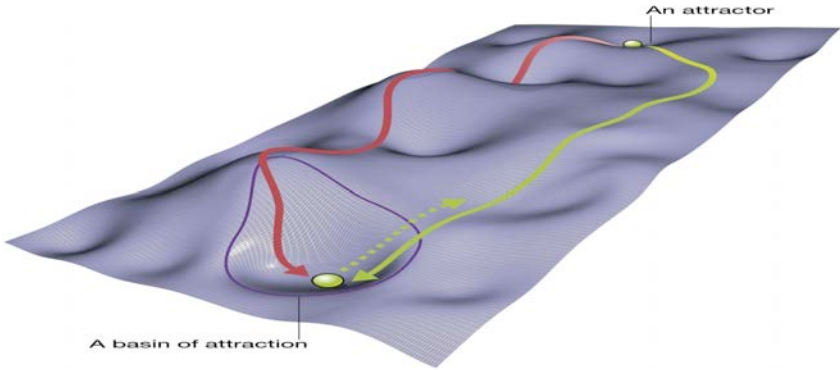
Figure 2: Landscape of Cell Phenotype. The basic idea is to project cellular phenotype or behaviour on a 2D plane, with the third dimension representing the energy of a cell displaying relevant behaviours (such as differentiation, dedifferentiation, proliferation, apoptosis, etc.), which are depicted by the depressions in the landscape. The hollows represent the most stable states, which the cells are likeliest to adopt. Mathematically, the hollows are 'attractors' signifying stable solutions to the equations capturing the dynamics of the concerned system. *Reproduced with kind permission from Ref. 23 © (2009) Elsevier.*

In the cellular context this means that the cell, based on endogenous or exogenous variables, has acquired a particular phenotype. There may have been multiple phenotypes (attractors)

that the cell could have acquired, but that is dependent upon its genotype and/or exogenous signals. Although quite promising, Complexity science, which considers dynamic systems in the light of attractors, is still developing.

Single-cell models form a part of the discrete mathematical approach that traces the evolution of a cellular system arising from a combination of cellular interactions and exogenous stimuli. Consequently, they are a better ontological choice (in other words a better form of representation) for the systems being modelled. Cellular automata and agent-based modelling are two of the most common frameworks employed to simulate tissue development based on interactions between (stem) cells and their microenvironment. The problem, however, with the single-cell based approach is that the models tend to be computationally expensive. Other approaches include Chemical equilibrium dynamics[24-26], Activation energy[27], Boolean networks[28], and Kinetic models[29,30].

While there exist a plethora of computational techniques that can be employed to model stem cell behaviour, the accuracy of model-based predictions is heavily reliant on the choice of approach as well as the system being modelled. For example, if the modeller is trying to understand the proliferative rate of stem cells, the statistical approach might seem the best alternative – these models are empirical and computationally fast. However, stem cell proliferation is not an isolated event and, based on the system under consideration, depend upon a variety of parameters: nutrient concentration *in vitro*, presence of growth factors or other relevant protein molecules *in vivo*, etc. Here, the population-based approach, which would consider, and model, the surrounding environment (perfusion rate of a nutrient medium, for example) might capture the system more accurately and provide more reliable predictions.

But what if stem cell proliferation and differentiation additionally depended upon the location of these cells? In this case, the cell-based approach will fare better in predicting system outcomes (and answer the modeller's questions) more accurately as the approach considers each cell individually in determining their response to their immediate environment.

It may appear from the discussion above that statistical models may not be very useful. This is not the case, for while the population- and cell-based approaches capture the underlying physics more accurately than their statistical counterpart, the statistical approach may be utilised if the system being modelled is highly complex (while modelling multiple scales, for example) and/or speed of computation is important. A hybrid model integrating either of the aforementioned approaches can also be employed if accuracy is of the utmost importance. Combining the continuum approach with cell-based approach, as an example, where the cell-based approach accounts for cellular behaviour and continuum approach captures the variations observed in the environment, is arguably a most suitable ontology for simulating stemness, as emerging from the stem cell niche. This hybrid approach can also be utilised if stemness is considered as an endogenous trait. The last few years have witnessed a rapid increase in the application of hybrid models towards modelling biological systems.

In conclusion, there seems to be no single right approach to model stem cell behaviour and stemness. As such, the choice of modelling method(s) depends upon (among other desirables) the complexity of the system being modelled, desired accuracy, speed of outcome, and the questions being asked of the model. However, one needs to be mindful that each of the aforementioned techniques have merits and drawbacks. Yet, each of these

techniques can yield valuable information into the dynamics of stem cell behaviour that may not be apparent from direct observation.

REFERENCES

1. Martin, G. R. ISOLATION OF A PLURIPOTENT CELL-LINE FROM EARLY MOUSE EMBRYOS CULTURED IN MEDIUM CONDITIONED BY TERATOCARCINOMA STEM-CELLS. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **78**, 7634-7638, doi:10.1073/pnas.78.12.7634 (1981).

2. Barrilleaux, B., Phinney, D. G., Prockop, D. J. & O'Connor, K. C. Review: Ex vivo engineering of living tissues with adult stem cells. *Tissue Engineering* **12**, 3007-3019, doi:10.1089/ten.2006.12.3007 (2006).

3. De Coppi, P. *et al.* Isolation of amniotic stem cell lines with potential for therapy. *Nature Biotechnology* **25**, 100-106, doi:10.1038/nbt1274 (2007).

4. Passier, R. & Mummery, C. Origin and use of embryonic and adult stem cells in differentiation and tissue repair. *Cardiovasc Res* **58**, 324-335 (2003).

5. Bongso, A. & Richards, M. History and perspective of stem cell research. *Best Pract Res Clin Obstet Gynaecol* **18**, 827-842, doi:10.1016/j.bpobgyn.2004.09.002 (2004).

6. Scadden, D. T. The stem-cell niche as an entity of action. *Nature* **441**, 1075-1079, doi:10.1038/nature04957 (2006).

7. Drummond-Barbosa, D. Stem Cells, Their Niches and the Systemic Environment: An Aging Network. *Genetics* **180**, 1787-1797, doi:10.1534/genetics.108.098244 (2008).

8. Ashton, R. S., Keung, A. J., Peltier, J. & Schaffer, D. V. Progress and Prospects for Stem Cell Engineering. *Annual Review of Chemical and Biomolecular Engineering, Vol 2* **2**, 479-502, doi:10.1146/annurev-chembioeng-061010-114105 (2011).

9. Schofield, R. RELATIONSHIP BETWEEN SPLEEN COLONY-FORMING CELL AND HEMATOPOIETIC STEM-CELL - HYPOTHESIS. *Blood Cells* **4**, 7-25 (1978).

10. Pander, C. H. *Beitra\0308ge zur Entwickelungs-Geschichte des Hu\0308hnchens im Eye. L.P.* (Wu\0308rzburg, 1817).

11. Spemann, H. Uber die Determination der ersten Organanlagen des Amphibienembryonen. *Zool. Jahr. Supp* **15**, 1-48 (1918).

12. Spemann, H. & Mangold, H. Uber induktion von Embryonalagen durch Implantation ArtfremderOrganisatoren. *Arch. Mikr. Anat. Entw. Mech.* **100**, 599-638 (1924).

13. Bissell, M. J., Hall, H. G. & Parry, G. HOW DOES THE EXTRACELLULAR-MATRIX DIRECT GENE-EXPRESSION. *Journal of Theoretical Biology* **99**, 31-68, doi:10.1016/0022-5193(82)90388-5 (1982).

14. Nelson, C. M. & Bissell, M. J. Of extracellular matrix, scaffolds, and signaling: Tissue architecture regulates development, homeostasis, and cancer. *Annual Review of Cell and Developmental Biology* **22**, 287-309, doi:10.1146/annurev.cellbio.22.010305.104315 (2006).

15. Hansen, R. K. & Bissell, M. J. Tissue architecture and breast cancer: the role of extracellular matrix and steroid hormones. *Endocrine-Related Cancer* **7**, 95-113, doi:10.1677/erc.0.0070095 (2000).

16. Kaul, H., Cui, Z. & Ventikos, Y. A multi-paradigm modeling framework to simulate dynamic reciprocity in a bioreactor. *PLoS One* **8**, e59671, doi:10.1371/journal.pone.0059671 (2013).

17. Yin, T. & Li, L. H. The stem cell niches in bone. *Journal of Clinical Investigation* **116**, 1195-1201, doi:10.1172/jci28568 (2006).

18. Fisher, O. Z., Khademhosseini, A., Langer, R. & Peppas, N. A. Bioinspired Materials for Controlling Stem Cell Fate. *Accounts of Chemical Research* **43**, 419-428, doi:10.1021/ar900226q (2010).

19. Khademhosseini, A. & Langer, R. Microengineered hydrogels for tissue engineering. *Biomaterials* **28**, 5087-5092, doi:10.1016/j.biomaterials.2007.07.021 (2007).

20. Burdick, J. A., Khademhosseini, A. & Langer, R. Fabrication of gradient hydrogels using a microfluidics/photopolymerization process. *Langmuir* **20**, 5153-5156, doi:10.1021/la049298n (2004).

21. Kaul, H. & Ventikos, Y. Investigating biocomplexity through the agent-based paradigm. *Brief Bioinform*, doi:10.1093/bib/bbt077 (2013).

22. Viswanathan, S. & Zandstra, P. W. Towards predictive models of stem cell fate. *Cytotechnology* **41**, 75-92 (2003).

23. Enver, T., Pera, M., Peterson, C. & Andrews, P. W. Stem Cell States, Fates, and the Rules of Attraction. *Cell Stem Cell* **4**, 387-397, doi:10.1016/j.stem.2009.04.011 (2009).

24. Furusawa, C. & Kaneko, K. Complex organization in multicellularity as a necessity in evolution. *Artificial Life* **6**, 265-281, doi:10.1162/106454600300103638 (2000).

25. Furusawa, C. & Kaneko, K. Origin of complexity in multicellular organisms. *Physical Review Letters* **84**, 6130-6133, doi:10.1103/PhysRevLett.84.6130 (2000).

26. Furusawa, C. & Kaneko, K. Theory of robustness of irreversible differentiation in a stem cell system: Chaos hypothesis. *Journal of Theoretical Biology* **209**, 395-416, doi:10.1006/jtbi.2001.2264 (2001).

27. Waddingt.Ch & Robertso.E. SELECTION FOR DEVELOPMENTAL CANALISATION. *Genetical Research* **7**, 303-& (1966).

28. Preisler, H. & Kauffman, S. A Proposal regarding the mechanism which underlies lineage choice during hematopoietic differentiation. *Leukemia Research* **23**, 685-694 (1999).

29. Peng, C. A., Koller, M. R. & Palsson, B. O. Unilineage model of hematopoiesis predicts self-renewal of stem and progenitor cells based on ex vivo growth data. *Biotechnology and Bioengineering* **52**, 24-33, doi:10.1002/(sici)1097-0290(19961005)52:1<24::aid-bit3>3.0.co;2-0 (1996).

30. Audet, J., Miller, C. L., Eaves, C. J. & Piret, J. M. Common and distinct features of cytokine effects on hematopoietic stem and progenitor cells revealed by dose-response surface analysis. *Biotechnology and Bioengineering* **80**, 393-404, doi:10.1002/bit.10399 (2002).

Are Tuition Fees Inevitable in Expanding Higher Education Systems?

An assessment of tuition fee regimes in the USA, the UK, the Nordic counties, and Scotland

KRISTEN GARTSIDE

Kristen is a Master's student of Higher Education at St Anne's College. With a background in psychology, she is interested in the epistemological development of study abroad opportunities. This interest has also led her to explore global trends in fee structures and how variations between overseas and local student tuition fees factor into the dramatic increase in or an introduction of tuition fee structures seen in some countries, while others like Scotland have remained fee-free. This article discusses the rationale and conditions behind increases in tuition fees and the measures the UK and the USA have taken to collect fees. It contrasts this analysis with the political, economic, and societal national contexts of Nordic countries and Scotland that have allowed them to cope without the introduction of tuition fees. From there, the article weighs the pros and cons of various loan schemes and speculates which is most fair to students, universities, and national governments for further policy presentation.

Major leading nations in the economic world have placed a strong emphasis on widening access and increasing participation in higher education. Therefore, they have contributed to a worldwide phenomenon of mass higher education. As the system of higher educating expands, however, practical issues of financing have become problematic. The United States, at one end of the spectrum, has consistently relied heavily on its students to bear the burden of their own tuition costs. On the other side of the spectrum, some countries continue to provide free higher education despite a dramatic increase in participation rates in recent years. Which countries have the right idea? In a world of mass higher education, is it inevitable that the student has to pay or are there other options? Looking at the shifts in higher education

sectors and their funding systems in the US, Nordic countries, Scotland, and the UK will help to clarify whether student fees are inevitable and which methods of charging students are most efficient.

Reasons Behind the Increase in Student Tuition Fees

The seemingly global change in favour of students paying for higher education is a response to many factors that have affected most industrialized nations in recent years. Firstly, higher education participation rates have climbed rapidly. Unprecedented growth in higher education has caused financial crises for national governments, as higher education competes for government funding with other worthwhile causes such as healthcare, primary education and etc. (Marcucci and Johnston 2007: 27). Secondly, most governments believe that having students pay tuition is fair. According to them, the private returns to students are higher than the social returns (Mayhew 2013). Research has shown that, in fact, college-educated individuals benefit economically from their education while higher education rates are not necessarily linked directly to a country's economic growth (Mayhew 2013). Thirdly, the introduction of tuition fees will, in theory, force universities to become more responsive to their students' needs (Marcucci and Johnston 2007: 27). Finally, it has been found that free higher education, despite being free, is used much more by middle and high socio-economic status students while the taxes that make that free education possible potentially burden low socio-economic families (Marcucci and Johnston 2007: 27). Such reasoning has encouraged the raising of student tuition fees in the UK, Australia, Germany, China, and many countries in Latin America and East Asia (Johnston 2004). "In fact, since 1995 more than half of the 25 OECD countries with available data on higher education have

overhauled their college tuition policies at public institutions, with many adding or raising fees" (Phillips 2013).

Tuition Fees in the United States and the United Kingdom

Looking exclusively at the USA and the UK, the transition towards mass higher education has led to students being asked to cover more of their university tuition costs. In the USA, it is understood that students will bear the burden of their education. High interest loans taken out to pay significantly high up-front tuition fees coupled with the requirement to begin repayment immediately causes student debt right after graduation. The UK's higher education funding system is easier on students than in the United States due to lower fees and loan schemes that are income-contingent with relatively lower interest rates. The UK's income-contingent deferred loan scheme has vast benefits for students as opposed to the mortgage loan system that is used primarily in the USA. The fact that repayments do not begin until the student is making a salary of 21,000 pounds is fairer to the student because a return on education investment is necessary before repayment begins. This makes higher education less of a gamble.

The UK, which used to provide free higher education, is one of many examples of countries submitting to economic pressures and introducing tuition fees. However, its loan schemes seem far superior to those of the USA. With countries that still offer free higher education in existence, however, the question remains as to whether or not student fees really are inevitable. Will countries providing free higher education eventually introduce fees or have they found a way to sustain free higher education that the rest of the world is missing out on entirely?

Countries Where Tuition Fees Do Not Exist

Denmark, Sweden, Norway, and Finland have succeeded in providing free higher education to their students despite economic pressures (Aamodt and Kyvik 2005: 123). In large part, these countries have been able to sustain free higher education through their record high taxation rates (OECD 2007). However, there have been some changes in the Nordic higher education funding systems that have caused debate. Denmark, Sweden, and Finland have all introduced tuition fees at their universities for international students (Grove 2013). The recent changes have caused much speculation about the future directions of higher education funding in Nordic countries now that they have broken their long-standing tradition of free higher education for all (Grove 2013). This development could be interpreted as a step towards the use of fees or as an effort by Denmark, Sweden, and Finland to protect themselves from an influx of international students exhausting the countries' resources. However, even for their own students, the Nordic higher education systems are also far from perfect. Despite the lack of tuition fees, students still graduate with hefty debts (Phillips 2013). Due to extremely high living costs and lack of student living cost support from the government and parents, students must take out loans anyway (Phillips 2013). Thus, even though higher education is free in these countries, a hidden price tag is still attached and students come out with debts comparable to the debts of some students in the US (Phillips 2013).

Scotland, as part of the UK, has also retained free tuition for their students and other EU students while slapping tuition fees onto their English neighbours and non-EU students. A portion of Scottish universities' funding comes from international fees, postgraduate tuition fees, lodging and catering (Bryce et al. 2013:

166). As in the Nordic system of higher education, while universities are free, there exist hidden student costs. Scottish universities and politicians are extremely dedicated to retaining their system of free higher education with almost all political parties campaigning against England's higher education reforms of 2012 (Bryce et al. 2013: 165). However, the economic climate might not allow these politicians to keep their word. "If there is no improvement in economic prospects, it is doubtful if the present position can be maintained without risk of Scottish institutions finding themselves at a serious disadvantage" (Bryce et al. 2013: 96). There are however, alternatives to re-introducing tuition fees in a poor economic climate (Finlay 2013). Surveys have shown that citizens would prefer higher taxation to the introduction of tuition fees (Finlay 2013). For Scotland, tuition fee policies will be tested very soon by how the country adapts to changing economic situations with UK-wide cuts in teaching funding and Scotland's potential independence from the UK (Finlay 2013). Scotland, Finland, Norway, Denmark, and Sweden are a few of the countries in the modern world that manage free higher education for their students, showing that perhaps student tuition fees are not inevitable after all. These countries are unique because they are capable of providing free education while in others it would be infeasible. This is largely due to their relatively small sizes, egalitarian attitudes towards socio-economic attainment, smaller wealth gaps among citizens, high rates of higher education participation, and commitment to the public sector. Because these countries have smaller populations relative to the USA or England, they are able to reach decisions with more ease and also have fewer students to fund. They are also able to increase taxation to fund higher education without causing massive revolts because such policy measures are more likely to help a majority of the

population in countries with small wealth gaps. Denmark and Norway are cited as enjoying the smallest wealth gaps (Deen 2011). In other words, introducing taxes in these countries will not necessarily impose on the poor while only benefitting the rich since citizens are more likely to have equal financial situations and there are higher rates of participation in higher education in the first place. If free higher education were to be introduced in the USA on the other hand, it would be completely impractical due to the sheer size of the nation, the massive wealth gap, and the general capitalistic ideology of the nation. So perhaps student fees in higher education are only inevitable in certain types of atmospheres and nations. It is clear, however, that even free higher education systems often depend on funding sources that can indirectly make the citizens pay through taxation, living costs, etc.

Fees Bound to National Contexts

The inevitability of student fees in the mass higher education system is something that applies to large industrialized and diverse nations in the modern world. This description, however, could be used for a majority of countries in the world today, leading to a prevailing thought that tuition fees are unavoidable. In the USA, the UK, and many other countries, this certainly seems to be the case. However, in specific nations and economic scenarios, students may not necessarily have to pay for higher education. That being said, however, many countries that offer free higher education have hidden costs and enroll international students with overseas fees in an attempt to sustain themselves and survive in global competition. If economic conditions continue to worsen, especially for Scotland which faces difficulties ahead, it would not be surprising for fees to be introduced, but that is not seen as an upcoming issue in the Nordic

countries. One contributing factor towards the Nordic countries' resistance of fees is their attitude towards education. In these countries, education is seen as a public right and a social necessity and if fees were suddenly imposed, they would be completely incongruent with the values of the nation. Education is thought of very highly in the Nordic countries. In fact, in Finland even teachers of primary school must have a Master's degree in order to be considered as qualified to teach (Childs and Mender 2013: 113). For countries that charge students tuition fees, it seems clear that among the existing options, deferred income-contingent loans are a far superior option to mortgage loans for charging students tuition, though perhaps they might need to be altered in the future to be remain economically feasible for governments in the long-term.

How Should Students be Charged?

With all of this in mind about the counties that succeed in achieving fee-free higher education, it is clear that a vast majority of modern industrialized nations are increasing and/or introducing fees. Most of these governments simply cannot afford to finance higher education with so many other competing demands for funding and the number of HE students increasing so rapidly. How, then, should students be charged?

In order to maintain equal access and keep up participation, means-tested grants and academic scholarships are necessary to attract low socio-economic status students into higher education despite its heavy price tag. Both the USA and the UK have such grants in place. Beyond simple grants and scholarships, however, the best system for students to pay their tuition fees is in deferred income-contingent loans. These loans are another way to keep higher education open to people of low socio-economic status and create a system where the economic returns that the

students receive are directly related to how much they pay for their education. This is a fair arrangement for students and governments. After all, universities should be graduating students of the best academic quality who can contribute in substantial ways to the nation through their graduate jobs.

While this system of deferred income contingent loans is ideal for students, however, it has the potential to become an economic burden to the government if a majority of students avoid repayment of their loans. There are ways that governments can avoid this type of repayment avoidance by lowering the minimum income at which students begin repayment of their loans. In the UK specifically, a graduate must be earning £21,000 annually before repayment begins (Browne 2010). By lowering this standard of minimum annual income for repayment even slightly while also maintaining a low rate of repayment, the government would be able to get some of its money back more promptly. Also, the UK has a scheme currently wherein any debts still owed after thirty years would be written off (Browne 2010). Eliminating this would undoubtedly be upsetting for graduates who would have to pay off their debts in the long term, but would help the government collect its money while still holding true to the notion of income-contingent loans and fairness to students.

The system of direct, straightforward deferred income-contingent loans that have been established in the UK and many other European countries is vastly superior to the mortgage loan system in the USA that paralyzes students with debt in their early 20's and unrealistically expects immediate repayment. It is also superior to covert methods of raising student fees through actions such as increasing living and catering fees, making grants less available, or having a dual track tuition scheme. Income-contingent

loans are transparent and understandable, easily repaid as part of income tax, and fair to both students and universities.

Another possible manner in which governments might increase the long-term feasibility of income-contingent loans would be to create the options for students to pay tuition fees in a variety of ways. For example, Australia has a tuition payment scheme that allows for students to either opt for the deferred income-contingent loans or pay the tuition upfront at a slightly discounted rate (Marcucci and Johnston 2007: 31). This may be the most efficient system because it provides the most options and allows the government to get access to some funds immediately from those students who opt for upfront payment. It would allow students from low socio-economic status to be included in higher education through the income-contingent loan scheme and encourage better and more prompt funding flow to the government from higher socio-economic status students who can perhaps afford upfront tuition fees.

Conclusion

As an interest in higher education enrollment sweeps the globe, national tuition fee schemes have been affected. While some national contexts such as the environment in the Nordic countries and Scotland have been resistant to international trends towards introduction of, and increases in, tuition fees, it has been possible largely due to their sizes, high rates of taxation, egalitarian ideals, and high national value of education. In larger countries with wider wealth gaps, capitalistic ideals, and mass student enrolments, free higher education quickly becomes economically infeasible. This is certainly true of the USA and has become true of the UK in recent years, the latter having slowly succumbed to the idea of high tuition fees. The UK has, however, in the introduction of high

fees, attempted to institute a fairer system of payment in a bid to avoid disadvantaging lower socio-economic status students and create a sustainable model for education delivery for both the nation and its students. At the same time, there will likely be more reforms in this system, as it currently slightly disadvantages the government in collecting back the finances necessary to sustain the higher education sector. Overall, however, the income-contingent loan scheme seems the best option for students and can also be beneficial to national governments with relatively easy adjustments to specific contexts.

REFERENCES

Aamodt, P.O. and Kyvik, S. (2005) Access to Higher Education in the Nordic Countries. In: Tapper, T. and Palfreyman, D. (eds.). *Understanding Mass Higher Education: Comparative Perspectives on Access.* London: RoutledgeFalmer, 121-138.

Browne, J. (2010) *Securing a Sustainable Future for Higher Education: An Independent Review of Higher Education Funding & Student Finance.* [Online] Available from:
https://www.gov.uk/government/uploads/system/uploads/attachment _data/file/31999/10-1208-securing-sustainable-higher-education-browne-report.pdf [Accessed 13 January 2014].

Bryce, T.G.K., et al. (eds.) (2013) *Scottish Education: Fourth Edition: Referendum.* Edinburgh: Edinburgh University Press.

Childs, A. and Mender, I. (2013) Teacher Education in 21st Century England. A Case Study in Neoliberal Public Policy. *Revista Española de Educación Comparada,* **22** (1), 93-116.

Deen, Mark (2011) *Wealth Gap Widens in Developed Nations: OECD.* [Online] Available from: http://www.bloomberg.com/news/2011-12-05/rich-poor-divide-is-widening-oecd-says.html [Accessed 16 November 2013]

Finlay, I. (2013) *Yes there is an alternative: Reflections on the Scottish Higher Education System.* [lecture]. Held on 19 November, Oxford University.

Grove, J. (2011) Last of the Free: will Norway's universal no-fee policy endure? [Online]. *The Times Higher Education.* Available from: http://www.timeshighereducation.co.uk/417752.article [Accessed 16 November 2013]

Johnston, B. (2004) The Economics and Politics of Cost Sharing in Higher Education: Comparative Perspectives. *Economics of Education Review*, **23** (4), 403-410.

Marcucci P.N. and Johnstons B. (2007) Tuition Fee Policies in a Comparative Perspective: Theoretical and Political Rationales. *Journal of Higher Education Policy and Management*, **29** (1), 25-40.

Mayhew, K. (2013) *Who Should Pay for Higher Education?* [lecture]. Held on 24 October, Oxford University.

OECD (2007) Denmark, Sweden still the highest-tax OECD. [Online] Available from:
http://www.oecd.org/general/denmarkswedenstillthehighest-taxoecdcountries.htm [Accessed 16 November 2013]

Phillips, M. (2013) The High Price of Free College in Sweden. [Online] *The Atlantic*. Available from:
http://www.theatlantic.com/international/archive/2013/05/the-high-price-of-a-free-college-education-in-sweden/276428/ [Accessed 16 November 2013]

Andro Linklater's *Why Spencer Perceval had to die: The Assassination of a British Prime Minister*

EDWARD HICKS

Edward is in the first year of his DPhil in history at St Anne's College, Oxford. His DPhil focuses on the life of Spencer Perceval, Britain's only assassinated Prime Minister, an often overlooked politician yet leader during the war against Napoleon and whose religious beliefs profoundly shaped his politics and policies. This review of the most recent publication dealing with Perceval, written by the prolific and well-respected author, the late Andro Linklater, considers the controversial notion that there was more to Perceval's death than just an insane bankrupt merchant.

On the 11th May 1812 the British Prime Minister Spencer Perceval walked into the lobby of the House of Commons. His path was blocked; a shot rang out, Perceval staggered and fell. Hurried into an adjoining room, he was pronounced dead. He passed into history as Britain's only assassinated Prime Minister. Two American presidents, slain in Ford's Theatre and in Dallas, have passed into immortality. Two others, Garfield and McKinley, descended into obscurity. Perceval served roughly the same length of time as Premier as Kennedy was President, and oversaw a war equal or greater in scope and importance than the American Civil War, against Napoleonic France. Yet he too has receded into the historical shadows, the preserve of pub quiz specialists and niche British historians. His assassin, John Bellingham, commands in Britain no notoriety equivalent to Lee Harvey Oswald. A merchant imprisoned in Russia for debt, Bellingham had unsuccessfully tried to claim compensation from the British Government for its inadequate support of his cause. He decided to assassinate Perceval

as the only way to secure 'justice'. So he did, and was hanged a week later for the murder.

Originally published in the bicentenary year of Perceval's assassination, this new book by Andro Linklater re-examines the circumstances surrounding the murder. As my dissertation topic is on Spencer Perceval and books about him are rare I was intrigued. What new insights would Linklater offer about the assassination? What new light would be shone on Perceval's life? I will begin by looking at Linklater's treatment of Perceval's assassination, before turning to how Linklater interprets Perceval himself.

Linklater comes to a startlingly conclusion about the assassination. His conspiracy theory is the justification for the book. He argues that Bellingham was the unwitting instrument of Liverpool merchants, exasperated at how Perceval's tough action against the slave trade (through commercially restrictive measures called Orders-in-Council) had both ruined their illegal profits and destroyed trade with America, thereby bringing Britain and America to the brink of war. They provided Bellingham with the money necessary to support him while living in London in the months building up to the assassination.

This conclusion seems based around first adopting a 'follow the money' approach and then asking *quo bono*? As with all conspiracy theories it picks out inconsistencies: the different temperament of Bellingham in 1812 from 1810 (when he first tried to secure compensation); the contradiction between his high spending in London and his impoverished state in Liverpool; that Bellingham was half-an-hour late in going to Parliament to assassinate Perceval; Liverpool MP Isaac Gascoyne's (a pro-slave trade MP) contradictory statements about his actions during and after Perceval's murder, among others. Then Linklater relies on supposition as to who this mystery backer might be: American

metal merchant Elisha Peck. Peck wished to avoid economically ruinous war between Britain and America (Linklater: 207, 209-10), so he allegedly gave Bellingham the necessary financial support and cajoled him into going ahead with the assassination.

Linklater raises interesting questions about the anomalies that any shocking event generates due to the sheer complexity of life and the fallibility of recollection. He offers a useful insight into life in Liverpool in the early nineteenth century. However I am unconvinced. He relies more on conjectures rather than evidence for his main thesis of a conspiracy. Probability may be fine in mathematics, less so in history.

Let us now turn from the murder to the victim. Perceval was forty-nine years old in 1812. The second son of the second wife of the Earl of Egmont he had had to make his own way in the world. His rise initially came through his legal career where he toiled away after leaving Trinity College Cambridge in 1783 (Thorne: 764). His financial prospects were so poor that Perceval's father-in-law had refused his suit, and his wife Jane had to elope in order to tie the knot. Their first home was above a carpet shop, not the abode normally associated with the son of a peer of the realm. Thereafter Perceval's family and career prospered together – by 1812 he had twelve children (six of each). After serving in the two main legal offices of Solicitor-General and Attorney-General, in 1807 Perceval became Chancellor of the Exchequer under the figurehead premiership of the Duke of Portland. Two years later, in October 1809, he ascended to be Prime Minister, praised by George III as "the most straightforward man he had almost ever known" (Thorne: 769).

His premiership was eventful. The previous government had collapsed with a duel on Putney Heath between two cabinet ministers. An inquiry into the disastrous expedition to Walcheren

threatened to immediately bring down Perceval's ministry. Thereafter a series of political storms battered his administration: the war against Napoleon, particularly that being fought in the Iberian Peninsula; the connected need to provide bullion to pay for Wellington's troops; an attempt to return Britain to the gold standard which Perceval had to defeat, and above all, the madness of George III and his replacement by the seemingly hostile Prince Regent. Yet Perceval survived all of these challenges. The current British Foreign Secretary William Hague has opined that "Had he lived, Spencer Perceval, would probably have become one of the longest serving Prime Ministers in British history" (Hague: 396).

Linklater's depicts Perceval as a pious, obstinately determined and ruthless politician whose single-mindedness, in the best classical tradition, is destroyed by the similar single-mindedness of his nemesis Bellingham. There are commendable aspects to Linklater's portrayal. He is right to stress the importance of Perceval's religious beliefs on his politics and policies, a point previous historiography has rarely addressed – for example the last biography of Perceval, in 1963, devoted one standalone chapter to the subject (Gray: 15-27.). Linklater is right to emphasise Perceval's role in tackling the slave trade and contributing to causing the Anglo-American war of 1812. However he tends to assert rather than provide solid evidence for such claims, rendering them interesting suggestions for investigation but not historically watertight interpretation. This is exacerbated by Linklater's inadequate referencing and mild predilection for relying on the accounts of radicals, opponents of the conservative Perceval, such as Henry Hunt and William Cobbett. I also disagree with the description of Perceval as ruthless. During the collapse of the Portland ministry that preceded Perceval's own ministry, Perceval had offered to serve under a third man as a compromise between

his claims to pre-eminence and those of the Foreign Secretary George Canning. Indeed Perceval's success in obtaining the premiership lay in avoiding the blatant ambition and unedifying arrogance that harmed his rival Canning (Throne: 769). Linklater skips over this incident without exploring this apparent counterpoint to his assertion.

It would be churlish to pass too harsh a judgement on this work. The conspiracy theory is interesting if unpersuasive. The assertions about Perceval are suggestive if occasionally unsubstantiated, exaggerated or erroneous. It is heartening that Mr. Linklater's book will spread knowledge of Perceval and his tragic end to a wider audience. Greater thoroughness and less sensationalism would however have made this a better work of history.

REFERENCES

Gillen, M. *Assassination of the Prime Minister; the shocking death of Spencer Perceval.* London: Sidgwick & Jackson Press, 1972.

Gray, D. *Spencer Perceval: The Evangelical Prime Minister.* Manchester: Manchester University Press, 1963.

Hague, W. *William Wilberforce*. London: HarperPress, 2007.

Hanrahan, D. *The assassination of the Prime Minister: John Bellingham and the murder of Spencer Perceval*. Stroud: Sutton Press, 2008.

Pentland, G. "'Now the great Man in the Parliament House is dead, we shall have a big Loaf!' Responses to the Assassination of Spencer Perceval". *The Journal of British Studies* (2012), 51.2, 340-363.

Thorne, R. "Perceval, Hon. Spencer", in R. Thorne (ed.), *The History of Parliament: The House of Commons 1790-1820*. 5 vols., London: Secker & Warburg Press, 1986), iv. [also accessible for free online on the *History of Parliament* website].

A Critical Analysis of Media Representations of Terrorism: A Case Study of the Madrid 2004 Train Bombings

LAURA PUGES

Laura completed an MSc in Criminology and Criminal Justice at Oxford in 2013, after finishing an undergraduate degree in Psychology and Organisational Behaviour at Lancaster University. She currently has a university research internship in New Zealand analysing methods of risk assessment and risk management of adolescent sex offenders with intellectual disabilities. Originally from Spain, she enjoys expanding her knowledge on crime-related issues and travelling around the world, and her efforts to combine both activities have taken her to the United Kingdom, Italy and New Zealand. Her aim, however, is to eventually find a job in Europe that allows her to combine her skills and interest in Psychology and Criminology, and to improve and develop them as well. This article, drawn from her Masters thesis, highlights the complex relationship between terrorism and the media in the Western World. In an attempt to report on terrorist acts the media risks inadvertently aiding terrorists in spreading alarm across a population, which can have various social and political consequences, particularly with regard to how governments respond to events. It then presents a case study of two Spanish newspapers' reports on the 11-M Madrid train bombings, which, despite some peculiarities of the Spanish press, supports the given analysis of the relationship in question.

Setting the Scene

If the darkness that arrives after absence was measured, the Cathedral would have been a pavilion for the blind, […] again that grandmother wiping away her tears with a handkerchief. If hugs were needed, there were arms coming from everywhere as if to stop a train, all trains, those trains. (*El Mundo 14*)

This quote is a translated excerpt from an article published by the Spanish newspaper *El Mundo* on the State Funeral for the victims of the 2004 Madrid Train Bombings. In the days following the attacks, many similarly sensationalised descriptions were used

to report on the consequences of the attacks by the newspapers *El Mundo* and *El País*. Undoubtedly, terrorist events will be tragic and any description of them will carry some dramatic elements. However, as many critics and academics have noted (Cooper 1976; Nacos 2007, among others), sensationalising events in such ways may be inappropriate, unethical and dangerous.

Terrorists aim to *communicate* a message to the world – to make their intended public know that they exist and that they can harm them. The media, in *communicating* with the public on terrorism in dramatized ways, can inadvertently help fulfil this goal (Wilkinson 1997; Weimann and Brosius 1991: 333-335). Wilkinson (1997) describes four main aims of terrorists; to instil fear among the wider population; make known their reasons and cause; to disrupt security agencies' activities; and to recruit committed newcomers to their cause (p.56). Indeed, as Nacos (2007) observes, we live in an era of 'mass-mediated terrorism', where the media often regretfully plays into the hands of terrorists by helping them meet their four aims, albeit unintentionally. It is fitting then, that the relationship between terrorism and the media has been described as 'symbiotic' (Wilkinson 1997: 52; Cooper 1976: 231).

Importantly, the landslide of reporting on terrorism that occurs after certain attacks is not always proportional to the actual threat. For example, in the one year and fifty days following 9-11, the *Washington Post* printed over 10,700 articles regarding terrorism (Rothe and Bower 2002, discussed in Rothe and Muzzatti 2004: 338). However, a database of world-wide terrorist incidents held by the US State Department clarifies that fewer attacks occurred between 2002 and 2003 than during the previous twenty years (Lewis 2012). Nonetheless, "the notion that the attacks on 11 September 2001 signalled the start of a new era of terrorist attacks became conventional in the blink of an eye, bypassing any sober

analysis" (*ibid*: 259). The thought is that events of terrorism have come to be over-reported, and so the public will consider them to be of the highest urgency even when this evaluation is inaccurate (Nacos 2007; Lewis 2012).

Over-reporting occurs because media workers want to report on events that are newsworthy and interesting to the public, thereby growing their audience. Reporters simply cannot report on everything that happens and must select what becomes 'news' among a wide range of options (MacDougall 1968, discussed in Hall *et al* 2010: 240). As Chibnall (2010, *originally 1977*) discovered in observing and interviewing crime reporters, eight 'professional news imperatives of journalism' increase an event's likelihood of being portrayed in the media, helping media outlets further attract audiences; immediacy, dramatization, personalization, novelty, conventionalism, simplification, structured access and titillation (p.205). In a more recent account, Jewkes (2010: 215-225, *originally 2004*) has added to Chibnall's list violence, proximity, perceived risk and visual imagery, among others.

In this context, it is easy to see why terrorism against Westerners is so widely and vigorously reported. It allows for the satisfaction of many of the imperatives above, including drama, visual images and violence, and can therefore be exploited by the media (Lewis 2012: 257). In this sense, the events on September 11, 2001 were extraordinarily newsworthy; "hijacked planes crashing into the World Trade Center and the Pentagon – a devastating attack on the centre of Western military, economic and cultural power – was such an excessive moment in terms of news value that it almost stretched credulity" (Lewis 2012: 258).

A vast amount of research has explored the negative social and political effects of such coverage and over-reporting of terrorism by the mass media. The first of these consequences is the

spread of fear among the target public of a terrorist group. In the years following September 11 2001 for example, "levels of public concern [about terrorism] roughly reflected the volume of terrorism coverage in the major TV networks" (Kern, Just and Norris 2003, discussed in Nacos 2007: 180). Fear of terrorism following a recent event is natural. However, it is dangerous for the media to *over*-report on the issue in sensationalist ways, as this may lead to *disproportionate* levels of fear among the public (Nacos 2007; Pozas and Toral 2004). Deciding whether fear is disproportionate can seem controversial, but there are some clear signs. For instance, throughout more than four years after 9-11, over 60 per cent of Americans persistently feared another terrorist attack in the United States 'a great deal' or 'somewhat' (Nacos 2007: 182). However, as I mentioned above, in the early 21ˢᵗ Century terrorism is in fact *not* as much of a threat as many believe it to be (Lewis 2012). It is unfortunate that the mass media's over-reporting contributed to such elevated and disproportionate concern.

Relatedly, frightened people are more easily manipulated into supporting certain political ends (Signorielli 1990: 102, discussed in Reiner 2007: 321). Rothe and Muzzatti (2004: 337-340) note that the growth in public anxieties after the reports of heightened terror alerts following 9-11 made it easier for the Bush administration to pursue an 'imperialist' political agenda, gathering support for a 'war on terror' (*ibid*: 339). Arguably, this course of action by the government was disproportionate to its impetus, but it received relatively little public criticism (Nacos 2007: 143-159).

However it would be remiss to ignore what the mass media can do to help combat terrorism. Appropriate reporting of a terrorist act can sensitize people and heighten vigilance, making it more difficult for terrorism to be carried out successfully. People can even be encouraged to cooperate with the authorities in

investigations (Wilkinson 1997: 60), as was seen recently in 2013 after the Boston Bombings, where many people who heard of the events helped the victims by donating blood and aiding with the police investigation (*The Guardian Online* 2013). Such extremely beneficial effects are brought about by a mass media that we are often quick to criticise. Thus, the ideal solution would be to provide proper safeguards for the reporting of terrorism, while still upholding freedom of the press and capitalizing on benefits the mass media can provide (Wilkinson 1997: 50-52).

Though by no means exhaustive, this overview has shed light on some of the main issues regarding the controversial relationship between terrorism and the mass media. We shall now focus on a specific case study of the news representation of the 2004 Madrid Train Bombings to give context to the issues highlighted – noting meanwhile how the Spanish press differed from traditional terrorism reporting.

The Madrid Train Bombings

On the morning of 11 March 2004, three days before the Spanish General Elections, ten bombs exploded on four commuter trains that were heading to Madrid's Atocha Station, killing 191 people and injuring over 1800. The Madrid Bombings, which have been described by the media as "the worst terrorist attack in mainland Europe since World War II" (Valdeón 2009: 67) were attributed to Islamic fundamentalist terrorists. No *direct* Al Qaeda link has been demonstrated through legal proceedings. The attacks had a variety of unexpected social and political consequences – including the ousting of the governing party, the centre-right-wing Partido Popular (PP), in favour of the centre-left-wing opposition party Partido Socialista Obrero Español (PSOE) during the elections, contrary to the polls' predictions (Pozas and Toral 2004).

11-M, as the day has come to be known, received widespread media attention across Spain and internationally (Casero 2005: 22). Given its effects, it serves as an interesting subject for a case study delving into the relationship between the media and terrorism.

To carry out the case study, I conducted a qualitative discourse and image analysis on 28 lead stories relating to 11-M published in two of the main broadsheet newspapers in Spain, *El País* (centre-left) and *El Mundo* (centre-right), during the two weeks after the attacks. I analysed the 14 images related to the attacks printed on the front page of either newspaper during the two-week period.

Newsworthiness

"By its very nature, a terroristic act is meant to be impressive" (Cooper 1976: 226). It is essentially a dramatic event, and 'dramatization' is one of the key factors news reports rely on to make their portrayals of events interesting and attractive. Chibnall (2010) argues that news workers aim to make an "impact" when presenting a report (p.206). By its very nature, an account of the 11-M Madrid Bombings achieves this. The risk is that the dramatic element of an event is over-emphasised in news portrayals. On Friday 12 March 2004, headlines such as "Terrorist Inferno in Madrid" (El País 2004) and "The Day of Infamy" (El Mundo 2004) were powerful and emotive. Reports that day were filled with phrases such as "the mortal sequence" (*El País 1*) of the train bombings – bombings that left a "trail of death on the railway line" (*ibid*) – which helped to increase the news hype around an already tragic situation. Sensationalist accounts of the attacks continued throughout the following two weeks. The articles focused on the human tragedy and the broken families (*El Mundo*

12; El País 8, among others), and many of them used similarly dramatized language.

Related to 'dramatization' is the news value of 'spectacle and graphic imagery' (Jewkes 2010: 222). Carrabine (2012) notes that in a recent 'visual turn' in criminology (p.463) it is now vital to analyse not only what crime is like, but also how it is represented (p.463-487). The photographs *El País* and *El Mundo* published of the 11-M bombings and their aftermath caused widespread critique, and rightly so. Vara (2006) complained that the images of the attacks were not selected with prudence, and they sensationalised the events in unethical ways that even violated the privacy of the victims and their families by, for example, showing recognisable wounded faces (p. 11-12). Notably, out of the fourteen images analysed here, seven contained images of dead (or injured) people or recognisable people who were visibly upset.

The dangers of such reporting have already been described, and include the possibility of enhanced and disproportionate fear and anxiety among the public (Cooper 1976), as well as a vulnerability to be manipulated for certain political ends (Signorielli 1990: 102, discussed in Reiner 2007: 321). Vivid and visual reports may sell, but the risks associated with this style of reporting must also be considered.

Consequences of Media Terrorism Reporting and How Spain Differed

Disproportionate Concern

Previous literature shows that levels of concern, and to an extent fear of terrorism, are related to the amount and form of media reports on terrorist acts (Nacos 2007: 180). The reports on 11-M, their abundance and their dramatic characteristics helped

spread panic and distress across the Spanish population (Vara 2006; Casero 2005). It is normal for alarm to spread after a terrorist act. The point is that the media can inadvertently help spread *disproportionate* concern, and this can have various undesirable effects. Three months before 11-M, 35.7 per cent of Spaniards believed terrorism was the greatest threat against Spain. This number increased to 60.9 per cent following the bombings (Instituto Opina 2004, discussed in Canel and Sanders 2010: 449). Terrorism may have been something to be worried about to an extent – but as Lewis (2012: 259-260) has argued, it is not actually the largest international threat in the early 21st Century. Although the increased levels of fear may not have been solely caused by the media's reporting on the events of 11 March 2004, the reports did contribute to the spread of alarm across the population. To reiterate, the consequences of these disproportionate levels of fear can be harmful to the appropriate functioning of society.

Political Consequences

We saw above how media reports can engender a "climate of public support, apathy or anger" (Cohen-Almagor 2000: 252), depending on how they are portrayed. Sometimes after a terrorist attack, governments can utilize the media to influence public sentiment and engender support for specific political goals (Rothe and Muzzatti 2004). Government administrations are able to leverage their role of valuable 'information-holders' and 'primary sources' in times of national crisis (Miller and Sabir 2012: 77; Chibnall 2010, *originally 1977*). This privileged role creates conditions in which news reporters and the public in general will pay particular attention to information they reveal and will be more likely to accept it unquestioningly than under normal circumstances (Nacos 2007). In the case of 11-M, the Aznar Administration was

very aware of the upcoming general election. The sitting conservative government (led by the *Partido Popular*) engaged in a media strategy that has been criticised by many, where it initially claimed that ETA, the Basque separatist terrorist organisation, was behind the attacks, even though the evidence was far from conclusive (Valdeón 2009: 67; Canel and Sanders 2010). This led *El Mundo*'s online articles to report on 'ETA's most bloody attacks' a few hours after the bombings (*ibid*: 71-72); information which was later found to be incorrect.

Why would the sitting government want to highlight ETA's, and not Islamic fundamentalists', involvement? One of the reasons could be that Islamic terrorism would likely relate, at least in part, to Spain's support for the Iraq war, which the PP sitting government had supported and the PSOE left-wing opposition party wanted to opt out of (Canel and Sanders 2010: 449). Three days before the election, terrorist attacks by Islamic jihadists could put the PPs lead in the polls (*ibid*: 449) in serious jeopardy by enhancing fear about a new attack for similar reasons. So, the government made an effort to portray its staunch belief that ETA was behind the events (*El País 1, 2, 3, 6, 13; El Mundo 1, 2, 3, 8*), and "government actions continued to hold ETA as responsible even after [...] the discovery of a van with a videotape with verses from the Koran" (Canel and Sanders 2010: 455).

However the Spanish newspapers, particularly *El País*, were not fooled. When the controversy over who was actually behind the attacks escalated, *El País* was quick to highlight the dispute and say evidence pointed to Islamic terrorism, and even *El Mundo* denounced the issue (*El Mundo 3*), even if in more subdued ways (*El Mundo 1; El Mundo 8*). Notably, the government's media strategy was not succeeding. In total, 14 of the 28 articles discussed the polemic over the presumed authorship of the bombings. This

contrasts greatly with the mostly-supportive accounts of government initiatives given after 9-11 (Rothe and Muzzatti 2004: 338-340) and 7-7 (Canel and Sanders 2010) by the US and UK press respectively, which hints at some particularities of the Spanish Press (Canel and Sanders 2010).

If the dramatized articles and images were not enough to influence voters, the critiques of the sitting government influenced the public even more. Pozas and Toral (2004) have argued that the PSOE party won the elections as a result of the emotions attached to the 11-M bombings and the fear they created about possible future attacks (p.13). Furthermore, the Spanish Press did not support the government's views, implying instead that the government was manipulating information (*El País 10, 13*). Consequently, "the issue [among the public] shifted from the question of who had killed to the question of who had lied" (Canel and Sanders 2010: 458), and "the government became the enemy (not the terrorists); and the remedy, a new government" (*ibid*: 458). It was thus that the left-wing PSOE, on March 14 2004, won the elections with nearly 11 million votes (*El País 4; El Mundo 4*). This electoral result was not solely the product of the media's reports on the bombings, but the reports did play a role in the formation of public opinion.

Having established this political consequence, it is important we now look back and examine whether the terrorists achieved what they wanted. The government changed as a direct result of the bombings. This is no place for deep political debates about the virtues of the left-wing or the promises of the right-wing, and I will take no stance regarding the particular political and social values the two main parties in Spain advocated throughout their campaigns. What is paramount to the discussion here, however, is that the PSOE party ended up withdrawing the Spanish troops

from Iraq. In fact Spain was criticized by various international officials for being a "nation that has succumbed to the threats of terrorism […]" (*El Mundo 7*). While the withdrawal from Iraq was likely not all the terrorists aimed for (Dannenbaum 2011: 323), it was certainly a significant aim. Thus, the terrorists did achieve at least *part* of their objectives. It appears then that sometimes terrorism does to an extent work (Rose, Murphy and Abrahms 2007). If terrorism is achieving any of its desired effects, and if any of these effects are being encouraged inadvertently by the media, then media outlets must be doing something wrong, even if this is done unintentionally.

Positive Consequences

The articles analysed depicted the terrorists as assassins and awoke in people a sense of not only concern but also unity. They highlighted a form of what Durkheim called 'collective conscience' (Erikson 1966: 4), where the law-abiding 'us' was pitted against the murdering, boundary-crossing 'other'. Partly due to the widespread reports of the bombings and their consequences, and the calls for help by authorities, thousands of citizens in Madrid took to the streets to help the victims and their families (*El País 1*), and many were willing to cooperate with the police investigation as much as possible (*El Mundo 6*).

It cannot be ignored then that the two newspapers' and other media outlets' reporting of the 11-M bombings had some deep benefits we would be careless to ignore.

Conclusion

In light of what has been said, what can be done to ensure that societies reap the benefits of media reporting while preventing the media from playing into the hands of terrorists? Scholars like

Wilkinson (1997), among others, largely agree that appropriate safeguards and self-regulation by the media are necessary so that that the media does not sensationalise terrorism and so that neither the government nor the terrorists manipulate the press (Wilkinson 1997: 51).

In contemporary democracies, media censorship and excessive statutory regulation would inevitably raise complaints regarding the violation of two of the main ideological pillars of democracy: freedom of the press and freedom from undue regulation (Wilkinson 1997: 60-61). Nonetheless, the risks posed by the effects of unrestrained media reporting of terrorism must be addressed. In this context, theorists have argued for "voluntary self-restraint" by the mass media (Wilkinson 1997: 51). Certainly, news-workers in all media outlets need to balance the desire to report and inform the public about newsworthy events, with the prudence not to sensationalise situations that already cause heightened emotions in themselves, and with the care not to play into the hands of terrorists by giving them too much publicity (Paletz and Tawney 1992: 110). Essentially, what we need is a variant of Green's (2012) 'public journalism' (p.281), where information in the press is provided in a rational, cool-headed manner, avoiding the "shallow, sensationalised, overly simplified presentations found in some segments of the press [that] serve only to tap into inchoate insecurities, inflaming them, without providing the informational tools necessary for the public to assess and to face them" (p. 282). In the case of the Madrid Bombings, as has been evidenced in this article, this form of professional self-restraint by the press would have been extremely beneficial.

In a world where the reach of the 'global mediasphere' (Ferrell and Greer 2009: 5) is expanding, and where terrorism

becomes more and more of a social preoccupation, we must find ways for the media not to be used in ways that benefit terrorists.

REFERENCES

Canel, M.J., and Sanders, K. (2010), 'Crisis Communication and Terrorist Attacks: Framing a Response to the 2004 Madrid Bombings and 2005 London Bombings', in T.W. Coombs and S.J. Holladay, eds., *The Handbook of Crisis Communication,* 1 edn., 449-466. Oxford: Blackwell Publishing Ltd.

Carrabine, E. (2012). 'Just Images: Aesthetics, Ethics and Visual Criminology', *British Journal of Criminology,* 52/3: 463-489.

Casero, A. (2005), 'Cobertura Periodística del 11-M: la Teoría del 'Caso Excepcional'', *VII Congreso Español de Ciencia Política y de la Administración, Democracia y Buen Gobierno*, available at http://www.aecpa.es/uploads/files/congresos/congreso_07/area04/GT 14/CASERO-RIPOLLES Andreu(URV).pdf, last accessed May 27th 2013.

Chibnall, S. (2010), 'Press Ideology: The Politics of Professionalism (1977)', in C. Greer, ed., *Crime and Media: A Reader,* 1 edn., 203-214. Abingdon: Routledge.

Cohen-Almagor, R. (2000), 'The Terrorists' Best Ally: The Quebec Media Coverage of the FLQ Crisis in October 1970', *Canadian Journal of Communication,* **25**(2), pp. 251-284.

Cooper, H.H.A. (1976), 'Terrorism and the Media', *Chitty's Law Journal,* 24/7: 226-232.

Dannenbaum, T. (2011), 'Bombs, Ballots, and Coercion: The Madrid Bombings, Electoral Politics, and Terrorist Strategy', *Security Studies,* 20: 303-349.

Erikson, K. (1966), *Wayward Puritans: A Study in the Sociology of Deviance,* 1 edn. New York: John Wiley and Sons.

Ferrell, J., and Greer, C. (2009), 'Global Collapse and Cultural Possibility', *Crime, Media, Culture: A International Journal,* 5: 5-7.

Green, D.A. (2012), *When Children Kill Children: Penal Populism and Political Culture,* 1 edn. Oxford: Oxford University Press.

Hall, S., Critcher, C., Jefferson, T., Clarke, J., and Roberts, B. (2010), 'The Social Production of News (1978)', in C. Greer, ed., *Crime and Media: A Reader.* 1 edn., 239-250. Abingdon: Routledge.

Jewkes, Y. (2010), 'The Construction of Crime News (2004)', in C. Greer, ed., *Crime and Media: A Reader,* 1 edn., 215-227. Abingdon: Routledge

Lewis, J. (2012), 'Terrorism and News Narratives', in D. Freedman and D.K. Thussu, eds., *Media & Terrorism: Global Perspectives.* 1 edn., 257-270. London: SAGE Publications Ltd.

Miller, D., and Sabir, R. (2012), 'Propaganda and Terrorism', in: D. FREEDMAN and D.K. THUSSU, eds., *Media & Terrorism: Global Perspectives.* 1 edn., London: SAGE Publications Ltd., pp. 77-94.

Nacos, B.L. (2007), '*Mass-Mediated Terrorism: The Central Role of the Media in Terrorism and Counterterrorism',* 2 edn., Lanham, Maryland: Rowman and Littlefield Publishers.

Paletz, D.L., and Tawney, L.L. (1992), 'Broadcasting Organizations' Perspectives', in D.L. PALETZ and A.P. SCHMID, eds., *Terrorism and the Media: How Researchers, Terrorists, Government, Press, Public, Victims View and Use the Media,* 1 edn., 105-110. London: SAGE Publications Ltd.

Pozas, S., and Toral, G. (2004), 'La gestión de la comunicación institucional entre el 11 y el 14 M', *Zer - Revista de Estudios de Comunicación,* 9/17: 11-23.

Reiner, R. (2007), 'Media Made Criminality: The Representation of Crime in the Mass Media', in M. Maguire, R. Morgan and R. Reiner, eds., *The Oxford Handbook of Criminology.* 4 edn., 302-337. Oxford: Oxford University Press.

Rose, W., Murphy, R., and Abrahms, M. (2007), 'Does Terrorism Ever Work? The 2004 Madrid Train Bombings', *International Security,* 32/1: 185-192.

Rothe, D., and Muzzatti, S.L. (2004), 'Enemies Everywhere: Terrorism, Moral Panic, and US Cvivil Society', *Critical Criminology,* 12/3: 327-350.

The Guardian Online. (2013), 'Boston Marathon blasts: three dead and more than 100 injured – as it happened', *The Guardian Online*, 15 April (written by T. McCarthy), available at http://www.guardian.co.uk/world/2013/apr/15/boston-marathon-explosion-live, last accessed May 29th 2013.

Valdeón, R.A. (2009), 'Discursive Constructions of Terrorism in Spain: Anglophone and Spanish Media Representations of ETA', *International Journal of Applied Linguistics,* 19/1: 66-83.

Vara, A. (2006), 'Las Sombras del Periodismo Español en el 11-M', in A. Vara, J. Rodríguez, E. Giménez and M. Díaz, eds., *Cobertura Informativa del 11-M,* 1 edn., 11-14. Pamplona: Ediciones Universidad de Navarra S.A.

Weimann, G., and Brosius, H. (1991), 'The Newsworthiness of International Terrorism', *Communication Research,* 18(3), pp. 333-354.

Wilkinson, P. (1997), 'The Media and Terrorism: A Reassessment', *Terrorism and Political Violence,* 9/2: 51-64.

Appendix A: Articles Analysed

El País 1: Romero, J. M. (2004), 'Cuatro Atentados Simultaneos Causan una Matanza en Trenes de Madrid', *El País*, 12 March, available at

http://elpais.com/diario/2004/03/12/espana/1079046001_850215.html
, last accessed 22 May 2013.

El País 2: Rodríguez, J. A., Mercado, F., Yoldi, J., and Duva, J. (2004), 'La Furgoneta con una Grabación en Árabe Fue Utilizada para Transportar las Bombas', *El País*, 13 March, available at
http://elpais.com/diario/2004/03/13/espana/1079132401_850215.html
, last accessed 22 May 2013.

El País 3: Moreno, R. (2004), 'Detenidos Tres Marroquíes y Dos Indios por Supueso Apoyo a los Autores de la Matanza', *El País*, 14 March, available at
http://elpais.com/diario/2004/03/14/espana/1079218801_850215.html
, last accessed 22 May 2013.

El País 4: Romero, J. M. (2004), 'Zapatero Vence con casi 11 Millones de Votos', *El País*, 15 March, available at
http://elpais.com/diario/2004/03/15/espana/1079305201_850215.html
, last accessed 22 May 2013.

El País 5: Díez, A. (2004), 'Zapatero Atribuye su Victoria a las "Ganas de Cambio en España" y No al Atentado del 11-M', *El País*, 16 March, available at
http://elpais.com/diario/2004/03/16/espana/1079391601_850215.html
, last accessed 23 May 2013.

El País 6: Rodríguez, J. A., and Comas, S. (2004), 'La Dinamita Usada en los Atentados se Fabricó en Febrero en Burgos', *El País*, 17 March, available at
http://elpais.com/diario/2004/03/17/espana/1079478038_850215.html
, last accessed 23 May 2013.

El País 7: Rodríguez, J. A. (2004), 'El Gobierno Lanza un Vasto Plan de Seguridad por la "Persistencia del Riesgo Terrorista"', *El País*, 18 March, available at
http://elpais.com/diario/2004/03/18/espana/1079564402_850215.html
, last accessed 23 May 2013.

El País 8: Duva, J., and Rodríguez, J. A. (2004), 'Un Trozo del Móvil de la Bomba Desactivada, Hallado en el Locutorio de Zougam en Lavapiés', *El* País, 19 March, available at
http://elpais.com/diario/2004/03/19/espana/1079650805_850215.html
, last accessed 23 May 2013.

El País 9: Yoldi, J. (2004), 'Cuatro Magrebíes Robaron los Explosivos en Asturias Guiados por un Ex Minero Español', *El País*, 20 March, available at
http://elpais.com/diario/2004/03/20/espana/1079737203_850215.html
, last accessed 24 May 2013.

El País 10: Alcaide, S., Ortega, P., Cantalapiedra, F., Cuartas, J., Hermida, X., Manresa, A., Martín, A., and Martínez, I. C. (2004), 'Más de 200.000 Personas Contra la Guerra', *El País*, 21 March, available at http://elpais.com/diario/2004/03/21/espana/1079823604_850215.html , last accessed 24 May 2013.

El País 11: Ordaz, P., Peregil, F., and Gómez, L. (2004), 'Cuatro Trenes, 14 Bombas, 6.000 Viajeros', *El País*, 22 March, available at http://elpais.com/diario/2004/03/22/espana/1079910017_850215.html , last accessed 24 May 2013.

El País 12: Duva, J., and Rodríguez, J. A. (2004), 'Cuatro Nuevos Detenidos por el 11-M, Uno de los Cuales ya Estaba Preso en Salamanca', *El País*, 23 March, available at http://elpais.com/diario/2004/03/23/espana/1079996414_850215.html , last accessed 24 May 2013.

El País 13: Mercado, F. (2004), 'Una Mochila Bomba Pasó Varias Horas en Ifema y Recorrió 25 Kilómetros Hasta Ser Desactivada', *El País*, 24 March, available at http://elpais.com/diario/2004/03/24/espana/1080082818_850215.html , last accessed 25 May 2013.

El País 14: Valdecantos, C. (2004), 'El Funeral Por las Víctimas del 11-M Congrega a Europa en la Catedral de Madrid', *El País*, 25 March, available at http://elpais.com/diario/2004/03/25/espana/1080169205_850215.html , last accessed 25 May 2013.

El Mundo 1: Lázaro, F. (2004), 'El Día de la Infamia', *El Mundo*, 12 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo 2 : Lázaro, F. (2004), 'Acebes Convencido de que las Nuevas Pistas Probaran la Autoría de ETA', *El Mundo*, 13 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo 3: Lázaro, F. (2004), 'Las Primeras Detenciones Vinculan La Masacre con el Terrorismo Islámico', *El Mundo*, 14 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 26 May 2013.

El Mundo 4: Garea, F. (2004), 'Zapatero Será Presidente Tras el Hundimiento del PP en un Vuelco Político Sin Precedences', *El Mundo*, 15 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 26 May 2013.

El Mundo 5: Deiros-Bronte, T., and Lázaro, F. (2004), 'Casablanca y Madrid: En el Mismo Punto de Mira', *El Mundo*, 16 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 26 May 2013.

El Mundo 6: Lázaro, F. (2004), 'La Policía Busca a 20 Sospechosos', *El Mundo*, 17 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 26 May 2013.

El Mundo 7: Fresneda, C. (2004), 'Kerry Solicita a Zapatero que no Retire las Tropas Españolas de Irak', *El Mundo*, 18 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 27 May 2013.

El Mundo 8: Segovia, C. (2004), 'El Gobierno Desclasifica Documentos Secretos Para Probar que "No Ha Mentido" Sobre el 11-M', *El Mundo*, 19 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 27 May 2013.

El Mundo 9: Ramírez de Ganuza, C. (2004), 'Zougam al Llegar a la Audiencia: '¿Quién Ha Ganado las Elecciones?', *El Mundo*, 20 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 27 May 2013.

El Mundo 10: Olmedo, I., and De la Cal, J. C. (2004), 'Jamal "El de Tánger"', *El Mundo*, 21 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 27 May 2013.

El Mundo 11: Valera, A. S. (2004), 'Racismo en la Universidad', *El Mundo*, 22 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 28 May 2013.

El Mundo 12: Ramírez de Ganuza, C. (2004), 'El Español Detenido Entra en Prisión Como Autor del 11-M', *El Mundo*, 23 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 28 May 2013.

El Mundo 13: Lázaro, F. (2004), 'Abderrahim Zbackn, 'El Químico', Presunto Autor de la Fabricación de las Bombas del 11-M', *El Mundo*, 24 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 28 May 2013.

El Mundo 14: Simon, P. (2004), 'La Mañana de los Pañuelos Mojados', *El Mundo*, 25 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/Buscador.aspx (past the pay wall), last accessed 28 May 2013.

Front Page Headlines

El País. (2004), 'Front Page', *El País*, 12 March, available at http://elpais.com/diario/2004/03/12/ , last accessed 22 May 2013.

El Mundo. (2004), 'Front Page', *El Mundo*, 12 March, available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 25 May 2013.

Images Analysed

El País Image 1 : *Decenas de heridos permanecen junto a las vías.* (12 March 2004), n.d. [image online], available at http://elpais.com/diario/2004/03/12/, last accessed 23 May 2013.

El País Image 2 : Espana se echa a la calle. (13 March 2004), n.d. [image online], available at http://elpais.com/diario/2004/03/13/ , last accessed 23 May 2013.

El País Image 3 : Associated Press. (2004), *'Entierro de una víctima de la matanza de Madrid'*, 14 March, [image online], available at http://elpais.com/diario/2004/03/14/ , last aaccessed 24 May 2013.

El País Image 4 : Álvarez, C. (2004), *'Detención de personas relacionadas con el 11-M',* 19 March, [image online], available at : http://elpais.com/diario/2004/03/19/ , last accessed 24 May 2013.

El País Image 5 : Reuters. (2004), *'Los musulmanes muestran su rechazo a los atentados del 11-M',* 22 March, [image online], available at http://elpais.com/diario/2004/03/22/ , last accessed 24 May 2013.

El País Image 6 : EFE/POOL. (2004), *'Funeral en la Almudena por las víctimas del 11-M',* 25 March, [image online], available at http://elpais.com/diario/2004/03/25/ , las accessed 24 May 2013.

El Mundo Image 1 : Marcou, P. (2004), *'Tres bomberos intentan liberar a dos pasajeros muertos del cercanías que sufrió ayer un atentado cuando entraba en la estación de Atocha (Madrid)',* 12 March, [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo Image 2 : Cordón, B. (2004), '*Imagen de la manifestación celebrada ayer en Madrid bajo una intensa lluvia, a su paso por la estación de Atocha*', 13 March, [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo Image 3 : Para, K. (2004), *'Los manifestantes, portando carteles con la palabra « paz », se agolpaban ayer frente a la sede del PP en la calle Génova'*, 14 March, [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo Image 4 : **No Caption** (17 March 2004) n.d.[image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo Image 5 : Sinova, D. (2004), *'Lágrimas en la última estación del vía crucis'*, 19 March, [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 25 May 2013.

El Mundo Image 6 : El Sospechoso 'Número Uno' (20 March 2004), n.d. [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 26 May 2013.

El Mundo Image 7 : De León, M. H. (2004), *'El Rey da el pésame a un familiar de las víctimas del 11-M tras el funeral celebrado ayer en La Almudena'*, 25 March, [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 27 May 2013.

El Mundo Image 8 : Vera, S. (2004), *'Varias personas lloran mientras siguen el funeral de Estado por una pantalla de TV situada fuera de la catedral'*, 25 March, [image online], available at http://quiosco.elmundo.orbyt.es/Hemeroteca/plantilla_binario.aspx (past the pay wall), last accessed 27 May 2013.

The effect of immediate feedback on the perception of Mandarin lexical tones by non-native speakers of Mandarin

CONG ZHANG

Cong Zhang is a first-year DPhil student in Linguistics. Her research interest in the prosody of Mandarin Chinese has motivated this paper on the effectiveness of immediate feedback on the perception of lexical tones by non-tonal language speakers with no previous knowledge of Mandarin. Lexical tone is one of the most difficult issues in learning Mandarin as a foreign language. Various efforts have been made by training non-native speakers to improve the perception of Mandarin lexical tones. Immediate feedback, as an essential and efficient way of perceptual learning, however, has been understudied. An AX discrimination task is used to test whether the participants' perception of Mandarin lexical tones improves after being given immediate feedback. The result shows an evident effect of immediate feedback on the perception of Mandarin lexical tones, both within the experiment groups as well as between the experiment group and the control group.

---

1. Introduction

Lexical tone, an important feature in Mandarin Chinese, is one of the most difficult features for non-native speakers of Mandarin to acquire (e.g. Francis et al. 2008; Wang 1999) due to the differences in pitch patterns, functions and distributions between English and Mandarin (Chen 1974; White 1981). Efforts have been made to improve either the perception or production of Mandarin lexical tones by previous studies, with various training methods used. However, hardly any study has examined the effect of immediate feedback on tone perception or production. The present study therefore will investigate the effect of immediate feedback on the perception of Mandarin lexical tones by non-native speakers of Mandarin, aiming at exploring a more effective

and efficient perceptual learning method and acting as a pilot study to prepare for better-controlled future research on tone perception.

Tones belong to suprasegmental phonology (Leben 1973; Yip 1980), which is different from segmental phonology, such as consonants and vowels. Goldsmith (1979) advanced the Autosegmental Theory, in which he contended that tonal and segmental phonology are separate. Zeng & Mattys (2011) investigated the separability of Mandarin tones and rhymes by way of the perceptual-migration paradigm and supported Autosegmental Theory by concluding that native Mandarin speakers automatically perceive lexical tones separately from the tone-bearing rhyme, while speakers of non-tonal languages cannot. Therefore, difficulty in perceiving lexical tones for non-native speakers of Mandarin is due to their inability to separate suprasegmental tones from segmental vowels and consonants.

1.1 Perceptual learning

Non-native speakers of Mandarin, however, are not tone-deaf. One way of improving their perception is through perceptual learning. Perceptual learning is "extracting previously unused information" (Gibson & Gibson1955), which can be divided into four different mechanisms: "stimulus imprinting", "differentiation", "unitization" and "attentional weighting". (Goldstone 1998) The last mechanism, "attentional weighting", means that more "weight" can be given to relevant features as experience increases. (Samuel & Kraljic 2009) For example, non-tonal language speakers do not pay attention to lexical tones because tones do not make any difference in their native language; nevertheless, after having more experience, such as receiving training or feedback, their perception of tones will improve. Training, as the most common perceptual learning method, has

been employed in many studies that investigated the effect of perceptual learning on music perception and speech perception, especially on segmental phonology perception (e.g. Bradlow et al. 1997; Iverson et al. 2003; Kraljic and Samuel 2006). Recently, a few studies have also been done on training non-native speakers of Mandarin to perceive or produce Mandarin lexical tones (e.g. Francis et al. 2008; Wang et al. 1999; Wang et al. 2003). However, feedback, as a more efficient perceptual learning method, has been understudied.

1.2 Feedback

Feedback can be divided into two types according to the timing of the feedback: immediate feedback, and delayed feedback. Scholars have been arguing about which is more effective since the 1920s.

The scholars who supported the immediate feedback argued that errors should be corrected before the students remember them (Pressey1932; Mason and Bruning 2001) and the correct response should be reinforced immediately (Skinner 1954; Renner 1964); the scholars who supported delayed feedback, however, believed that delayed feedback could reduce proactive interference so that the incorrect information could be forgotten before inputting the correct information (Kulhavy and Anderson 1972).

Some studies, mostly early studies, were in favour of delayed feedback. For example, Kulhavy and Anderson (1972) claimed that delayed feedback was superior to immediate feedback with evidence from a multiple-choice test. Kulhavy (1977) again supported his previous claim with evidence from a writing task. Bardwell (1981) also found delayed feedback was more effective than immediate feedback in terms of a school related learning.

1.3 Present study

Based on all the aforementioned studies, the current study focuses on evidencing that immediate feedback is efficient. The research question therefore is: will immediate feedback improve the perception of Mandarin lexical tones by non-native speakers of Mandarin? This study hypothesizes that the error rate of the perception of Mandarin lexical tones by non-native speakers of Mandarin will decrease as more feedback is received.

In order to exclude the influence of the native languages of the participants from the results, participants with two different L1 backgrounds were chosen, namely, native English speakers and native Arabic speakers. The participants were asked to accomplish an AX discrimination task and the error rate of each participant was calculated. The results showed that the average error rate of the participants who received feedback was lower than that of the participants who did not. Moreover, as the participants received more feedback, the error rate kept decreasing. These results clearly showed that immediate feedback did affect tone perception.

2. Background: a brief introduction to Mandarin phonology

Traditionally, a Mandarin syllable is considered to consist of three parts: the initial, the final and the tone (Cheng 1973; Huang 1992). However, more and more scholars tend to categorize them as onset, rhyme and tone. (Duanmu 2007) The difference between these two categorization methods lies on the glide. In the first categorization, the glide belongs to the final, while it belongs to the onset in the second categorization (Triskova 2011). Regardless of the names, the onset (or the initial) and the rhyme (or the final) are composed of consonants and vowels. They are therefore segmental features. The tone, which is indisputably

separate in both methods, is a suprasegmental feature. The present study will adopt the names of "onset" and "rhyme" for the segments.

2.1 Segmental features of Mandarin syllables

Mandarin is a monosyllabic language. A Mandarin syllable usually has an onset and a rhyme. The onset is either a consonant (C) or it is omitted; the rhyme can be monophthongs, diphthongs, or vowel (V) + nasal ([n] or [ŋ]) or vowel+ liquid ([ɻ]). Therefore, the syllable structure of Mandarin could be V, VC (nasal/ liquid), CV, CVV, CVC (nasal/ liquid), or CVVC (nasal/ liquid).

2.1.1 Onsets

Most of the onsets in Mandarin exist in English, for example, [m], [n], [pʰ], [kʰ], etc. Some others have subtle differences with those in English, such as /f/ is pronounced as [fʰ] in Chinese, which is different from [f] in English. These subtle differences are on the level of allophones, which will not influence the perception of the syllable. However, there are a few onsets that the English speakers find difficult to perceive and produce, for instance, the alveolar-palatals ([tɕ], [tɕʰ], [ɕ]), the fricatives ([ts], [tsʰ]) and the affricates ([tʂ], [tʂʰ], [ʂʰ], [ʐ]). These sounds will be excluded in the present study since the focus is on the suprasegmental features rather than the segmental features. /w/ and /j/, considered by the scholars who support the Onset-Rhyme Model as glides, are also excluded.

2.1.2 Rhymes

Mandarin has a small repertoire of vowel phonemes, /a/, /o/, /i/, /u/ and /y/. Except for [y], the other four phonemes all

exist in English, including their allophones. As for the V + nasal and V+ liquid, such combinations are legal in English phonology. The most difficult rhymes for the English speakers are the apicals, [ɻ] and [ɻ].Therefore, in the design of present study, [y], [ɻ] and [ɻ] are excluded. /u/ and /i/ are also excluded, but only when they are in positions in which the scholars supporting the Initial-Final Model consider them as glides.

2.2 Suprasegmental features of Mandarin syllables

The lexical tone is the only suprasegmental feature on syllable level in Mandarin. There are five lexical tones, high level, rising, dipping, falling, and neutral (Li and Thompson 1977). The first four are the most important ones for comprehension; the neutral tone, however, only occurs when the syllable is unstressed. Chao (1930) designed a tone letter system that could represent the tones according to their relative pitch levels. In this system, the lowest level is "1" and "5" is the highest. The pitch levels of the four tones are 55, 35, 214 and 51 respectively. When the syllables are pronounced in isolated syllables under ideal circumstances, the first four tones will have the contours as shown in Figure 1. The

Figure 1: Four Mandarin lexical tones (Chao 1930)
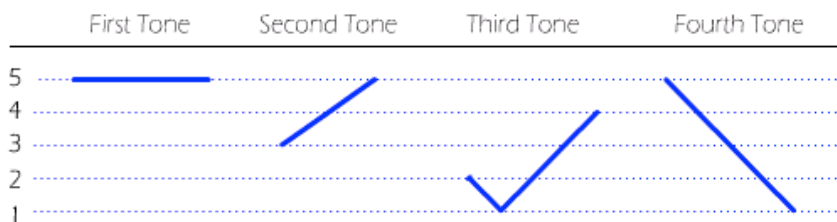
four tones resemble music notes to some extent, but the pitches of music tones are absolute, while those of Mandarin lexical tones are

comparatively relative. In this study, neutral tone, which is commonly referred to as "no tone", is excluded, due to its many allotones, which vary according to the tone of the preceding syllable.

Tone plays an important role in the semantics of a syllable. In Mandarin, one character may have one or several pronunciations. But it is always monosyllabic no matter which pronunciation it adopts. Each syllable has one of the five lexical tones. Change of lexical tones of a syllable may result in either referencing another character, or varying the meaning of the same character altogether. Different characters may have the same syllable and the same tone. The corresponding relationship between the tones and characters may be many-to-one or one-to-many. For example, in Table 1, a syllable of a certain tone can have more than one corresponding characters; one character can also have different tones with the same syllable, such as "吗" has the second, third and neutral tones; one character can have different syllables with different tones as well, for instance, "抹" can also be pronounced as *mǒ* [mo]214 ('erase'), or *mò* [mo]51 ('plaster').

Table 1: Example of the relationship between tone and character in Mandarin

(*Pinyin*: Romanized representation of Mandarin pronunciation)

Pronunciation		Chinese characters
Pinyin	IPA	
mā	[A̠] 55	妈('mother'), 抹(wipe)[2]
má	[A̠] 35	麻('hemp'), 吗('what')[1]
mǎ	[A̠] 214	马('horse'), 吗('morphine')[1]
mà	[A̠] 51	骂('scold')

| ma | [A̪] | 吗(a particle used at the end of questions)[1] |

Yang et al. (1988) calculated that there are approximately 1300 syllable sounds in Mandarin, while more than 50,000 characters are used in the writing system. If there are no tones, only 411 syllable sounds are left in Mandarin. Therefore, perceiving the lexical tones correctly is an important task in learning Mandarin.

2.3 Summary

A Mandarin syllable is composed of segmental features and suprasegmental features. Segmental features include onsets and rhymes. In the current study, some segmental sounds in both onset and rhyme positions that may cause too much perceptual difficulty for English speakers are excluded. The suprasegmental feature on syllable level only consists of lexical tones. Lexical tones are closely related to the meaning of Mandarin characters; therefore, the perception of lexical tones is a very important issue.

3. Experiment design

3.1 Participants

12 native speakers of British English and 12 native speakers of Arabic, all with no known hearing or speech problems, participated in the experiment. The randomness of the choice of participants was assured since the recruitment was through various means such as in a campus activity or through emails. All participants were undergraduate or post-graduate students at Newcastle University. None had any knowledge of or experience with Mandarin or any other tonal language.

The participants were divided into four groups: six English speakers and six Arabic speakers formed two experiment groups respectively, while the other six English speakers and six Arabic speakers acted as two control groups. The experiment groups received feedback after each judgment, while the control groups did not.

3.2 Stimuli and apparatus

A total of 80 pairs of monosyllabic Mandarin morphemes were selected, every one of which were composed of an onset and a rhyme. In order to avoid the influence of non-native syllable structure or non-native phonology on the perception, phonemes that do not exist in English were excluded. The rhymes included both monophthongs and diphthongs, which were evenly numbered and distributed in the stimuli list (see Appendix). The full list of onsets and rhymes is as shown below.

Table 2: full list of onsets and rhymes in *pinyin*

onsets	b p m f d t n g k h
rhymes	a o e i u ao ai ou ei an en in ang eng ing ong

The stimuli were divided into two categories according to their onset contexts in the experiment, i.e. stimuli with the same consonants but different rhymes (SCDR) and stimuli with different consonants and different rhymes (DCDR). In each category, 20 of the stimuli pairs were of the same tone, and the other 20, different. Since the third tone (214) was the most difficult tone to perceive

for non-Mandarin speakers (Zeng 2008), this experiment included more contrasts with the third tone, with those between the second and the third doubling others as previous studies suggested that the non-native speakers of Mandarin were most confused between the second tone and the third tone in both perception and production (Lee et al 2010; Zeng 2008). Table 3 shows the stimuli design and the numbers of each tone pair. "55", "35", "214" and "51" are pitch values of the four mandarin lexical tones.

Table 3: numbers and categories of stimuli

	Same tone (20)				Different tones (20)		
CD R (40)	55+55 (5)	35+35 (5)	214+214 (5)	51+51 (5)	55+214 (5)	35+214 (10)	51+214 (5)
CD R (40)	55+55 (5)	35+35 (5)	214+214 (5)	51+51 (5)	55+214 (5)	35+214 (10)	51+214 (5)

All the stimuli pairs were recorded by a female native Mandarin speaker with Sony ICD-PX312 digital voice recorder. An ASUS U36J laptop was used for playing the stimuli, and the judgments were taken down with pen and paper.

3.3 Procedures

The experimenter first explicitly introduced what lexical tones were in Mandarin, using "a 55", "a 35", "a 214" and "a 51" as examples and then compared the lexical tones to musical notes.

Having had a thorough idea of what Mandarin lexical tones were, the participants were asked to perform an AX discrimination task. The participants heard a trial of two stimuli at a time and they needed to make a judgment on whether the tones of the two stimuli were the same, ignoring the consonants and vowels.

The experimenter gave the participants in the experiment groups a feedback after each judgment by telling the participants whether the judgment was "right" or "wrong". For the control groups, however, the experimenter would immediately play the next trial without giving any feedback. The experimenter took notes of the correctness of the judgments of both the experiment groups and the control groups for data analysis.

## 4. Results

4.1 Average error rate

The average error rate of the English control group was 16.67% while that of the English experiment group was 7.71% (Figure 2). The significant improvement indicated that feedback did have an effect on the perception of Mandarin lexical tones. The Arabic experiment group, with an error rate of 23.96%, also showed significant improvement over the Arabic control group, whose error rate was 35.21%.

The Arabic speakers made more mistakes on average than the English speakers. The Arabic experiment group, which made fewer mistakes than the Arabic control group, still had a higher error rate than the English speakers.

### 4.2 Average error rates by blocks

With the stimuli divided into four blocks (i.e. First Block: the first 20 pairs of stimuli; Second Block: the second 20 pairs of

Figure 2: Average error rates
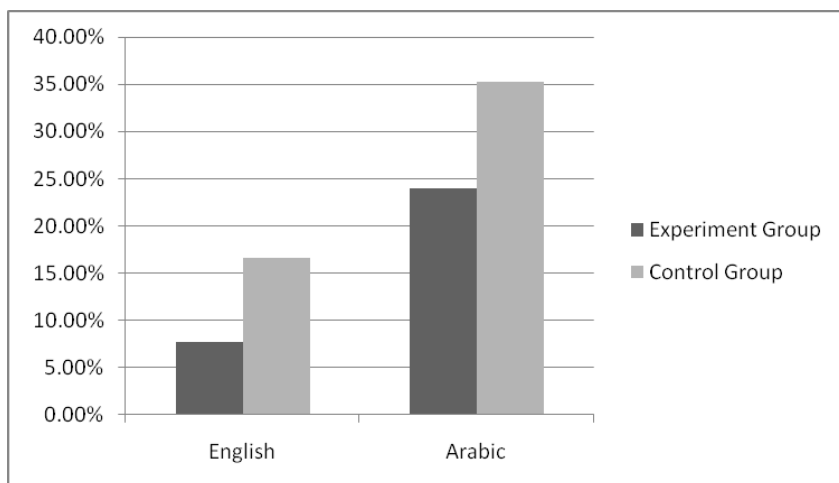
stimuli; Third Block: the third 20 pairs of stimuli; Fourth Block: the fourth 20 pairs of stimuli), the average error rates by blocks of the English control group were 21.7%, 19.2%, 11.7% and 14.2% respectively while those of the English experiment group were15.0%, 10.0%, 1.7% and 4.2% respectively (Figure 3). These results indicated that as receiving more feedback, the perception

Figure 3: Average error rates by blocks (English speakers)
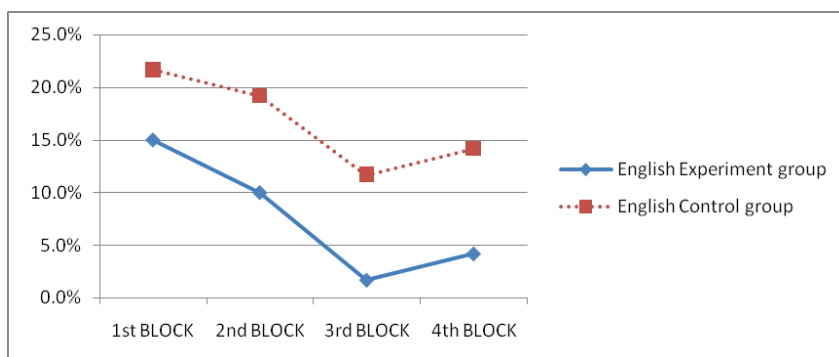
abilities of both groups were improving; moreover, the experiment group improved at a higher speed and reached a higher level than the control group. Further discussion will be made on the fourth block rebound.

For the Arabic groups, the average error rates by blocks of the Arabic control group were 39.2%, 38.3%, 37.5% and 25.8% respectively while those of the Arabic experiment group were 31.7%, 20.8%, 18.3% and 25.0% respectively (Figure 4).

Figure 4: Average error rates by blocks (Arabic speakers)
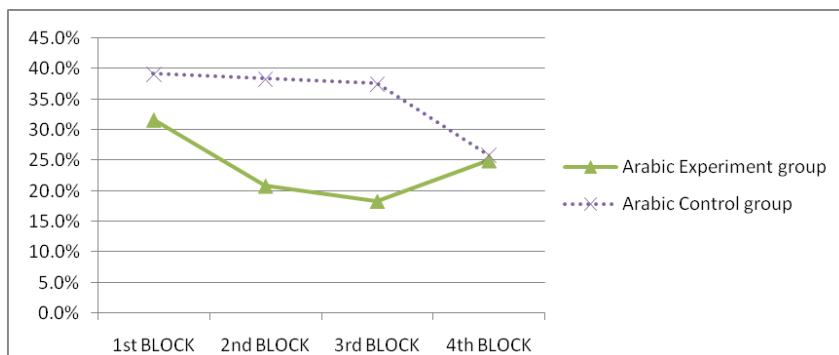
The English control group and the English experiment group showed similar curve: the error rates decreased from Block 1 to Block 3 and rebounded in Block 4. The Arabic experiment group revealed similar trend. However, for the Arabic control group, the error rate basically remained at the same level while in Block 4, it had a sudden plunge.

5. Discussion

The results were overall in accordance with the hypothesis: the error rates generally decreased within the experiment groups themselves as the participants received more feedback; the error

rates of the experiment groups, who received feedback, were lower than the control groups, who did not receive feedback. These results indicated that feedback had some positive effects on tone perception.

However, there were three unexpected results, namely the low error rates compared with the data from previous studies, the fourth-block rebound for the experiment groups and the much higher error rate of the Arabic speakers than the English speakers.

The data from Zeng (2008) showed that English speakers generally made more than 50% errors and even the native Mandarin speakers had an average error rate of 21.5%. However, in this study, the English control group, who made more mistakes than the English experiment group, only had 16.67% of errors; the Arabic groups made higher error rates, i.e. 35.21% for the Arabic control group and 23.96% for the Arabic experiment group, which were still much lower than expected. This issue can be attributed to the different experiment methods used in the literature and this pilot study. In Zeng (2008), the AX discrimination task was performed on a computer with DMDX software, the stimuli durations, stimuli display time and intervals between trials were strictly controlled; whereas in the current pilot study, a simple AX discrimination task without controlling the abovementioned factors. The participants sometimes thought about the decision for more than 10 seconds, which made the judgments more like processing rather than perception. Therefore, the future researches should have stricter control over such factors as the stimuli durations, stimuli display time, intervals between trials and so on.

The second unexpected result was the rebound in the fourth block. The error rates of the fourth block were expected to be the lowest among all four blocks; however, both experiment groups had a fourth-block rebound: the English experiment group

increased from the 1.7% in the third block to 4.2% in the fourth block, while the Arabic experiment group increased from 18.3% in the third block to 25.0% in the fourth block, which was almost as high as the 25.8% of Arabic control group. The rebound of the English speakers' error rate might be due to the fixed stimuli order since the control group had the same curve as shown in Figure 3. However, the curves of the two Arabic groups were different. The error rate of the control group remained at the same level while decreased in the fourth block, which showed normal cognition and perception process. The fact that Arabic experiment group revealed decreasing trend in the first three blocks and performed better than the Arabic control group in general displayed the effect of the immediate feedback. Other than the fixed stimuli order, the unexpected rebound might be attributed to the concentration span of the participants, small sample size of this pilot study, or a possible ceiling effect. To avoid all these factors, the future researches should randomize the stimuli order, give breaks after each block, recruit more participants and detect the ceiling effect.

That the Arabic speakers made notably more incorrect judgments than the English speakers was the third unexpected result. Both Arabic and English are non-tonal and have intonations while the results showed that Mandarin lexical tones were more difficult for the Arabic speakers than the English speakers. There were studies on comparing the perceptions of tones by tonal language speakers and non-tonal language speakers (e.g. Kaan et al. 2008) and the results showed that tonal language speakers achieved better results than non-tonal language speakers. However, the perception difference of speakers from different language backgrounds within non-tonal languages was understudied. The current pilot study therefore has provided a research topic that is worth studying in the future.

REFERENCES

Bradlow, A. et al. 1997. 'Training Japanese listeners to identify English/r/and IV: IV. Some effects of perceptual learning on speech production'. *Journal of the Acoustical Society of America* 101: 2299-23.

Bruning, R. H., G. J. Schraw and R. R. Ronning. 1999. *Cognitive psychology and instruction*. Prentice Hall.

Chao, Y. R. 1930. 'A system of tone letters'. *Le Maitre Phonetique* 45: 24-27.

Chen, G. 1974. 'The pitch range of English and Chinese speakers'. *Journal of Chinese Linguistics* 2: 159-171.

Cheng, C. C. 1973. *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.

Cincotta, C. M. and C. A. Seger. 2007. 'Dissociation between striatal regions while learning to categorize via feedback and via observation'. *Journal of cognitive neuroscience* 19: 249-265.

Duanmu, S. 2007. *The phonology of standard Chinese*. Oxford: Oxford University Press.

Francis, A. L. et al. 2008. 'Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers'. *Journal of Phonetics* 36: 268-294.

Gibson, J. and E. J. Gibson. 1955. 'Perceptual learning: differentiation or enrichment?'. *Psychological review* 62: 32-41.

Goldsmith, J. 1979. *Autosegmental Phonology*. New York: Garland Publishing.

Goldstone, R. L. 1998. 'Perceptual learning'. *Annual Review of Psychology* 49: 585-612.

Holland, J. G. 1960. 'Teaching machines: An application of principles from the laboratory'. *Journal of the experimental analysis of behavior* 3: 275-287.

Huang, L. M. 1992. 'Remarks on the phonological structure of Mandarin Chinese'. *Bul. Nat 1 Taiwan Nor. Univ* 37: 363-383.

Iverson, P. et al. 2003. 'A perceptual interference account of acquisition difficulties for non-native phonemes'. *Cognition* 87: B47-B57.

Kaan, E., C. et al. 2008. 'Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study.' *BMC neuroscience* 9: 53.

Kimura, M. and A. Graybiel. 1995. 'Role of basal ganglia in sensory motor association learning'. *Functions of the Cortico-Basal Ganglia Loop. Springer, Tokyo*: 2-17.

Kraljic, T. and A. G. Samuel. 2006. 'Generalization in perceptual learning for speech'. *Psychonomic bulletin & review* 13: 262-268.

Kulhavy, R. W. 1977. 'Feedback in written instruction'. *Review of Educational Research* 47: 211–232.

Kulik, J. A. and C-L C. Kulik. 1988. 'Timing of Feedback and Verbal Learning'. *Review of educational research* 58: 79-97.

Leben, W. 1973. *Suprasegmental phonology*. Unpublished PhD dissertation, MIT.

Lee, C. Y., L.Tao and Z. Bond, 2010. 'Identification of acoustically modified Mandarin tones by non-native listeners'. *Language and Speech* 53: 217-243.

Li, C. N. and S. A. Thompson. 1977. 'The acquisition of tone in Mandarin-speaking children'. *Journal of Child Language* 4: 185-199.

Nagata, N. and M. V. Swisher. 1995. 'A study of consciousness-raising by computer: The effect of metalinguistic feedback on second language learning'. *Foreign Language Annals* 28: 337-347.

Pressey, S. L. 1950. 'Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction'. *The Journal of Psychology* 29: 417-447.

Reynolds, J. N. J. and J. R. Wickens. 2002. 'Dopamine-dependent plasticity of corticostriatal synapses'. *Neural Networks* 15: 507-521.

Rutherford, W. E. 1987. *Second language grammar: Learning and teaching*. Longman.

Rutherford, W. E. and M. Sharwood. 1985. 'Consciousness-raising and universal grammar'. *Applied linguistics* 6: 274-282.

Samuel, A. G. and T. Kraljic. 2009. 'Perceptual learning for speech'. *Attention, Perception, & Psychophysics* 71: 1207-1218.

Skinner, B. F. 1954. *The science of learning and the art of teaching*. USA: Cambridge Mass.

Triskova, H. 2011. 'The structure of the Mandarin syllable: Why, when and how to teach it'. *Archiv orientální* 79: 99-134.

Wang, Y. et al. 1999. 'Training American listeners to perceive mandarin tones'. *Journal of the Acoustical Society of America* 106: 3649-3658.

Wang, Y. et al. 2003. 'Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training'. *Journal of the Acoustical Society of America* 113: 1033.

White, C.M. 1981. 'Tonal perception errors and interference from English intonation'. *Journal of* the *Chinese Language Teachers Association* 16: 27-56.

Yang, W. J., J. C. Lee, Y. C. Chang and H. C. Wang. 1988. 'Hidden markov model for Mandarin lexical tone recognition'. *Acoustics, Speech and Signal Processing, IEEE Transactions on* 36: 988-992.

Yip, M. 1980. The tonal phonology of Chinese. PhD, MIT.

Zeng, B. 2008. *Tone Perception and Migration by English speakers*. Unpublished PhD dissertation, Bristol University.

Zeng, B. and S. Mattys. 2011. 'Separability of tones and rhymes in Chinese speech perception: evidence from perceptual migration'. *17th International Congress of Phonetic Sciences*, Hong Kong, China: ICPhS.

Appendix: Pilot Stimuli List (Fixed order)

No. 1	mo55	mi55
No. 2	gang55	gong214
No. 3	mu214	fa214
No. 4	mo214	ga51
No. 5	ge35	gu214
No. 6	du55	ta55
No. 7	fei51	fen214
No. 8	bo55	ning214
No. 9	po35	pin214
No. 10	du35	da35
No. 11	dou51	hao51
No. 12	nin35	da35
No. 13	du51	di51
No. 14	nan55	mao55
No. 15	hou35	di214
No. 16	du55	dou55
No. 17	ba35	hu214
No. 18	ha214	hou35
No. 19	kou214	hai35
No. 20	ken214	kai214
No. 21	ke55	ka55
No. 22	bi55	tou55
No. 23	heng35	han214
No. 24	pi214	po51
No. 25	ka214	mai214
No. 26	nong35	nao35
No. 27	mo214	ge35

No. 28	ku51	kan51
No. 29	hao214	hu214
No. 30	mo35	ti35
No. 31	mu214	mo35
No. 32	bao51	pa51
No. 33	de55	ti214
No. 34	fo35	fang35
No. 35	fan35	ning214
No. 36	dong51	pao214
No. 37	fang35	ping35
No. 38	bing214	hu35
No. 39	ge214	gen35
No. 40	pin214	pan55
No. 41	hu214	he55
No. 42	tao51	ting51
No. 43	ni35	na214
No. 44	fang55	deng55
No. 45	mai35	mei35
No. 46	he35	ba35
No. 47	pa51	ku51
No. 48	nei214	nin35
No. 49	di214	ge214
No. 50	te51	ti214
No. 51	ba214	bi214
No. 52	gen35	ku214
No. 53	ni214	nan55
No. 54	tu214	pin214
No. 55	hang55	heng55
No. 56	kan214	bei51
No. 57	tu214	na35
No. 58	ba214	gu55
No. 59	fan214	fou214

No. 60	mao55	kai214
No. 61	tan35	ni35
No. 62	na51	fu214
No. 63	bao35	bing214
No. 64	dei214	dan51
No. 65	mei214	bang214
No. 66	pi51	pai51
No. 67	nan51	gen51
No. 68	ti55	ting55
No. 69	pu214	geng51
No. 70	kou51	ke214
No. 71	gei214	dao35
No. 72	nen214	ni214
No. 73	tong214	pin55
No. 74	mi35	meng214
No. 75	hou55	bu55
No. 76	men35	mo35
No. 77	bo55	bi214
No. 78	men51	fu51
No. 79	mo51	ma51
No. 80	fen214	pi35